

# REVISTA BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS

*Vol. 13, n. 3, 2019*

- INTER-INDUSTRY RELATIONS AND AGGLOMERATION ECONOMIES IN BRAZIL: AN APPROACH BASED ON FIRM CREATION** 309-339  
 Patrícia Araújo Amarante, Magno Vamberto Batista da Silva, Paulo Aguiar do Monte
- IMPACTOS DO CRÉDITO RURAL NO MERCADO DE TRABALHO DA AGROPECUÁRIA DOS ESTADOS BRASILEIROS** 340-356  
 João Paulo Rios e Silva, Elano Ferreira Arruda
- WIND POWER AND THE LABOR MARKET IN THE BRAZILIAN NORTHEAST: A SPATIAL PROPENSITY SCORE MATCHING APPROACH** 357-378  
 Thiago Pastorelli Rodrigues, Solange Ledi Gonçalves, André Squarize Chagas
- FLUXOS MIGRATÓRIOS SOB A PERSPECTIVA DA CENTRALIDADE DAS CIDADES: UMA INVESTIGAÇÃO PARA OS CHEFES DE FAMÍLIAS (2000-2010)** 379-404  
 Claubert Eduardo Marchezan Scherer, Alysson Lorenzon Portella
- ESTRUTURA FAMILIAR E PADRÕES ESPACIAIS DA ESCOLHA RESIDENCIAL** 405-421  
 Tatiana Kolodin Ferrari, Antônio Miguel Vieira Monteiro, Pedro Vasconcelos Maia do Amaral
- ANÁLISE ESPACIAL DAS AGLOMERAÇÕES DA INDÚSTRIA DE TRANSFORMAÇÃO NOS SEGMENTOS MODERNO E TRADICIONAL NO ESTADO DO PARANÁ** 422-442  
 Andréia Ferreira Prestes, Fernanda Mendes Bezerra, Talita Egevardt de Castro
- A EVOLUÇÃO DA INTERAÇÃO UNIVERSIDADE-EMPRESA NO RIO GRANDE DO SUL: O CASO DAS INDÚSTRIAS EXTRATIVAS E DE TRANSFORMAÇÃO NO PERÍODO 2008-2014** 443-464  
 Rodrigo Morem da Costa

# REVISTA BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS

ISSN 2447-7990

A **Revista Brasileira de Estudos Regionais e Urbanos (RBERU)** é uma publicação oficial da **Associação Brasileira de Estudos Regionais e Urbanos (ABER)**. A revista tem por objetivo divulgar trabalhos que representam contribuição científica para o campo do conhecimento sobre os fenômenos regionais e urbanos com ênfase para a realidade brasileira. A perspectiva multidisciplinar é valorizada pela revista e os artigos publicados podem representar contribuições teóricas, aplicadas ou de política.

Os temas de interesse da revista incluem, de forma não necessariamente exclusiva: teoria e metodologia em economia regional e urbana, localização e concentração das atividades econômicas, análise estatística espacial, desenvolvimento regional e crescimento econômico, competitividade regional, infraestrutura regional e urbana, estudos urbanos e metropolitanos, meio ambiente e recursos naturais, turismo e indústrias culturais, desenvolvimento urbano e uso da terra, empreendedorismo e inovação, políticas sociais, pobreza, migração e mercado de trabalho regional e finanças públicas locais.

A revista é publicada pela Associação Brasileira de Estudos Regionais e Urbanos (ABER) com objetivo de divulgar estudos socioeconômicos com destaque para as temáticas regional e urbana.

## EDITORES

Alexandre Alves Porsse, UFPR.  
Weslem Rodrigues Faria, UFJF.

## EDITOR ASSISTENTE

Inácio Fernandes de Araújo Junior, USP.

## CONSELHO EDITORIAL

Alexandre Rands, UFPE.  
Carlos Roberto Azzoni, USP.  
Cássio Rolim, UFPR.  
Eduardo Haddad, USP.  
Fernando Salgueiro Perobelli, UFJF.  
Geoffrey J. D. Hewings, UIUC.  
Jan Oosterhaven, University of Groningen.  
João Francisco de Abreu, PUC-MG.  
Joaquim José Martins Guilhoto, USP.  
Juan R. Cuadrado-Roura, Universidad de Alcalá.  
Marcos Holanda, UFCE.  
Mauro Borges, UFMG.  
Michael L. Lahr, Rutgers.  
Patricio Aroca, Universidad Adolfo Ibáñez.  
Raul da Mota Silveira Neto, UFPE.  
Roberta Capello, Politecnico di Milano.  
Sabino da Silva Porto Júnior, UFRGS.  
Tatiane Almeida de Menezes, UFPE.  
Werner Baer - *In memoriam*.

**INTER-INDUSTRY RELATIONS AND AGGLOMERATION ECONOMIES IN BRAZIL:  
AN APPROACH BASED ON FIRM CREATION\***

**Patrícia Araújo Amarante**

Professora Adjunta do Departamento de Ciências Sociais Aplicadas da Universidade Federal da Paraíba (UFPB),  
Campus III.

E-mail: patyamarante@yahoo.com.br

**Magno Vamberto Batista da Silva**

Professor Associado do Departamento de Economia da Universidade Federal da Paraíba (UFPB), Campus I.

E-mail: magnovbs@gmail.com

**Paulo Aguiar do Monte**

Professor Associado do Departamento de Economia da Universidade Federal da Paraíba (UFPB), Campus I.

E-mail: pauloaguiardomonte@gmail.com

**ABSTRACT:** This study analyzes the effects of the agglomeration mechanisms proposed by Marshall (1920) on the intensity of the economies of localization and urbanization in the manufacturing industry, based on the creation of companies in Brazilian municipalities between 2011 and 2013. For this purpose, the following two-stage procedure was conducted: 1) estimation of regression models for count data to identify the type of agglomeration economy that contributes to new firm location and 2) use of the coefficients obtained in the first stage to assess the industry characteristics related to Marshall's agglomeration mechanisms that may contribute to differences in the strength of those economies. Due to the results obtained, localization economies are more intense in the industries that employ workers with specific skills of the sector, and thus, can share a pooling of skilled workers. Similarly, knowledge spillovers act in a way to boost the economies of localization, providing an environment conducive to the transmission of knowledge and new ideas. However, localization economies are less intense in industries that have a greater dependence on manufactured inputs and the primary and energy sectors. Urbanization economies, in turn, are negatively related to labor market pooling and knowledge spillovers. On the other hand, the effects of input sharing and dependence on primary inputs on urbanization economies are positive and statistically significant.

**Keywords:** Firm creation; Agglomeration mechanisms; Localization economies; Urbanization economies.

**JEL Codes:** L60; R12; R30.

**RELAÇÕES INTERINDUSTRIAIS E AS ECONOMIAS DE AGLOMERAÇÃO NO  
BRASIL: UMA ABORDAGEM A PARTIR DO NASCIMENTO DE FIRMAS**

**Resumo.** Este estudo analisa os efeitos dos mecanismos de aglomeração propostos por Marshall (1920) sobre a intensidade das economias de localização e urbanização na indústria de transformação, com base na criação de empresas em municípios brasileiros entre 2011 e 2013. Para esse fim, o seguinte procedimento em duas etapas foi realizado: 1) estimativa de modelos de regressão para dados de contagem para identificar o tipo de economia de aglomeração que contribui para a nova localização da empresa e 2) uso dos coeficientes obtidos no primeiro estágio para identificar as características setoriais relacionadas aos mecanismos de aglomeração de Marshall que possam contribuir para as diferenças na intensidade dessas economias. Conforme os resultados obtidos, as economias de localização são mais intensas nas indústrias que empregam trabalhadores com habilidades específicas do setor e, portanto, podem compartilhar um conjunto de trabalhadores qualificados. Da mesma forma, os *spillovers* de conhecimento atuam de maneira a impulsionar as economias de localização, proporcionando um ambiente propício à transmissão de conhecimento e novas ideias. No entanto, as economias de localização são menos intensas nas indústrias que dependem mais dos insumos manufaturados e dos setores primário e energético. As economias de urbanização, por sua vez, estão negativamente relacionadas com o *pooling* do mercado de trabalho e os *spillovers* de conhecimento. Por outro lado, os efeitos do compartilhamento e da dependência de insumos primários nas economias de urbanização são positivos e estatisticamente significativos.

**Palavras-Chave:** Nascimento de firmas; Mecanismos de aglomeração; Economias de localização; Economias de urbanização.

**Classificação JEL:** L60; R12; R30.

## 1. Introduction

One of the issues addressed in the literature regarding agglomeration economies is the possibility for companies and workers to take advantage of positive externalities derived from the geographic concentration of economic activities, predominantly in large urban centers, wherein they find an environment that is more diversified and prone to knowledge -sharing -prone environment (GLAESER; KERR, 2009). In this line of research, the following two types of agglomeration economies are analyzed: localization economies (MARSHALL, 1920), which derive from the benefits generated by the presence of same-industry firms in a specific region, and urbanization economies (JACOBS, 1969), wherein economic diversity benefits firms, particularly in large cities (JOFRE-MONSENY et al., 2011).

Among the methods most commonly used to quantify the effects of agglomeration economies, the research strategy based on firm creation is an underused approach, especially because firm location choice involves many determinants that are often difficult to measure, such as the technological level of firms and worker efficiency (COMBES; GOBILLON, 2014). Those studies hypothesize that entrepreneurs locate their firms based on a profit maximization analysis. In this case, the existence of evidence for agglomeration economies, which lead to productivity gains, favors concentration of firms in specific geographical locations (ROSENTHAL; STRANGE, 2004).

Dumais et al. (2002), for example, analyze the process of geographical concentration of manufacturing firms in the United States based on a dynamic approach that decomposes aggregate concentration changes into portions attributable to plant creations, expansions, contractions and closures. Conversely et al. (2003) address the geographical and organizational nature of agglomeration economies in the United States by analyzing firm creation and their employment levels.

Despite the need to assess the presence of agglomeration economies, theoretical and empirical studies have more recently focused on identifying the determinants of this phenomenon. Thus, in line with the precepts of Marshall (1920), who identifies one of the determinants of agglomeration of firms and people in external economies, this study aims to associate agglomeration economies with the following factors that affect the geographical distribution of economic activities: 1) the formation of clusters of specialized workers and the production of new ideas (labor market pooling); 2) the physical conditions of the region, such as climate and soil types, and the availability of and ease of access to production inputs and consumer markets (input sharing); and 3) the existence of knowledge spillovers.

In this line of research, studies addressing theoretical and empirical aspects have been published in the economic literature (ROSENTHAL; STRANGE, 2001; GLAESER; KERR, 2009; ELLISON et al., 2010; JOFRE-MONSENY et al., 2011; JOFRE-MONSENY et al., 2014). In the Brazilian literature, studies regarding agglomeration economies have predominantly addressed industrial employment growth and effects on wages. However, no Brazilian studies assessing those mechanisms based on firm creation have been published thus far. From the standpoint of companies, Brazilian studies have focused on industry clustering and geographical location, albeit without assessing the importance of agglomeration economies for this process. Garcia (2001), Britto and Albuquerque (2002), Rezende et al. (2012) and Sobrinho and Azzoni (2015) are examples of such studies.

In this context, the study addresses the location of new manufacturing firms in Brazilian municipalities, created during the period from 2011 to 2013, as a function of the agglomeration mechanisms proposed by Marshall (1920): labor market pooling, customer-supplier relationship and knowledge spillovers. For this purpose, a two-stage estimation procedure is conducted, following the method of Jofre-Monseny et al. (2014). First, Poisson and negative binomial regression models are estimated for count data to identify, for each sector, the type of agglomeration economy that contributes to new firm locations. Then, the industry characteristics related to Marshall's agglomeration mechanisms (labor market pooling, customer-supplier relationship and knowledge spillovers) that determine differences in the incidence of those economies are assessed using the

coefficients obtained in the first stage and the generalized method of moments (GMM) with endogeneity control.

Accordingly, the present study improves on that of Jofre-Monseny et al. (2014), which has limitations, particularly regarding the empirical analysis method. The study mentioned used the weighted least squares method (WLSM) to analyze the effect of Marshall's agglomeration mechanisms on agglomeration economies, disregarding possible endogeneity problems. Diversely, the analysis performed in this study allows assessing the relative importance of each of Marshall's agglomeration mechanisms, considering its incidence in different geographic scales using sample exclusion criteria. Furthermore, we use a 100% formal database, the *Relação Anual de Informações Sociais Identificada - RAIS identificada* (identified RAIS - Annual Social Information Report), with no limit to the size of municipalities.

Thus, the present study contributes to the empirical literature regarding agglomeration mechanisms shaping the geography of economic activities by adopting an underused strategy, especially in Brazil: assessing agglomeration economies based on firm creation. In the analysis of agglomeration economies, the basic hypothesis of this approach takes into account the choice of the place where the entrepreneurs install their plants as a function of a profit maximization analysis. If agglomeration economies are present, they generate productivity gains that may induce a disproportionate concentration of firms at specific points in space. Thus, the location and spatial concentration of new firms become potential sources of investigation of agglomeration economies. In addition, the studies of agglomeration economies through the birth of firms do without data on capital, which is subject to high imprecision and whose access is notoriously difficult. Another quality is that new establishments are not restricted by previous decisions. They decide where to locate and how much labor to hire based on the current economic environment, which is exogenously given (ROSENTHAL; STRANGE, 2004). Another novel aspect of this study is the association of agglomeration economies with determinants of the geographical distribution of economic activities according to Marshall's agglomeration mechanisms. Furthermore, localization and urbanization agglomeration economies are initially measured separately for the different sectors of the manufacturing industry to allow a sector-level analysis of variability in those effects.

The results indicate that localization economies are more intense in the industries employing workers with specific sectoral skills. Similarly, knowledge spillovers act in a way to boost the localization economies, providing an environment conducive to the transmission of knowledge and new ideas. On the other hand, localization economies are less intense in industries that have a greater dependence on manufactured inputs and the primary and energy supply sectors. Urbanization economies, in turn, are negatively related to labor market pooling and knowledge spillovers. However, the effects of input sharing and dependence on primary inputs on urbanization economies are positive and statistically significant.

In addition to this introduction, the study is organized as follows: Section 2 describes the empirical strategy, Section 3 presents the database and the treatments performed, Section 4 describes how the variables representative of agglomeration mechanisms were measured, Section 5 discusses the empirical results, and Section 6 presents the final considerations.

## 2. Empirical strategy

The central idea of the present study is to analyze the relative importance of each of Marshall's agglomeration mechanisms – labor market pooling, customer-supplier relationship and knowledge spillovers – by analyzing new firm locations in Brazilian manufacturing industries. For this purpose, based on the method proposed by Jofre-Monseny et al. (2014), this study initially assesses, for each industry, the possible effects of (localization and urbanization) agglomeration economies on the explanation for firm creation. Subsequently, based on the coefficients obtained in those estimations, Marshall's agglomeration mechanisms that may explain differences in the incidence of those economies are analyzed.

However, studying agglomeration economies based on firm creation has advantages and limitations. One of the main advantages is the fact that this type of analysis usually requires no data regarding capital (for example, capital, labor, equipment and land stock), to which access is notoriously difficult. Moreover, new firms are not limited by previous decisions. Thus, their location choices and decisions about how much labor to hire, based on the current economic environment, are determined exogenously.

The following are the main disadvantages: 1) the possibility of an absence of firm creation in specific locations in the period and sectors analyzed, which may limit the choice of econometric methods; and 2) the tendency of new firms to locate in areas with a preexisting concentration of industrial activities. These facts limit the choice of more restrictive econometric methods because they cause the presence of censored observations in the sample (ROSENTHAL; STRANGE, 2004; COMBES; GOBILLON, 2014). Rosenthal and Strange (2003), for example, analyze the geographical scope of agglomeration economies by concurrently measuring firm creation and associated job levels using tobit and probit models. More recently, Jofre-Monseny et al. (2011, 2014) include count data in the analysis of new firm creation using Poisson regression models.

The firm creation process can be expressed using the random profit function approach proposed by Carlton (1983), assuming the following linear expected profit function:

$$\pi_{kic} = \beta_{loc}^i \cdot emp_{ic} + \beta_{urb}^i \cdot emp_{-ic} + \chi'_{ic} \gamma' + \varepsilon_{kic} \quad (1)$$

where  $\pi_{kic}$  denotes the profit of firm  $k$  in industry  $i$  and municipality  $c$ . Profit is determined by the following variables: 1) localization economies, which are measured as the logarithm of the employment level in industry  $i$  of municipality  $c$  ( $emp_{ic}$ ); 2) urbanization economies, which are measured as the logarithm of the overall employment level in municipality  $c$  minus the employment in industry  $i$  ( $emp_{-ic}$ ); 3)  $\chi'_{ic}$ , the vector with the other control variables; and 4)  $\varepsilon_{kic}$ , the random error term.

In practice,  $\pi_{kic}$  cannot be observed. However, a conditional logit model can be constructed by assuming a Weibull distribution for  $\varepsilon_{kic}$  (DUMAIS et al, 2002). Therefore, the probability that firm  $k$  is located in the geographical unit  $c$  has the following conditional logit form:

$$\Pr(\text{firm } k \text{ is located in } c) = \frac{\exp(\beta_{loc}^i \cdot emp_{ic} + \beta_{urb}^i \cdot emp_{-ic} + \chi'_{ic} \gamma^i)}{\sum_c (\beta_{loc}^i \cdot emp_{ic} + \beta_{urb}^i \cdot emp_{-ic} + \chi'_{ic} \gamma^i)} \quad (2)$$

However, Guimarães et al. (2003) warn that using the conditional logit method to model firm location decisions brings problems when working with complex choice scenarios with a large number of spatial alternatives. In this case, the conditional logit model can be estimated regardless of the number of choices, based on an equivalence relation between the conditional logit likelihood function and the Poisson regression, as follows:

$$E(N_{ic}) = \exp(\beta_{loc}^i \cdot emp_{ic} + \beta_{urb}^i \cdot emp_{-ic} + \chi'_{ic} \gamma') \quad (3)$$

where the dependent variable ( $N_{ic}$ ) is the number of new firms in industry  $i$  and municipality  $c$ .

The Poisson distribution is appropriate when the dependent variable has only non-negative integer values and can be used to model the number of occurrences of an event (CAMERON; TRIVEDI, 2009). The Poisson distribution for a specific observation has the following probability of occurrence of a count  $y$  in a given period, area or region:

$$\Pr [Y = y] = e^{-\mu} \mu^y / y!, \quad y = 0, 1, 2, \dots \quad (4)$$

where  $\mu$  is the expected number of occurrences or the estimated mean incidence rate of a specific phenomenon for a given exposure and  $y$  is a non-negative number that represents the number of occurrences of the dependent variable.

Because there is only one parameter in the Poisson distribution, the mean and variance are equal, that is,  $E(Y) = \mu$  and  $Var(Y) = \mu^1$ . This property is known as Poisson equidispersion.

Although the Poisson regression model is an alternative to count data estimation, the restrictive data dispersion assumption (the Poisson equidispersion assumption) is not usually satisfied because the variance of count data tends to exceed the mean<sup>2</sup>. In this case, for this type of data, the variance of the dependent variable usually exceeds its mean. This characteristic is termed overdispersion, and it may be attributed to the presence of undetected data heterogeneity. Disregarding the existence of overdispersion when estimating a Poisson model may generate biased parameters (CAMERON; TRIVEDI, 2005).

Alternatively, the negative binomial model may be used to address the overdispersion problem. This model incorporates a random component ( $\alpha$ ), with a Gamma distribution, into the Poisson model, treating the overdispersion process as a result of undetected heterogeneity (WOOLDRIDGE, 2002). In this case, the density function of the negative binomial distribution is as follows:

$$h[y|\mu, \delta] = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(\alpha^{-1} + y)} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left( \frac{\mu}{\mu + \alpha^{-1}} \right)^y \quad (5)$$

where  $\alpha$  is equal to  $1/\delta$  and represents the degree of data scattering and  $\Gamma$  indicates a standard Gamma function. The negative binomial model is non-linear and commonly estimated using the maximum likelihood estimation method.

Thus, the possible effects of agglomeration economies on new firm creation were estimated using the Poisson and negative binomial models to enable assessing the sensitivity of the results considering different count models.

A goodness-of-fit chi-square test was performed to assess the fit of the Poisson regression model (CAMERON; TRIVEDI, 2009). A non-significant test indicates that the model has a good fit to the Poisson distribution. Conversely, a significant test strongly suggests the existence of a data scattering problem, and the negative binomial model should be used in this case. Moreover, as previously mentioned, the second research stage aims to identify the industry characteristics that may explain inter-industry differences in the importance of localization and urbanization economies based on Marshall's agglomeration theory. For this purpose, the following estimated equation was used:

$$\hat{\beta}_{Loc\ ou\ Urb} = \alpha + \beta_i X_i + \varphi_i \gamma_i + \varepsilon_i \quad (6)$$

where  $X_i$  is a vector of explanatory variables that measure Marshall's agglomeration factors, that is, labor market pooling, knowledge spillovers and input sharing;  $\gamma_i$  is a set of control variables; and  $\varepsilon_i$  is a random error term. This equation was estimated for the following dependent variables: the coefficients of the localization ( $\hat{\beta}_{loc}^i$ ) and urbanization ( $\hat{\beta}_{urb}^i$ ) economies estimated using Equation (3).

However, as noted by Combes and Gobillon (2014), virtually all local variables explaining firm location choices may be considered endogenous, which remains overlooked in empirical research studies on this subject. In Ellison et al. (2010), for example, instrumental variables representative of the measures of inter-industry relations are constructed based on United Kingdom data and are then used to define the corresponding variables for the United States. However, the authors concede that

<sup>1</sup> The standard mean parameterization is  $\mu = \exp(x'\beta)$  to ensure that  $\mu > 0$ . This implies that based on the property of equidispersion of the Poisson distribution, the model is intrinsically heteroscedastic (CAMERON; TRIVEDI, 2009).

<sup>2</sup> Overdispersion has consequences that are qualitatively similar to those of failure of the homoscedasticity property in linear regression models (CAMERON; TRIVEDI, 2005).

such instruments will only reduce the simultaneity bias if the agglomeration mechanisms similarly affect firm location choice in the United Kingdom and in the United States, which is unlikely.

In this context, the generalized method of moments (GMM) was used to estimate those parameters to correct possible endogeneity problems between the coefficients of localization and urbanization economies, the dependent variables of the model, and the Marshall's agglomeration mechanisms, i.e., the explanatory variables. The values of those explanatory variables lagged by one or two periods ( $t - 1, t - 2$ ) were used as measuring instruments of Marshall's agglomeration mechanisms. As highlighted by Wooldridge (2002), GMM is used to control for endogeneity and belongs to a class of estimators that combine the main methods of generalized estimation of the parameters of linear and nonlinear models, thereby overcoming autocorrelation, heteroscedasticity and multicollinearity problems.

Notwithstanding the aforementioned advantages of using the GMM, it should be noted that the efficiency of the GMM is subject to the following set of constraints: 1) the included instrumental variables must be correlated with the endogenous variables and 2) the instrumental variables must be orthogonal to the random error term. In this context, the endogeneity test using the C statistic, or "difference-in-Sargan" statistic, and the Sargan-Hansen robustness test (1982) were used to assess the validity of the instrumental variables (CAMERON; TRIVEDI, 2009).

### 3. Database and description of variables

The database used was constructed from secondary data extracted from the *Relação Anual de Informações Sociais Identificada - RAIS identificada* (Annual Social Information Report - identified RAIS), published by the Ministry of Labor and Employment (Ministério do Trabalho e Emprego), in 2011, 2012 and 2013. Identified RAIS enables monitoring firms across years using the National Registry of Legal Entities (Cadastro Nacional da Pessoa Jurídica – CNPJ) and provides key data for the present study, including opening and closing dates, location, labor stock and activity sector.

In relation to the timeframe used, specifically the final year is justified by the availability of information, being the most recent year at the time of writing the article. Thus, since a three-year period is the minimum time span necessary for the signature counting process and, as such, has been defined for the counting of new forms, 2011 is the initial year. In addition, although the data for the calculation of the creation of firms are available for several years, the information necessary for calculations of other variables, such as the input-output matrix, are not available for every year.

Some exclusion criteria were used to reduce the sample size. Initially, companies with more than one subsidiary were removed from the sample, thereby excluding firm creation resulting from the expansion of preexisting companies. Firms with statutory employment relationships were also removed, thereby excluding government-owned companies from the analysis. Also, only new firms active in the reference year were kept in the sample.

Because the second stage of the analysis involved constructing variables representative of agglomeration mechanisms and input sharing between them, it required performing a matching<sup>3</sup> procedure between the industry sectors of the National Classification of Economic Activities (Classificação Nacional de Atividades Econômicas – CNAE) and the activities of the input-output matrix, according to the correspondence between the CNAE 2.0 Groups and the list of activities provided by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE). This matching procedure resulted in a total of 30 industrial sectors, as outlined in Table A1 in the Appendix.

To estimate parameters indicative of localization and urbanization economies, the dependent variable is defined as the number of firms created in 2011, 2012 and 2013 by industry sector and location. The definition of new firms was based on the identified RAIS, which allows monitoring companies over the years using their CNPJ number, including opening and closing dates, location,

<sup>3</sup> The compatibilization was necessary for the calculation of the input sharing variable, since the activities defined in the input-output matrix are more disaggregated in relation to the division presented in the CNAE for the manufacturing industry.



work stock and activity sector. The variable “new firms” is defined as the number (count) of firms created and active in the reference years, by industry and location, and filtered by the exclusion criteria mentioned above.

The data collected from the RAIS database refer to the manufacturing industry at the group level, that is, a three-digit code, according to the CNAE 2.0. To avoid simultaneity, the explanatory variables are measured using 2010 data. Furthermore, all explanatory variables are expressed as logarithms.

As shown in Equation 3, the vector of explanatory variables consists of three dimensions: 1) localization economies, which are measured as the logarithm of the employment level in industry  $i$  and municipality  $c$ ; 2) urbanization economies, which are measured as the logarithm of the overall employment level in municipality  $c$  minus the employment level in industry  $i$ ; and 3) a set of control variables for geographical or location characteristics.

The control variables include the logarithm of the municipality area and regional dummies (North, Northeast, Southeast, South and Central-West)<sup>4</sup>. For some industry-municipality pairs, the number of employees is zero, which precludes logarithmic transformation of the variables. To circumvent this problem, the method proposed by Crépon and Duguet (1997) was adopted. In this approach, a dummy variable with a value of 1 is created if the employment of a specific industry-municipality pair is zero, and it is assigned a value of 0 otherwise. The dummy variable is then added to the employment level, and subsequently, this sum is log-transformed. In the estimates, the dummy variable created is included as a regressor.

Equation 3 is estimated separately for the 30 industry sectors analyzed, thereby obtaining estimates of the coefficients of localization ( $\hat{\beta}_{loc}^i$ ) and urbanization ( $\hat{\beta}_{urb}^i$ ) economies for each sector. After estimating the effect of localization and urbanization economies on new firm location, the next stage is to assess how some industry characteristics affect the patterns of localization and urbanization economies in the geographical space of new firms.

For this purpose, the estimates of coefficients of localization ( $\hat{\beta}_{loc}^i$ ) and urbanization ( $\hat{\beta}_{urb}^i$ ) economies are regressed on the variables that measure the characteristics of the industry. Specifically, the hypothesis that the following industries have strong localization and/or urbanization economies is tested: 1) industries that employ workers with industry-specific skills (labor market pooling); 2) industries that extensively use manufactured inputs (input sharing); and 3) industries that extensively use knowledge (knowledge spillovers). This method makes it possible to identify which of these three industry characteristics are more important to explain why firms prefer to locate where the density of their industry sector is high (localization economies) or where more diversified industry environments are found (urbanization economies) (JOFRE-MONSENY et al., 2014).

#### 4. Measurement of agglomeration mechanisms

Variables representative of industry characteristics describe Marshall’s agglomeration mechanisms and are calculated as follows:

Labor market pooling (*pooling*): This variable is representative of the advantages to firms and workers obtained by locating in concentrated markets. Thus, if labor market pooling is a relevant agglomeration theory, the incentive for companies and workers of the same industry sector to locate near each other, given the existence of workers with sector-specific skills, will be stronger (Rosenthal & Strange, 2001).

Similarly to Bacolo et al. (2010), Ellison et al. (2010), Jofre-Monseny et al. (2014), the proxy variable for labor market pooling was calculated using occupation to represent the skills of workers. For this purpose, the 2002 Brazilian Classification of Occupations (Classificação Brasileira de

<sup>4</sup> The inclusion of regional dummies is important since the methodology used in the Jofre-Monseny et al. (2014) study was initially applied to Spain, a country quite different from that analyzed in this article, Brazil. Unlike Spain, Brazil is a continental country with great heterogeneity, which can lead to different impacts of the economies of urbanization and location in each region.

Ocupações – CBO; CBO 2002) of the IBGE, which is available in the RAIS statistical records, was used at the subgroup level, with 192 occupations, considering its distribution by labor stock. According to the IBGE, the CBO describes and ranks the occupations within a hierarchical structure, thereby allowing aggregation of information about the workforce based on occupational characteristics, such as functions, tasks and obligations, and on the content of the work, that is, knowledge, skills and personal attributes, among others.

Thus, the objective is to create a proxy that measures the specificity of workers in each industry in terms of skills. This was achieved by calculating a measure of occupational dissimilarity between this industry sector and the rest of the economy, the index of dissimilarity. Also known as the Duncan Segregation Index, because it was proposed by Duncan and Duncan (1955), this index compares the occupational structure of each industry with the rest of the economy as follows:

$$Pooling_{ic} = \frac{1}{2} \sum_o \left| \frac{L_{oi}}{L_i} - \frac{L_{o-i}}{L - i} \right| \quad (7)$$

where  $o$  is the occupation index;  $L_{oi}$  is the number of workers in occupation  $o$  of industry  $i$ ;  $L_i$  is the number of workers in industry  $i$ ;  $L_{o-i}$  is the number of workers in occupation  $o$  minus the total workforce of industry  $i$ ; and  $L - i$  is total workforce of the economy minus the workforce of industry  $i$ . This variable ranges from 0 to 1 and may be interpreted as the fraction of workers in industry  $i$  that must change occupation to simulate the occupational structure of the economy. Thus, the more specialized the occupational structure is, the closer to 1 the values of this variable will be, that is, the more specific the industry will be in terms of worker skills (JOFRE-MONSENY et al., 2014).

Knowledge spillovers (*spillovers*): This variable refers to the ease with which companies and workers located nearby establish informal relations, thereby enabling the exchange of knowledge and ideas relevant to the improvement of industrial products and processes (Marshall, 1920). The present study is based on the hypothesis that knowledge spillovers are more important in knowledge-intensive industries, that is, in industries with greater human capital needs (JOFRE-MONSENY et al., 2014).

To test this agglomeration mechanism, two complementary lines of research can be followed: 1) the relationship between city growth and innovation (DURANTON; PUGA, 2001; JAFFE et al., 1993; CARLINO et al., 2007); and 2) growth and skills in cities (GLAESER; RESENGER, 2010). However, these strategies are difficult to measure, especially regarding the calculation of variables representative of innovations and skills. Therefore, the knowledge spillovers variable is measured as the number of workers with a university degree employed in the industry. The variable was constructed based on the distribution of workers by level of education, which was available in the RAIS and measured as follows:

$$spillovers_{ic} = \frac{\text{Total number of college graduates}_i}{\text{Total number of workers}_i} \quad (8)$$

The hypothesis that same-sector industries co-locate to benefit from knowledge spillovers is consistent with the idea that location effects are more important in industries with high fractions of skilled workers. Similarly, high levels of urbanization economies in knowledge-intensive industries suggest that companies co-locate in large urban areas to benefit from some form of tacit knowledge of workers (JOFRE-MONSENY et al., 2014).

Input sharing (*Input Sharing*): This variable refers to the advantages of firm concentration associated with sharing a larger base of suppliers and, simultaneously, customers. According to this premise, firms will be more efficient when located near input suppliers because they will be able to minimize their transportation costs. Furthermore, the greater the dependence on inputs is, the greater the incentive for these firms to locate in areas with a high firm concentration. The 2010 Brazilian System of Input-Output Matrices, with 68 industry sectors or activities and 128 products, was used to quantify this variable.

This matrix was constructed based on the National Accounts published by the IBGE following the method proposed by Guilhoto and Sesso Filho (2005; 2010). As previously mentioned, this required performing a matching procedure between the CNAE industry sectors and the input-output matrix activities based on the match between the CNAE 2.0 Groups of the IBGE. Furthermore, this matrix was chosen because it includes more updated data compatible with the study period.

The variable measuring the relative importance of manufactured inputs for the industry was constructed following Rosenthal and Strange (2001) and Jofre-Monseny et al. (2014), dividing the purchased manufactured inputs by the total number of sales of each industry sector analyzed:

$$\text{Input Sharing}_{ic} = \frac{\text{Total number of Inputs}_i}{\text{Total number of Sales}_i} \quad (9)$$

Thus, if industries with evidence of localization economies use manufactured inputs intensively, one of the reasons for these industries to locate in environments that are economically more specialized is easier sharing of input suppliers. Conversely, if industries with strong urbanization economies tend to use manufactured inputs intensively, input sharing will be a reason for these firms to locate in large cities (JOFRE-MONSENY et al., 2014).

In addition to variables representative of Marshall's agglomeration mechanisms, natural advantage was used as a control variable. Natural advantage is a set of factors that affect the location of economic activities, such as the abundance of natural resources and cost of energy (ELLISON; GLAESER, 1999). Intuitively, if a specific location can provide easier access to natural resources and energy, industries that use these factors of production more intensively will tend to concentrate at that location.

Using again the 2010 Brazilian System of Input-Output Matrices, the natural advantage variable is defined as the sum of inputs purchased from the primary and energy sectors divided by the total sales of each industry sector analyzed. To calculate this variable, the following input-output matrix sectors are analyzed: agriculture, including agricultural and postharvest support; animal husbandry, including animal husbandry support; wood industry; fisheries and aquaculture; coal and nonmetallic mining; oil and natural gas extraction, including support activities; iron ore mining, including beneficiation and agglomeration; and non-ferrous mining, including beneficiation. The electricity, natural gas and other utilities sector was considered as the power supply sector.

A set of binary variables characteristic of the location, large regions, sector, firm size and technological intensity was also used. The technological intensity variable was calculated according to the classification of the Organization for Economic Co-operation and Development (OECD), which groups the manufacturing industry sectors according to their technological intensity: high, medium-high, medium-low and low. Thus, the variable has a value of 0, 1, 2 or 3 if the industry has low, medium-low, medium-high and high technological intensity, respectively. Table A2 in Appendix outlines the classification according to the technological intensity of the sectors analyzed.

## 5. Results

### 5.1 Descriptive analysis of the sample

The sample used is formed by 106,395 firms (creations) between 2011 and 2013, distributed by sector, as outlined in Table 1. Table 1 presents the maximum number of firms created in each sector, considering their geographical location and the percentage of municipalities with zero creations. In the industries of biofuel production and oil refining and coking plants, for example, no new firms were created in approximately 99% of Brazilian municipalities during the study period.

**Table 1 - Brazil: Firms created in the manufacturing industry, 2011-2013**

Sectors	New Firms	New Firms (%)	Maximum	Municipalities with zero firm creations (%)
1. Meat packing and production of meat products, including dairy and fishery products	2,092	1.97	38 (Brasília-DF)	78.53
2. Sugar manufacture and refining	54	0.05	4 (Catanduva-SP)	99.16
3. Other food products	10,447	9.82	355 (São Paulo-SP)	52.45
4. Manufacture of beverages	422	0.40	33 (Brasília -DF)	94.47
5. Manufacture of tobacco products	44	0.04	8 (Santa Cruz do Sul-RS)	99.46
6. Manufacture of textiles	3,376	3.17	265 (São Paulo-SP)	83.23
7. Manufacture of apparel and accessories	23,466	22.06	3933 (São Paulo-SP)	58.62
8. Manufacture of footwear and leather goods	4,114	3.87	547 (Franca-SP)	87.40
9. Manufacture of wood products	3,473	3.26	64 (São Paulo-SP)	74.52
10. Manufacture of cellulose, paper and paper products	898	0.84	100 (São Paulo-SP)	92.15
11. Printing and copying	4,776	4.49	853 (São Paulo-SP)	79.89
12. Oil refining and coking plants	32	0.03	2 (Cascavel-PR, Fortaleza-CE, Diadema-SP and Araucária-PR)	99.46
13. Biofuel production	47	0.04	3 (Jataí-GO)	99.26
14. Production of other organic and inorganic chemicals, resins and elastomers	339	0.32	15 (São Paulo-SP)	95.99
15. Manufacture of pesticides, disinfectants, paints and various chemicals	513	0.48	18 (São Paulo-SP)	94.25
16. Manufacture of cleaning products, cosmetics/perfumes and toilet preparations	553	0.52	51 (São Paulo-SP)	93.84
17. Manufacture of pharmaco-chemical and pharmaceutical products	79	0.07	10 (São Paulo-SP)	98.94
18. Manufacture of rubber and plastic products	2,228	2.09	192 (São Paulo-SP)	86.34
19. Manufacture of non-metallic mineral products	6,801	6.39	125 (Cachoeiro de Itapemirim-ES)	60.93
20. Production of pig iron/ ferrous alloys, steel and seamless steel tubes	148	0.14	17 (São Paulo-SP)	97.97
21. Metallurgy of nonferrous metals and metal smelting industry	376	0.35	26 (São Paulo-SP)	95.87
22. Manufacture of metal products, excluding machinery and equipment	12,676	11.91	556 (São Paulo-SP)	57.25
23. Manufacture of computers, electronics and optical products	628	0.59	77 (São Paulo-SP)	95.90
24. Manufacture of electrical machinery and equipment	679	0.64	98 (São Paulo-SP)	94.90
25. Manufacture of machinery and mechanical equipment	2,172	2.04	142 (São Paulo-SP)	87.78
26. Manufacture of cars, trucks and buses, excluding parts	399	0.38	20 (Goiânia-GO)	95.42
27. Manufacture of parts and accessories for motor vehicles	736	0.69	80 (São Paulo-SP)	93.84
28. Manufacture of other transport equipment, excluding motor vehicles	270	0.25	25 (Manaus-AM)	96.86
29. Manufacture of furniture and products of various industries	11,695	10.99	635 (São Paulo-SP)	61.11
30. Maintenance, repair and installation of machinery and equipment	12,862	12.09	1463 (São Paulo-SP)	69.79

Source: Authors, based on RAIS data.

The data also indicate that new firms were predominantly created in the following sectors: manufacture of apparel and accessories (23,466); maintenance, repair and installation of machinery and equipment (12,862); manufacture of metal products, excluding machinery and equipment (12,676); manufacture of furniture and products of various industries (11,695); and other food products (10,447).

Conversely, the following sectors had the lowest number of firm creations: oil refining and coking plants (32); manufacture of tobacco products (44); biofuel production (47), sugar manufacture and refining (54); and manufacture of pharmaco-chemical and pharmaceutical products (79). The geographical concentration of new firm creation is noteworthy. The municipality of São Paulo, for example, has the highest number of creations in virtually all sectors.

Table 2 presents descriptive statistics for the variables used to measure the role of agglomeration mechanisms in explaining the sources of localization and urbanization economies. The data indicate that the highest percentages of firms created are found in the Southeast and South regions, with 45.25% and 27.74%, respectively, and that 97.60% of industries have from 0 to 19 workers and are predominantly located in large cities, with populations greater than 100,000 inhabitants.

**Table 2 - Brazil: Descriptive statistics from the database: mean, minimum and maximum values (2011-2013)**

Variables	Mean	Minimum Value	Maximum Value
<b>Agglomeration mechanisms</b>			
Pooling	0.61	0.49	0.77
Spillovers	0.05	0.01	0.31
Input Sharing	0.28	0.09	0.38
<b>Natural factors of agglomeration</b>			
Natural Advantage	0.03	0.00	0.22
<b>Firm characteristics</b>			
From 0 to 19 workers*	97.60	0	1
From 20 to 99 workers*	2.02	0	1
From 100 to 499 workers*	0.31	0	1
500 or more workers*	0.08	0	1
Technological Intensity	0.46	0	3
<b>Location characteristics</b>			
North*	3.03	0	1
Northeast*	15.23	0	1
Southeast*	45.25	0	1
South*	27.74	0	1
Central-West*	8.75	0	1
Up to 5.000*	34.27	0	1
From 50,001 to 100,000*	13.96	0	1
From 100,001 to 500,000*	30.65	0	1
More than 500.000*	27.52	0	1
Observations	106,39		

Source: Authors, based on RAIS data.

Note: \*Represents a binary variable. For binary variables, the mean should be interpreted as the ratio.

Labor market pooling reflects the possible advantages for companies and workers because they are located in regions with dense labor markets. In this case, if this is a relevant agglomeration mechanism, industries employing similar workers should be concentrated, thereby benefiting from

knowledge transfers generated by worker mobility between industries. The closer to one labor market pooling is, the stronger the pooling effects on the generation of agglomeration economies will be. The mean labor market pooling is 0.61, and manufacture of footwear and leather goods is the industry sector that employs workers with the most specific skills, with pooling of 0.77. Conversely, oil refining and coking plants is the sector that employs the least specialized workers, with pooling of 0.49.

The mean fraction of graduate workers, a proxy for the variable knowledge spillovers, which is the agglomeration mechanism that measures the transfer of knowledge and ideas that may occur between companies and workers close to each other, is 0.05. The industry of manufacture of pharmaco-chemical and pharmaceutical products has the highest fraction of qualified workers, with spillover of 0.31, whereas manufacture of footwear and leather goods has the lowest fraction, 0.01

Regarding the variable input sharing, which reflects the advantages derived from a greater possibility of sharing suppliers and customers provided by economically concentrated environments, the mean value is 0.28. Manufacture of cars, trucks and buses, excluding parts, is the industry with the highest use of manufactured inputs, 0.38, whereas sugar manufacture and refining has the lowest use, 0.09. Furthermore, the use of primary and energy inputs in relation to sales, the natural advantage, is highest for the industry of production of organic and inorganic chemicals, resins and elastomers, 0.22, and has a mean value of 0.03.

### 5.2 New firm location and agglomeration economies

The next stage aims to identify the effects of (localization and urbanization) agglomeration economies on each sector explaining new firm creation. Table 3 outlines the 30 estimates made for each industry sector studied by analyzing the results of the Poisson and negative binomial models. The dependent variable is new firm count by sector and location. The explanatory variables are employment in industry (localization economies) and employment in industry minus employment in the sector (urbanization economies). Control variables include regional fixed effects (dummies for each of the five Brazilian regions) and the area of the municipality in km<sup>2</sup>.

The magnitude of the parameters estimated varies slightly between the first two specifications, but their significance and signs do not. Table A3 in the Appendix describes the robust estimates and standard errors of location and urbanization economies of all sectors, calculated using the Poisson model, while Table A4 describes the complete results of the negative binomial model. A goodness-of-fit chi-square test was performed to assess the fit of the Poisson regression model (Cameron & Trivedi, 2009), which was significant for all study sectors, thus indicating the existence of significant differences between the predicted and the observed values of the dependent variable, firm creations; that is, the model does not fit the data well. Therefore, the estimates made using the negative binomial model should be more robust because this model is more indicated in the presence of overdispersion.

Thus, considering the estimates from the negative binomial model controlled for area and region, the mean value of the parameters of localization economies is 0.0262, with a standard deviation of 0.2205. For 28 of the 30 industry sectors, the estimated coefficient is significant and mostly positive; that is, localization economies may positively affect firm creation. The mean location elasticity estimates range from -0.2852, in the industry sector of maintenance, repair and installation of machinery and equipment, to 0.3609, in the sector of manufacture of apparel and accessories.

Regarding the manufacturing industries of meat packing and production of meat products, including dairy and fishery products; other food products; manufacture of beverages; manufacture of non-metallic mineral products; manufacture of metal products, excluding machinery and equipment; manufacture of machinery and mechanical equipment; manufacture of furniture and products of various industries; and maintenance, repair and installation of machinery and equipment, the localization economies had a significantly negative effect on new firm location decisions. Regarding the sectors of manufacture of wood products and production of organic and inorganic chemicals, resins and elastomers, the localization economies had a non-significant effect on firm creation.

Conversely, the mean value of the parameters of urbanization economies is 0.6728, with a standard deviation of 0.2222. Furthermore, the estimated coefficient of urbanization economies is positive and significant for all industry sectors analyzed. The lowest estimated coefficient is 0.3331, for the manufacture of apparel and accessories industry sector, whereas the three highest coefficients are found for the industries of maintenance, repair and installation of machinery and equipment (1.0058), manufacture of metal products, excluding machinery and equipment (0.8715), and manufacture of machinery and mechanical equipment (0.7702). Thus, on average, urbanization economies are stronger than localization economies and have a deeper effect on new firm location.

**Table 3 - Brazil: Descriptive statistics regarding the estimated coefficients of agglomeration economies (2011-2013)**

Localization economies		
	Poisson	Binomial
Mean	0.1703	0.0262
Standard Deviation	0.3740	0.2205
Minimum	-0.6944	-0.2852
Maximum	0.5954	0.3609
Coefficient Significant at 5% probability	29/30	28/30
Urbanization economies		
	Poisson	Binomial
Mean	0.6751	0.6728
Standard Deviation	0.4082	0.2222
Minimum	0.2032	0.3331
Maximum	1.6235	1.0058
Coefficient Significant at 5% probability	30/30	30/30

Source: Authors, based on RAIS data.

However, Brazil is a continental country with great heterogeneity between regions. Thus, it is possible that the economies of urbanization and location had different impacts in their different geographic spaces. To test this hypothesis, Table 4, below, shows negative binomial regressions, controlled by area size in km<sup>2</sup>, for each Brazilian region<sup>5</sup>: Northeast, North, Southeast, South and Central-West. In general, it is possible to observe that the results are similar to those found previously in the more aggregated analysis presented in Table 3, especially with regard to the significance and average magnitude of the estimated parameters.

Regarding localization economies, the average value of the parameters for the Northeast region is -0.0300, with a standard deviation of 0.2197. For 29 of the 30 industries, the coefficient is statistically significant. The average estimates of location elasticity range from -0.3640, in the sector of meat packing and production of meat products, including dairy and fishery products, to 0.5120, in the manufacture of apparel and accessories. In the North region, the mean value of the parameters is 0.0722, with a standard deviation of 0.1054. For 28 of the 30 industries, the coefficient is statistically significant. The average estimates of location elasticity range from -0.4331, in the sugar manufacturing and refining industry to 0.1610, in the sector of other food products. For the Southeast region, the estimated coefficient for all industries analyzed is statistically significant, varying from -0.3512, in the maintenance, repair and installation of machinery and equipment industry, to 0.3335, in the manufacture of apparel and accessories. The mean value of the parameters is 0.0434, with a standard deviation of 0.1664. For the South region, the estimated coefficient for all industries

<sup>5</sup> Tables A5, A6, A7, A8 and A9 in the appendix present the estimates, as well as the robust standard errors, of the economies of localization and urbanization for all industries and for the Northeast, North, Southeast, South and North regions, respectively, obtained by means of the negative binomial model.

analyzed is statistically significant, ranging from -0.3062, in the maintenance, repair and installation of machinery and equipment industry, to 0.2411, in the manufacture of apparel and accessories sector. The mean value of the parameters is -0.0223, with a standard deviation of 0.1245. Finally, for the Central-West region, the estimated coefficient was statistically significant in 29 of the 30 industries analyzed. The mean value of the parameters is 0.1177, with a standard deviation of 0.2136. The average estimates of location elasticity range from -0.3177 in the manufacture of metal products, excluding machinery and equipment, to 0.5904 in the oil refining and coking plants industry.

As for urbanization economies, it is possible to observe that the magnitude of the parameters estimated for all regions is higher than those verified for the economies of localization. In addition, for all regions, the estimated coefficient for the 30 industries analyzed is statistically significant. Specifically, for the Northeast, average estimates of urbanization elasticity range from -0.0450, in the manufacture of apparel and accessories industry, to 0.7933, in the meat packing and production of meat products industry, including dairy and fishery products. For the North region, the lowest elasticity of urbanization, 0.2698, is verified in the maintenance, repair and installation of machinery and equipment industry, while the largest, 0.4372, in the manufacture of non-metallic mineral products. In the Southeast region, average estimates of urbanization elasticity vary from 0.3933, in the manufacture of apparel and accessories industry, to 1.1232, in the sector of maintenance, repair and installation of machinery and equipment. For the South region, the manufacture of textiles industry presents the lowest elasticity of urbanization, 0.3544, while the greater, 0.9044, is found in the industry of maintenance, repair and installation of machinery and equipment. Finally, for the Central-West, average estimates of urbanization elasticity vary from 0.2770, in the manufacture of apparel and accessories, to 1.0029, in the manufacture of metal products, excluding machinery and equipment.

**Table 4 - Descriptive statistics regarding the estimated coefficients of agglomeration economies for the Brazilian regions (2011-2013)**

Localization economies					
Regions	Northeast	North	Southeast	South	Central-West
Mean	-0.0300	0.0722	0.0434	-0.0223	0.1177
Standard Deviation	0.2197	0.1054	0.1664	0.1245	0.2136
Minimum	-0.3640	-0.4331	-0.3512	-0.3062	-0.3177
Maximum	0.5120	0.1610	0.3335	0.2411	0.5904
Coefficient Significant at 5% probability	29	28	30	30	29
Urbanization economies					
Regions	Northeast	North	Southeast	South	Central-West
Mean	0.5596	0.3376	0.6998	0.6221	0.5972
Standard Deviation	0.1717	0.0368	0.6998	0.1080	0.1528
Minimum	-0.0450	0.2698	0.3933	0.3544	0.2770
Maximum	0.7933	0.4372	1.1232	0.9044	1.0029
Coefficient Significant at 5% probability	30	30	30	30	30

Source: Authors, based on RAIS data.

A possible negative correlation between the combined effects of localization and urbanization economies has been published in the literature (DURANTON; PUGA, 2000; JOFRE-MONSENY et al., 2014). A possible explanation for this negative correlation is the fact that specialized cities (with stronger localization economies) are usually smaller in area. Therefore, to some extent, firms must choose between gains in specialization (localization economies) and city size (urbanization economies), that is, the magnitude of the effects of those economies would vary. Indeed, the present



findings indicate that the coefficients estimated for the localization economies of the three sectors with the strongest urbanization economies are negative.

The results reported herein suggest the validity of the three findings previously reported in the empirical literature regarding agglomeration economies and reviewed by Rosenthal and Strange (2004) and Combes and Gobillon (2014). First, agglomeration economies may be considered a key determinant of firm location decisions and of their productivity. Second, both types of agglomeration economies, localization and urbanization, are key determinants of new firm location choice. Finally, the importance of localization and urbanization economies may vary considerably between industry sectors.

### *5.3 The role of agglomeration mechanisms in explaining the sources of localization and urbanization economies*

Here, based on the coefficients obtained in the previous subsection, Marshall's agglomeration mechanisms that may explain differences in the strength of localization and urbanization economies are analyzed. Tables 5 and 6 indicate which industry characteristics may explain the interindustry differences in the importance of localization and urbanization economies, respectively. In those tables, columns 1 and 2 present the results obtained using the ordinary least squares (OLS) method, whereas column 3 outlines the GMM estimates controlled for endogeneity.

Models with different specifications were estimated. Based on the most parsimonious specification, column 1 outlines the regression results obtained when using the different agglomeration mechanisms simultaneously, without using control variables. The results of the OLS and GMM models are found in columns 2 and 3, respectively, which include the control variables for the use of natural resources, technological intensity, mean firm size and regions. The analysis of the last specification reveals that the estimates obtained using OLS and the GMM indicate no considerable variation in terms of magnitude, parameter sign and standard deviation. However, as previously discussed, the literature warns of the likely endogeneity between agglomeration economies and agglomeration mechanisms, which suggests the need for an estimation method considering that characteristic, such as the GMM. Furthermore, the robustness of estimates obtained by the GMM is demonstrated by the results of the endogeneity tests, which confirm the exogeneity of the model variables after including instrumental variables, and those of the Sargan–Hansen test, which suggest not rejecting the null hypothesis of exogeneity of the instrumental variables used.

In Table 5, when the dependent variable is agglomeration economies, the variable labor market pooling has a positive value for all specifications, with an estimated coefficient of 2.4400 in the GMM model. This result indicates that the effects of localization economies are usually more important in industries that employ workers with sector-specific skills; that is, companies in the same sector tend to cluster towards sharing a pool of skilled workers. Indeed, labor market pooling has been one of the most relevant agglomeration mechanisms, as shown, for example, in Rosenthal and Strange (2001), Gleaser and Kerr (2009) and Jofre-Monseny et al. (2011).

Similarly, knowledge spillovers are also significantly and positively related to localization economies, with a coefficient of 1.9378 in the GMM model. In this case, firms in the same sector would also tend to locate in specialized economic environments, thus sharing the possible benefits of knowledge transfer and new ideas. Regarding the relationship between the effects of localization economies, input sharing and dependence on primary inputs (primary and energy sectors), the results reveal a negative and significant effect, thus indicating weak localization economies in industries using manufactured inputs and inputs from the primary and energy supply sectors more intensely. This result suggests that firms do not locate in specialized environments to share inputs. Regarding the other control variables, the results indicating that localization economies tend to weaken in more technologically intensive industries and to strengthen with firm size in terms of number of workers should also be noted.

Regarding knowledge spillovers and input sharing, empirical studies have obtained different results. Jofre-Monseny et al. (2011), for example, find evidence that the effect of knowledge

spillovers on new firm creation is relatively weak and restricted to a more limited geographical scale, whereas the effect of input sharing is more intense. Ellison et al. (2010) also confirm the importance of input sharing as an agglomeration mechanism. Conversely, in a more recent study, Jofre-Monseny et al. (2014) suggest that knowledge spillovers negatively affect the location decisions of firms of the same sector, whereas input sharing exhibits no significant effect.

**Table 5 - Brazil: Localization economies in the manufacturing industry, 2011-2013**

	OLS		GMM
	(1)	(2)	(3)
Pooling	2.6855*** (0.0075)	2.4384**** (0.0101)	2.4400*** (0.0101)
Knowledge spillovers	0.3709*** (0.0224)	1.9326*** (0.0276)	1.9378*** (0.0277)
Input sharing	-1.8711*** (0.0133)	-1.1998*** (0.0125)	-1.2027*** (0.0125)
Input sharing (Primary and Energy Sectors)		-0.8715*** (0.0201)	-0.8716*** (0.0201)
Technological Intensity		-0.1060*** (0.0010)	-0.1061*** (0.0010)
From 0 to 19 workers		0.1581*** (0.0184)	0.1582*** (0.0184)
From 20 to 99 workers		0.1478*** (0.0185)	0.1478*** (0.0185)
From 100 to 499 workers		0.1074*** (0.0203)	0.1073*** (0.0203)
Northeast		0.0301*** (0.0019)	0.0301*** (0.0019)
Southeast		0.0240*** (0.0018)	0.0240*** (0.0018)
Sul		0.0152*** (0.0019)	0.0152*** (0.0019)
Central West		0.0177*** (0.0021)	0.0177*** (0.0021)
R <sup>2</sup>	0.7388	0.8023	0.8023
Observations	106.393	106.393	106.393
Endogeneity Test			chi2(3) = 5.3323 (p = 0.1490)
Sargan-Hansen Test			chi2(3) = 5.84337 (p = 0.1195)

Source: Authors, based on RAIS data.

Notes: 1. \*\*\*Significant values at 1% probability. 2. Robust standard errors are highlighted in parentheses.

The results outlined in Table 6 suggest that urbanization economies are negatively related to labor market pooling and knowledge spillovers, thereby indicating that firms in sectors employing workers with industry-specific skills and that employ high fractions of qualified workers usually have weak urbanization economies, which are typically stronger in large cities. Considering labor market

pooling, this result corroborates the findings of Jofre-Monseny et al. (2014) for Spain. In contrast, the same authors find positive effects of knowledge spillovers on urbanization economies.

**Table 6 - Brazil: Urbanization economies in the manufacturing industry, 2011-2013**

	OLS		GMM	
	(1)	(2)	(3)	(3)
Pooling	-2.5325*** (0.0078)	-2.2573*** (0.0110)	-2.2591*** (0.0110)	-2.2591*** (0.0110)
Knowledge spillovers	-0.0947*** (0.0229)	-1.4359*** (0.0291)	-1.4412*** (0.0292)	-1.4412*** (0.0292)
Input sharing	1.9863*** (0.0135)	1.3230*** (0.0128)	1.3268*** (0.0128)	1.3268*** (0.0128)
Input sharing (Primary and Energy Sectors)		0.9900*** (0.0217)	0.9901*** (0.0217)	0.9901*** (0.0217)
Technological Intensity		0.1002*** (0.0011)	0.1002*** (0.0011)	0.1002*** (0.0011)
From 0 to 19 workers		-0.1643*** (0.0194)	-0.1635*** (0.0194)	-0.1635*** (0.0194)
From 20 to 99 workers		-0.1514*** (0.0196)	-0.1505*** (0.0196)	-0.1505*** (0.0196)
From 100 to 499 workers		-0.1089*** (0.0214)	-0.1078*** (0.0214)	-0.1078*** (0.0214)
Northeast		-0.0333*** (0.0020)	-0.0333*** (0.0022)	-0.0333*** (0.0022)
Southeast		-0.0248*** (0.0019)	-0.0249*** (0.0019)	-0.0249*** (0.0019)
Sul		-0.0155*** (0.0019)	-0.0156*** (0.0019)	-0.0156*** (0.0019)
Central West		-0.0208*** (0.0021)	-0.0209*** (0.0021)	-0.0209*** (0.0021)
R <sup>2</sup>	0.7211	0.7842	0.7843	0.7843
Observations	106.393	106.393	106.393	106.393
Endogeneity Test			chi2(3) = 7.42953 (p = 0.0594)	
Sargan-Hansen Test			chi2(3) = 7.52626 (p = 0.0569)	

Source: Authors, based on RAIS data.

Notes: 1. \*\*\*Significant values at 1% probability. 2. Robust standard errors are highlighted in parentheses.

Regarding the effects of input sharing and dependence on primary inputs (the primary and energy sectors) on urbanization economies, Table 6 highlights a positive and significant effect, thus indicating the presence of strong urbanization economies in industries using manufactured inputs and inputs from the primary and energy supply sectors more intensely. This result suggests that firms locate in economic environments that are more diversified for the purpose of input sharing. Regarding the other control variables, the results indicate that urbanization economies tend to strengthen among more technologically intensive industries and among smaller firms, in terms of number of workers.

Lastly, Table 7 demonstrates the robustness of the results of alternative specifications, which assess the effects of agglomeration mechanisms on localization and urbanization economies, considering the population size. Furthermore, the analysis performed enables quantifying the relative importance of each agglomeration mechanism proposed by Marshall (1920), considering their incidence in different geographical scales using different exclusion criteria to reduce the sample size. In this case, the estimates presented were obtained using the GMM. Furthermore, the overall results reveal no differences from previously performed analyses of overall samples. This result corroborates, for example, the findings of Jofre-Monseny et al. (2011), who analyzed the effects of agglomeration mechanisms on new firm location in Spain using different population exclusion criteria.

Regarding localization economies, the results outlined in Table 7 indicate that pooling and knowledge spillover are apparently the most relevant agglomeration mechanisms, enhancing the effects of those economies. Regarding urbanization economies, the variables related to input supply and input sharing (the primary and energy sectors) enhance the effects of these economies. The results also demonstrate that the effects of agglomeration mechanisms may strengthen with increased population density of the economic environment.

Indeed, significant effects of Marshall's agglomeration mechanisms on new firm location have been reported in the literature, although they differ in terms of the magnitude and specific contribution of each agglomerative factor. In Rosenthal and Strange (2001), for example, evidence that labor market pooling is the strongest effect, positively affecting firm location decisions, is reported. Conversely, Ellison, Glaeser and Kerr (2010) find significant evidence for the three agglomeration mechanisms, among which input sharing has the strongest effect. Similarly, Jofre-Monseny et al. (2011) find significant evidence for the three agglomeration mechanisms, and labor market pooling is the most relevant agglomeration mechanism. In this line of research, analyzing localization and urbanization economies, the present study finds significant evidence for the three Marshall's agglomeration mechanisms by analyzing new firm location, a research strategy underused in studies on this subject at the national and even international level.

## 6. Final considerations

This study aimed to assess, based on an analysis of firm creation in the manufacturing industry of Brazilian municipalities, how the agglomeration mechanisms proposed by Marshall (1920) – labor market pooling, customer-supplier relationship or input sharing and knowledge spillovers – determine localization and urbanization economies. For this purpose, the analysis followed the method proposed by Jofre-Monseny et al. (2014), performing a two-stage estimation procedure: 1) estimation of Poisson and negative binomial regression models for count data to identify, for each industry sector, the type of agglomeration economy that contributes to new firm location and 2) use of the coefficients obtained in the first stage to assess the industry characteristics related to Marshall's agglomeration mechanisms that may contribute to differences in the incidence of those economies by applying the GMM and controlling for endogeneity.

Initial evidence indicates that both agglomeration economies, localization and urbanization, can be considered key determinants of new firm location decisions and firm productivity. Furthermore, the relative importance of those economies may vary considerably between industry sectors. Localization economies are stronger in industry sectors that employ workers with industry-specific skills, thus enabling specialized labor market pooling. Similarly, knowledge spillovers strengthen localization economies, thereby providing an environment conducive to knowledge transfer and new ideas. In contrast, input sharing and dependence on primary inputs (the primary and energy sectors) have positive and significant effects on urbanization economies, thus suggesting the presence of strong urbanization economies in industries that use manufactured inputs and inputs from the primary and energy supply sectors more intensely. In this case, firms would tend to locate in more diversified economic environments for input sharing, thus reducing input and freight costs.

**Table 7 - Brazil: Localization and urbanization economies by population size, 2011-2013**

	Up to 50,000		From 50,001 to 100,000		From 100,001 to 500,000		More than 500,000	
	Localization	Urbanization	Localization	Urbanization	Localization	Urbanization	Localization	Urbanization
Pooling	2.3443*** (0.0176)	-2.1377*** (0.0193)	2.1886*** (0.0260)	-1.9696*** (0.0283)	2.4529*** (0.0177)	-2.2761*** (0.0193)	2.7332*** (0.0216)	-2.6080*** (0.0233)
Knowledge spillovers	1.7880*** (0.0484)	-1.2429*** (0.0519)	1.7022*** (0.0708)	-1.2185*** (0.0749)	2.0524*** (0.0490)	-1.5909*** (0.0504)	2.1042*** (0.0627)	-1.6447*** (0.0657)
Input sharing	-0.9930*** (0.0219)	1.0798*** (0.0223)	-1.1775*** (0.0327)	1.2859*** (0.0342)	-1.3698*** (0.0239)	1.5084*** (0.0245)	-1.3214*** (0.0255)	1.4700*** (0.0258)
Input sharing (Primary and Energy Sectors)	-0.8089*** (0.0417)	0.9353*** (0.0417)	-0.9755*** (0.0566)	1.1305*** (0.0566)	-0.8002*** (0.0339)	0.9029*** (0.0339)	-0.7699*** (0.0378)	0.8548*** (0.0378)
R <sup>2</sup>	0.7668	0.7366	0.7693	0.7384	0.8039	0.7871	0.8693	0.8690
Observations	34.268	34.268	13.958	13.958	30.653	30.653	27.516	27.516
Endogeneity test	chi2(3) = 3.89403 (p = 0.2731)	chi2(3) = 3.91368 (p = 0.2709)	chi2(3) = 3.32955 (p = 0.3436)	chi2(3) = 3.11314 (p = 0.3745)	chi2(3) = 2.49053 (p = 0.4770)	chi2(3) = 4.83065 (p = 0.1846)	chi2(3) = 3.50365 (p = 0.3203)	chi2(3) = 2.01861 (p = 0.5686)
Sargan–Hansen test	chi2(3) = 2.90461 (p = 0.4066)	chi2(3) = 5.60533 (p = 0.1325)	chi2(3) = 4.18619 (p = 0.2420)	chi2(3) = 4.2818 (p = 0.2326)	chi2(3) = 4.88921 (p = 0.1801)	chi2(3) = 4.29918 (p = 0.2309)	chi2(3) = 1.87791 (p = 0.5981)	chi2(3) = 3.98397 (p = 0.2632)

Source: Authors, based on RAIS data.

Notes: 1. \*\*\*Significant values at 1% probability. 2. Robust standard errors are highlighted in parentheses. 3. Estimates were controlled by the same covariates as those of Tables 5 and 6.

It is noteworthy that the results are robust to alternative specifications which assess the effects of agglomeration mechanisms on localization and urbanization economies. Thus, regarding localization economies, labor market pooling and knowledge spillover are apparently the most relevant agglomeration mechanisms, enhancing the effects of those economies. Regarding urbanization economies, the variables that most enhance their positive effects are related to input supply and input sharing (the primary and energy sectors).

However, it is possible to highlight some limitations in the present study. First, the results found for the manufacturing industry may be different in other types of industries or economic sectors. Second, for the Brazilian case, the need to use some explanatory variables makes it impossible to use longer data panels to count the birth of firms, which could mitigate the participation of censored observations in the sample. Thus, future studies could explore the effect of cluster economies on the creation of firms considering different sectors, as well as analyze this creation using a longer time span.

Furthermore, this study contributes to the empirical literature regarding agglomeration mechanisms that shape the geography of economic activities, whose understanding may ultimately help to apply local development policies. The findings demonstrated that the effects of localization economies are strongest for industries with intensive knowledge spillovers and with marked labor market pooling. For example, the results suggest that public initiatives aimed at promoting local specialization of economic activities would be more effective if they were directed towards traditional sectors, for which local expertise can bring advantages by generating a shared pool of skilled workers, or knowledge-intensive activities. Conversely, incentives for urbanization economies should consider improvements in urban infrastructure and accessibility which enable extending the benefits generated by input sharing and customer-supplier relationships.

## Referências

- BACOLO, M.; BLUM, B. S.; STRANGE, W. C. Elements of skills: traits, intelligences, education, and agglomeration. *Journal of Regional Science*, v. 50, n. 1, p. 245-280, 2010.
- BRITTO, J.; ALBUQUERQUE, E. M. Clusters industriais na economia brasileira: uma análise exploratória a partir de dados da RAIS. *Estudos Econômicos*, v. 32, n. 1, p. 71-102, 2002.
- CAMERON, A.; TRIVEDI, P. *Microeconometrics: methods and applications*. New York: Cambridge University Press, 2005.
- CAMERON, A.; TRIVEDI, P. *Microeconometrics using Stata*. Texas: StataCorp LP, 2009.
- CARLINO, G.; CHATTERJEE, S.; HUNT, R. Urban density and the rate of invention. *Journal of Urban Economics*, v. 61, n. 3, p. 389-419, 2007.
- CARLTON, D. W. The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables. *Review of Economics and Statistics*, v. 65, n. 3, p. 440-449, 1983.
- COMBES, P. P.; GOBILLON, L. The empirics of agglomeration economies. In: DURANTON, G.; HENDERSON, V.; STRANGE, W. (eds.). *Handbook of Urban and Regional Economics*, v. 5A, p. 247-348. Amsterdam: North-Holland, 2014.
- CRÉPON, B.; DUGUET, E. Estimating the innovation function from patent numbers: GMM on count panel data. *Journal of Applied Econometrics*, v. 12, n. 3, p. 243-263, 1997.
- DUMAIS, G.; ELLISON, G.; GLAESER, E. Geographic Concentration as a dynamic process. *The Review of Economics and Statistics*, v. 84, n. 2, p. 193-204, 2002.
- DUNCAN, O.; DUNCAN, B. A methodological analysis of segregation indexes. *American Sociological Review*, v. 20, n. 2, p. 210-217, 1955.

- DURANTON, G.; PUGA, D. Diversity and specialization in cities: why, where and when does it matter? *Urban Studies*, v. 37, n. 3, p. 533-555, 2000.
- DURANTON, G.; PUGA, D. Nursery cities: urban diversity, process innovation, and the life-cycle of products. *American Economic Review*, v. 91, n. 5, p. 1454-1477, 2001.
- ELLISON, G.; GLAESER, E. L. The geographic concentration of an industry: does natural advantage explain agglomeration? *The American Economic Review*, v. 89, n. 2, p. 311-316, 1999.
- ELLISON, G.; GLAESER, E. L.; KERR, W. R. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, v. 100, n. 3, p. 1195-1213, 2010.
- GARCIA, R. C. Aglomerações industriais na indústria brasileira de calçados: identificação e delimitação a partir dos dados da RAIS. *Revista Econômica do Nordeste*, v. 32, n. 1, p. 505-523, 2001.
- GLAESER, E. L.; KERR, W. R. Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain? *Journal of Economics and Management Strategy*, v. 18, n. 3, p. 623-663, 2009.
- GLAESER, E.; RESENGER, M. The complementarity between cities and skills. *Journal of Regional Science*, v. 50, n. 1, p. 221-244, 2010.
- GUILHOTO, J. J.; SESSO FILHO, U. Estimação da matriz insumo-produto utilizando dados preliminares das contas nacionais: aplicação e análise de indicadores econômicos para o Brasil em 2005. *Economia & Tecnologia*, v. 6, n. 4, p. 1-10, 2010.
- GUILHOTO, J. J.; SESSO FILHO, U. A. Estimação da Matriz Insumo-Produto a Partir de Dados Preliminares das Contas Nacionais. *Economia Aplicada*, v. 9, n. 2, p. 277-299, 2005.
- GUIMARÃES, P.; FIGUEIREDO, O.; WOODWARD, D. A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, v. 85, n. 1, p. 201-204, 2003.
- JACOBS, J. *The economy of cities*. New York: Vintage, 1969.
- JAFFE, A. B.; TRAJTENBERG, M.; HENDERSON, R. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, v. 108, n. 3, p. 577-598, 1993.
- JOFRE-MONSENY, J.; MARÍN-LÓPES, R.; VILADECANS-MARSAL, E. The mechanisms of agglomeration: evidence from the effect of inter-industry relations on the location of new firms. *Journal of Urban Economics*, v. 70, n. 2-3, p. 61-74, 2011.
- JOFRE-MONSENY, J.; MARÍN-LÓPES, R.; VILADECANS-MARSAL, E. The determinants of localization and urbanization economies: evidence from the location of new firms in Spain. *Journal of Regional Science*, v. 54, n. 2, p. 313-337, 2014.
- MARSHALL, A. *Principles of economics*. New York: Mcmillan, 1920.
- REZENDE, A. C.; CAMPOLINA, B.; PAIXÃO, A. N. Clusterização e Localização da Indústria de Transformação no Brasil entre 1994 e 2009. *Revista Econômica do Nordeste*, v. 43, n. 4, p. 27-49, 2012.
- ROSENTHAL, S. S.; STRANGE, W. C. The determinants of agglomeration. *Journal of Urban Economics*, v. 50, n. 2, p. 191-229, 2001.
- ROSENTHAL, S. S.; STRANGE, W. C. Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, v. 85, n. 2, p. 377-393, 2003.

ROSENTHAL, S. S.; STRANGE, W. C. Evidence on the nature and sources of agglomeration economies. In: HENDERSON, J. V.; THISSE, J. F. (eds.). *Handbook of Urban and Regional Economics*, v. 4, p. 2119-2172. Amsterdam: North Holland, 2004.

SOBRINHO, E. M.; AZZONI, C. R. Aglomerações industriais relevantes do Brasil em 2010. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 9, n. 1, p. 1-18, 2015.


WOOLDRIDGE, J. M. *Econometric analysis of cross section and panel data*. Cambridge: The MIT, 2002.

## ORCID

Patrícia Araújo Amarante  <https://orcid.org/0000-0002-0542-9223>

Magno Vamberto Batista da Silva  <https://orcid.org/0000-0001-5376-3771>

Paulo Aguiar do Monte  <https://orcid.org/0000-0001-5376-3771>

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.



## Appendix

Table A1 - Matching between RAIS sectors and Input-Output Matrix Activities

CNAE Code					Activities Code	Activities			
	10.1	10.2	10.5		1091	Meat packing and production of meat products, including dairy and fishery products			
			10.7		1092	Sugar manufacturing and refining			
10.3	10.4	10.6	10.8	10.9	1093	Other Food Products			
			11.1	11.2	1100	Manufacture of beverages			
			12.1	12.2	1200	Manufacture of tobacco products			
13.1	13.2	13.3	13.4	13.5	1300	Manufacture of textiles			
			14.1	14.2	1400	Manufacture of apparel and accessories			
15.1	15.2	15.3	15.4		1500	Manufacture of footwear and leather goods			
			16.1	16.2	1600	Manufacture of wood products			
17.1	17.2	17.3	17.4		1700	Manufacture of cellulose, paper and paper products			
		18.1	18.2	18.3	1800	Printing and copying			
			19.1	19.2	1991	Oil refining and coking plants			
				19.3	1992	Biofuel production			
20.1	20.2	20.3	20.4		2091	Production of other organic and inorganic chemicals, resins and elastomers			
		20.5	20.7	20.9	2092	Manufacture of pesticides, disinfectants, paints and various chemicals			
				20.6	2093	Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations			
			21.1	21.2	2100	Manufacture of pharmaco-chemical and pharmaceutical products			
			22.1	22.2	2200	Manufacture of rubber and plastic products			
23.1	23.2	23.3	23.4	23.9	2300	Manufacture of non-metallic mineral products			
			24.1	24.2	24.3	2491	Production of pig iron/ ferrous alloys, steel and seamless steel tubes		
				24.4	24.5	2492	Metallurgy of nonferrous metals and metal smelting industry		
25.1	25.2	25.3	25.4	25.5	25.9	2500	Manufacture of metal products, excluding machinery and equipment		
26.1	26.2	26.3	26.4	26.5	26.6	26.7	26.8	2600	Manufacture of computers, electronics and optical products
	27.1	27.2	27.3	27.4	27.5	27.9		2700	Manufacture of electrical machinery and equipment
	28.1	28.2	28.3	28.4	28.5	28.6		2800	Manufacture of machinery and mechanical equipment
		29.1	29.2	29.3				2991	Manufacture of cars, trucks and buses, excluding parts
			29.4	29.5				2992	Manufacture of parts and accessories for motor vehicles
	30.1	30.3	30.4	30.5	30.9			3000	Manufacture of other transport equipment, excluding motor vehicles
31.0	32.1	32.2	32.3	32.4	32.5	32.9		3180	Manufacture of furniture and products of various industries
			33.1	33.2				3300	Maintenance, repair and installation of machinery and equipment

Source: Authors, based on IBGE data.

**Table A2 - Sector classification by technological intensity**

Sectors	Technological Intensity
Meat packing and production of meat products, including dairy and fishery products	Low
Sugar manufacturing and refining	Low
Other Food Products	Low
Manufacture of beverages	Low
Manufacture of tobacco products	Low
Manufacture of textiles	Low
Manufacture of apparel and accessories	Low
Manufacture of footwear and leather goods	Low
Manufacture of wood products	Low
Manufacture of cellulose, paper and paper products	Low
Printing and copying	Low
Oil refining and coking plants	Medium-low
Biofuel production	Medium-low
Production of other organic and inorganic chemicals, resins and elastomers	Medium-high
Manufacture of pesticides, disinfectants, paints and various chemicals	Medium-high
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	Medium-high
Manufacture of pharmaco-chemical and pharmaceutical products	High
Manufacture of rubber and plastic products	Medium-low
Manufacture of non-metallic mineral products	Medium-low
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	Medium-low
Metallurgy of nonferrous metals and metal smelting industry	Medium-low
Manufacture of metal products, excluding machinery and equipment	Medium-low
Manufacture of computers, electronics and optical products	High
Manufacture of electrical machinery and equipment	Medium-high
Manufacture of machinery and mechanical equipment	Medium-high
Manufacture of cars, trucks and buses, excluding parts	Medium-high
Manufacture of parts and accessories for motor vehicles	Medium-high
Manufacture of other transport equipment, excluding motor vehicles	Medium-high
Manufacture of furniture and products of various industries	High
Maintenance, repair and installation of machinery and equipment	Medium-low

Source: Authors, based on OECD data.

**Table A3 - Brazil: Poisson estimates of the parameters of localization and urbanization economies. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	0.0083*	(0.0050)	0.8395***	(0.0034)
Sugar manufacturing and refining	0.0298***	(0.0033)	0.8552***	(0.0043)
Other Food Products	0.0915***	(0.0093)	0.7271***	(0.0075)
Manufacture of beverages	-0.1190***	(0.0037)	0.9749***	(0.0053)
Manufacture of tobacco products	0.3747***	(0.0087)	0.6788***	(0.0044)
Manufacture of textiles	0.4148***	(0.0034)	0.4233***	(0.0034)
Manufacture of apparel and accessories	0.5954***	(0.0045)	0.2032***	(0.0043)
Manufacture of footwear and leather goods	0.3503***	(0.0034)	0.5659***	(0.0035)
Manufacture of wood products	0.3017***	(0.0070)	0.6180***	(0.0042)
Manufacture of cellulose, paper and paper products	0.4319***	(0.0048)	0.5098***	(0.0039)
Printing and copying	0.2978***	(0.0100)	0.4872***	(0.0107)
Oil refining and coking plants	0.0628***	(0.0039)	0.9391***	(0.0053)
Biofuel production	0.3993***	(0.0041)	0.6396***	(0.0033)
Production of other organic and inorganic chemicals, resins and elastomers	0.1915***	(0.0058)	0.7670***	(0.0036)
Manufacture of pesticides, disinfectants, paints and various chemicals	0.2078***	(0.0039)	0.6495***	(0.0046)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	0.2036***	(0.0043)	0.6688***	(0.0055)
Manufacture of pharmaco-chemical and pharmaceutical products	0.2502***	(0.0042)	0.6762***	(0.0039)
Manufacture of rubber and plastic products	0.3321***	(0.0047)	0.5692***	(0.0047)
Manufacture of non-metallic mineral products	-0.0134	(0.0090)	0.8583***	(0.0062)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	-0.0362***	(0.0032)	0.9207***	(0.0051)
Metallurgy of nonferrous metals and metal smelting industry	0.2484***	(0.0050)	0.6971***	(0.0037)
Manufacture of metal products, excluding machinery and equipment	0.1741***	(0.0093)	0.6440***	(0.0076)
Manufacture of computers, electronics and optical products	0.0463***	(0.0054)	0.8563***	(0.0056)
Manufacture of electrical machinery and equipment	0.1518***	(0.0053)	0.7032***	(0.0047)
Manufacture of machinery and mechanical equipment	0.2324***	(0.0054)	0.6647***	(0.0040)
Manufacture of cars, trucks and buses, excluding parts	0.1140***	(0.0038)	0.8025***	(0.0035)
Manufacture of parts and accessories for motor vehicles	0.1734***	(0.0035)	0.7310***	(0.0036)
Manufacture of other transport equipment, excluding motor vehicles	0.0914***	(0.0034)	0.8161***	(0.0049)
Manufacture of furniture and products of various industries	0.1994***	(0.0095)	0.6302***	(0.0073)
Maintenance, repair and installation of machinery and equipment	-0.6944***	(0.0062)	1.6235***	(0.0082)

Source: Authors, based on RAIS data.

\*Significant values at 10% probability; \*\* Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: Robust standard errors are highlighted in parentheses.

**Table A4 - Brazil: Estimates of the parameters of localization and urbanization economies by Negative Binomial: specification with controls. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	-0.0656***	(0.0033)	0.7413***	(0.0030)
Sugar manufacturing and refining	0.0789***	(0.0042)	0.6606***	(0.0024)
Other Food Products	-0.0965***	(0.0050)	0.7725***	(0.0049)
Manufacture of beverages	-0.0094**	(0.0041)	0.7159***	(0.0036)
Manufacture of tobacco products	0.2257***	(0.0043)	0.5972***	(0.0027)
Manufacture of textiles	0.3072***	(0.0025)	0.3704***	(0.0031)
Manufacture of apparel and accessories	0.3609***	(0.0030)	0.3331***	(0.0031)
Manufacture of footwear and leather goods	0.2009***	(0.0026)	0.5492***	(0.0027)
Manufacture of wood products	-0.0038	(0.0045)	0.7169***	(0.0037)
Manufacture of cellulose, paper and paper products	0.1932***	(0.0036)	0.5559***	(0.0038)
Printing and copying	0.0348***	(0.0054)	0.6714***	(0.0064)
Oil refining and coking plants	0.1727***	(0.0047)	0.6448***	(0.0031)
Biofuel production	0.2709***	(0.0035)	0.6215***	(0.0023)
Production of other organic and inorganic chemicals, resins and elastomers	0.0057	(0.0037)	0.7609***	(0.0037)
Manufacture of pesticides, disinfectants, paints and various chemicals	0.0778***	(0.0037)	0.6388***	(0.0046)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	0.0475***	(0.0037)	0.7017***	(0.0042)
Manufacture of pharmaco-chemical and pharmaceutical products	0.1356***	(0.0025)	0.6472***	(0.0034)
Manufacture of rubber and plastic products	0.0707***	(0.0043)	0.6486***	(0.0048)
Manufacture of non-metallic mineral products	-0.1024***	(0.0055)	0.7596***	(0.0048)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	0.0484***	(0.0032)	0.6534***	(0.0032)
Metallurgy of nonferrous metals and metal smelting industry	0.0819***	(0.0038)	0.6730***	(0.0039)
Manufacture of metal products, excluding machinery and equipment	-0.1745***	(0.0053)	0.8715***	(0.0065)
Manufacture of computers, electronics and optical products	0.0247***	(0.0031)	0.7045***	(0.0043)
Manufacture of electrical machinery and equipment	0.0163***	(0.0032)	0.7162***	(0.0041)
Manufacture of machinery and mechanical equipment	-0.0250***	(0.0038)	0.7702***	(0.0046)
Manufacture of cars, trucks and buses, excluding parts	0.0087***	(0.0029)	0.7176***	(0.0033)
Manufacture of parts and accessories for motor vehicles	0.0550***	(0.0029)	0.7107***	(0.0037)
Manufacture of other transport equipment, excluding motor vehicles	0.1279***	(0.0033)	0.6268***	(0.0033)
Manufacture of furniture and products of various industries	-0.0431***	(0.0042)	0.7246***	(0.0047)
Maintenance, repair and installation of machinery and equipment	-0.2852***	(0.0046)	1.0058***	(0.0051)

Source: Authors, based on based on RAIS data.

\*Significant values at 10% probability; \*\*Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: Robust standard errors are highlighted in parentheses.

**Table A5 - Northeast region: Estimates of the parameters of localization and urbanization economies by Negative Binomial: specification with controls. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	-0.3640***	(0.0089)	0.7933***	(0.0084)
Sugar manufacturing and refining	-0.0366***	(0.0057)	0.5454***	(0.0043)
Other Food Products	-0.1140***	(0.0139)	0.5868***	(0.0135)
Manufacture of beverages	-0.0502***	(0.0117)	0.6123***	(0.0090)
Manufacture of tobacco products	0.3481***	(0.0190)	0.5595***	(0.0071)
Manufacture of textiles	0.3308***	(0.0089)	0.1496***	(0.0073)
Manufacture of apparel and accessories	0.5120***	(0.0060)	-0.0450***	(0.0065)
Manufacture of footwear and leather goods	-0.0876***	(0.0064)	0.5519***	(0.0067)
Manufacture of wood products	-0.2820***	(0.0117)	0.7384***	(0.0108)
Manufacture of cellulose, paper and paper products	0.1909***	(0.0092)	0.6021***	(0.0087)
Printing and copying	-0.0009	(0.0142)	0.4364***	(0.0156)
Oil refining and coking plants	0.2954***	(0.0055)	0.5833***	(0.0070)
Biofuel production	-0.0342***	(0.0090)	0.5160***	(0.0043)
Production of other organic and inorganic chemicals, resins and elastomers	-0.0698***	(0.0093)	0.6933***	(0.0071)
Manufacture of pesticides, disinfectants, paints and various chemicals	-0.2490***	(0.0121)	0.6106***	(0.0096)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	-0.2987***	(0.0096)	0.7133***	(0.0096)
Manufacture of pharmaco-chemical and pharmaceutical products	0.0782***	(0.0107)	0.4855***	(0.0079)
Manufacture of rubber and plastic products	-0.2838***	(0.0114)	0.7163***	(0.0104)
Manufacture of non-metallic mineral products	-0.1620***	(0.0115)	0.5867***	(0.0098)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	-0.1547***	(0.0107)	0.4741***	(0.0068)
Metallurgy of nonferrous metals and metal smelting industry	0.0825***	(0.0129)	0.5222***	(0.0091)
Manufacture of metal products, excluding machinery and equipment	-0.2724***	(0.0104)	0.6769***	(0.0126)
Manufacture of computers, electronics and optical products	-0.1486***	(0.0093)	0.6350***	(0.0094)
Manufacture of electrical machinery and equipment	0.1245***	(0.0084)	0.5201***	(0.0090)
Manufacture of machinery and mechanical equipment	-0.2196***	(0.0113)	0.7761***	(0.0100)
Manufacture of cars, trucks and buses, excluding parts	0.1530***	(0.0102)	0.5275***	(0.0086)
Manufacture of parts and accessories for motor vehicles	0.0982***	(0.0095)	0.5592***	(0.0098)
Manufacture of other transport equipment, excluding motor vehicles	0.0285***	(0.0099)	0.4366***	(0.0087)
Manufacture of furniture and products of various industries	-0.0362***	(0.0104)	0.4729***	(0.0112)
Maintenance, repair and installation of machinery and equipment	-0.2790***	(0.0104)	0.7505***	(0.0111)

Source: Authors, based on based on RAIS data.

\*Significant values at 10% probability; \*\*Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: Robust standard errors are highlighted in parentheses.

**Table A6 - North region: Estimates of the parameters of localization and urbanization economies by Negative Binomial: specification with controls. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	-0.0195**	(0.0088)	0.4251***	(0.0079)
Sugar manufacturing and refining	-0.4331***	(0.0996)	0.3688***	(0.0086)
Other Food Products	0.1610***	(0.0135)	0.2841***	(0.0138)
Manufacture of beverages	0.1114***	(0.0097)	0.3071***	(0.0126)
Manufacture of tobacco products	0.1493***	(0.0121)	0.3068***	(0.0103)
Manufacture of textiles	0.0848***	(0.0102)	0.3397***	(0.0114)
Manufacture of apparel and accessories	0.1175***	(0.0143)	0.3329***	(0.0150)
Manufacture of footwear and leather goods	0.0388***	(0.0097)	0.3721***	(0.0108)
Manufacture of wood products	0.0992***	(0.0076)	0.3498***	(0.0104)
Manufacture of cellulose, paper and paper products	0.0999***	(0.0079)	0.3170***	(0.0105)
Printing and copying	0.1247***	(0.0125)	0.3096***	(0.0155)
Oil refining and coking plants	0.0793***	(0.0100)	0.3124***	(0.0095)
Biofuel production	0.0231	(0.0271)	0.3829***	(0.0089)
Production of other organic and inorganic chemicals, resins and elastomers	0.1346***	(0.0152)	0.3186***	(0.0105)
Manufacture of pesticides, disinfectants, paints and various chemicals	0.0957***	(0.0097)	0.3158***	(0.0110)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	0.1575***	(0.0112)	0.2982***	(0.0113)
Manufacture of pharmaco-chemical and pharmaceutical products	0.1289***	(0.0127)	0.3240***	(0.0099)
Manufacture of rubber and plastic products	0.0773***	(0.0100)	0.3710***	(0.0139)
Manufacture of non-metallic mineral products	0.0018	(0.0148)	0.4372***	(0.0121)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	0.0513***	(0.0115)	0.3378***	(0.0108)
Metallurgy of nonferrous metals and metal smelting industry	0.0925***	(0.0096)	0.3243***	(0.0111)
Manufacture of metal products, excluding machinery and equipment	0.0708***	(0.0136)	0.3487***	(0.0159)
Manufacture of computers, electronics and optical products	0.0367***	(0.0052)	0.3331***	(0.0102)
Manufacture of electrical machinery and equipment	0.0550***	(0.0072)	0.3274***	(0.0098)
Manufacture of machinery and mechanical equipment	0.0807***	(0.0090)	0.3219***	(0.0120)
Manufacture of cars, trucks and buses, excluding parts	0.1313***	(0.0147)	0.3719***	(0.0110)
Manufacture of parts and accessories for motor vehicles	0.0967***	(0.0095)	0.3371***	(0.0118)
Manufacture of other transport equipment, excluding motor vehicles	0.0648***	(0.0061)	0.3268***	(0.0098)
Manufacture of furniture and products of various industries	0.1081***	(0.0148)	0.3571***	(0.0150)
Maintenance, repair and installation of machinery and equipment	0.1452***	(0.0122)	0.2698***	(0.0150)

Source: Authors, based on based on RAIS data.

\*Significant values at 10% probability; \*\*Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: Robust standard errors are highlighted in parentheses.

**Table A7 - Southeast region: Estimates of the parameters of localization and urbanization economies by Negative Binomial: specification with controls. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	-0.0269***	(0.0060)	0.6959***	(0.0049)
Sugar manufacturing and refining	0.1113***	(0.0047)	0.6842***	(0.0033)
Other Food Products	-0.2232***	(0.0084)	0.8791***	(0.0079)
Manufacture of beverages	-0.0403***	(0.0070)	0.7290***	(0.0051)
Manufacture of tobacco products	0.2432***	(0.0084)	0.5857***	(0.0040)
Manufacture of textiles	0.2587***	(0.0042)	0.4356***	(0.0051)
Manufacture of apparel and accessories	0.3335***	(0.0058)	0.3933***	(0.0058)
Manufacture of footwear and leather goods	0.3067***	(0.0027)	0.4853***	(0.0032)
Manufacture of wood products	-0.0231***	(0.0078)	0.7372***	(0.0064)
Manufacture of cellulose, paper and paper products	0.1861***	(0.0048)	0.5450***	(0.0053)
Printing and copying	0.1031***	(0.0075)	0.6237***	(0.0085)
Oil refining and coking plants	0.1250***	(0.0062)	0.6678***	(0.0046)
Biofuel production	0.3146***	(0.0054)	0.6096***	(0.0041)
Production of other organic and inorganic chemicals, resins and elastomers	-0.0212***	(0.0049)	0.7865***	(0.0053)
Manufacture of pesticides, disinfectants, paints and various chemicals	0.1109***	(0.0050)	0.6225***	(0.0064)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	0.0241***	(0.0051)	0.7477***	(0.0062)
Manufacture of pharmaco-chemical and pharmaceutical products	0.0963***	(0.0036)	0.6855***	(0.0047)
Manufacture of rubber and plastic products	0.2243***	(0.0057)	0.4979***	(0.0071)
Manufacture of non-metallic mineral products	-0.1425***	(0.0081)	0.7905***	(0.0073)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	0.0427***	(0.0047)	0.6568***	(0.0048)
Metallurgy of nonferrous metals and metal smelting industry	0.0664***	(0.0049)	0.6778***	(0.0052)
Manufacture of metal products, excluding machinery and equipment	-0.1997***	(0.0077)	0.9206***	(0.0089)
Manufacture of computers, electronics and optical products	-0.0140***	(0.0043)	0.7717***	(0.0061)
Manufacture of electrical machinery and equipment	-0.0832***	(0.0044)	0.8445***	(0.0060)
Manufacture of machinery and mechanical equipment	-0.1005***	(0.0059)	0.8453***	(0.0068)
Manufacture of cars, trucks and buses, excluding parts	-0.0173***	(0.0043)	0.7198***	(0.0047)
Manufacture of parts and accessories for motor vehicles	-0.0237***	(0.0046)	0.7973***	(0.0065)
Manufacture of other transport equipment, excluding motor vehicles	0.1164***	(0.0044)	0.6581***	(0.0048)
Manufacture of furniture and products of various industries	-0.0938***	(0.0075)	0.7761***	(0.0082)
Maintenance, repair and installation of machinery and equipment	-0.3512***	(0.0067)	1.1232***	(0.0069)

Source: Authors, based on based on RAIS data.

\*Significant values at 10% probability; \*\*Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: Robust standard errors are highlighted in parentheses.

**Table A8 - South region: Estimates of the parameters of localization and urbanization economies by Negative Binomial: specification with controls. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	-0.0955***	(0.0041)	0.6799***	(0.0045)
Sugar manufacturing and refining	0.0224***	(0.0058)	0.6425***	(0.0041)
Other Food Products	0.0489***	(0.0083)	0.5792***	(0.0082)
Manufacture of beverages	-0.0431***	(0.0052)	0.6231***	(0.0056)
Manufacture of tobacco products	-0.0508***	(0.0059)	0.6621***	(0.0042)
Manufacture of textiles	0.2282***	(0.0036)	0.3544***	(0.0045)
Manufacture of apparel and accessories	0.2411***	(0.0040)	0.4280***	(0.0042)
Manufacture of footwear and leather goods	0.1668***	(0.0039)	0.5256***	(0.0048)
Manufacture of wood products	-0.0337***	(0.0062)	0.6280***	(0.0050)
Manufacture of cellulose, paper and paper products	0.1919***	(0.0051)	0.3912***	(0.0054)
Printing and copying	-0.0549***	(0.0072)	0.6768***	(0.0087)
Oil refining and coking plants	-0.2360***	(0.0112)	0.6308***	(0.0040)
Biofuel production	-0.0565***	(0.0113)	0.6308***	(0.0034)
Production of other organic and inorganic chemicals, resins and elastomers	-0.1909***	(0.0056)	0.7597***	(0.0053)
Manufacture of pesticides, disinfectants, paints and various chemicals	-0.0257***	(0.0060)	0.6402***	(0.0068)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	0.0126**	(0.0064)	0.6308***	(0.0065)
Manufacture of pharmaco-chemical and pharmaceutical products	-0.1248***	(0.0064)	0.7447***	(0.0050)
Manufacture of rubber and plastic products	0.0654***	(0.0054)	0.5201***	(0.0066)
Manufacture of non-metallic mineral products	-0.0166**	(0.0065)	0.6264***	(0.0063)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	-0.1709***	(0.0058)	0.7046***	(0.0055)
Metallurgy of nonferrous metals and metal smelting industry	-0.0486***	(0.0057)	0.6437***	(0.0061)
Manufacture of metal products, excluding machinery and equipment	0.0150**	(0.0070)	0.6160***	(0.0086)
Manufacture of computers, electronics and optical products	-0.0641***	(0.0053)	0.6610***	(0.0074)
Manufacture of electrical machinery and equipment	0.0289***	(0.0046)	0.5498***	(0.0063)
Manufacture of machinery and mechanical equipment	-0.0911***	(0.0056)	0.6927***	(0.0071)
Manufacture of cars, trucks and buses, excluding parts	-0.0964***	(0.0037)	0.7233***	(0.0051)
Manufacture of parts and accessories for motor vehicles	0.0330***	(0.0043)	0.6134***	(0.0060)
Manufacture of other transport equipment, excluding motor vehicles	-0.0511***	(0.0070)	0.5962***	(0.0048)
Manufacture of furniture and products of various industries	0.0346***	(0.0059)	0.5848***	(0.0064)
Maintenance, repair and installation of machinery and equipment	-0.3062***	(0.0059)	0.9044***	(0.0075)

Source: Authors, based on based on RAIS data.

\*Significant values at 10% probability; \*\*Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: rRobust standard errors are highlighted in parentheses.



**Table A9 - Central West region: Estimates of the parameters of localization and urbanization economies by Negative Binomial: specification with controls. 2011-2013**

Sectors	Localization ( $\hat{\beta}_{Loc}^i$ )		Urbanization ( $\hat{\beta}_{Urb}^i$ )	
Meat packing and production of meat products, including dairy and fishery products	0.0719***	(0.0091)	0.6059***	(0.0086)
Sugar manufacturing and refining	-0.3035***	(0.0092)	0.4855***	(0.0068)
Other Food Products	-0.1824***	(0.0163)	0.8510***	(0.0144)
Manufacture of beverages	0.1437***	(0.0075)	0.6985***	(0.0104)
Manufacture of tobacco products	0.4648***	(0.0411)	0.4783***	(0.0084)
Manufacture of textiles	0.2775***	(0.0119)	0.4346***	(0.0112)
Manufacture of apparel and accessories	0.4336***	(0.0077)	0.2770***	(0.0083)
Manufacture of footwear and leather goods	0.1833***	(0.0096)	0.4966***	(0.0109)
Manufacture of wood products	0.1681***	(0.0127)	0.6140***	(0.0118)
Manufacture of cellulose, paper and paper products	-0.0772***	(0.0163)	0.6052***	(0.0108)
Printing and copying	0.2396***	(0.0178)	0.3767***	(0.0196)
Oil refining and coking plants	0.5904***	(0.0131)	0.4825***	(0.0099)
Biofuel production	-0.1182***	(0.0077)	0.6161***	(0.0070)
Production of other organic and inorganic chemicals, resins and elastomers	0.1292***	(0.0146)	0.7337***	(0.0105)
Manufacture of pesticides, disinfectants, paints and various chemicals	-0.0906***	(0.0116)	0.8454***	(0.0130)
Manufacture of cleaning products, cosmetics/ perfumes and toilet preparations	0.2099***	(0.0112)	0.5503***	(0.0132)
Manufacture of pharmaco-chemical and pharmaceutical products	0.1793***	(0.0073)	0.4955***	(0.0117)
Manufacture of rubber and plastic products	-0.0359**	(0.0175)	0.6502***	(0.0140)
Manufacture of non-metallic mineral products	-0.0657***	(0.0157)	0.7629***	(0.0135)
Production of pig iron/ ferrous alloys, steel and seamless steel tubes	0.012	(0.0149)	0.5900***	(0.0071)
Metallurgy of nonferrous metals and metal smelting industry	0.3494***	(0.0190)	0.5300***	(0.0170)
Manufacture of metal products, excluding machinery and equipment	-0.3177***	(0.0183)	1.0029***	(0.0173)
Manufacture of computers, electronics and optical products	0.1360***	(0.0141)	0.3961***	(0.0107)
Manufacture of electrical machinery and equipment	0.1172***	(0.0130)	0.6320***	(0.0147)
Manufacture of machinery and mechanical equipment	0.0858***	(0.0162)	0.6888***	(0.0138)
Manufacture of cars, trucks and buses, excluding parts	0.1524***	(0.0136)	0.6400***	(0.0133)
Manufacture of parts and accessories for motor vehicles	0.3511***	(0.0106)	0.4773***	(0.0122)
Manufacture of other transport equipment, excluding motor vehicles	0.3093***	(0.0155)	0.5931***	(0.0116)
Manufacture of furniture and products of various industries	0.0566***	(0.0209)	0.5863***	(0.0216)
Maintenance, repair and installation of machinery and equipment	0.0618***	(0.0147)	0.7209***	(0.0161)

Source: Authors, based on based on RAIS data.

\*Significant values at 10% probability; \*\*Significant values at 5% probability; \*\*\* Significant values at 1% probability.

Note: Robust standard errors are highlighted in parentheses.

**IMPACTOS DO CRÉDITO RURAL NO MERCADO DE TRABALHO DA  
AGROPECUÁRIA DOS ESTADOS BRASILEIROS\***

**João Paulo Rios e Silva**

Doutorando em Economia na Universidade Federal do Ceará

E-mail: joaopaulorios92@gmail.com

**Elano Ferreira Arruda**

Professor do Departamento de Economia Aplicada da Universidade Federal do Ceará

E-mail: elano@ufc.br

**RESUMO:** Este trabalho examina as repercussões do crédito rural sobre o mercado de trabalho da agropecuária dos estados brasileiros. Especificamente, são analisados os impactos desse instrumento de política sobre o nível de emprego e sobre os salários reais desse setor. Para o alcance dos objetivos, foi utilizado um painel com dados no período compreendido entre 2002 e 2015 para as 26 unidades da federação e o Distrito Federal, com as variáveis: “salários reais da agropecuária”, “número de empregos formais na agropecuária”, “valor adicionado da agropecuária” e “volume de crédito rural”. As evidências foram extraídas das funções de impulso resposta acumuladas obtidas de modelos *Panel Vector Autoregression* (PVAR). Os resultados evidenciam que os salários reais e o valor adicionado da agropecuária respondem positivamente a choques no crédito rural e que, em condições semelhantes, o número de empregos apresenta redução nos modelos agregados. Todavia, considerando modelos para subamostras regionais, há evidências de impactos positivos do crédito rural sobre o volume de empregos nos estados das regiões Norte e Nordeste, com agropecuária intensiva em trabalho, e de repercussões negativas, considerando os estados das regiões Centro-Oeste, Sul e Sudeste, ou seja, aqueles mais intensivos em capital.

**Palavras-chave:** Crédito rural; Mercado de trabalho; Salários reais; Emprego; PVAR.

**Classificação JEL:** Q14; C23; J00.

**IMPACT OF RURAL CREDIT ON THE AGRICULTURAL JOB MARKET OF  
BRAZILIAN STATES**

**ABSTRACT:** This paper analyses the repercussions of rural credit on agricultural labour market of Brazilian states, particularly on wages and employment level of this sector. For this a panel data was built with information on real wages, formal jobs created, rural credit and agricultural GDP between 2002-2015 for all 26 Brazilian states. Evidence was withdrawn from cumulative impulse response functions obtained from Panel Vector Autoregressive (PVAR) models. The results indicated that real wages and agricultural GDP respond positively to shocks in rural credit and that, under similar conditions, the employment level presents a slightly reduction in the aggregate model. However, considering models for regional subsamples, there is evidence of positive impacts of agricultural credit on employment in the states of the North and Northeast regions, with labor-intensive farming, whereas negative repercussions on employment were found in Midwest, South and Southeast states, which are more capital intensive.

**Keywords:** Agricultural credit; Labour market; Real wages; Employment; PVAR.

**JEL Codes:** Q14; C23; J00.

## 1. Introdução

A agropecuária consiste em um setor de fundamental importância para a economia brasileira. Em 2015, por exemplo, durante um período de forte retração na economia do país (na magnitude de -3,55%), o setor agropecuário cresceu 3,33%, segundo dados do IBGE. Nesse mesmo ano, a participação da agropecuária brasileira no PIB subiu para 23%.

Diante disso, o estudo das repercussões de políticas voltadas para a agropecuária brasileira se reveste de importância. A atuação do Governo, nesse setor, pode ser observada tanto de forma direta, por meio de políticas de crédito rural, seguro safra e projetos de auxílio à agricultura familiar, por exemplo, como via investimentos em infraestrutura, energia, educação e nos demais setores da economia, gerando *spillovers* para a agropecuária, impactando o nível de atividade econômica e o mercado de trabalho.

Em vasta revisão da literatura internacional sobre os impactos do crédito rural sobre a agropecuária, Yadav e Sharma (2015) observaram que seus efeitos não são consensuais. Existem estudos que destacam a efetividade e significância estatística do crédito rural sobre o PIB agropecuário em diversos países (SALEEM; JAN, 2011; SIAL et al., 2011; HARTARSKA et al., 2015) e aqueles que obtêm evidências de que esses efeitos não são conclusivos (SRIRAM, 2007; AHMAD, 2011). Considerando o mercado de trabalho, Buera et al. (2012) e Hartarska et al. (2017) observaram resultados que indicaram que o crédito rural tende a impactar positivamente os salários desse setor.

Em relação à literatura para a economia brasileira, são observados resultados igualmente não consensuais. Santos e Braga (2013) mostraram que o crédito não foi efetivo para aumentar a produtividade dos fatores no setor agrícola, exceto para a região Nordeste do Brasil. Capoiango et al. (2014) encontraram impactos positivos do crédito sobre o PIB agropecuário da região de Pirapora.

Nesse contexto, o crédito rural tem sido uma política pública importante para o setor agropecuário brasileiro. Nas regiões Sul, Sudeste e Centro-Oeste, principalmente, verifica-se uma rápida ascensão de novas tecnologias no desenvolvimento de diversas culturas e na criação de gado de forma geral. Esse processo se mostra mais lento nas regiões Norte e Nordeste, embora o desenvolvimento tecnológico tenha possibilitado a expansão e intensificação da pecuária na primeira e o cultivo da soja na região conhecida como Matopiba (Maranhão, Tocantins, Piauí e Bahia) tenha se destacado no Nordeste.

Portanto, o presente estudo examina as repercussões do crédito rural sobre o mercado de trabalho da agropecuária dos estados brasileiros. Foram analisados, especificamente, os impactos desse instrumento de política sobre o nível de emprego e sobre os salários reais. Para o alcance deste objetivo, foi utilizado um painel com dados entre 2002 e 2015 para as 26 unidades da federação e o Distrito Federal, com as variáveis: “salários reais da agropecuária”, “número de empregos formais na agropecuária”, “valor adicionado da agropecuária” e “volume de crédito rural”<sup>1</sup>. Os resultados foram observados a partir das funções de impulso resposta acumuladas (AIRF) obtidas de modelos Panel Vector Autoregression (PVAR).

Apesar de a maioria dos estudos terem como foco análises para painéis de países, Pappa (2009) adverte que esse tipo de investigação pode se mostrar mais precisa considerando dados intranacionais, ou de unidades federativas, uma vez que, nesse contexto, a política monetária e o regime cambial podem ser tomados como dados. Rickman (2010) argumenta que os modelos econômicos regionais são construídos a partir de seus congêneres nacionais e que técnicas macroeconômicas como VAR, VAR em painéis de dados, VAR estrutural (SVAR) e os modelos de Equilíbrio Geral Dinâmico Estocásticos (DSGE) têm sido subutilizados em economia regional. Bénétrix (2012) defende que esse tipo de estudo deve se pautar em dados anuais, pois esses dados são menos afetados por alterações sazonais e minimizam os chamados efeitos de antecipação dos choques no crédito.

---

<sup>1</sup> O estudo considera apenas as operações de crédito endereçadas às empresas do setor, dado que o objetivo do trabalho é examinar a repercussão desse indicador sobre os salários reais e empregos formais. Portanto, as operações ligadas ao Programa Nacional de Fortalecimento da Agricultura Familiar (PRONAF) não foram consideradas.

O presente trabalho inova ao aplicar um modelo PVAR para os estados brasileiros de modo a examinar, em um contexto regional (RICKMAN, 2010), os efeitos do crédito rural sobre os salários reais e o nível de emprego da Agropecuária Brasileira. Adicionalmente, foram considerados exercícios de robustez regionais, com o objetivo de identificar possíveis assimetrias desses impactos entre as regiões. Essa técnica se mostra conveniente, dada a possibilidade de considerar a heterogeneidade não observada no processo de estimação.

Além desta introdução, este trabalho possui mais quatro seções. A próxima consiste em destacar os aspectos teóricos e empíricos que versam sobre os impactos do crédito rural no setor agropecuário. A terceira seção reserva-se à apresentação do banco de dados e da estratégia econométrica empregada. Os resultados dos modelos são apresentados e discutidos na quarta seção. Por fim, são tecidas as considerações finais do estudo.

## 2. Revisão de literatura

O impacto do crédito rural sobre o PIB do setor agropecuário e sobre o mercado de trabalho é alvo de debate na literatura empírica. Yadav e Sharma (2015), em uma revisão da literatura sobre essa temática, concluíram que os estudos podem ser segregados em duas categoriais: *i*) O crédito agrícola possui impacto positivo e fortemente significativo sobre o produto (SALEEM; JAN, 2011; SIAL et al., 2011; HARTARSKA et al., 2015); *ii*) O impacto do crédito agrícola sobre o produto agrícola não pode ser diretamente estabelecido, isto é, não há como definir a ordem do impacto entre o crédito rural e o produto (SRIRAM, 2007; AHMAD, 2011).

Sriram (2007), através de revisão das principais medidas para estímulo à produção agrícola, dentre elas a imposição de um teto para as taxas de juros aplicadas aos produtores rurais e a duplicação da oferta de crédito rural na Índia, não encontrou uma relação causal bem definida entre essas variáveis e o produto agrícola.

Ahmad (2011) estimou um modelo ARDL para investigar os efeitos do crédito direto e indireto no setor agropecuário do Paquistão. Para isso, utilizou dados anuais entre 1972 e 2008 de produto agrícola, área plantada, força de trabalho, volume de água disponível e uma série de variáveis *dummy* para anos de estiagem. Posteriormente, utilizou testes de causalidade de Granger aumentado para investigar a relação de causalidade entre as variáveis. O autor não encontrou uma relação direta entre o crédito rural e o produto agrícola, encontrando, contudo, um efeito indireto, positivo e significativo do primeiro sobre a aquisição de maquinário, fertilizantes e sementes.

Das et al. (2009) investigaram o impacto do crédito rural sobre o PIB do setor agropecuário indiano, seja por via direta, isto é, através da oferta direta de crédito à agricultura e às atividades conjuntas, seja por via indireta, por meio de sistemas de cooperativas de fazendeiros. Para isso, utilizaram um modelo de análise dinâmica de dados em painel com variáveis instrumentais, estimado por regressão de Arellano-Bond. A análise sugeriu que o crédito direto possui impacto positivo significativo sobre o produto agregado do setor agropecuário do país, com efeito imediato, com o crédito indireto seguindo a mesma direção de impacto, mas com uma defasagem de um ano.

Sial et al. (2011), ao investigarem o impacto de diferentes variáveis sobre a produção agrícola no Paquistão, estimaram uma função de produção Cobb-Douglas por meio de uma regressão de MQO, com as variáveis transformadas na razão de hectares cultivados. Os autores utilizaram dados em séries temporais entre o período de 1972 a 2008. Os resultados apontaram efeitos positivos sobre a produção agrícola do crédito rural e da força de trabalho agrícola, dentre outras variáveis.

Tal resultado foi reforçado em Salem e Jan (2011). Os autores usaram dados compreendendo o período de 1990 a 2008 para identificar o impacto do crédito rural no PIB agrícola do distrito de Dera Ismail Khan. Foi aplicado um modelo de regressão linear em uma função Cobb-Douglas, encontrando uma correlação forte entre o PIB agrícola e o crédito direcionado à compra de sementes, fertilizantes, além de crédito direcionado a investimentos em irrigação e compra de tratores.

Ao investigarem os impactos de longo prazo do crédito direcionado a pequenos agricultores no norte da Etiópia em seu consumo e melhorias habitacionais, Berhane e Gardebroeck (2011), por meio de estimação de modelos de efeitos fixos e aleatórios com tendência, encontraram impacto

significativo e positivo do crédito em melhorias habitacionais e no nível de consumo. De acordo com os dados da pesquisa, verificou-se que empréstimos sucessivos também geram efeitos cumulativos sobre essas variáveis no longo prazo.

É fundamental, contudo, analisar os efeitos de políticas direcionadas do crédito e do microfinanciamento sobre diferentes ambientes macroeconômicos. Buera et al. (2012) avaliaram quantitativamente o impacto agregado e distributivo de programas de crédito direcionado em pequenos negócios. Através de um modelo de equilíbrio geral com dados de vários países, os autores encontram que o efeito redistributivo do crédito direcionado foi observado mais fortemente em equilíbrio geral. Além disso, a Produtividade Total dos Fatores (TFP) se eleva com o crédito subsidiado sob equilíbrio geral.

Ainda segundo Buera et al. (2012), quando efeitos de equilíbrio geral são levados em conta, a ampliação do programa de microfinanciamento leva apenas a um pequeno incremento na renda *per capita*. Isso ocorre porque o aumento da TFP é contrabalanceado pela menor acumulação de capital resultante da redistribuição de renda de grandes para pequenos poupadores. No entanto, a grande maioria da população será positivamente afetada pelo microfinanciamento através do aumento dos salários de equilíbrio.

Ao redor do mundo, o crédito rural tem sido ofertado por agências de crédito governamentais, por vezes interligadas a um sistema nacional de crédito rural. Turvey e Wang (2012) tentaram precificar o valor da garantia de títulos emitidos por agências de financiamento pertencentes ao Sistema de Crédito Rural dos Estados Unidos, através de um modelo Black-Scholes, com o objetivo de extrair a volatilidade dos títulos garantidos e substituí-la pela volatilidade no nível de preços da terra. Os resultados evidenciaram que a perda dessa garantia implícita resultaria em um impacto mínimo nos títulos de curto prazo, porém com resultados bem mais expressivos em longo prazo. Os autores concluíram que a perda da garantia governamental de títulos de crédito rural afetaria o equilíbrio competitivo entre os ofertantes de crédito.

Ammani (2012) investigou a relação entre a produção agrícola e a oferta formal de crédito na Nigéria. A metodologia empregada no estudo envolveu o desenvolvimento e a estimação de três modelos de regressão simples relacionando a produção agrícola ao crédito formal, utilizando dados anuais de produção da agricultura, pesca e pecuária nigeriana, compreendendo o período de 1991 a 2009 e mantendo-se outras variáveis explicativas constantes. Os resultados do trabalho indicaram que o crédito formal está positivamente e significativamente relacionado com a produtividade dos setores agrícola, pecuário e pesqueiro da agricultura nigeriana.

Por meio da aplicação de restrição ao crédito e utilizando o método de Propensity Score Matching para microdados do Censo Agropecuário de 2006, Santos e Braga (2013) mostraram que o crédito não foi efetivo para aumentar a produtividade dos fatores no setor agrícola, exceto para a região nordeste do Brasil. Os resultados apontaram para a necessidade de melhorias na política de crédito rural no Brasil. Capobianco et al. (2014) utilizaram dados em painel compreendendo o período 2000-2004, com o objetivo de verificar os efeitos econômicos do crédito rural na região de Pirapora. Eles encontraram impactos positivos deste sobre o PIB agropecuário, além de efeitos subjacentes em outros setores da economia da região.

Buscando medir a eficiência técnica relativa das associações de fomento e bancos pertencentes ao Sistema de Crédito Rural americano, Dang et al. (2014) calcularam a fronteira estocástica de produção dessas instituições. Os autores encontraram uma grande diferença entre grandes e pequenas associações e bancos, com estes últimos possuindo maior eficiência do que as primeiras. A crise financeira não demonstrou nenhum efeito significativo sobre o crescimento dos rendimentos das fazendas e da utilização de bioenergia.

Utilizando um sistema de equações simultâneas para dados em corte transversal e séries temporais coletados entre abril e junho de 2013, Rosari et al. (2014) verificaram que a alocação de crédito nas Pequenas Ilhas de Sonda Orientais, na Indonésia, proporcionou aumento na criação de gado, além de elevações no nível de consumo e investimento, ressaltando que as políticas de crédito direcionado aos agricultores familiares daquela região deveriam ser melhor ajustadas aos seus parâmetros.

Nnmamocho e Eke (2015) investigaram o efeito do crédito bancário na produção agrícola na Nigéria usando um Modelo de Correção de Erros (ECM). Dados anuais (1970-2013) obtidos do Banco Central da Nigéria evidenciaram que todas as variáveis foram integradas de ordem um (1) e que há relação de longo prazo entre elas. No entanto, embora o crédito bancário e a produção industrial tenham contribuído muito para a produção agrícola na Nigéria no longo prazo, apenas a produção industrial influenciou a produção agrícola no curto prazo.

Avaliando a relação entre o crédito ofertado por grandes agentes pertencentes ao Sistema de Crédito Rural americano e o crescimento econômico do setor agropecuário entre o período de 1991-2010, por meio de modelos de efeitos fixos aplicados em dois bancos de dados em painel, Harstarska et al. (2015) encontraram uma associação positiva entre crédito e PIB agropecuário por residente rural. Dados regionais confirmam esse resultado. Witte et al. (2015) modelaram o volume de crédito rural em função da distância das agências de crédito, usando tal modelo para realizar projeções a respeito do impacto da alocação de agências em áreas carentes no condado de Oklahoma. Os resultados evidenciaram que a distância das agências ainda é um fator central para o acesso ao crédito rural.

Garcias e Kassouf (2016) avaliaram o impacto de restrições ao crédito rural sobre a produtividade da terra e a produtividade do trabalho dos agricultores familiares do Brasil, utilizando dados do Censo agropecuário brasileiro de 2006 por município, aplicados a um modelo de *Propensity Score Matching*. Os resultados apontaram para um efeito positivo do crédito sobre a produtividade do trabalho e da terra. Porém, o valor desse impacto é distinto entre diferentes níveis de mercantilização dos agricultores, requerendo, segundo os autores, políticas distintas.

Hartarska et al. (2017) revisitaram o tema dos efeitos do crédito rural nos rendimentos das fazendas tomadoras de crédito, provendo evidências do impacto positivo do crédito sobre os rendimentos. Os autores construíram um painel para o período 1991-2010 com as demonstrações financeiras pertencentes ao Sistema de Crédito Rural americano e avaliaram como os empréstimos concedidos pelas instituições pertencentes ao sistema afetaram os rendimentos agrícolas e a produção agrícola. Os autores utilizaram estimativas de efeitos fixos e controlaram efeitos da oferta de crédito de outros credores agrícolas, bem como os efeitos de estoque de capital, preços e taxas de juros. Os resultados mostraram uma relação positiva entre crédito rural e rendimento das fazendas, mostrando que essa relação era mais concisa nos anos 90 do que nos anos 2000.

Gasques et al. (2017) analisaram os impactos do crédito rural em diversas variáveis, como valor produto da produção, produto interno bruto da agropecuária, produto interno bruto do agronegócio e produtividade total dos fatores. Foram ajustados modelos de Função de Transferência, cuja variável de interesse é explicada por componentes autorregressivos e de média móvel, além de variáveis relacionadas. Os autores encontraram efeitos significativos sobre as variáveis de desempenho da agricultura, bem como a produtividade total dos fatores, relacionada à mudança tecnológica.

Apesar da vasta literatura observada, percebe-se certa ausência de análises em perspectivas intranacionais, ou regionais, linha adotada por este estudo, conforme sugere Rickman (2010). Além disso, o presente trabalho teve como foco os efeitos de choques no crédito rural sobre o mercado de trabalho da agropecuária dos estados brasileiros, considerando os impactos sobre salários reais e o nível de emprego do setor.

### 3. Aspectos metodológicos

#### 3.1 Descrição dos dados

Para cotejar as repercussões do crédito rural sobre o mercado de trabalho da agropecuária dos estados brasileiros, foi utilizado um painel de dados com informação das 26 Unidades Federativas e do Distrito Federal entre 2002 e 2015. O Quadro 1 apresenta uma síntese das variáveis utilizadas.

Os dados de salário real médio da agropecuária por estado foram obtidos junto à Relação Anual de Informações Sociais (RAIS). Para o nível de emprego, empregou-se uma variável *proxy* dada pelo número de vínculos estabelecidos em um ano, cujos dados também foram coletados da RAIS.

**Quadro 1 - Descrição das variáveis**

Variável	Definição	Fonte
Credito rural	Crédito Rural: volume de crédito rural disponibilizado no Sistema de Operações do Crédito Rural e Proagro.	BCB
Nível de Emprego na Agropecuária	Emprego: utiliza-se o número de vínculos da agropecuária como proxy para o nível de emprego do setor.	RAIS
Salário Real Médio	Salário Real Médio por Trabalhador do setor agropecuário: representa a média de rendimentos mensais recebidos por trabalhador em cada ano.	RAIS
PIB da Agropecuária	PIB da Agropecuária: utiliza-se como proxy o Valor Adicionado Bruto da Agropecuária por estado.	IBGE

Fonte: Elaboração própria.

Nota: 1. As variáveis de PIB e Crédito Rural foram deflacionadas pelo Índice Geral de Preço - Disponibilidade Interna (IGP-DI) e o salário real médio foi deflacionado pelo deflator de rendimentos da PNAD.

Dados anuais sobre o valor adicionado bruto da agropecuária dos estados brasileiros foram obtidos junto ao banco de dados do Instituto Brasileiro de Geografia e Estatística (IBGE) disponibilizado pelo Instituto de Pesquisa Econômica Aplicada (IPEADATA).

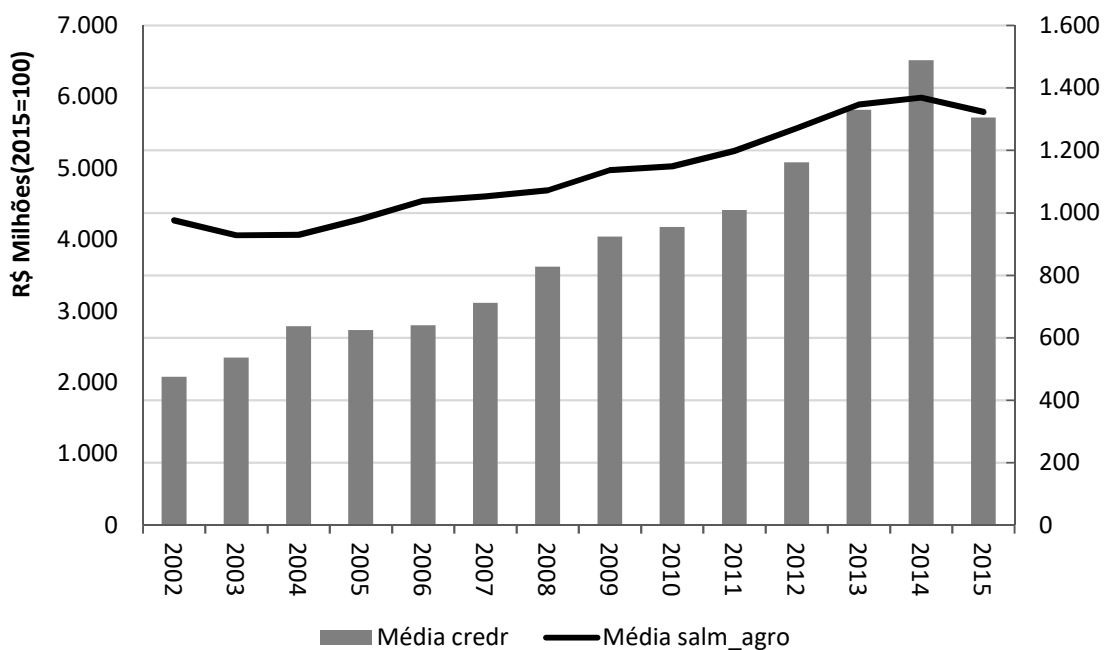
As informações a respeito da oferta de crédito rural foram retiradas do Anuário Estatístico do Crédito Rural, disponibilizado pelo Banco Central do Brasil (BCB). Vale ressaltar que foram consideradas apenas as operações de crédito endereçadas às empresas do setor, dado que o objetivo do trabalho é examinar a repercussão desse indicador sobre os salários em empregos formais. Portanto, os dados de crédito do Programa Nacional de Fortalecimento da Agricultura Familiar (PRONAF) não foram considerados por esse se tratar de um programa voltado ao atendimento da demanda por crédito para agricultores familiares e, portanto, sem vínculo empregatício.

É possível elencar benefícios decorrentes do uso de dados para unidades da federação (UFs) dentro de um mesmo país. Com dados dessa natureza, a política monetária pode ser tomada como dada na análise e, tendo em vista que dados comparáveis internacionais sobre variáveis creditícias são difíceis de encontrar, esse tipo de análise pode caracterizar de forma transversal a dinâmica de transição do choque no crédito rural e dar algumas indicações do que se deveria esperar encontrar em grandes regiões econômicas de mesma moeda (PAPPA, 2009). Outra vantagem no uso de dados anuais é o fato de serem menos afetados por mudanças sazonais (BÉNÉTRIX, 2012).

A Figura 1 e 2 apresentam a evolução média anual do volume de crédito e dos salários reais do setor agropecuário, além do PIB agropecuário e do número de vínculos. Há uma clara tendência crescente do volume médio anual de crédito rural ofertado, bem como dos salários reais e do PIB agropecuário. Já em relação ao número de vínculos, verifica-se uma desaceleração, a partir do ano de 2008.

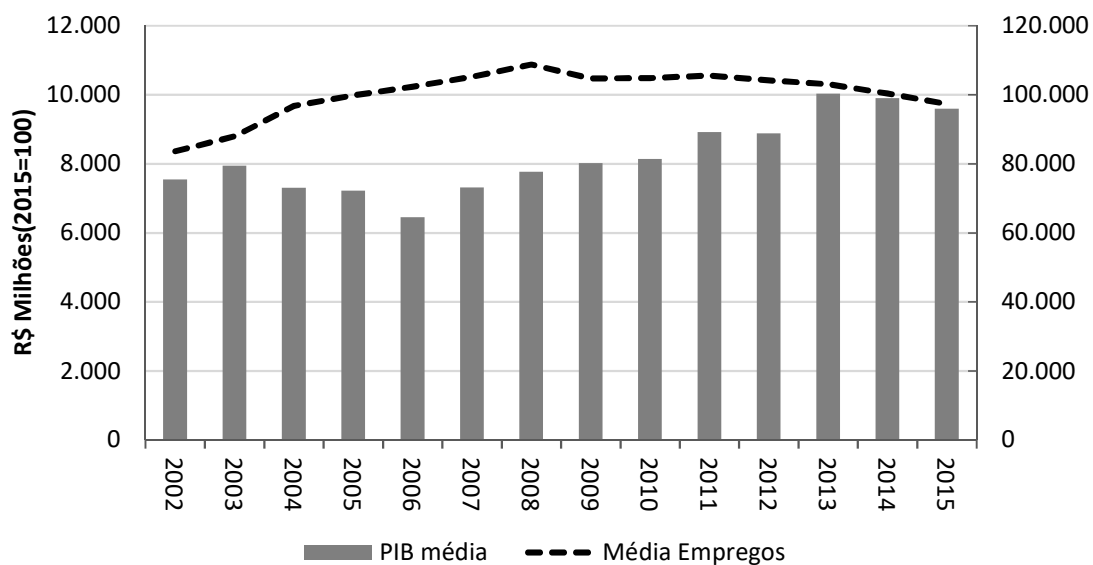
As estatísticas descritivas das variáveis estão expostas na Tabela 1. É bastante notória a grande amplitude existente entre as observações de crédito, PIB e número de empregos. O salário real teve média de R\$ 1.126,56, sendo a variável de menor desvio padrão. O crédito rural apresentou média de R\$ 3,94 bilhões, possuindo um desvio padrão de R\$ 5,98 bilhões de reais. A variável de PIB agropecuário teve média de aproximadamente R\$ 8,2 bilhões, com desvio padrão de R\$ 8,7 bilhões, enquanto o número de vínculos variou entre 569 e 804.888, com média de 100.320 empregos formais.

**Figura 1 - Evolução da média dos indicadores empregados: salário**



Fonte: Elaboração própria.

**Figura 2 - Evolução da média dos indicadores empregados: emprego**



Fonte: Elaboração própria.

**Tabela 1 - Estatísticas descritivas**

Variável	Média	Desvio-Padrão	Amplitude	
			Mínimo	Máximo
Crédito	3.944.460.316,94	5.980.222.410,57	6.359.957,21	28.149.659.385,74
PIB Agro	8.214.045.577,16	8.773.690.644,73	73.687.984,69	39.821.655.058,95
Emprego	100.320,00	157.425,00	569,00	804.888,00
Salário Agro	1.126,56	263,28	653,79	2.353,73

Fonte: Elaboração própria.



Além disso, em virtude de o Brasil ser um país de dimensões continentais, com grande heterogeneidade e disparidade regional, faz-se necessária uma análise descritiva em uma perspectiva regional com vistas a contribuir com informações importantes para a análise dos resultados dos modelos econométricos.

Nesse sentido, levando em consideração as distintas características estruturais concernentes ao mercado de trabalho agropecuário das regiões brasileiras, a amostra foi dividida em duas macrorregiões: a primeira, chamada Norte-Nordeste (NN), compreendendo estados do Norte e Nordeste do país, onde se verificam menores rendimentos e PIB agropecuários; a segunda, chamada Centro-Sul (CS), contendo estados das regiões Sul, Sudeste e Centro-Oeste, com maiores rendimentos e uso intensivo de tecnologia na produção agrícola. As Tabelas 2 e 3 expressam as estatísticas descritivas de cada macrorregião.

**Tabela 2 - Estatísticas descritivas da subamostra NN**

Variável	Média	Desvio-Padrão	Amplitude	
			Mínimo	Máximo
Crédito	802.925.209,65	1.031.683.180,82	6.359.957,21	6.675.962.698,85
PIB Agro	4.012.341.183,96	4.008.140.816,93	73.687.984,69	17.835.814.000,00
Emprego	34.892,00	43.132,00	569,00	188.556,00
Salário Agro	1.049,22	201,85	653,79	1.619,96

Fonte: Elaboração própria.

É possível observar que os menores valores para PIB, volume de crédito rural, salários reais e vínculos da amostra completa, expostos na Tabela 1, estão presentes na macrorregião NN, enquanto os maiores valores encontram-se na região CS. Além disso, a macrorregião CS apresenta um maior grau de dispersão em todas as variáveis, quando comparadas aos seus pares da subamostra NN.

**Tabela 3 - Estatísticas descritivas da subamostra CS**

Variável	Média	Desvio-Padrão	Amplitude	
			Mínimo	Máximo
Crédito	8.513.965.927,53	7.141.305.057,22	135.942.892,2	28.149.659.385,7
PIB Agro	14.325.615.603,6	10.126.563.230,01	365.102.687,8	39.821.655.058,9
Emprego	195.488,00	206.979,00	7.883,00	804.888,00
Salário Agro	1.239,07	299,20	693,92	2.353,73

Fonte: Elaboração própria.

### 3.2 Estratégia econométrica

Para examinar os efeitos dos choques no crédito rural sobre os salários reais e o número de vínculos da agropecuária nos estados brasileiros, foi utilizada uma adaptação do modelo empírico PVAR nos moldes de Bénétrix (2012).

$$Z_{i,t} = B(L)Z_{i,t-1} + DX_{i,t} + u_{it} \quad (1)$$

A forma estrutural desse modelo é dada por  $A_0Z_{i,t} = A(L)Z_{i,t-1} + CX_{i,t} + \varepsilon_{it}$ , onde  $A(L) = A_0B(L)$ ,  $C = A_0D$ ,  $\varepsilon_{i,t} = A_0u_{i,t}$  e, ainda,

$$Z_{i,t} = \begin{pmatrix} R_{i,t} \\ Y_{i,t} \\ W_{i,t} \end{pmatrix}; A_0 = \begin{pmatrix} 1 & -a_{yr} & -a_{wr} \\ -a_{ry} & 1 & -a_{wy} \\ -a_{rw} & -a_{yw} & 1 \end{pmatrix}; X_{i,t} = \begin{pmatrix} c_i \\ t_t \\ d_t \end{pmatrix}; \varepsilon_{i,t} = \begin{pmatrix} \varepsilon_{i,t}^g \\ \varepsilon_{i,t}^y \\ \varepsilon_{i,t}^w \end{pmatrix} \quad (2)$$

onde  $i$  e  $t$  denotam o Estado e o ano, respectivamente.  $Z_{i,t}$  é o vetor de variáveis endógenas, quais sejam: o volume de crédito rural,  $R_{i,t}$ , o PIB agropecuário estadual,  $Y_{i,t}$  e o salário real,  $W_{i,t}$ . A matriz  $A_0$  captura as relações contemporâneas entre os regressores. A matriz  $A(L)$  é a matriz que incorpora os efeitos das variáveis endógenas e suas defasagens. Todas as variáveis estão expressas em logaritmo.

Como se utiliza um modelo com dados em painel, é importante levar em consideração a heterogeneidade não observada dos estados. Por isso, foi utilizado o estimador de efeitos fixos, por mínimos quadrados com variáveis *dummies* (LSDV) para dados em painel, em que  $c_i$ , no vetor  $X_{i,t}$ , incorpora o efeito específico de cada unidade *cross-section*. Adicionalmente, no vetor  $X_{i,t}$ , podem ser incorporadas tendências temporais,  $t_t$ , e, ainda, *dummies* indicativas para mudanças de conjuntura na economia,  $d_t$ . Por fim,  $\varepsilon_{i,t}$  contém os choques estruturais ortogonais.

Em todo modelo sistêmico multiequacional, é importante definir uma estratégia para a identificação do modelo. Existem diversas formas de se buscar a identificação de modelos dessa natureza, desde as que simplesmente impõem estruturas nos parâmetros, de modo *ad hoc*, a outras que fazem uso da teoria econômica para tal. A estratégia empregada neste estudo mescla esses dois elementos e é tradicionalmente utilizada em modelos que avaliam efeitos da política fiscal na atividade econômica (BLANCHARD; PEROTTI, 2002; PEROTTI, 2005; MONACELLI; PEROTTI, 2010) e no mercado de trabalho (BÉNÉTRIX, 2012); ou seja, o presente estudo faz uma adaptação dessa estratégia de identificação para a política de crédito rural (que tem seu mecanismo de transmissão para a atividade econômica e mercado de trabalho via investimentos).

Portanto, a estratégia de identificação adotada no processo de estimação foi o ordenamento recursivo. Assume-se que o crédito rural não é afetado pelas demais variáveis dentro do mesmo ano, que o PIB agropecuário é explicado pelo crédito rural e que os salários reais são influenciados por ambos os indicadores. Noutros termos, admite-se que  $a_{yr} = a_{wr} = a_{wy} = 0$ . Assim, o canal de atuação do crédito rural no PIB agropecuário e no mercado de trabalho se dá via aumento dos investimentos no setor.

Após a estimação, as evidências das respostas dos salários reais e do número de vínculos aos choques no crédito rural serão extraídas a partir das funções de impulso resposta acumuladas (AIRF). A opção pela versão acumulada reside no fato de que ela representa o efeito permanente do choque analisado. Os intervalos de confiança para as AIRF foram obtidos por meio do procedimento de *bootstrap* proposto por Hall (1992).

Em suma, a estratégia econométrica pode ser assim resumida: inicialmente, foi analisada a ordem de integração das séries. Em seguida, procederam-se as estimações dos modelos PVAR. Nesta etapa, foram estimados dois modelos, um contendo as variáveis de salários reais, PIB do setor agropecuário e crédito rural e outro com as variáveis de empregos, PIB agropecuário e crédito rural, utilizando o painel com observações das 26 unidades federativas mais o Distrito Federal no período de 2002 a 2015, de onde foram extraídas as funções de impulso resposta acumuladas (AIRF). Por fim, exercícios regionais foram analisados em duas subamostras: Norte-Nordeste (NN) e Centro-Sul (regiões Sul, Sudeste e Centro-Oeste) (CS), com vistas a identificar possíveis assimetrias regionais. As tabelas com os resultados de todos os modelos estimados encontram-se disponíveis em apêndice.

#### 4. Resultados

Para verificar os efeitos dos choques no crédito rural sobre os salários reais do setor agropecuário, foi analisada, inicialmente, a ordem de integração das séries utilizadas através dos testes de Levin et al. (2002), que tinham como hipótese nula a presença de uma raiz unitária comum a todas as unidades

no painel e Im et al. (2003), que testaram a presença de raiz unitária individual nos *cross-section*. Os resultados estão sintetizados na Tabela 4.

**Tabela 4 - Resultado dos testes de raiz unitária**

Variável	Teste		
	Levin, Lin e Chu	Im, Pesaran e Shin	Resultado
Ln(Credr)	-6,94* (0,00)	-1,93* (0,02)	I(0)
Ln(PIB_agro)	-8,36* (0,00)	-1,86* (0,03)	I(0)
Ln(wagro)	-4,96* (0,00)	-5,65* (0,00)	I(0)
Ln(Emp)	-11,49* (0,00)	-2,34* (0,01)	I(0)

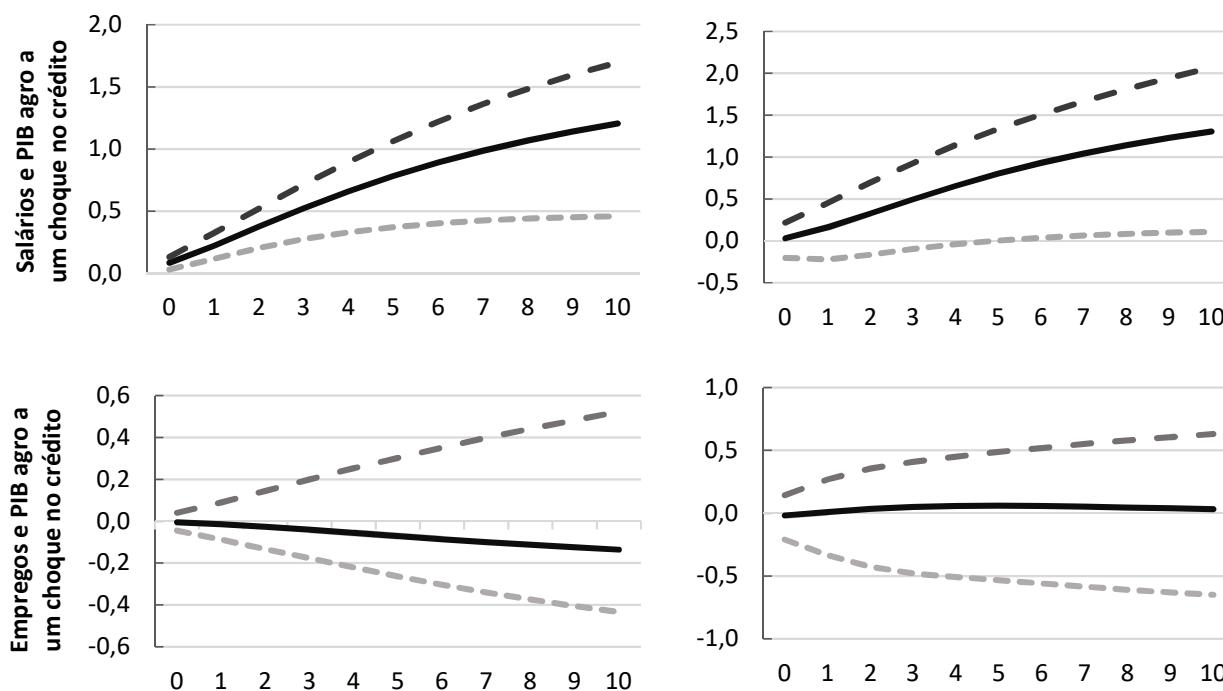
Fonte: Elaboração própria a partir dos resultados obtidos.

Nota: Valor p entre parênteses. \* Significante ao nível de 5%.

Nota 2: O teste de raiz unitária para a *proxy* de PIB agropecuário foi realizado com tendência.

Os testes de Levin et al. (2002) e Im et al. (2003) indicaram que todas as variáveis se mostraram estacionárias em nível. Portanto, foram estimados dois modelos PVAR, como exposto na seção anterior. Em seguida, partiu-se para a análise das AIRF's em todos os modelos. Os resultados estão dispostos na Figura 3. A primeira linha de gráficos descreve as AIRF do modelo estimado com salários reais e a segunda apresenta os resultados para o modelo com o nível de emprego.

**Figura 3 - AIRF's dos modelos de salários reais e nível de emprego na agropecuária**



Fonte: Elaboração própria a partir dos resultados obtidos.

Pode-se observar no modelo que leva em consideração os salários reais do setor agropecuário, primeira linha de gráficos, um efeito acumulado positivo em dez períodos, tanto para os salários reais

do setor agropecuário quanto para o PIB agropecuário, com as duas variáveis apresentando um efeito contemporâneo positivo. Já em relação ao modelo que leva em consideração o nível de emprego, ele evidencia um efeito acumulado levemente negativo para essa variável e um impacto positivo sobre o PIB da agropecuária.

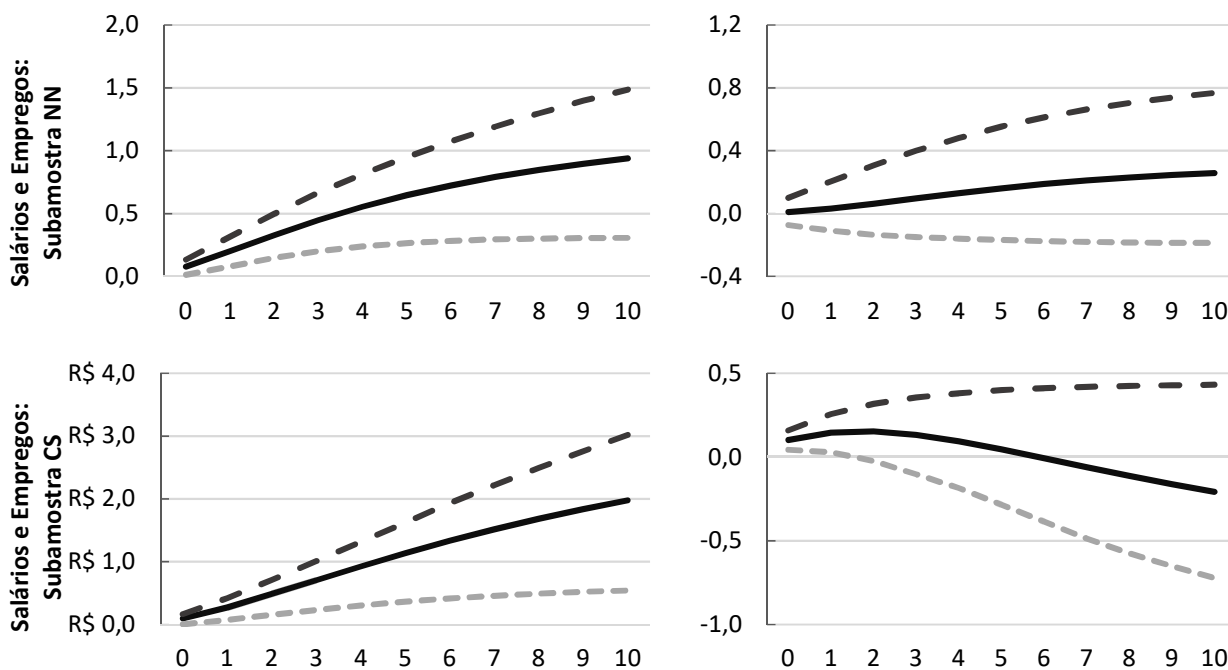
Vale destacar que autores como Buera et al. (2012) e Hartarska et al. (2017) também apontaram para impactos positivos do crédito rural sobre os salários reais desse setor e que Saleem e Jan (2011), Sial et al. (2011), Capobiango et al. (2012) e Hartarska et al. (2015) também encontraram efeitos positivos do crédito rural sobre o PIB agropecuário.

Todavia, autores como Alvarez (2020) advertem que os efeitos do crédito rural sobre o mercado de trabalho desse setor podem ser distintos/assimétricos, dependendo da intensidade tecnológica empregada na região. Em outras palavras, em áreas com maior intensidade tecnológica pode ocorrer um efeito substituição de trabalhadores por máquinas, podendo isso refletir em um impacto negativo sobre o nível de emprego, ainda que os salários reais tendam a ser maiores, dada a maior produtividade. Em regiões que empregam menos tecnologia, os recursos tendem a ser aplicados em maiores contratações, uma vez que, nesse contexto, o setor se mostra intensivo em mão de obra.

Portanto, dado o resultado agregado levemente negativo observado anteriormente, faz-se necessária uma análise para subamostras regionais, tendo em vista a grande disparidade tecnológica existente na atividade agropecuária entre as diferentes regiões do país.

Nesse sentido, foram estimados mais quatro modelos. Os dois primeiros consideram apenas os estados das regiões Norte e Nordeste (NN) (para salários reais e vínculos), enquanto os dois últimos promovem análise semelhante para os estados das regiões Sul, Sudeste e Centro-Oeste (Centro Sul - CS). A escolha dessa subdivisão se deve pelo fato de que, via de regra, os estados que compõem a subamostra CS são mais tecnologicamente intensivos em suas economias do que os pertencentes à subamostra NN. Os resultados estão sintetizados na Figura 4. Os gráficos à esquerda reportam a resposta dos salários reais a choques no crédito rural; já os gráficos à direita apresentam as mesmas evidências para o indicador de nível de emprego. A primeira linha de gráficos refere-se aos modelos para a macrorregião NN e a segunda aos resultados para a macrorregião CS.

**Figura 4 - AIRF's das subamostras regionais**



Fonte: Elaboração própria a partir dos resultados obtidos.

É possível observar que, embora a resposta acumulada em dez períodos dos salários reais ao crédito seja positiva nas estimações para as duas subamostras, a resposta acumulada para o nível de emprego para a macrorregião NN é positiva, enquanto para CS é negativa.

Uma das possíveis razões para a ocorrência de uma resposta negativa do número de vínculos na macrorregião Centro-Sul a um choque positivo no volume de crédito pode ser um grau maior de mecanização da agropecuária nos estados que a compõem. Uma elevação no crédito pode gerar uma intensificação do efeito substituição do trabalho por capital, o que reduz a demanda por trabalho no setor. Considerando essa hipótese, tem-se como corolário que o aumento dos salários reais nessa macrorregião seja melhor explicado por ganhos de produtividade, quando comparada com a outra subamostra (ALVAREZ, 2020).

Em suma, os resultados apontam de modo robusto e alinhado com a literatura que o crédito rural influencia positivamente o PIB agropecuário e os salários reais do setor. Entretanto, considerando o nível de emprego, esses efeitos se mostram assimétricos e dependentes do grau tecnológico empregado na região; ou seja, regiões com agropecuária mais mecanizada tendem a substituir mão de obra; já áreas intensivas em mão de obra tendem a contratar mais trabalhadores.

## 5. Considerações finais

O presente estudo investigou os efeitos dos choques no crédito rural sobre os salários reais e o nível de emprego da agropecuária em uma perspectiva regional para os estados brasileiros, por meio de dados entre 2002-2015 e modelos VAR em painel, estimados por mínimos quadrados com variáveis *dummies* (LSDV).

Em termos gerais, os resultados apontaram de modo consistente que o crédito rural apresenta impacto positivo sobre os salários reais e sobre o PIB da agropecuária dos estados brasileiros. A primeira evidência também foi observada por Buera et al. (2012) e Hartarska et al. (2017) e autores como Saleem e Jan (2011), Sial et al. (2011), Capobianco et al. (2012) e Hartarska et al. (2015) também reportam efeitos benéficos do crédito rural sobre o PIB agropecuário.

Os resultados dos modelos agregados apontaram ainda uma resposta negativa do nível de emprego em resposta ao aumento do crédito rural. Porém, estudos como o de Alvarez (2020) indicam que esses efeitos parecem depender da intensidade tecnológica empregada no setor. Diante disso, análises regionais foram realizadas e os resultados apontaram para a existência de assimetrias nos efeitos do crédito rural sobre o emprego nas regiões brasileiras. As evidências reportam impactos positivos no emprego para os estados das regiões Norte e Nordeste, com agropecuária intensiva em trabalho, e repercussões negativas considerando os estados das regiões Centro-Oeste, Sul e Sudeste, ou seja, aqueles mais intensivos em capital.

Portanto, as evidências deste estudo sugerem que o crédito rural para empresas se mostra uma ferramenta estratégica para o crescimento da agropecuária dos estados brasileiros. Aponta-se para a importância da melhoria do capital humano do setor, de modo a produzir empregos com maiores salários.

Cabe ressaltar que as linhas de crédito rural são designadas para fins distintos. As atividades financiadas pelo crédito rural para empresas examinadas nesse trabalho compreendem o custeio de despesas ligadas à produção, ao investimento em máquinas e equipamentos necessários à produção agropecuária e à comercialização de produtos agrícolas. Assim, os resultados indicam que o crédito rural vem contribuindo com o estímulo ao investimento no setor, para a geração de renda e de empregos formais (na macrorregião NN). Todavia, o presente estudo não tem a ambição de exaurir todos os efeitos de políticas de crédito rural; ou seja, sugere-se, em trabalhos futuros, uma avaliação dessa natureza para os efeitos do PRONAF que também é um importante instrumento nesse sentido, focado na agricultura familiar.

**Referências**

- AHMAD, N. Impact of institutional credit on agricultural output: a case study of Pakistan. *Theoretical and Applied Economics*, v. 10, n. 10, p. 99-120, 2011.
- ALVAREZ, J. A. The agricultural wage gap: evidence from Brazilian micro-data. *American Economic Journal: Macroeconomics*, v. 12, n. 1, p. 153-173, 2020.
- AMMANI, A. A. An investigation into the relationship between agricultural production and formal credit supply in Nigeria. *International Journal of Agriculture and Forestry*, v. 2, n. 1, p. 46-52, 2012.
- BÉNÉTRIX, A. S. Fiscal shocks and real wages. *International Journal of Finance and Economics*, v. 17, n. 3, p. 203-220, 2012.
- BERHANE, G.; GARDEBROEK, C. Does Microfinance reduce rural poverty? Evidence based on household panel data from Northern Ethiopia. *American Journal of Agricultural Economics*, v. 93, n. 1, p. 43-55, 2011.
- BLANCHARD O.; PEROTTI R. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics*, v. 117, n. 4, p. 1329-68, 2002.
- BUERA, J.; KABOSKI, P.; SHIN, Y. The macroeconomics of microfinance. *National Bureau of Economic Research*, Cambridge, MA, 2012. (Working Paper, n. 17905).
- CAPOBIANGO, R. P.; BRAGA, M. J.; SILVEIRA, S. F. R.; COSTA, C. C. M. Análise do Impacto econômico do crédito rural na microrregião de Piraropora. *Revista de Economia e Sociologia Rural*, v. 50, n. 4, p. 631-644, 2012.
- DANG, T.; LEATHAM, D.; MCCARL, B. A. Wu, X. Measuring the efficiency of the farm credit system. *Agricultural Finance Review*, v. 74, n. 1, p. 38-54, 2014.
- DAS, A.; SENAPATI, M.; JOHN, J. Impact of agricultural credit on agriculture production: an empirical analysis in India. *Reserve Bank of India Occasional Papers*, v. 30, n. 2, p. 75-107, 2009.
- GARCIAS, M.; KASSOUF, A. L. Assessment of rural credit impact on land and labor productivity for Brazilian family farmers. *Revista de Política Agrícola*, v. 26, n. 3, p. 721-746, 2016.
- GASQUES, J. C.; BACCHI, M.; BASTOS, E. Impactos do crédito rural sobre as variáveis do agronegócio. *Revista de Política Agrícola*, v. 26, n. 4, p. 132-140, 2017.
- HALL, P. *The bootstrap and Edgeworth expansion*. New York: Springer, 1992.
- HARTARSKA, V.; NADOLNYAK, D.; SHEN, X. Agricultural credit and economic growth in rural areas. *Agricultural Finance Review*, v. 75, n. 3, p. 302-312, 2015.
- HARTARSKA, V.; NADOLNYAK, D.; SHEN, X. Farm income and output and lending by the farm credit system. *Agricultural Finance Review*, v. 77, n. 3, p. 125-136, 2017.
- IM, K. S.; PESARAN, M. H.; SHIN, Y. Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, v. 115, n. 1, p. 53-74, 2003.
- LEVIN, A.; LIN, C. F.; CHU, C. S. J. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, v. 108, n. 1, p. 1-24, 2002.
- MONACELLI, T.; PEROTTI, R. Fiscal policy, the real exchange rate and traded goods. *The Economic Journal*, v. 120, n. 544, p. 437-461, 2010.

- NNAMOCHA, P. N.; EKE, C. N. Bank credit and agricultural output in Nigeria (1970 - 2013): an error correction model (ECM) approach. *British Journal of Economics, Management & Trade*, v. 10, n. 2, p. 1-12, 2015.
- PAPPA, E. The effects of fiscal shocks on employment and the real wage. *International Economic Review*, v. 50, n. 1, p. 217-244, 2009.
- PEROTTI R. Estimating the effects of fiscal policy in OECD countries. Centre for Economic Policy Research, 2005. (CEPR Discussion Paper, n. 4842).
- RICKMAN, D. S. Modern macroeconomics and regional economic modeling. *Journal of Regional Science*, v. 50, n. 1, p. 23-41, 2010.
- ROSARI, B. B. de; SINAGA, B. M.; KUSNADI, N.; SAWIT, M. H. The impact of credit and capital supports on economic behavior of farm households: a household economic approach. *International Journal of Food and Agricultural Economics*, v. 2, n. 6, p. 81-90, 2014.
- SALEEM, M. A.; JAN, F. A. The impact of agricultural credit on agricultural productivity in Dera Ismail Khan (District) Khyber Pakhtonkhawa Pakistan. *European Journal of Business and Management*, v. 3, n. 2, p. 38-44, 2011.
- SANTOS, R. B. N.; BRAGA, M. J. Impactos do crédito rural na produtividade da terra e do trabalho nas regiões brasileiras. *Economia Aplicada*, v. 17, n. 3, p. 299-324, 2013.
- SIAL, M. H.; AWANM, S.; WAQAS, M. Role of institutional credit on agricultural production: a time series analysis of Pakistan. *International Journal of Economics and Finance*, v. 3, n. 2, p. 126-132, 2011.
- SRIRAM, M. S. Productivity of rural credit: a review of issues and some recent literature. *International Journal of Rural Management*, v. 3, n. 2, p. 245-268, 2007.
- TURVEY, C. G.; WANG, Y. The effects of government sponsored enterprise (GSE) status on the pricing of bonds issued by the federal farm credit banks funding corporation (FFCB). *Agricultural Finance Review*, v. 72, n. 3, p. 488-506, 2012.
- WITTE, T.; DEVUYST, E.; WHITACRE, B.; JONES, R. Modeling the impact of distance between officer and borrowers on agricultural loan volume. *Agricultural Finance Review*, v. 75, n. 4, p. 848-498, 2015.
- YADAV, P.; SHARMA, A. K. Agricultural credit in developing economies: a review of relevant literature. *International Journal of Economics and Finance*, v. 7, n. 12, p. 219-244, 2015.


## Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro para a realização desta pesquisa.

## ORCID

João Paulo Rios e Silva  <https://orcid.org/0000-0003-1392-0933>

Elano Ferreira Arruda  <https://orcid.org/0000-0002-4207-2921>

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.

## Apêndice

Tabela 5 - Modelos estimados com amostra total

Modelo com salários reais				Modelo com número de empregos			
Efeitos Contemporâneos				Efeitos Contemporâneos			
	$Crédito_t$	$PIB_t$	$w_t$		$Crédito_t$	$PIB_t$	$Emp_t$
$Crédito_t$	1	0	0	$Crédito_t$	1	0	0
$PIB_t$	0,029	1	0	$PIB_t$	0,018	1	0
$w_t$	0,084	0,074	1	$Emp_t$	-0,005	0,0008	1
Equações com Efeitos Defasados				Equações com Efeitos Defasados			
	$Crédito_t$	$PIB_t$	$w_t$		$Crédito_t$	$PIB_t$	$Emp_t$
$Crédito_{t-1}$	0,646* (0,00)	0,08** (0,04)	0,077* (0,00)	$Crédito_{t-1}$	-0,563* (0,00)	0,0345 (0,27)	-0,004 (0,57)
$PIB_{t-1}$	0,224* (0,01)	0,416* (0,00)	0,013 (0,51)	$PIB_{t-1}$	0,144** (0,02)	0,382* (0,00)	-0,048* (0,00)
$w_{t-1}$	0,125 (0,28)	4,23* (0,00)	0,72* (0,00)	$Emp_{t-1}$	0,301* (0,00)	0,339* (0,00)	0,95* (0,00)
$R^2$	0,98	0,98	0,89	$R^2$	0,98	0,98	0,99
$F(3,321)$	223,55* (0,00)	97,44* (0,00)	388,27* (0,00)	$F(3,321)$	250,71* (0,00)	107,72* (0,00)	3778,01* (0,00)
Obs.	351	351	351	Obs.	351	351	351
Grupos	27	27	27	Grupos	27	27	27
Obs. por Grupo	13	13	13	Obs. por Grupo	13	13	13

Fonte: Elaboração própria a partir dos resultados obtidos.

Nota 1: Crédito – Crédito Rural em logaritmo;  $PIB$  – valor adicionado bruto, em logaritmo;  $w$  – Salário Real, em logaritmo;  $Emp$  – Número de Empregos em logaritmo.

Nota 2: P-Valor entre parênteses. \* Significante a 1%; \*\* Significante a 5%; \*\*\* Significante a 10%.



**Tabela 6 - Modelos estimados com subamostra norte-nordeste**

Modelo com salários reais				Modelo com número de empregos			
Efeitos Contemporâneos				Efeitos Contemporâneos			
	$Crédito_t$	$PIB_t$	$w_t$		$Crédito_t$	$PIB_t$	$Emp_t$
$Crédito_t$	1	0	0	$Crédito_t$	1	0	0
$PIB_t$	-0,02	1	0	$PIB_t$	-0,03	1	0
$w_t$	0,07	0,075	1	$Emp_t$	0,009	0,074	1
Equações com Efeitos Defasados				Equações com Efeitos Defasados			
	$Crédito_t$	$PIB_t$	$w_t$		$Crédito_t$	$PIB_t$	$Emp_t$
$Crédito_{t-1}$	0,599*	0,068**	0,069*	$Crédito_{t-1}$	-0,589*	0,079**	0,018
	(0,00)	(0,04)	(0,00)		(0,00)	(0,03)	(0,32)
$PIB_{t-1}$	0,337*	0,491*	0,027	$PIB_{t-1}$	0,321*	0,514*	0,074**
	(0,00)	(0,00)	(0,32)		(0,00)	(0,00)	(0,03)
$w_{t-1}$	-0,022	0,375*	0,69*	$Emp_{t-1}$	0,027	0,147**	0,679*
	(0,88)	(0,00)	(0,00)		(0,76)	(0,03)	(0,00)
$R^2$	0,96	0,97	0,89	$R^2$	0,96	0,97	0,99
$F(3,189)$	103,09*	60,78*	172,71*	$F(3,189)$	103,15*	56,95*	300,46*
	(0,00)	(0,00)	(0,00)		(0,00)	(0,00)	(0,00)
Obs.	208	208	208	Obs.	208	208	208
Grupos	16	16	16	Grupos	16	16	16
Obs. por Grupo	13	13	13	Obs. por Grupo	13	13	13

Fonte: Elaboração própria a partir dos resultados obtidos.

Nota 1: *Crédito* – Crédito Rural em logaritmo; *PIB* – valor adicionado bruto, em logaritmo; *w* – Salário Real, em logaritmo; *Emp* – Número de Empregos em logaritmo.

Nota 2: P-Valor entre parênteses. \* Significante a 1%; \*\* Significante a 5%; \*\*\* Significante a 10%.

**Tabela 7 - Modelos estimados com subamostra centro-sul**

Modelo com salários reais				Modelo com número de empregos			
Efeitos Contemporâneos				Efeitos Contemporâneos			
	<i>Crédito<sub>t</sub></i>	<i>PIB<sub>t</sub></i>	<i>w<sub>t</sub></i>		<i>Crédito<sub>t</sub></i>	<i>PIB<sub>t</sub></i>	<i>Emp<sub>t</sub></i>
<i>Crédito<sub>t</sub></i>	1	0	0	<i>Crédito<sub>t</sub></i>	1	0	0
<i>PIB<sub>t</sub></i>	0,18	1	0	<i>PIB<sub>t</sub></i>	0,15	1	0
<i>w<sub>t</sub></i>	0,097	0,071	1	<i>Emp<sub>t</sub></i>	0,10	0,04	1
Equações com Efeitos Defasados				Equações com Efeitos Defasados			
	<i>Crédito<sub>t</sub></i>	<i>PIB<sub>t</sub></i>	<i>w<sub>t</sub></i>		<i>Crédito<sub>t</sub></i>	<i>PIB<sub>t</sub></i>	<i>Emp<sub>t</sub></i>
<i>Crédito<sub>t-1</sub></i>	0,818*	0,177*	0,111*	<i>Crédito<sub>t-1</sub></i>	0,86*	0,217*	-0,034**
	(0,00)	(0,00)	(0,00)		(0,00)	(0,00)	(0,04)
<i>PIB<sub>t-1</sub></i>	-0,094	0,219**	-0,027	<i>PIB<sub>t-1</sub></i>	-0,08	0,200**	0,00006
	(0,22)	(0,01)	(0,35)		(0,25)	(0,01)	(0,99)
<i>w<sub>t-1</sub></i>	0,274***	0,347***	0,715*	<i>Emp<sub>t-1</sub></i>	0,226***	0,474*	0,795*
	(0,09)	(0,06)	(0,00)		(0,06)	(0,00)	(0,00)
<i>R<sup>2</sup></i>	0,99	0,98	0,95	<i>R<sup>2</sup></i>	0,96	0,97	0,99
<i>F(3,129)</i>	254,72*	39,82*	300,08*	<i>F(3,129)</i>	256,49*	46,12*	119,46*
	(0,00)	(0,00)	(0,00)		(0,00)	(0,00)	(0,00)
Obs.	143	142	143	Obs.	143	143	143
Grupos	11	11	11	Grupos	11	11	11
Obs. por Grupo	13	13	13	Obs. por Grupo	13	13	13

Fonte: Elaboração própria a partir dos resultados obtidos.

Nota 1: *Crédito* – Crédito Rural em logaritmo; *PIB* – valor adicionado bruto, em logaritmo; *w* – Salário Real, em logaritmo; *Emp* – Número de Empregos em logaritmo.

Nota 2: P-Valor entre parênteses. \* Significante a 1%; \*\* Significante a 5%; \*\*\* Significante a 10%.

**WIND POWER AND THE LABOR MARKET IN THE BRAZILIAN NORTHEAST: A SPATIAL PROPENSITY SCORE MATCHING APPROACH\***

**Thiago Pastorelli Rodrigues**

Energy Research Office (EPE) and University of São Paulo (USP)

E-mail: rodrigues.t@gmail.com

**Solange Ledi Gonçalves**

Federal University of São Paulo (UNIFESP)

E-mail: slgoncalves@unifesp.br

**André Squarize Chagas**

University of São Paulo (USP)

E-mail: achagas@usp.br

**ABSTRACT:** Debates on the importance of renewable energy sources have arisen in the context of climate change. Wind power is a renewable and clean energy source with minimal environmental impact. Also, the existing literature indicates that the installation of wind farms can be an important channel for local development through job and income generation. In this context, this study aims to analyze the relationship between the presence of wind farms and the labor market outcomes of the municipalities of the Northeast region of Brazil, which contains around 80% of the installed capacity of this technology in the country. The database of this study is built from several public data sources. The methodological approach taken was a combination of propensity score matching and the estimation of the average treatment effect. However, we consider the existence of spatial dependence of the municipalities in the estimation of the propensity score. The results suggest that the presence of wind power plants is positively related to increases in the total wage of the municipalities, mostly in sectors such as construction, transportation, and logistics. In addition, the presence of these plants allows the shifting of resources to the agricultural sector.

**Keywords:** Wind energy; Spatial propensity score matching; Local development.

**JEL Codes:** Q42; C31; R58.

**ENERGIA EÓLICA E O MERCADO DE TRABALHO NA REGIÃO NORDESTE DO BRASIL: UMA ABORDAGEM DE PAREAMENTO POR ESCORE DE PROPENSÃO ESPACIAL**

**RESUMO:** O debate acerca do potencial da energia eólica surge no contexto das discussões sobre mudanças climáticas. A energia eólica é uma fonte de energia renovável, limpa e de baixo impacto ambiental. Além disso, a literatura indica que a implantação de usinas eólicas pode ser um importante canal de desenvolvimento local ao gerar empregos e renda. Nesse sentido, este estudo busca analisar a relação entre a instalação de parques eólicos nos municípios da região Nordeste do Brasil, a qual concentra cerca de 80% da capacidade instalada dessa tecnologia no país, e os resultados de mercado de trabalho da região. A base de dados deste trabalho é construída a partir de várias fontes de dados públicos. A metodologia empregada é o pareamento por escore de propensão, além do cálculo do efeito médio de tratamento nos tratados. Em decorrência da natureza espacial das unidades observadas, a existência de dependência espacial é considerada na estimação. Os resultados sugerem que a implantação de usinas eólicas pode elevar a massa salarial nos setores de construção, transporte e logística. Além disso, a presença dessas usinas pode deslocar recursos para o setor agropecuário.

**Palavras-chave:** Energia eólica; Pareamento por escore de propensão espacial; Desenvolvimento local.

**Classificação JEL:** Q42; C31; R58.

## 1. Introduction

The electricity sector is essential for the growth and development of any economy. For this reason, understanding the determinants of supply and demand in this sector is strategic in public policy formulation<sup>1</sup>. Debates on the importance of renewable energy sources have arisen in the context of climate change. The Kyoto Protocol, created at the United Nations Framework Convention on Climate Change (UNFCCC) in 1997 and came into force in 2005, was a relevant agreement in the promotion of the use of renewable energy. This agreement set targets for reducing greenhouse gas emissions, from which each country sought to implement the use of renewable energy according to its characteristics and potential (MORENO; LOPEZ, 2008). In December 2015, during the 21st Conference of the Parties in Paris (COP21), governments of around 190 countries presented their plans to reduce domestic greenhouse gas emissions. Brazil committed to reduce its emissions by 37% as of 2025 and by 43% as of 2030 from the 2005 levels (EPE, 2016).

Given this context, wind energy is deemed an important renewable source of energy since it has a high potential in reducing greenhouse gas emissions, although it is not equally distributed in space. However, this source of renewable energy is not widely used worldwide, notwithstanding that it could, theoretically, supply up to 20% of world electricity demand by 2050 (EDENHOFER et al., 2011). In addition to environmental impacts, some studies provide evidence demonstrating that advancing the production of renewable energy sources, especially wind energy, may be relevant to the local development of regions impacted by the installation of wind farms. Wei et al. (2010) argue that many local jobs can be generated, mainly at the construction and installation stages of such plants. Río and Burguillo (2008) also show that the implementation of renewable energy projects can contribute to local development, especially in rural areas. These authors further argue that during the construction phase, an increase in the demand for goods and services is expected, and local suppliers are benefited from increased income and temporary jobs, inside and outside the communities. In addition, changes in community structure can occur, such as improving public infrastructure, lower-cost energy supply, and more.

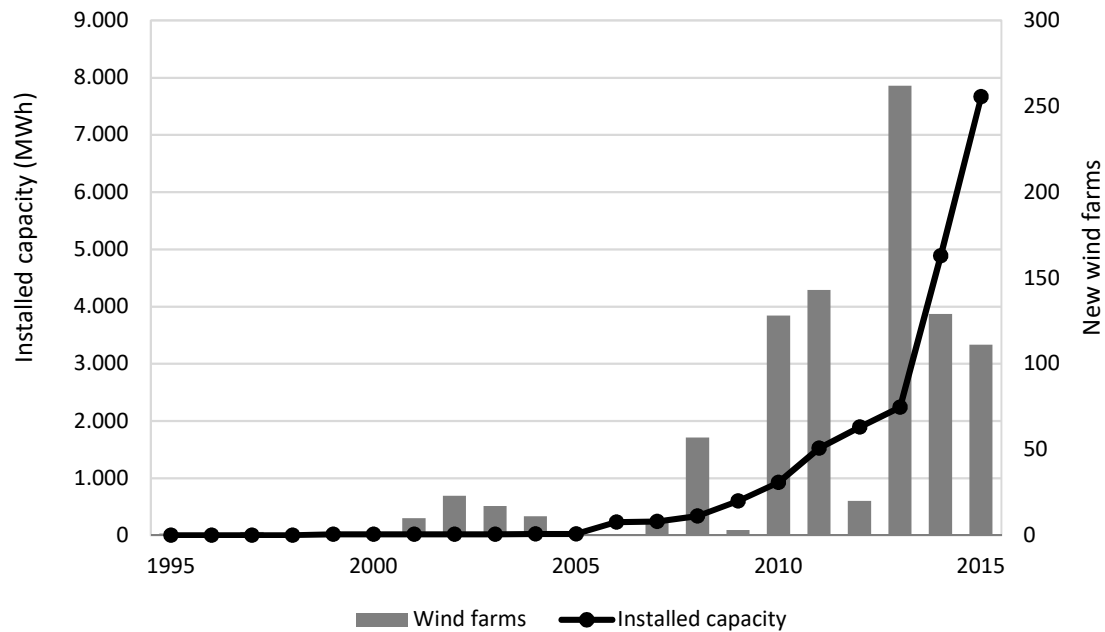
Despite the existence of many studies focused on the socioeconomic impacts of other renewable energy sources<sup>2</sup>, literature that specifically addresses the effects of wind energy on local development is merely recent and scarce. Blanco and Rodrigues (2009) showed that the deployment of wind energy in countries in the European Union creates a significant number of direct jobs, mainly in the wind turbine and component industry. Ejdemo and Söderholm (2015) have identified the impacts on employment in the Swedish construction sector. Concerning Brazil, Simas and Pacca (2014) have verified that the main contribution is in the construction phase and that, despite the low number of jobs created in operation and maintenance of the farms, these jobs have persisted over time.

According to data from the Brazilian Electricity Regulatory Agency (ANEEL), the installed capacity of wind energy in Brazil grew by approximately 430% between 2010 and 2014, from 927 MWh to 4,888 MWh. In addition, data from ANEEL indicate that wind farm records in Brazil began in 1995 and have grown significantly from 2010 (see Figure 1). According to the Global Wind Energy Council (GWEC) report, Brazil, in 2015, was the Latin American country with the largest installed capacity of wind power energy (5.96 MWh) and was ranked tenth in the global ranking for installed capacity, accounting for 2% of the global share of accumulated capacity this year (8.72 GW). In addition, Brazil was the country ranked fourth in the ranking of the new installed capacity of wind power generation (2.76 GW) in 2015.

---

<sup>1</sup> For a literature review on the conditions of electricity demand in Brazil, see Uhr et al. (2017 and 2019).

<sup>2</sup> For a literature review on the potential of various renewable energy sources in creating jobs, see Wei et al. (2010) and Lehr et al. (2008).

**Figure 1 - Wind farms and the installed capacity of wind energy in Brazil**

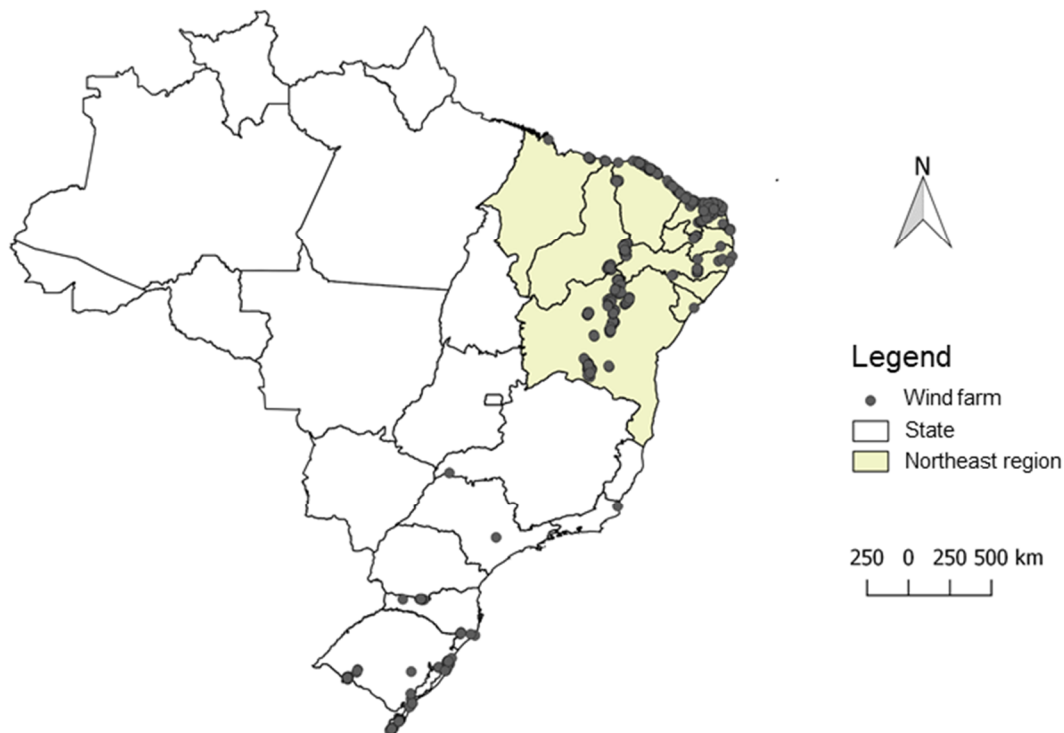
Source: Elaborated by the authors, based on data from ANEEL.

Given these facts and the small number of studies that have analyzed the impact of wind farms on local development, this article aims to understand the relationship between the implementation of wind farms and the labor market results of the municipalities of the Northeast region of Brazil, focusing on three main variables: (i) employment level, (ii) the total amount of wages, and (iii) the number of firms.

Brazil's Northeast region accounts for 54% of the country's wind potential. Moreover, in this region, the sources of wind and hydroelectric power are complementary, because the periods of drought, when the reservoirs of the dams are at the lowest level, coincide with the period of highest incidence and intensity of winds (COSTA et al., 2009). According to ANEEL, 80% of the municipalities with wind farms, in operation or under construction, are located in this region. Figure 2 shows the spatial distribution of wind farms in Brazil and illustrates this intense concentration in the Northeast region, highlighted in gray.

The main contribution of this paper is to relate the implementation of wind farms and the labor market outcomes of the municipalities, through the spatial propensity score matching method as well as the calculation of the average treatment effect on the treated. This methodological approach allows the determination of a control group of municipalities, statistically similar to the treatment group of municipalities, in terms of specific observable characteristics, with the presence of wind farms as the only factor that differentiates the groups. We considered the existence of spatial dependence in the estimate of the propensity score, due to the spatial nature of the units analyzed (CHAGAS et al., 2012).

We concentrated our analysis for the year 2013, and we built our database based on two ANEEL databases, the Georeferenced Information System of the Electricity Sector (SIGEL) and the Generation Information Data (BIG), in addition to the Annual Social Information Report (RAIS), from the Brazilian Ministry of Economy, and other variables from the Brazilian Institute of Geography and Statistics (IBGE) and the National Institute for Space Research (INPE).

**Figure 2 - Spatial distribution of wind farms in Brazil**

Source: Elaborated by the authors, based on data from ANEEL.

The results suggest that the total wages of the construction, transportation, and logistics sectors are higher in municipalities that have wind farms, either in operation or under construction. In addition, municipalities that have wind farms in operation or under construction contain more firms in the agricultural sector, suggesting that the presence of wind power plants can lead to the transfer of resources to the agricultural sector, boosting this activity in the local economy. For the number of jobs in the municipalities, no statistically significant results were found.

The article contains six sections, besides this introduction, appendices, and bibliographical references. The second section consists of a brief literature review on wind energy in Brazil. The third section presents the empirical strategy of the paper and the description of the sample. The fourth section shows the descriptive analyses we conducted. The fifth section presents the results of our study. Lastly, the last section contains our final considerations.

## 2. Wind farms and the local development

Wind energy has grown exponentially in the world since the late 1990s, especially in the European Union and the United States. Wind energy stands out among renewable energy sources due to its lower production costs, the security of supply, and environmental sustainability (GWEC, 2011).

Indeed, this energy source has some specific advantages: (i) the non-use of water and the absence of greenhouse gas emissions during the operation phase of wind turbines; (ii) the absence of danger of fuel leakage; (iii) the possibility of occupation of the land by agriculture or pasture, around the perimeter of the wind farm; (iv) the relatively small area occupied by wind turbines; and (5) the opportunity given to geographically isolated regions that have no available power grid but have strong winds, to generate energy without incurring high transmission costs (Costa, Casotti, and Azevedo, 2009)<sup>3</sup>.

<sup>3</sup> Disadvantages of wind power generation include: (i) the threat to the life of wild animals, especially birds; (ii) noise pollution; and (iii) visual pollution (Klick and Smith, 2010).

In Brazil, the emergence of the Programme of Incentives for Alternative Electricity Sources (Proinfa) has also been identified as a contributor to the growth of wind farms in recent years. The main objectives of this program were: the diversification of the Brazilian energy matrix; the security of supply; the reduction of greenhouse gas emissions; and the exploration of regional and local characteristics and potentials, thereby generating employment and training for the workforce. This program was aimed at increasing the participation of electricity through incentives to independent producers, based on wind sources, small hydroelectric plants (PCHs), and biomass, in the National Interconnected System.

Thus, Proinfa can be regarded as a relevant milestone in the dissemination of wind energy in Brazil, with 44% of the implemented plants of the program are wind power plants. Several studies have highlighted the importance of public programs in the promotion of wind energy. Ohl and Eichhornb (2010) have stated that government policy is an essential step towards the implementation of renewable energy, including the development of wind energy. Bird et al. (2005), in turn, highlighted the role of financial incentives in terms of wind energy in Europe and the United States.

Government incentives and policies may be further justifiable if, in addition to environmental impacts, we take into account the effects of wind farms on the local development of regions affected by the installation of wind farms. Several studies have pointed to the generation of employment and mastery of technology as factors relevant to environmental preservation and energy security in terms of wind energy (MARTINS et al., 2008; SZARKA, 2006).

In this context, Blanco and Rodrigues (2009) have shown that many of the jobs generated by wind energy, especially those related to specialized services as well as the operation and maintenance of the plants, tend to be offered by local firms. These authors also determined that about 59% of direct jobs related to wind energy in Europe are in the component and turbine industries. They have also concluded that because much of the wind energy activity takes place locally (construction, operation, and maintenance), there is a positive relationship between the location of a wind farm and the number of new jobs.

On the other hand, Simas and Pacca (2015) found that, in Brazil, the jobs generated in construction correspond to about 55% of direct jobs and 40% of the total jobs created from the wind energy. They thereafter concluded that this is the stage of greater potential in terms of generating new jobs.

### 3. Empirical strategy

In this article, we use the propensity score matching (PSM) method. The existence of spatial dependence is taken into account in the estimates due to the spatial nature of the units analyzed (i.e., the municipalities).

#### 3.1. Propensity score matching

Matching methods seek to obtain a control group statistically similar to the treatment group in terms of specific observable characteristics. Thus, in comparing two groups of municipalities with the same observable characteristics, the only factor that differentiates them is the existence or the expectation of the presence of a wind farm in its territory.

According to Khandker et al. (2010), under the assumptions of conditional independence and common support, the *average treatment effect on the treated* (ATT), for a subsample with observable characteristics,  $X$ , can be written as:

$$ATT = E[Y(1)|T = 1, X] - E[Y(0)|T = 1, X] \quad (1)$$

where  $Y(1)$  is the potential outcome of the variable of interest under treatment and  $Y(0)$  is the potential outcome of the variable of interest in the absence of treatment.

The PSM method seeks to summarize the effects of different observable variables on a single index. Thus, treated and control municipalities with similar propensity scores are compared. Using the PSM, it is possible to construct a comparison group based on the probability of a municipality being treated ( $T$ ), conditioned on the observed characteristics  $X$  or the propensity score:  $P(X) = P(T = 1|X)$ . In this article, we estimated the propensity score using a *probit* model, such that:

$$P(T = 1|X) = \Phi(X\beta) \quad (2)$$

where  $\Phi$  is the cumulative distribution function (CDF) of the standard normal distribution and  $\beta$  is the estimated parameter vector.

However, as the unit of interest in this study consists of the municipalities of the Northeast region of Brazil, the sample has a geographical nature; thus, the units analyzed may not be randomly distributed in space. Hence, the possibility of spatial dependence between the observed units should be considered in the propensity score estimation through the implementation of a spatial propensity score. Based on Chagas et al. (2012), the spatial dependence can be inserted into equation (2) in two ways: through a regressor in the form of a spatially lagged dependent variable ( $Wy$ ) or in the structure of the error term ( $Wu$ ). Formally, the spatial model is represented by:

$$\begin{aligned} y &= \rho Wy + X\beta + u \\ u &= \lambda Wu + \varepsilon \end{aligned} \quad (3)$$

where  $y$  is the vector of the dependent variable,  $W$  is a spatial weight matrix that characterizes the type of neighborhood relationship between municipalities,  $\rho$  and  $\lambda$  are the spatial parameters, and  $\varepsilon$  is the error term. If  $\rho \neq 0$  and  $\lambda = 0$ , the specification is known as the *spatial autoregressive model* (SAR), and if  $\rho = 0$  and  $\lambda \neq 0$ , the specification is called the *spatial error model* (SEM)<sup>4</sup>.

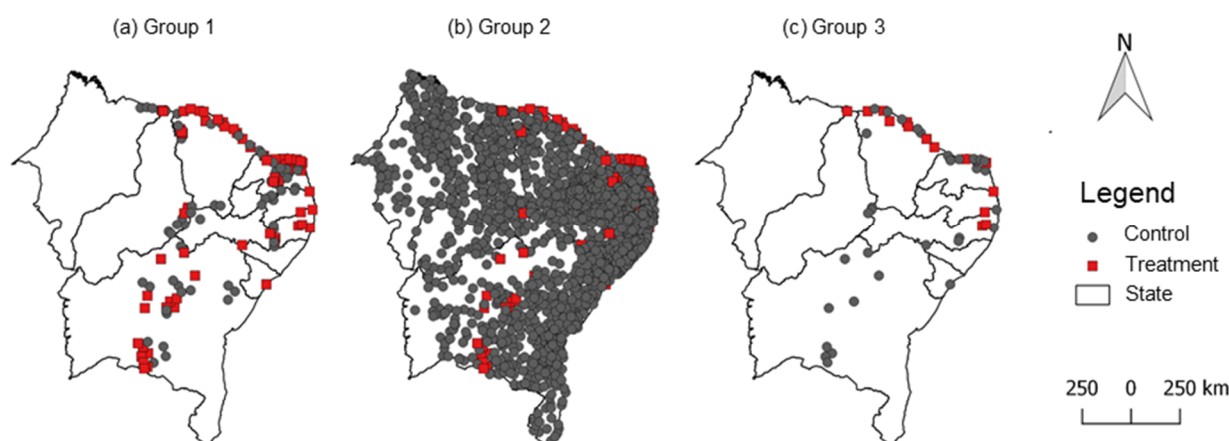
### 3.2. Treatment and control groups

According to the Georeferenced Information System of the Electricity Sector (SIGEL) from ANEEL, wind power plants can be classified as: (i) operation; (ii) construction with grant; (iii) construction not started; (iv) winner of the auction; (v) disabled; and (vi) *Despacho de Requerimento de Outorga* (DRO). Municipalities with wind farms, either in operation or under construction, with a grant, already have the minimum required infrastructure in the area. On the other hand, the other municipalities have only the expectation of the installation of a wind power plant. Thus, to analyze the relationship between the presence of wind farms in the municipalities in the Northeast region of Brazil and the variables that capture local development, three analysis groups were defined.

For Group 1, the treatment consists of the municipalities where there is at least one wind farm in operation or under construction with a grant and the control group, the municipalities with wind power plants at the project stage, that is, wind farms with statuses of construction not started, winner of the auction or DRO. For Group 2, the control group is expanded to include all other municipalities of the Northeast region, that is, those with no expectation of implementation of wind farms. Finally, Group 3 contains the municipalities with operating wind farms that are part of the Proinfa program in the treatment group, and the municipalities with operating wind farms that are not participating in this program in the control group. Figure 3 shows the spatial distribution of municipalities according to these groups. The red dots on the map are the treatment group municipalities, and the green dots are the control group municipalities. For all analyses, the reference year is 2013.

<sup>4</sup> For more details on spatial models and spatial econometrics, see Anselin (2001), LeSage and Pace (2009), and Almeida (2012).



**Figure 3 - Spatial distribution of the treatment and control groups**

Source: Elaborated by the authors.

Note: Municipalities of the treatment group are in red color and municipalities of the control group are in gray color.

#### 4. Database and variables

We build the database by aggregating two ANEEL databases, SIGEL and BIG. SIGEL presents georeferenced information from all registered or projected wind power plants. The data extracted from these sources are the power of the wind farm, the project registration date at ANEEL, the implementation stage of the wind farm, and the municipality where the wind farm is located. BIG presents information on wind farms that are at least in the construction phase with a grant. Although SIGEL has more detailed information on wind farms, BIG contains the beginning of the wind farm operation, information relevant for our analysis.

It is noteworthy that we have disregarded deactivated wind plants and wind plants with an installed capacity of 100 kilowatts (kW) or less since these wind power plants are not implemented for commercial purposes (ANEEL, 2014). Thus, we used these data to build the database at the municipal level and to define the treatment and control groups described in Section 3.2.

The objective of this paper is to analyze the relationship between the implementation of wind farms and variables that capture the level of welfare and local development of the municipalities of the Northeast region of Brazil. Thus, in this paper, the outcome or dependent variables, for which the average treatment effect on the treated is calculated, are the employment level, the total wages, and the number of firms in the municipalities analyzed.

We used RAIS data from the Brazilian Ministry of Economy in constructing the dependent variables of employment, wages, and the number of firms. We focused on workers with active employment relationship as of December 31, 2013. In addition to the total level of employment and wages in municipalities, we computed for the estimates using disaggregation based on (i) two schooling level of workers – completed and not completed higher education<sup>5</sup>; (ii) economic sectors defined by the Brazilian Classification of Occupations (CBO) – Agriculture, Industry and Maintenance, and Services and Commerce<sup>6</sup>; and (iii) three subcategories of the Industry and Maintenance sector – Construction, Electricity and Gas, and Transport and Logistics<sup>7</sup>. For wages, we

<sup>5</sup> For this disaggregation, we use the filter for schooling after 2005 (IBGE), present in RAIS data.

<sup>6</sup> For this disaggregation, we use the filter for CBO 2002 Large Groups, present in RAIS data. Agricultural, forestry, and fishery jobs are aggregated to obtain the level of employment in Agriculture. Jobs from the production of industrial goods and services of repair and maintenance are aggregated to get the level of employment of Industry and Maintenance. Finally, jobs in administrative services, other services, salespeople in shops, and markets are aggregated to obtain the number of jobs in Services and Commerce.

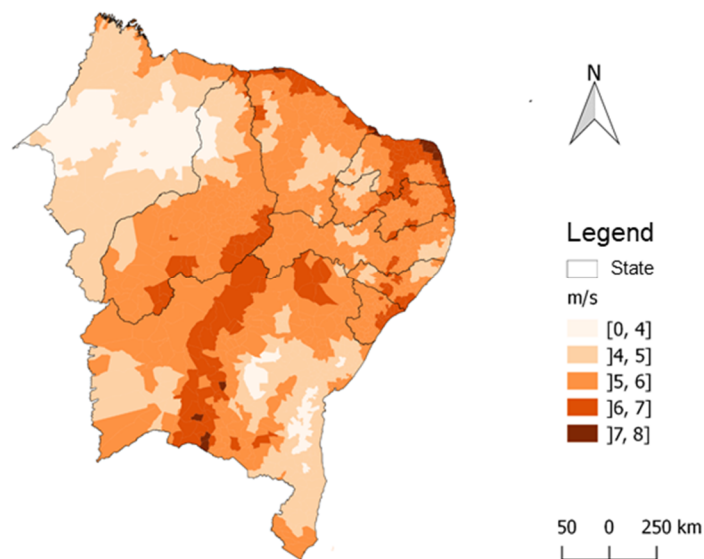
<sup>7</sup> For this disaggregation, in addition to the filter for the categories of CBO 2002, we use the filter for the workers of the production of industrial goods and services among the groups of CNAE (National Classification of Economic Activities) 2.0. Jobs from transportation, storage, and mail are aggregated to get the number of Transport and Logistics.

applied a natural logarithm. For the total number of firms, we used a different form of disaggregation, since firms are heterogeneous in terms of worker educational level. Another difference can be seen in the breakdown by economic sector, due to the number of firms, it is only possible to apply filters based on IBGE Groups. Thus, this procedure implies consolidating the Industry with the Construction sector.

The first step of the empirical strategy is the propensity score estimation, in which some explanatory variables are considered for the probability of the municipalities belonging to the treatment group. The explanatory variables are wind speed, the ratio of agricultural gross domestic product (GDP) to total GDP, area of municipalities, population, and a dummy variable for the metropolitan region.

The wind speed of a locality is the explanatory variable for which we expect a positive and high magnitude impact on the probability of having a wind farm, given that this is one of the main criteria evaluated by the Ministry of Mines and Energy (MME) for approval of participation in energy auctions of wind energy projects. This variable is considered at the municipal level, based on georeferenced information provided by the National Institute for Space Research (INPE), and is calculated as the average annual wind speed (in meters per second) at 50 meters in height<sup>8</sup>. Figure 4 presents the spatial distribution of this variable. The comparison of Figures 3 and 4 shows that the Northeast region of Brazil, in which most of the municipalities with wind power plants are concentrated, is also the area with the highest wind speed.

**Figure 4 - Average annual wind speed of the municipalities of the Northeast region of Brazil**



Source: Elaborated by the authors, based on data from INPE.

Note: Average annual wind speed in meters per second at 50 meters in height.

We used the natural logarithm of the number of inhabitants of the municipality as a population variable, obtained through data provided by IBGE. The inclusion of a control for the municipality size in terms of population density, which is highly correlated with the economically active population, improves the matching of the municipalities of the treatment and control groups in terms of economic dynamics. We used the municipal GDP data at current prices to calculate the ratio between the agricultural GDP and the total GDP of the municipalities. These data are available from IBGE. The inclusion of this variable seeks to capture the opportunity cost of land use for the presence

<sup>8</sup> According to Grubb and Meyer (1993), the density of the wind energy must be greater than or equal to 500 W/m<sup>2</sup> at the height of 50 meters for being considered technically usable, which requires a wind speed between 7 and 8 meters per second.

of wind farms, as one of the possible alternatives for land use in agricultural activities. Thus, a negative relationship is expected between this variable and the probability of existing wind farms in the localities. Another relevant explanatory variable is the total area of the municipality – in natural logarithm – whose inclusion aims to control for the availability of space for the construction of wind farms and allows the matching of the treatment and control municipalities with similar size. Therefore, we expect the relationship between this variable and the presence of wind farms in the localities to be positive. Finally, the inclusion of a variable that indicates whether the municipality is located in a metropolitan region aims to control for populational and economic dynamics, considering that urban areas have distinct characteristics from non-metropolitan areas.

## 5. Descriptive analysis

The descriptive analysis presents t-tests for the differences between the means of the dependent variables of the municipalities of the treatment and control groups. The objective of these tests is to verify, in an unconditional analysis, the existence of statistically significant differences in the average of the variables for total employment, wages, and the number of firms, before performing the econometric estimation. We also perform the tests disaggregated by educational level of workers and economic sectors. Finding positive and statistically significant differences for the variables that capture the level of local development is the first evidence that the municipalities with wind farms – treatment units – could perform better in terms of these labor market outcomes.

Tables 1 to 3 show the mean variables for the treatment and control groups of municipalities – as described in section 3.2 – and the statistical significance of the tests. Comparing the results presented in these tables allows us to conclude the feasibility of using these treatment and control groups for estimates.

The comparison between the total level of employment in municipalities with wind farms under construction and operation and the municipalities with projects of wind power plants (Table 1) allows us to verify the statistically significant differences in terms of the average employment of the agricultural, electricity, and gas sectors. In contrast, when the other municipalities of the Northeast region of Brazil are added to the control group (Table 2), it is precisely the employment in these two sectors that no longer presents statistically significant differences, where all other employment variables are different between treatment and control groups. In the case of Group 3, in which the comparison is between municipalities with operating wind farms participating and non-participating in Proinfa, no differences are found for the total level of employment and employment of the sectors and subsectors analyzed (Table 3).

Regarding the level of wages, we observed that the municipalities with wind plants in operation and under construction present a higher level of wages for all sectors and subsectors analyzed, except in the electricity and gas sectors, in comparison to the municipalities with wind farms at the project phase (Group 1). This result evidences the possibility of an increase in the wages and welfare of workers in municipalities that have some wind farm operating or in construction. Statistically significant results are also found from the tests when we added the other Brazilian municipalities to the control group (Group 2). In this case, the differences also occurred for the total level of wages and the wages of sectors and sub-sectors, except in the electricity and gas sectors.

Also, when we compared the municipalities with operating wind farms participating and non-participating in Proinfa, we found significant results for the total level of wages, the wages of higher-educated workers, and the wages of the three sectors analyzed: agriculture, industry and maintenance, and services and commerce. This result serves as preliminary evidence showing that participation in this program could further reinforce the positive relationship between the presence of wind farms and local development. However, no differences are found when we disaggregated according to subsectors.

**Table 1 - T-test for the differences between the means of the dependent variables - Group 1**

Employment	Treatment (T)	Control (C)	Difference (T-C)	<i>t</i> -statistic
Total	15,934	2,282	13,652	-1.14
Completed higher education	2,918	496	2,422	-1.07
Not completed higher education	13,015	1,786	11,230	-1.15
Agriculture	231	104	127	-1.69 *
Industry and Maintenance	3,329	490	2,839	-1.16
Construction	977	70	907	-1.17
Electricity and Gas	7	1	6	-1.82 *
Transport and Logistics	270	20	250	-1.20
Services and Commerce	7,139	943	6,196	-1.13
<b>Wages</b>				
Total	14.7	14.1	0.6	-2.49 **
Completed higher education	13.5	13.0	0.5	-2.06 **
Not completed higher education	14.2	13.7	0.5	-2.38 **
Agriculture	10.7	9.6	1.2	-3.47 ***
Industry and Maintenance	12.4	11.8	0.6	-2.04 **
Construction	10.6	9.5	1.1	-2.26 **
Electricity and Gas	10.4	10.0	0.4	-0.54
Transport and Logistics	10.3	9.2	1.0	-2.00 **
Services and Commerce	13.5	13.0	0.5	-1.92 *
<b>Number of firms</b>				
Total	1,065	184	881	-1.15
Agriculture	20	9	11	-2.69 ***
Industry and Construction	197	25	172	-1.13
Construction	77	7	70	-1.14
Electricity and Gas	1	0	1	-1.63
Transport and Logistics	0	0	0	0.00
Services and Commerce	848	150	698	-1.15
Observations	65	69		

Notes: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*; employment and number of firms are in levels; wages are in natural logarithm. The treatment group is composed of the municipalities in which there is a wind farm in operation or construction with a grant, whose base year is 2013, and the control group is composed of the municipalities with projects of wind farms, that is, with the status of construction not started, winner auction or DRO.

**Table 2 - T-test for the differences between the means of the dependent variables - Group 2**

Employment	Treatment (T)	Control (C)	Difference (T-C)	<i>t</i> -statistic
Total	15,934	4,564	11,370	-2.43 **
Completed higher education	2,918	806	2,112	-2.19 **
Not completed higher education	13,015	3,758	9,257	-2.48 **
Agriculture	231	173	59	-0.73
Industry and Maintenance	3,329	957	2,372	-2.64 ***
Construction	977	265	712	-1.97 **
Electricity and Gas	7	3	3	-0.88
Transport and Logistics	270	72	198	-2.29 **
Services and Commerce	7,139	2,063	5,075	-2.27 **
<b>Wages</b>				
Total	14.7	13.9	0.8	-4.36 ***
Completed higher education	13.5	12.7	0.8	-4.45 ***
Not completed higher education	14.2	13.5	0.7	-4.03 ***
Agriculture	10.7	9.9	0.9	-3.17 ***
Industry and Maintenance	12.4	11.5	1.0	-3.84 ***
Construction	10.6	9.9	0.7	-1.97 **
Electricity and Gas	10.4	10.8	-0.4	0.95
Transport and Logistics	10.3	9.4	0.8	-2.34 **
Services and Commerce	13.5	12.8	0.7	-3.81 ***
<b>Number of firms</b>				
Total	1,065	306	759	-2.77 ***
Agriculture	20	17	2	-0.42
Industry and Construction	197	43	154	-3.57 ***
Construction	77	17	60	-3.06 ***
Electricity and Gas	1	0	1	-4.64 ***
Transport and Logistics	0	0	0	0.00
Services and Commerce	848	245	603	-2.62 ***
Observations	65	1,729		

Notes: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*; employment and number of firms are in levels; wages are in natural logarithm. The treatment group is composed of the municipalities in which there is a wind farm in operation or construction with a grant, whose base year is 2013, and the control group is composed of all other municipalities of the Northeast region of Brazil.

**Table 3 - T-test for the differences between the means of the dependent variables - Group 3**

Employment	Treatment (T)	Control (C)	Difference (T-C)	<i>t</i> -statistic
Total	5,454	26,020	-20,566	0.57
Completed higher education	1,131	4,737	-3,606	0.53
Not completed higher education	4,323	21,283	-16,960	0.58
Agriculture	335	240	95	-0.55
Industry and Maintenance	1,091	5,499	-4,408	0.60
Construction	242	1,650	-1,408	0.61
Electricity and Gas	7	9	-2	0.18
Transport and Logistics	61	461	-400	0.64
Services and Commerce	2,544	11,632	-9,088	0.55
<b>Wages</b>				
Total	15.5	14.6	0.9	-1.79 *
Completed higher education	14.6	13.5	1.1	-2.14 **
Not completed higher education	15.0	14.2	0.8	-1.55
Agriculture	12.1	10.4	1.6	-2.88 ***
Industry and Maintenance	13.6	12.4	1.2	-1.92 *
Construction	10.4	11.1	-0.7	0.76
Electricity and Gas	10.1	10.8	-0.7	0.84
Transport and Logistics	9.9	10.8	-0.9	0.97
Services and Commerce	14.4	13.4	1.0	-1.82 *
<b>Number of firms</b>				
Total	485	1,679	-1,194	0.52
Agriculture	32	18	13	-1.42
Industry and Construction	66	324	-257	0.56
Construction	22	128	-105	0.57
Electricity and Gas	1	1	0	0.05
Transport and Logistics	0	0	0	0.00
Services and Commerce	388	1,338	-950	0.52
Observations	14	36		

Notes: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*; employment and number of firms are in levels; wages are in natural logarithm. The treatment group is composed of the municipalities with operating wind farms that are part of the Proinfra program, and the control group is composed of the municipalities with operating wind farms that are not participating in this program.

## 6. Results

The results are divided into two parts. In the first part, we present the propensity score estimates, and we discuss the possibility of spatial dependence and the existence of common support (i.e., the overlap between treatment and control groups). In the second part, we present the relationship between the implementation of wind farms and the local development of the municipalities, by calculating the average treatment effect on the treated.

### 6.1. Spatial propensity score

The propensity score for the probability of the municipality presenting at least one wind farm is estimated using equation (2). Under the evidence of spatial dependence, identified by Moran's I

statistic, spatial elements are incorporated into the estimates according to the spatial characteristics of the data, that is, the spatial propensity score is estimated.

Five spatial weight matrices are considered in verifying the existence of spatial dependence in the observed units: 100, 200, 300, 400, and 500 kilometers. These distances are calculated by geographical coordinates of the administrative center of the municipality, available from IBGE. Thus, the proper strategy for choosing the appropriate spatial weight matrix for each model is to identify the distances associated with the highest spatial correlation using Moran's I statistic<sup>9</sup>. Such an indicator measures the spatial autocorrelation based on the product of the deviations of a variable to its average (Anselin, 2001). Moran's I statistic is calculated for the dependent variable and for the residual of the estimation of equation (2), whose result is presented in Table 4. The choice of the estimation method for the propensity score is based on the existence of spatial dependence of the dependent variable and the residuals of equation (2).

**Table 4 - Moran's I statistic for spatial correlation**

Distance	Dependent variable					
	Group 1		Group 2		Group 3	
	Moran's I	p-value	Moran's I	p-value	Moran's I	p-value
50 km	-	-	-	-	-	-
100 km	-	-	0.098	0.000	-	-
200 km	0.027	0.120	0.049	0.000	-	-
300 km	0.005	0.269	0.027	0.000	0.100	0.023
400 km	0.009	0.162	0.014	0.000	0.089	0.012
500 km	-0.008	0.519	0.007	0.000	0.077	0.004
Distance	Residual					
	Group 1		Group 2		Group 3	
	Moran's I	p-value	Moran's I	p-value	Moran's I	p-value
50 km	-	-	-	-	-	-
100 km	-	-	0.033	0.000	-	-
200 km	-0.002	0.426	0.009	0.000	-	-
300 km	-0.008	0.509	0.008	0.000	0.010	0.306
400 km	0.004	0.237	0.005	0.000	0.015	0.227
500 km	-0.008	0.517	0.003	0.001	0.019	0.141

Note: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*.

We can, therefore, observe that there is no evidence of spatial dependence of the dependent variable for Group 1. For Group 2, it is not possible to reject the null hypothesis regarding the absence of spatial dependence for all cases considered. The 100-kilometer matrix has the highest value for Moran's I. Finally, for Group 3, there is evidence of spatial dependence only in considering the spatial lag in the dependent variable, where the matrix with the highest value for Moran's I is of 300 kilometers.

Table 5 shows the results of propensity score estimates. We found no evidence of spatial dependence for the dependent variable and residuals of the estimation for Group 1. Thereafter, the non-spatial probit model is estimated. For Group 2, which presents evidence of spatial dependence

<sup>9</sup> Stakhovych and Bijmolt (2009) point out that there are several procedures for identifying spatial dependence on data and the most common for spatial correlation are Moran's I statistic, the Kelejian-Robinson test, and the Lagrange Multiplier (LM) test. We adopt one of these tests (Moran's I), recognizing that each test has its advantages and disadvantages.

on both the dependent variable and the residuals of equation (2), the SAR probit and SEM probit models are estimated. For Group 3, which presents spatial dependence only for the dependent variable, the SAR probit is likewise estimated. The model specification includes the entire set of variables available that could be correlated to the presence of wind farms in the localities.

**Table 5 - Results of the propensity score estimations**

Group 1	Probit		SAR Probit		SEM Probit	
	Coefficient	SD	Coefficient	SD	Coefficient	SD
Constant	-6.373***	2.048	-	-	-	-
Wind	0.602***	0.201	-	-	-	-
Area	-0.174	0.137	-	-	-	-
Population	0.382**	0.152	-	-	-	-
GDP ratio	-0.396	1.412	-	-	-	-
MR	0.043	0.540	-	-	-	-
Rho	-	-	-	-	-	-
Lambda	-	-	-	-	-	-
Observations	134		-		-	
Akaike Inf. Crit.	181		-		-	
Group 2	Probit		SAR Probit		SEM Probit	
	Coefficient	SD	Coefficient	SD	Coefficient	SD
Constant	-12.287***	1.157	-9.482***	1.712	-1.363***	0.22652
Wind	1.246***	0.123	0.921***	0.193	0.137***	0.02121
Area	0.108	0.071	0.057	0.066	0.011	0.00843
Population	0.285***	0.079	0.300***	0.100	0.033***	0.01047
GDP ratio	-0.775	0.869	-0.707	0.841	-0.065	0.0901
MR	-0.562*	0.302	-0.535*	0.321	-0.075**	0.03656
Rho	-	-	0.359*	0.209	-	-
Lambda	-	-	-	-	0.068	0.14522
Observations	1,792		1,792		1,792	
Akaike Inf. Crit.	388		-		-	
Group 3	Probit		SAR Probit		SEM Probit	
	Coefficient	SD	Coefficient	SD	Coefficient	SD
Constant	-2.608	3.536	-2.411	3.629	-	-
Wind	-0.209	0.348	-0.238	0.354	-	-
Area	-0.427*	0.252	-0.439*	0.255	-	-
Population	0.552**	0.241	0.566**	0.236	-	-
GDP ratio	4.811**	2.454	5.024*	2.584	-	-
MR	-1.147	0.903	-1.322	0.849	-	-
Rho	-	-	0.128	0.404	-	-
Lambda	-	-	-	-	-	-
Observations	50		50		-	
Akaike Inf. Crit.	60		-		-	

Note: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*.

The results for the propensity score estimation suggest that wind and population are the main variables correlated with the presence of wind farms when using both Group 1 and Group 2. Wind



speed is positively associated with the presence of wind farms in the localities, as expected. Population size is also positively related to the presence of wind farms in the municipalities, indicating that the presence of these wind power plants is more likely to occur in denser areas. For Group 2, the binary variable for the metropolitan regions (MR) is also significant, but it negatively affects the probability of presenting wind farms. In the estimates for Group 3, in addition to the population, the area and the ratio between agricultural GDP and total GDP were deemed relevant variables explaining the probability of presenting at least one wind farm. The higher the participation of agricultural activities in the municipality's economic activities, the greater the opportunity cost of land use to set up a wind farm, and the lower the likelihood that the locality will have wind farms. On the other hand, unlike expected, the area of municipalities – an indicator of area availability to implement wind power plants – is negatively correlated with the presence of wind power plants.

The spatial parameter of the SAR probit model associated with Group 2 is statistically significant, thereby corroborating the spatial nature of the sample, as indicated already by Moran's I. Therefore, in the matching step, for Groups 1 and 3, the propensity score is estimated by the standard probit and, for Group 2, by the SAR probit.

In the matching step, for the three treatment and control groups analyzed, the existence of common support is verified. We found common support for all groups, and we excluded some observations out of support in the second estimation step of the paper. It is noteworthy that in the matching performed for Group 2, there was a higher concentration of control units found in lower propensity score values. Indeed, based on the definition of this group, the treatment group is the municipalities with wind farms in operation or construction, while the control group is broad and includes the municipalities with wind farm projects and other municipalities in the Northeast region of Brazil. It should also be noted that the matching procedure for the three groups analyzed ensures that the mean propensity score was not different for treated and non-treated municipalities<sup>10</sup>.

## 6.2. Average treatment effect on the treated

This paper aims to analyze the relationship between the implementation of wind power plants and the local development of municipalities. We estimated this relationship by calculating the average treatment effect on the treated (see Tables 6 to 8). Table 6 presents the results for the total level of employment and the employment of the sectors and subsectors analyzed. Tables 7 and 8 show the results for the natural logarithm of wages and the number of firms, respectively.

In Table 6, we were not able to verify any statistically significant result for the relationship between the implementation of wind farms and the level of employment in the municipalities, for any of the defined treatment and control groups. The lack of statistical significance for this relationship contradicts the results found in the literature, given that Blanco and Rodrigues (2009) found a positive relationship for employment in the industry and Simas and Pacca (2015) and Ejdemo and Söderholm (2015) identified a positive relationship for the employment of the construction sector. The difference in the results of our paper and the cited studies may be due to estimation methods, degree of robustness of each method, or the definition of the sample of localities.

Table 7 shows that wages in the construction and transportation and logistics subsectors are higher in municipalities that have wind farms in operation or construction (Group 1). This difference accounts for approximately 11% for the construction subsector and 9% for the transportation and logistics subsector. We found a negative correlation between the presence of wind power plants and the total level of wages for the electricity and gas subsectors when we added the other municipalities of the Northeast region of Brazil to the control group (Group 2). This result is counterintuitive since the demand-side effects of labor in this subsector should pressure and raise wages. However, in this paper, we obtained employment and wage data based on RAIS data for employee-employer matching; thus, we focused on the equilibrium results of the labor market. Thus, this expected increase in wages in the electricity and gas subsectors, through increased demand for labor in this subsector, could be

---

<sup>10</sup> The table presenting the t-tests performed before and after the matching procedure is in the Appendix of the paper.

offset by an increase in labor supply in the subsector, which could, in turn, decrease average wages per worker and total level of wages.

**Table 6 - Results for the level of employment in the municipalities of the Brazilian Northeast**

Group 1	Treatment	Control	Difference	
			Mean	SD
Total	15,933	2,863	13,069	12,384
Completed higher education	2,918	615	2,303	2,345
Not completed higher education	13,015	2,248	10,766	9,701
Agriculture	231	233	-2,18	134
Industry and Maintenance	3,329	627	2,701	2,620
Construction	976	82	894	915
Electricity and Gas	6,52	0,57	5,95	3,69
Transport and Logistics	270	22,15	247	230
Services and Commerce	7,138	1,122	6,016	5,133
Observations	65	69		
Group 2				
Total	15,934	14,519	1,414	13,965
Completed higher education	2,918	2,722	196	2,410
Not completed higher education	13,015	11,797	1,218	8,932
Agriculture	231	408	-176	132
Industry and Maintenance	3,329	2,920	409	2,593
Construction	977	954	22	735
Electricity and Gas	6,52	12,49	-5,97	5,97
Transport and Logistics	270	223	47	239
Services and Commerce	7,139	6,865	273	6,272
Observations	65	1,729		
Group 3				
Total	5,453	103,622	-98,169	77,870
Completed higher education	1,131	19,339	-18,209	13,424
Not completed higher education	4,323	84,283	-79,960	59,899
Agriculture	335	506	-171	276
Industry and Maintenance	1,091	21,032	-19,942	20,442
Construction	242	6,609	-6,367	5,047
Electricity and Gas	7,21	26,53	-19,31	30,05
Transport and Logistics	61,21	1,783	-1,722	1,331
Services and Commerce	2,544	47,114	-44,560	46,474
Observations	14	36		

Note: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*; level of employment in absolute numbers.

**Table 7 - Results for the level of wages in the municipalities of the Brazilian Northeast**

Group 1	Treatment	Control	Difference	
			Mean	DP
Total	14.69	14.44	0.24	0.20
Completed higher education	13.50	13.27	0.23	0.25
Not completed higher education	14.23	14.02	0.22	0.21
Agriculture	10.74	9.97	0.77	0.52
Industry and Maintenance	12.45	12.34	0.11	0.26
Construction	10.62	9.46	1.16 **	0.57
Electricity and Gas	10.40	9.86	0.54	0.42
Transport and Logistics	10.25	9.35	0.90 *	0.50
Services and Commerce	13.48	13.31	0.17	0.22
Observations	65	69		
Group 2				
Total	14.69	14.62	0.06	0.19
Completed higher education	13.50	13.34	0.15	0.21
Not completed higher education	14.23	14.22	0.01	0.29
Agriculture	10.74	10.35	0.40	0.30
Industry and Maintenance	12.45	12.43	0.02	0.32
Construction	10.62	10.52	0.10	0.50
Electricity and Gas	10.40	11.29	-0.90 *	0.52
Transport and Logistics	10.25	10.17	0.09	0.47
Services and Commerce	13.48	13.50	-0.02	0.21
Observations	65	1,729		
Group 3				
Total	15.51	15.15	0.36	0.81
Completed higher education	14.59	13.96	0.63	0.78
Not completed higher education	14.95	14.70	0.25	0.88
Agriculture	12.08	11.15	0.94	0.78
Industry and Maintenance	13.59	12.71	0.88	1.03
Construction	10.36	11.01	-0.65	1.61
Electricity and Gas	10.11	12.24	-2.13 *	1.17
Transport and Logistics	9.89	12.68	-2.79 **	1.40
Services and Commerce	14.36	14.12	0.23	0.87
Observations	14	36		

Note: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*; wages are in natural logarithm.

**Table 8 - Results for the number of firms in the municipalities of the Brazilian Northeast**

Group 1	Treatment	Control	Difference	
			Mean	DP
Total	1,065	222	843	631
Agriculture	19	10	9 **	4
Industry and Construction	197	31	166	166
Construction	77	8	68	62
Electricity and Gas	1	0	1	1
Transport and Logistics	0	0	0	0
Services and Commerce	848	181	667	649
Observations	65	69		
Group 2				
Total	1,065	971	94	1,061
Agriculture	19	30	-11	9
Industry and Construction	197	147	50	147
Construction	77	70	7	62
Electricity and Gas	1	1	1	0
Transport and Logistics	0	0	0	0
Services and Commerce	848	793	54	766
Observations	65	1,729		
Group 3				
Total	485	6,607	-6,122	4,771
Agriculture	31	31	1	19
Industry and Construction	66	1,312	-1,246	1,151
Construction	22	529	-507	435
Electricity and Gas	1	4	-3	4
Transport and Logistics	0	0	0	0
Services and Commerce	388	5,264	-4,876	3,970
Observations	14	36		

Note: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*; total of firms in absolute numbers.

In the estimates for Group 3 – treatment group is the municipalities with operating wind farms participating in the Proinfa program, and the control group is formed by the other municipalities with operating wind power plants – we also found a negative relationship for wages of the subsectors of electricity and gas and transportation and logistics. However, in this case, the results are not out of line, as it would be possible for municipalities with operating wind farms, but not participants in Proinfa, to have better outcomes for the labor market after the implementation of the wind farms. We also verified that the level of wages is about 21% higher for electricity and gas and 28% for transportation and logistics in municipalities not participating in Proinfa.

In Table 8, we observed that the municipalities having wind farms in operation or construction present 48% more firms in the agricultural sector, and this is the only statistically significant result that we found regarding the number of firms. This positive result for the number of firms in the agricultural sector may be related to the possibility of land leasing in the implementation of a wind

farm project in a locality. This mechanism allows the continuity of agricultural activities in the land and can strengthen the local development.

Therefore, the results of this paper suggest that the presence of wind farms in operation or construction may be related to a higher level of wages in the construction and transportation and logistics subsector. Moreover, the results suggest that the presence of wind power plants may shift resources to the agricultural sector, boosting this activity in the local economy. This latter result is in line with the study of Bergmann, Hanley, and Wright (2006) for Scotland, which shows that the most significant socioeconomic benefits of implementing renewable energy sources, especially wind energy, occur in rural areas.

Finally, regarding the Proinfa program, the results indicate that the effects of the program on employment, wages, and the number of firms can be limited since the presence of wind farms out of the program is related to a higher level of wages in the electricity, gas, transportation, and logistics subsectors.

## 7. Concluding remarks

This study aims to analyze the relationship between the implementation of wind farms and local development in the municipalities of the Northeast region of Brazil in 2013. We focused on labor market outcomes, such as employment, wages, and the number of firms. The main contribution of this article is to relate the presence of wind farms to the labor market outcomes through the use of the spatial propensity score matching method and the calculation of the average treatment effect on the treated.

We found that the total level of wages in the construction, transportation, and logistics sectors are higher in municipalities that have wind power plants in operation or construction. This difference accounts for about 11% in the construction subsector and 9% in the transportation and logistics subsectors. Also, municipalities that have wind farms in operation or construction contain 48% more firms in the agricultural sector, thereby suggesting that the presence of wind power plants leads to the shifting of resources to the agricultural sector, boosting this activity in the local economy.

In addition, wages are higher in the electricity and gas and transportation and logistics subsectors in municipalities with operating wind farms but not participating in the Proinfa program. This result serves as evidence that the impacts of this government program may have been limited in the case of wind farm implementation. Finally, no statistically significant results were found regarding the relationship between the presence of wind farms and the total level of employment.

The results of this paper suggest that the presence of wind farms may generate positive impacts on the local development of Brazilian municipalities. Thus, given the natural conditions of the country, especially in its Northeast region, it could be expected that the expansion in the number of projects of this type of renewable energy will contribute to the increase of the local welfare. We can also expect that once auctions for renewable energy projects continue to take place, as a process of transition to an increasingly cleaner energy matrix, these positive impacts can be even higher, thereby positively affecting jobs, income, and the number of firms in the region.

It is noteworthy that, despite the relevance of the correlations found in this paper, it is not the purpose of the study to identify and estimate the isolated effect of wind energy. For this purpose, it would be necessary to use longitudinal data and econometric panel methods or other methods, allowing achieving robustness of the effects

## References

- ALMEIDA E. *Econometria espacial*. Campinas: Alínea. 2012.
- ALVES, J. J. A. Análise regional da energia eólica no Brasil. *Revista Brasileira de Gestão e Desenvolvimento Regional*, v. 6, n. 1, p. 165-188, 2010.

- ANSELIN, L. Spatial econometrics. In: BALTAGI, B. H. (ed.). *A Companion to Theoretical Econometrics*. Oxford: Blackwell Publishing, 2001. Cap. 14, p. 310-330.
- ANEEL. *Micro e minigeração distribuída*. Sistema de Compensação de Energia Elétrica. Brasília, DF, Brasil: Centro de Documentação (Cedoc). 2014.
- BERGMANN, A.; HANLEY, N.; WRIGHT, R. Valuing the attributes of renewable energy investments. *Energy Policy*, v. 34, n. 9, p. 1004-1014, 2006.
- BIRD, L.; BOLINGER, M.; GAGLIANO, T.; WISER, R.; BROWN, M.; PARSONS, B. Policies and market factors driving wind power development in the United States. *Energy Policy*, v. 33, n. 11, p. 1397-1407, 2005.
- BLANCO, M. I.; RODRIGUES, G. Direct employment in the wind energy sector: An EU study. *Energy Policy*, v. 37, n. 8, p. 2847-2857, 2009.
- CHAGAS, A. L. S.; TONETO-JR, R.; AZZONI, C. R. A spatial propensity score matching evaluation of the social impacts of sugarcane growing on municipalities in Brazil. *International Regional Science Review*, v. 35, n. 1, p. 48-69, 2012.
- COSTA, R. A.; CASOTTI, B. P.; AZEVEDO, R. L. S. *Um panorama da indústria de bens de capital relacionados à energia eólica*. BNDES Setorial, Rio de Janeiro, n. 29, p. 1-50, 2009.
- DEL RÍO, P.; BURGUILLO, M. Assessing the impact of renewable energy deployment on local sustainability: Towards a theoretical framework. *Renewable and Sustainable Energy Reviews*, v. 12, n. 5, p. 1325-1344, 2008.
- EDENHOFER, O.; PICHS-MADRUGA, R.; SOKONA, Y.; SEYBOTH, K. *IPCC special report on renewable energy sources and climate change mitigation*. Prepared By Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 2011.
- EJDEMO, T.; SÖDERHOLM, P. Wind power, regional development and benefit-sharing: the case of Northern Sweden. *Renewable and Sustainable Energy Reviews*, v. 47, p. 476-485, 2015.
- EPE. Empresa de Pesquisa Energética. Nota Técnica: *O Compromisso do Brasil no Combate às Mudanças Climáticas: Produção e Uso de Energia*. Disponível em <http://www.epe.gov.br>. Rio de Janeiro/RJ. 2016.
- GWEC. Global Wind Energy Council. *Global wind statistics 2011 Report*. Brussels, Belgium: Global Wind Energy Council (GWEC), 2011.
- GWEC. Global Wind Energy Council. *Global Wind Report 2015: Annual Market Update*. Brussels, Belgium: Global Wind Energy Council (GWEC), 2015.
- GRUBB, M. J.; MEYER, N. I. Wind energy: resources, systems and regional strategies. In: JOHANSSON, T. B. (ed.). *Renewable Energy: Sources for Fuels and Electricity*. Washington, D.C.: Island Press, 1993.
- KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington: World Bank Publications, 2010.
- KLICK, H.; SMITH, E. Public understanding of and support for wind power in the United States. *Renewable Energy*, v. 35, n. 7, p. 1585-1591, 2010.
- LEHR, U.; NITSCH, J.; KRATZAT, M.; LUTZ, C.; EDLER, D. Renewable energy and employment in Germany. *Energy Policy*, v. 36, n. 1, p. 108-117, 2008.

- LESAGE, James; PACE, Robert Kelley. *Introduction to spatial econometrics*. Chapman and Hall/CRC, 2009.
- MARTINS, F. R.; GUARNIERI, R. A.; PEREIRA, E. B. O aproveitamento da energia eólica. *Revista Brasileira de Ensino de Física*, v. 30, n. 1, p. 1-13 2008.
- MORENO, B.; LOPEZ, A. J. The effect of renewable energy on employment. The case of Asturias (Spain). *Renewable and Sustainable Energy Reviews*, v. 12, n. 3, p. 732-751, 2008.
- OHL, C.; EICHHORN, M. The mismatch between regional spatial planning for wind power development in Germany and national eligibility criteria for feed-in tariffs - A case study in West Saxony. *Land Use Policy*, v. 27, n. 2, p. 243-254, 2010.
- SIMAS, M.; PACCA, S. Assessing employment in renewable energy technologies: A case study for wind power in Brazil. *Renewable and Sustainable Energy Reviews*, v. 31, p. 83-90, 2014.
- STAKHOVYCH, S.; BIJMOLT, T. H. A. Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, v. 88, n. 2, p. 389-408, 2009.
- SZARKA, J. Wind power, policy learning and paradigm change. *Energy Policy*, v. 34, n. 17, p. 3041-3048, 2006.
- UHR, D. A. P., CHAGAS, A. S., UHR, J. G. Z. Demand for residential energy in Brazil revisited: a dynamic panel data approach. *The Empirical Economics Letters*, v. 16, n. 8, p. 747-753, 2017.
- UHR, D. A. P., CHAGAS, A. S., UHR, J. G. Z. Estimation of elasticities for electricity demand in Brazilian households and policy implications. *Energy Policy*, v. 129, p. 69-79, 2019.
- WEI, M.; PATADIA, S.; KAMMEN, D. M. Putting renewables and energy efficiency to work: How many jobs can the clean energy industry generate in the US? *Energy Policy*, v. 38, n. 2, p. 919-931, 2010.


## ORCID

Thiago Pastorelli Rodrigues  <https://orcid.org/0000-0001-5762-8497>

Solange Ledi Gonçalves  <https://orcid.org/0000-0002-4607-9151>

André Squarize Chagas  <https://orcid.org/0000-0001-7870-8736>

The views and opinions expressed in this study are those of the authors and do not necessarily reflect the position of EPE.

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.

## Appendix

Table A1 - T-tests for the explanatory variables (prior and after the matching)

Group 1 Variables	Prior the matching			After the matching		
	T	C	<i>t</i> -statistic	T	C	<i>t</i> -statistic
Wind	6.35	6.05	-2.82 ***	6.22	6.19	-0.26
Area (log)	6.36	6.52	0.92	6.33	6.60	0.86
Population (log)	10.05	9.76	-1.67 *	9.88	9.99	0.42
Ratio Agriculture and total GDP	0.10	0.10	0.33	0.12	0.09	-1.34
MR	0.09	0.04	-1.13	0.04	0.04	0.03
Observations	65	69		65	69	
Group 2 Variables	Prior the matching			After the matching		
	T	C	<i>t</i> -statistic	T	C	<i>t</i> -statistic
Wind	6.35	5.17	-12.41 ***	6.59	6.57	-0.15
Area (log)	6.36	6.11	-1.75 *	6.50	6.44	-0.16
Population (log)	10.05	9.60	-3.67 ***	10.24	10.34	0.25
Ratio Agriculture and total GDP	0.10	0.14	2.76 ***	0.08	0.10	0.89
MR	0.09	0.08	-0.33	0.06	0.14	0.76
Observations	65	1,728		65	1,728	
Group 3 Variables	Prior the matching			After the matching		
	T	C	<i>t</i> -statistic	T	C	<i>t</i> -statistic
Wind	6.34	6.47	0.70	5.82	6.04	0.57
Area (log)	6.14	6.36	0.69	6.04	5.89	-0.11
Population (log)	10.51	10.05	-1.25	10.48	12.43	0.83
Ratio Agriculture and total GDP	0.13	0.08	-1.83 *	0.23	0.17	-0.32
MR	0.07	0.14	0.65	-	-	-
Observations	14	36		14	36	

Note: p-value < 0.10: \*, p-value < 0.05: \*\*, p-value < 0.01: \*\*\*. T-tests after the matching are performed for one of the blocks generated in the PSM. As this matching procedure ensures that the mean propensity score is not different for the treatment and control groups, the selection criterion for blocks was the maximum number of observations in each group.



**FLUXOS MIGRATÓRIOS SOB A PERSPECTIVA DA CENTRALIDADE DAS CIDADES:  
UMA INVESTIGAÇÃO PARA OS CHEFES DE FAMÍLIAS (2000-2010)\***

**Clauber Eduardo Marchezan Scherer**

Doutor em Economia pelo Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR) da Universidade Federal de Minas Gerais (UFMG). Assessor parlamentar na Câmara Municipal de Porto Alegre  
E-mail: clauberscherer@yahoo.com.br

**Alysson Lorenzon Portella**

Doutorando em Economia dos Negócios no Instituto de Ensino e Pesquisa (INSPER)  
E-mail: alyssonlp@al.insper.edu.br

**RESUMO:** Este trabalho investiga como a escolha de migração feita pelo chefe de família relaciona-se com a inserção hierárquica funcional das cidades brasileiras. Para isso, combinam-se os dados censitários de 2000 e 2010 referentes à migração de data fixa, juntamente com o arranjo da rede urbana brasileira proposta no REGIC (2008), aplicando o modelo logístico multinomial para analisar como a escolha de movimentos migratórios entre a hierarquias está correlacionado com características individuais e regionais dos migrantes. Como resultados, observou-se que migrações cujo destino é o Sudeste associam-se a uma mudança ascendente de hierarquias. Além disso, pessoas com ensino superior completo têm maior probabilidade de migração para cidades de menor hierarquia. Esses resultados são indícios de que, embora a polarização exercida pelo Sudeste e pelas capitais continue dominante, há uma mudança do perfil migratório em direção às centralidades intermediárias, apontando para uma ressignificação do papel funcional desempenhado por esses centros.

**Palavras-chave:** Hierarquias urbanas; Migração; Análise multinomial.

**Classificação JEL:** C21; J11; R23.

**MIGRATORY FLOWS FROM CENTRALITY OF CITIES PERSPECTIVE: AN  
INVESTIGATION FOR FAMILIES' HEAD (2000-2010)**

**ABSTRACT:** This paper investigates how the choice of migration made by the head of the household is related with the hierarchical structure of Brazilian cities. We combine data from the Brazilian Census of 2000 and 2010 related to migration of fixed date with the Brazilian urban network proposed by REGIC (2008), using multinomial logistic models to analyze how the choice of migratory between hierarchies is correlated with individual and regional characteristics of migrants. The results show that migrations towards the Southeast region tend to have to cities of higher hierarchy as destiny. Also, more qualified workers are more likely to migrate to cities of lower hierarchy. These results suggest that although the Southeast region is still a dominant center, there is a change in the migration profile towards cities of intermediate hierarchy, corroborating the idea that these centers are now playing a new role in the Brazilian network of cities.

**Keywords:** Urban hierarchy; Migration; Multinomial analysis.

**JEL Codes:** C21; J11; R23.

## 1. Introdução

Este trabalho propõe-se a estudar a concentração e desconcentração da rede urbana no Brasil nas últimas décadas a partir dos fluxos migratórios entre cidades de diferentes hierarquias, levando em conta as características individuais e regionais dos migrantes. Assim, busca-se entender os movimentos de migração entre cidades com distintas posições funcionais na rede, avaliando se há uma tendência à concentração (migração para cidades com hierarquias maiores na rede) ou desconcentração (migração para centros que ocupam hierarquias menores) e, se, em alguma medida, é possível identificar diferenças no perfil das pessoas que escolhem cidades de maior ou menor hierarquia como destino. Nesse processo, privilegiam-se dois aspectos: (i) em que medida as regiões de origem e destino afetam as chances de migrações para distintas hierarquias; (ii) se o tipo de migração está relacionado com o grau de escolaridade do migrante.

Para o planejamento público, essas questões são relevantes, pois, embora se espere que a migração equalize diferenciais salariais existentes entre diferentes locais, caso exista uma seletividade no perfil migratório, ela pode reforçar as disparidades regionais (GREENWOOD, 1997). Assim, aqui, analisa-se a (des)concentração da rede urbana no Brasil nas últimas décadas a partir dos fluxos migratórios entre cidades de diferentes hierarquias, levando em conta as características individuais dos migrantes. Conciliam-se, assim, duas noções distintas: uma que parte de uma perspectiva micro e procura explicar a decisão de migração individual e/ou familiar baseada em uma análise utilitarista de ganhos líquidos entre a origem e o destino da migração; a outra, macro, e que se concentra no entendimento das relações de trocas entre os centros urbanos a partir de uma análise focada no papel hierárquico desempenhado pelas cidades na rede urbana. Sendo assim, este trabalho diferencia-se pela abordagem pioneira, ao incluir ambos os aspectos no estudo da migração interna no Brasil.

Essa abordagem justifica-se na medida em que surgem evidências acerca de um aumento da articulação da rede urbana brasileira (IPEA/IGBE/NESUR, 2001; REGIC, 2008; SIMÕES; AMARAL, 2011), ao mesmo tempo em que os movimentos internos do país da segunda metade do século XX em diante são fortemente relacionados aos processos de urbanização e redistribuição espacial da população (BAENINGER, 2011). Assim, aumenta-se a complexidade de compreensão dos fluxos migratórios ocorridos no Brasil e, por isso, esforços para compreender essas novas relações contribuem para o entendimento dos deslocamentos populacionais e, de modo mais geral, para caracterização socioeconômica do país.

A partir disso, com auxílio dos dados censitários 2000 e 2010, juntamente com o REGIC (2008), neste trabalho são levantados questionamentos acerca da relação existente entre hierarquia das cidades e fluxos migratórios, investigando se o perfil do migrante interno brasileiro tende a ser diferente de acordo com os movimentos entre as hierarquias das cidades de origem e de destino, isto é, se existem características individuais relacionadas ao grau de escolaridade e grupos de indivíduos de acordo com a região de origem e/ou destino que estão mais associados à escolha de subir, descer ou se manter em cidade de mesma hierarquia.

Para explorar essas questões, o trabalho foi organizado em cinco seções: além desta introdução, na seção dois, são discutidas as ideias conceituais que norteiam o estudo, bem como é feita uma breve revisão acerca das principais características dos fluxos migratórios internos recentes no país; na seção três, é apresentada a metodologia utilizada para a manipulação dos dados censitários e sua integração com o REGIC, bem como a modelagem econométrica utilizada; a seção quatro apresenta os principais resultados da pesquisa, sendo seguida da seção cinco onde são apresentadas as considerações finais deste trabalho.

## 2. Revisão bibliográfica

Neste trabalho, procura-se investigar a existência de correlação entre características individuais de migrantes e hierarquias urbanas no Brasil. Por isso, é importante que, primeiramente, seja feita uma discussão sobre hierarquias urbanas e migração. Assim, nesta seção, estruturada em quatro tópicos, discute-se a Teoria do Lugar Central e sua tradição de uso pelo IBGE no Brasil, bem como

apresenta-se uma breve discussão teórica sobre os processos migratório, dando destaque para as discussões que focam em discussões de movimentos populacionais que são internos e concentram-se em aspectos regionais, para então apresentarem-se as tendências observadas para os fluxos migratórios no Brasil ao longo das últimas décadas.

### *2.1 A Teoria do Lugar Central: as concepções de Christaller e Lösch*

A Teoria do Lugar Central (TLC) foi inspirada no funcionamento das cidades da primeira metade do século XX. Por meio dessa, procura-se explicar como centros hierarquicamente diferenciados – medida dada pela complexidade dos bens ofertados em cada um dos lugares – surgem e se distribuem no espaço.

A origem dessa teorização está intimamente ligada aos trabalhos de Christaller (1966, [1933]) e Lösch (1954, [1944]). Ainda que cheguem a conclusões semelhantes, os caminhos percorridos por eles diferem, pois, enquanto Christaller parte de uma construção indutiva, utilizando por base o sistema urbano do sul da Alemanha, Lösch constrói um modelo dedutivo, originado de uma análise microfundamentada (PARR, 2002).

Christaller (1966) desenvolve em seu trabalho o princípio da centralidade, o que permite hierarquizar as cidades a partir de seu nível de importância econômica. Essa relevância seria avaliada de acordo com o número de atividades econômicas existentes e da sua diversidade de oferta. Isso porque, para o autor, os serviços e bens centrais, devido à sua complexidade, são produzidos e ofertados em poucos pontos, embora sejam consumidos em um espaço maior. Assim, uma maior gama de oferta de bens e serviços corresponde a uma maior centralidade que será tão maior quanto for a extensão do seu “domínio”, constituindo sua região de influência.

Lösch (1954), por sua vez, partindo do pressuposto fundamental de uma planície homogênea, mas com o uso de economias de escala e custos de transporte, desenvolve um modelo em que há uma distribuição não homogênea da produção e da oferta de bens e serviços, de modo a permitir que ocorram aglomerações e que essas sejam diferentes em termos de hierarquia. Isso porque, em função do aumento dos custos de transporte, à medida que vão se afastando dos centros de produção, os produtos ficam mais caros. Dessa forma, chega-se a um limite no qual se torna mais barato adquirir o mesmo produto de outro produtor.

Assim, as economias de escala atuam de modo a baratear o custo de produção médio do produto, possibilitando sua venda para um número maior de consumidores em mercados mais distantes do centro produtor. A interação dessas duas forças, uma centrífuga (custos de transporte) e outra centrípeta (escala) irá determinar a área de mercado. Lösch (1954) conclui que a forma ideal dessas áreas é um hexágono, que possibilita a cobertura de todas as áreas sem qualquer intersecção, evitando a perda de consumidores e maximizando o lucro das firmas.

Vale destacar ainda que a oferta de bens estará localizada nos pontos em que há uma maior densidade de demanda. Dessa forma, as diversas redes de oferta de produtos terão um ponto em comum, capaz de maximizar a demanda local. Em torno dela, irão se desenvolver outros centros menores, dos mais diversos tamanhos, mas hierarquicamente menores, configurando assim uma rede de cidades.

Ainda que geometricamente os sistemas espaciais dos dois autores forneçam resultados bastante similares, existem algumas diferenças importantes entre eles, como, por exemplo, o fato de que, enquanto no sistema de Christaller determinada área de mercado será atendida por um número *mínimo* de pontos de produção, no sistema de Lösch o resultado se dá com a alocação *ótima máxima* de oferta de bens para atender as áreas (McCANN, 2013).

Apesar de apresentar certa rigidez no que diz respeito às configurações que as redes tomam, as ideias trazidas pela Teoria do Lugar Central nas produções de Christaller (1966) e Lösch (1954) ainda são relevantes para o entendimento da evolução e estabelecimento das redes de cidades. Tanto é verdade que, no caso brasileiro, existe uma tradição de longa data no IBGE que utiliza essas ideias para a avaliação da rede nacional, fato que será discutido na próxima seção.

## 2.2 A tradição brasileira no estudo da rede de cidades

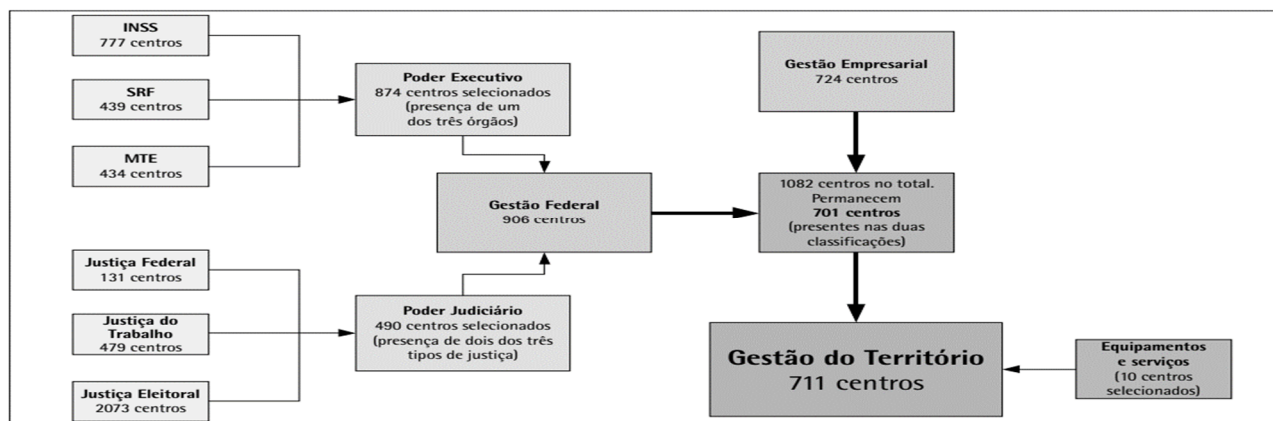
O estudo da rede urbana brasileira partindo de uma orientação christalleriana, com base na Teoria do Lugar Central (TLC), tem uma longa tradição em publicações conduzidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Iniciando com “*Divisão do Brasil em Regiões Funcionais Urbanas*” de 1972, seguiram-se mais três estudos que mantêm essa linha teórica, sendo a mais recente – que servirá de base neste trabalho – de 2008. Como essa é, muito provavelmente, a publicação mais robusta disponível acerca da rede urbana brasileira, configura-se como um subsídio indispensável em qualquer tentativa de caracterização da rede de cidades do país.

Nessa abordagem, parte-se da ideia de que a organização espacial do território é manifestada tanto por elementos fixos – cidades, indústrias, etc – quanto pelos fluxos que estabelecem interações entre eles. Dessa interação surgem as redes, nas quais os centros se articulam pelas funções que oferecem, de forma a gerar as configurações de hierarquias sobre o território (REGIC, 2008).

Para o estabelecimento da rede e seus diferentes níveis, inicialmente, é estabelecida uma classificação dos centros, privilegiando a função de gestão do território. Em um segundo momento, são investigados eventuais casos que, apesar de não apresentarem nível de gestão representativo, exercem uma centralidade chave em função das ofertas de equipamentos e serviços específicos nessas cidades. Por fim, na última fase, após a realização dessa dupla filtragem, são investigadas as ligações entre cidades de modo a delimitar as áreas de influência desses locais.

Para a definição dos centros de gestão do território, buscaram-se informações de subordinação administrativa no setor público federal, para gestão pública, e da localização das sedes e filiais de empresas, para a gestão empresarial. A identificação dos centros dotados de capacidade de oferta de equipamentos e serviços levou em conta o grau de complexidade da oferta<sup>1</sup>, ligações aéreas, internações hospitalares, cobertura televisiva, ensino superior, setor bancário e domínios de *internet*. Desse processo, na primeira etapa, foram definidos 701 centros<sup>2</sup>, sendo agregados a eles, na segunda etapa, mais dez cidades, conforme observado na Figura 1.

**Figura 1 - Etapas da definição dos centros de gestão do território**



Fonte: Adaptado a partir do REGIC (2008).

A partir da definição dos 711 centros de gestão do território, são definidas suas regiões de influência, ou seja, municípios e ACPs que são hierarquicamente subordinados a esses locais. Como produto do REGIC (2008), tem-se a classificação da rede urbana brasileira em cinco grandes níveis (metrópoles, capitais regionais, centros sub-regionais, centros de zona, centros locais), os quais são

<sup>1</sup> Para a definição das centralidades, a partir da oferta de bens e serviços, a fonte foi o Cadastro de Empresas 2004, do qual foram extraídas informações de emprego para atividades comerciais e de serviços nas classes do CNAE 1.0. Assim, calculou-se o percentual de classes presentes no município em relação ao total nos grupos.

<sup>2</sup> No REGIC (2008), incorporou-se nas análises a ideia de conurbações urbanas, por meio das Áreas de Concentração da População (ACP), em que todos os municípios membros de uma ACP são interpretados como uma única unidade.

subdivididos em dois ou três grupos. Vale notar ainda que o REGIC (2008) tem como base de dados informações coletadas no ano de 2007.

### 2.3 Breve fundamentação teórica sobre migrações

O que determina a migração de um indivíduo? A compreensão dos mecanismos por trás dessa decisão não é nada simples, de modo que entender as motivações que levam uma pessoa a mover-se no espaço torna-se uma questão intrigante – e longe de consenso – aos pesquisadores do tema. Conforme destacado por Golgher (2004), à migração estariam associados fatores *push* e *pull*. Os fatores *push* seriam aqueles que fazem com que as pessoas não desejem mais permanecer em seus locais de moradia e os fatores *pull* seriam aqueles responsáveis por tornar o local de destino mais atrativo ao potencial imigrante.

Nas interpretações econômicas tradicionais acerca do tema, a decisão de mudança se dá sob uma perspectiva de maximização de utilidade, cuja migração resulta de um comportamento voltado para o futuro, que visa maximizar o bem-estar de um indivíduo ou de uma família ao longo de um horizonte de tempo definido. Sendo assim, variáveis observáveis para os migrantes, como as expectativas de diferenças salariais reais ou probabilidades de melhores oportunidades de emprego, constituiriam as principais forças da migração regional (TODARO, 1969; MASSEY et al., 1998; CUSHING; POOT, 2003; RODRÍGUEZ-POSE et al., 2015).

Todaro (1969) e, posteriormente, Haris e Todaro (1970) apresentam modelagem canônica na visão neoclássica – que vê a migração como uma decisão baseada na maximização intertemporal. Por essa modelagem, a migração é causada por diferenças geográficas na oferta e demanda de mão de obra. As regiões com oferta de trabalho relativamente limitada em relação ao capital geralmente terão um salário mais alto que atrairá uma grande entrada de mão de obra das regiões de baixos salários. Esse influxo do fator de produção e trabalho é espelhado por uma saída do fator capital.

Borjas (1990) apresenta um refinamento, dando uma perspectiva microeconômica ao modelo de Haris e Todaro (1970). No modelo de Borjas (1990), os migrantes comparam o custo de se mudar para locais alternativos, maximizando o retorno líquido – diferença entre os ganhos esperados líquidos correspondentes às habilidades individuais na região de origem e na região de destino – em um determinado horizonte de tempo. Ao subtrair os custos esperados de mudança, encontra-se o retorno líquido esperado da migração.

A principal diferença entre as duas abordagens está no fato de que o movimento da população entre regiões é determinado pelos diferenciais nos salários e nas taxas de desemprego em um contexto macro, enquanto o retorno líquido esperado da migração é a principal consideração para a decisão individual de mudar para outra região. Em outras palavras, os diferenciais salariais dados em função da alocação dos fatores de produção determinam a migração em um modelo macro, enquanto os retornos líquidos esperados dão a motivação na perspectiva micro (PARIKH et al., 2003). Algumas extensões desse modelo são importantes, como o caso de Massey (1993), que analisa a decisão de migração sob a perspectiva de ser essa uma decisão da família, de modo que se maximiza uma função utilidade conjunta e não individual. Por essa ótica, é dada maior relevância para a migração enquanto movimento populacional do que propriamente como movimento de fator trabalho.

De modo geral, esses trabalhos são fundamentais na delimitação de aspectos motivadores da migração como fenômeno econômico. Nesse contexto, aos propósitos deste trabalho, cabe focar nas migrações internas, ou seja, nos movimentos populacionais permanentes que acontecem dentro das fronteiras geográficas de um país. Esses movimentos, que dentro da ótica neoclássica têm entre suas causas alterações demográficas, fatores macroeconômicos, de mercado de trabalho e desenvolvimentos tecnológicos (GREEN, 2017), são fundamentais para o funcionamento eficiente das economias locais, pois, a partir do movimento do fator trabalho, permitem a equalização de desequilíbrios regionais na alocação de fatores de produção (LUCAS, 2015).

Porém, a migração tem diversas facetas (pendular, internacional, regional), além de apresentar variabilidade para sua ocorrência entre países e, dentro desses, até mesmo entre suas regiões. Assim,

aos aspectos gerais devem ser agregadas características e nuances que fazem esse processo ser único em cada local, o que aumenta sensivelmente a complexidade de sua análise.

Bell et al. (2015), por exemplo, identificam um padrão singular na migração interna dos países da América Latina, onde percebe-se uma coluna de alta mobilidade nos Andes declinando rapidamente para o leste e norte do continente em direção à América Central. Ainda, Bernard et al. (2017) observam que, na América Latina, o principal fator propulsor das mudanças demográficas vem das migrações internas, já que, nesses países, percebem-se taxas de fertilidade estáveis e baixo índice de migração internacional – exceção para o caso recente observado na Venezuela.

Diferenças regionais de salários reais, disponibilidade e qualidade de amenidades urbanas, chances de encontrar um bom emprego, redes sociais já estabelecidas, distância e restrições financeiras são todos fatores que têm seus custos considerados quando tomada a decisão de migração (LUCAS, 2015). Embora se espere que a migração iguale os diferenciais salariais regionais, na presença de heterogeneidade nos movimentos migratórios – quem vai e para onde vai –, ela pode reforçar as disparidades (GREENWOOD, 1997).

Assim, além das características pessoais, as locais também aparecem na tomada de decisões sobre migração interna. Forças de convergência e divergência, atuando por meio dos efeitos de aglomeração, tendem a aprofundar as diferenças de renda real entre os locais e, conseqüentemente, exercer influência diferenciada nas decisões de migração (BAFURI, 2018). Se considerada a tendência de que trabalhadores qualificados se beneficiam mais da migração para regiões ricas, enquanto trabalhadores pouco qualificados tendem a migrar para regiões pobres para diminuir seus custos de vida (MORETTI, 2012), a migração inter-regional torna-se um fenômeno importante na compreensão das dinâmicas regionais.

Nesse sentido, uma ampla literatura tem encontrado resultados nos quais as grandes cidades parecem ser, *ceteris paribus*, mais atraentes do que as cidades pequenas, especialmente para aquelas pessoas mais qualificadas (FRANKLIN, 2003; WINTERS, 2011; HAAPANEN; TERVO, 2012; BUCH et al., 2014). Esses resultados apontam para a importância de amenidades que estão ligados também ao tamanho da cidade, fator que é intimamente ligado com a posição hierárquica ocupada pelos centros na rede.

Disso tudo, percebe-se que, a partir de desenvolvimentos teóricos iniciais que priorizavam aspectos individuais, tem se avançado cada vez mais no sentido de entender nuances mais específicas associadas aos processos migratórios. Nesse contexto é que são incorporados aspectos da localização como determinantes na atração e retenção de migrantes. É com esse viés de interconexão entre a decisão individual e as variações que ocorrem em função das diferentes percepções do local de destino que este trabalho pretende inserir-se.

Assim, realiza-se aqui um exercício pioneiro na literatura nacional, em que se combina a literatura que enfatiza o comportamento micro – focado nas características do indivíduo, ou de uma família, e as influências que podem desencadear uma mudança para outro local – com os aspectos macro – literatura concentrada em “lugares” ao invés de “pessoas”. Isso é feito para compreender melhor os fluxos e as características migratórias observadas para o Brasil no período recente.

#### 2.4 Fluxos migratórios no Brasil contemporâneo

Para o Brasil, não é possível compreender seu curso migratório interno sem considerar o processo de rápida urbanização pelo qual passou o país. Como pode ser observado através da Tabela 1, em um período de 60 anos, o país saiu de uma condição em que 36,16% da população era urbana, para uma realidade em que 84,36% da população se encontra nessa situação.

Esse resultado ocorre, em um primeiro momento, a partir de um processo de migração do campo para a cidade, principalmente em direção aos grandes centros da rede e, a partir da década de 1970, deixam de ser do tipo campo-cidade e passam a ser dominadas por movimentos inter-regionais (MATOS, 2013). Nesse sentido, entre 1980 e 1990, os tradicionais destinos de migrantes (sobretudo a região Sudeste) começam a perder sua relevância, aumentando inclusive o peso das migrações de retorno (LIMA, 2013). Dentre os fatores apontados para isso, encontra-se o processo de espraiamento

industrial que estimulou o adensamento da rede urbana brasileira e os vínculos de interdependência e complementaridade entre cidades estratégicas nas diferentes partes do sistema (MATOS, 2013).

**Tabela 1 - Distribuição da população brasileira rural/urbana (1950-2010)**

Ano	Urbanização (%)	População Total		População Urbana		População Rural	
		Total	Δt (%)	Total	Δt (%)	Total	Δt (%)
1950	36,16	51.944.397	-	18.782.891	-	33.161.506	-
1960	45,08	70.992.343	36,67	32.004.817	70,39	38.987.526	17,57
1970	55,98	94.508.583	33,13	52.904.744	65,30	41.603.839	6,71
1980	67,70	121.150.573	28,19	82.013.375	55,02	39.137.198	-5,93
1991	75,47	146.917.459	21,27	110.875.826	35,19	36.041.633	-7,91
2000	81,25	169.799.170	15,57	137.953.959	24,42	31.845.211	-11,64
2010	84,36	190.755.799	12,34	160.925.804	16,65	29.829.995	-6,33

Fonte: Elaboração própria a partir dos Censos.

Além disso, em função do esgotamento da capacidade de absorção dos fluxos migratórios das grandes cidades, percebe-se um processo de desconcentração, em que cidades menores e de menor hierarquia começam a ganhar maior relevância. Destaca-se neste processo a emergência e a representatividade assumidas pelas cidades médias, geralmente caracterizadas por uma população que vai de 100 a 500 mil habitantes (ANDRADE et al., 2000).

Nesse cenário, na década seguinte – entre 1990 e 2000 –, as mudanças nos padrões migratórios brasileiros são estimuladas muito mais pelo aumento da capacidade de atração/retenção de determinadas regiões, especialmente as cidades médias, do que pela perda de dinamismo de seus tradicionais destinos migratórios – São Paulo e Rio de Janeiro. Esses movimentos podem ser vistos como reflexos do amadurecimento da rede urbana brasileira, em que se percebe o surgimento de centros funcionalmente representativos, para além das metrópoles, tornando-se assim lugares de interesse na escolha de transição dos potenciais migrantes, indicando um reforço da tendência de novos espaços de migração de curta distância, os quais estão mais relacionados ao âmbito de suas próprias regiões (BAENINGER, 2012).

Somadas a essas características, no início do século XXI – entre 2000 e 2010 –, a maior dinâmica da economia interna, impulsionada pela expansão de renda das famílias, fez com que aumentasse a dispersão dos fluxos migratórios internos (SILVEIRA NETO, 2008). Nesse processo, foi reforçada a importância, como regiões absorvedoras de migrantes no país, de locais que têm cidades médias atuando como centros polarizadores de porções da rede de cidades, rompendo assim com o padrão de migração nordeste-sudeste, até então dominante (LIMA, 2013). Além disso, nota-se ainda a crescente expressividade tomada pelas migrações intrametropolitanas (AMARAL, 2013). Nesse contexto, se até algum tempo os fluxos migratórios eram controlados sobretudo pela primazia exercida pela região Sudeste – em especial, pelas cidades do Rio de Janeiro e de São Paulo – e pelos processos de ocupações de fronteiras, sua compreensão, no contexto atual, torna-se tarefa mais desafiadora (BRITO; CARVALHO, 2010; BAENINGER, 2011).

Conforme mostra a Tabela 2, as migrações inter-regionais do país têm diminuído em volume ao longo do tempo. Porém, longe de significar um caminho à estagnação das migrações, esses resultados denotam novos arranjos da própria migração interna, com novas modalidades de deslocamentos populacionais em âmbitos locais e regionais. Conforme destacado por Baeninger (2012), essa constatação reforça o fato de que as dinâmicas regionais, e, em especial, as de curta distância, passam a imprimir especificidades às migrações urbanas-urbanas nos movimentos populacionais recentes do país.

**Tabela 2 - Saldos migratórios de data fixa inter-regionais (1980-2010)**

Região	Período			
	1975/1980	1986/1991	1995/2000	2005/2010
Norte	316.464	9.758	76.480	36.482
Nordeste	-870.640	-869.909	-743.188	-701.078
Sudeste	1.005.055	680.561	546.723	325.495
Sul	-627.707	-167.680	33.465	76.294
Centro-Oeste	287.148	300.143	288.740	262.809

Fonte: Elaboração própria a partir dos Censos.

Nesse sentido, Nascimento, Silva e Lucas (2017) mostram que, apesar de haver uma diminuição das chances de migração intermunicipal nos quinquênios 1995/2000 e 2005/2010, percebe-se o estabelecimento de municípios para os quais as probabilidades de migrar são maiores. Da mesma forma, Baptista (2017), analisando os mesmos períodos, a partir das Taxas Líquidas Migratórias (TLMs), encontra que a maioria das capitais (exceção para a região Norte e outras poucas nas demais regiões) apresentou TLMs negativas. Entretanto, as migrações em seus municípios adjacentes foram positivas, corroborando a ideia de espalhamento para os municípios periféricos das grandes cidades.

Os resultados encontrados por Barufi (2018) mostram que os fluxos migratórios parecem aumentar as disparidades regionais, com os maiores benefícios sendo capturados pelas regiões mais desenvolvidas que são os principais destinos do grupo de trabalhadores mais educados. Ao analisar a escolha migratória feita por pessoas com maior grau de qualificação, outros estudos mostram que os fluxos migratórios de migrantes qualificados entre os municípios no Brasil são em grande parte impulsionados pela maior renda esperada no destino, população, PIB *per capita*, grau de industrialização e amenidades urbanas (MATA et al., 2007; TAVEIRA; ALMEIDA, 2014). Resultado muito similar aos encontrados por Freguglia, Gonçalves e Silva (2014) ao analisarem os determinantes da saída dos trabalhadores qualificados no nível estadual.

Conforme notado por Lima et al. (2016), a compreensão da dinâmica de interação entre desenvolvimento e fluxos de capital/trabalho segue sendo crucial para analisar as trajetórias de crescimento regionais, de modo que estudos voltados para o entendimento da interação entre rede urbana e migração são centrais para uma melhor compreensão do caso brasileiro. Assim, o esforço realizado aqui, ao investigar, ainda que de forma exploratória, a relação entre fluxos migratórios e hierarquias das cidades brasileiras, nos períodos 1995-2000 e 2005-2010, mostra-se bastante pertinente frente aos desafios postos para compreensão dos fluxos migratórios internos no país, sobretudo na busca de avançar a compreensão envolvendo a migração e seu aspecto regional no Brasil.

### 3. Metodologia

#### 3.1 Bases de dados

Como principal objetivo do trabalho, pretende-se investigar se características pessoais afetam a decisão de migração entre hierarquias. Para isso, a partir dos Censos Demográficos – 2000 e 2010 – e das Regiões de Influência das Cidades (REGIC), são estabelecidos os grupos de migrantes, suas características e as hierarquias das cidades na rede urbana brasileira.

Primeiramente, foram selecionadas, em cada um dos Censos Demográficos (2000 e 2010), informações de migrações de data fixa, as quais, a partir dessas bases, permitem obter um retrato dos movimentos migratórios de cinco em cinco anos. Assim, em ambos os Censos, parte-se como filtro da pergunta “Você sempre morou neste município?”. Aqueles que respondem negativamente são considerados migrantes e a partir disso foram analisadas as perguntas de referência “Município de Residência em 31 de julho de 1995” e “Município de Residência em 31 de julho de 2005”, para os



Censos 2000 e 2010, respectivamente. Por fim, foram mantidos os declarantes que responderam ser os responsáveis principais do domicílio, para 2000, a pergunta feita foi “Relação com Responsável pelo Domicílio” e, para 2010, “Relação de Parentesco ou de Convivência com a Pessoa Responsável pelo Domicílio”.

Dadas as opções feitas, algumas ressalvas são importantes. Em primeiro lugar, ao definir migrantes como data fixa e não última etapa, é utilizada uma arbitrariedade na definição de migrantes – aqueles que alteraram residência em até cinco anos – além de não ser possível uma distinção clara do que acontece entre esse período, por exemplo, mais de uma migração. Da mesma forma, a análise centrada unicamente no chefe tem suas limitações por não considerar a estrutura em que esse se insere. Deve-se notar ainda que, ao tratar das características dos indivíduos no *destino*, como relacionado ao processo migratório da *origem*, há um problema de precedência temporal na análise, de forma que as interpretações causais devem ser feitas com bastante cuidado.

Uma das principais vantagens deste quesito é que ele permite o cálculo de todas as medidas convencionais de migração: imigrantes, emigrantes e saldo migratório. Além disso, os lugares de origem e destino são conhecidos, o período dentro do qual ocorre a migração é bem determinado e o conceito de migrante é facilmente compreendido (RIGOTTI, 2011). Ainda, o manual sobre migrações da ONU, *United Nations Population Division* (1970), destaca que, caso haja necessidade de escolher apenas um quesito para o estudo de migração, o de data fixa seria o mais adequado.

É importante notar ainda que o Censo 2000 não tem informação de município de última etapa, apenas a unidade da federação, o que, para os objetivos, comprometeria a realização da análise. Além disso, a migração de última etapa torna difícil estabelecer um parâmetro de comparabilidade, justamente pelo fato de o critério “tempo” ser variável.

Outra ressalva a ser feita é que foram excluídos da análise econométrica os migrantes “pendulares”, isto é, aqueles migrantes que moram na mesma cidade de cinco anos atrás, mas que no interim se mudaram para uma segunda cidade. A sua exclusão se dá pela razão de o movimento entre hierarquias já estar determinado para esses migrantes<sup>3</sup>. Como será visto adiante, esses migrantes compõem cerca de 25% dos chefes de família que migraram nos períodos considerados. As tabelas de contingência e de fluxos migratórios incluem essas observações, sendo que esses migrantes foram ignorados apenas para a construção da amostra analítica final.

Em uma segunda etapa, a partir do estudo Região de Influência das Cidades – REGIC –, que utiliza uma abordagem reconhecidamente christalliana e que tem por principal objetivo acompanhar o desenvolvimento da rede urbana nacional a partir de seus papéis hierárquicos-funcionais, os migrantes foram classificados conforme as hierarquias de suas cidades de origens/destinos. Seguindo esse método de hierarquização proposto – e que será aqui utilizado –, as cidades brasileiras foram divididas em cinco grandes grupos:

- Metrôpoles: divididas em três grupos – grande metrópole nacional, metrópoles nacionais e metrópoles – compreendem os 12 principais centros do País;
- Capitais Regionais: divididas em três grupos – níveis A, B e C – têm capacidade de gestão no nível imediatamente inferior ao das metrópoles, com sua área de influência notória no âmbito regional;
- Centros Sub-Regionais: divididos em dois grupos – A e B – têm área de atuação mais reduzida e seus relacionamentos mais intensos com centros externos à sua própria rede ocorrem, em geral, apenas com as Metrôpoles;
- Centros de Zona: divididos em dois grupos – A e B – com área de atuação restrita à sua área imediata, com populações, em geral, não superiores a 40 mil;
- Centros Locais: demais cidades cujas centralidades não extrapolam os limites do seu município. Em geral, com população inferior a 10 mil habitantes.

<sup>3</sup> Nesse sentido, seria muito mais interessante, por exemplo, estudar se algumas características desses migrantes pendulares os distinguem dos outros migrantes.

É importante notar que, para o Brasil, os Censos, que são realizados a cada 10 anos, constituem-se como única fonte extensiva a todo território nacional da qual é possível obter informações ao nível municipal para fluxos migratórios. Por sua vez, o REGIC (2008) é o estudo mais completo para hierarquização da rede de cidades que, como demonstrado no próprio estudo, é bastante estável, de modo que a mudança entre hierarquias, quando essas ocorrem, se deem de forma lenta e gradual. Assim, ainda que temporalmente não sejam exatamente coincidentes, acredita-se que, da combinação dessas duas fontes, é possível construir uma análise robusta com os propósitos feitos nesse estudo.

Feitas essas alocações, foi possível identificar os migrantes a partir da hierarquia de sua cidade de origem e também de sua cidade de destino, tornando possível explorar a existência de possíveis relações entre a escolha migratória e o perfil individual do migrante, naquele que configura o questionamento principal deste trabalho.

### 3.2 Estratégia econométrica

O objetivo principal deste trabalho é entender como se deram os padrões migratórios das últimas décadas em relação às hierarquias urbanas, avaliando a existência de processos de concentração ou desconcentração. Ao mesmo tempo, busca-se uma melhor qualificação desses processos, por meio da identificação de características dos indivíduos que estariam associados a eles. Por exemplo, os indivíduos que migram para cidades de maior hierarquia, fortalecendo o processo de concentração, podem ser mais novos e com maior educação, enquanto os migrantes que buscam cidades de hierarquia mais baixa podem ser homens mais velhos que retornam para sua cidade de origem. Em particular, além de investigar como os padrões de migração inter-hierárquicos diferem para as diferentes regiões, busca-se avaliar se o nível educacional do chefe de família está relacionado com a decisão de migração em direção a cidades de maior ou menor hierarquia.

Com o objetivo acima em mente e assumindo que os indivíduos já tenham tomado a decisão de migrar, busca-se modelar a escolha de uma entre as seguintes três alternativas:

- i) migrar para um município de mesma hierarquia;
- ii) migrar para um município de hierarquia maior;
- iii) migrar para um município de hierarquia menor.

São duas as razões para trabalhar com esse pequeno número de alternativas. Em primeiro lugar, modelar apenas três escolhas, ao invés de todas as possíveis escolhas entre as hierarquias, torna o entendimento dos resultados e sua apresentação muito mais simples. Em segundo lugar, como o interesse está em melhor compreender o processo de concentração e desconcentração da rede de cidades no Brasil, não necessariamente identificar quais tipos de cidades estão atraindo mais ou menos migrantes, optou-se por trabalhar apenas com a direção dos movimentos, ao invés de tentar entender a escolha ou não por cada uma das hierarquias, condicional à hierarquia de origem. Finalmente, como o interesse recai em entender as características dos indivíduos associadas com as decisões de movimento entre as hierarquias, e não nas características de cada alternativa que tornam essa escolha mais ou menos provável<sup>4</sup>, foi tomada a escolha de utilizar o Modelo Logístico Multinomial (MNL), no qual não são incluídas as características das alternativas, apenas dos indivíduos<sup>5</sup>.

Neste modelo, seguindo Greene (2007), a probabilidade de escolha da alternativa  $j$ , condicional às características dos indivíduos apenas, é dada por:

<sup>4</sup> Além disso, não é claro quais as variáveis relativas às alternativas que poderiam ser utilizadas quando se agregam as opções como direções de movimento entre hierarquias. Por exemplo, embora a nível de município variáveis como salários médios e taxa de desemprego sejam escolhas óbvias, não é evidente a forma como tais características poderiam ser calculadas para a variável dependente utilizada neste trabalho.

<sup>5</sup> Note que esse modelo é usado para o caso no qual não há um ordenamento das alternativas. Destacamos que, embora haja uma ordem clara entre as hierarquias urbanas, não há uma ordem para as escolhas de migrar para uma cidade de hierarquia menor, maior ou igual. Assim, o MNL deve ser empregado ao invés do Modelo Logit Ordenado.

$$P_{ij} = P(Y_i = j|x_i) = \frac{e^{x'_i\beta_j}}{1 + \sum_{k=1}^J e^{x'_i\beta_k}}, \text{ para } j = 0, 1, \dots, J \quad (1)$$

em que os coeficientes são específicos para cada uma das alternativas, sendo necessária a restrição  $\beta_1 = 0$  para garantir a identificação do modelo. Nesse sentido, os efeitos das variáveis sobre as alternativas são expressos em relação a uma categoria base, neste caso, mudar para um município de mesma hierarquia ( $j = 1$ ).

Aqui, embora a fórmula para os efeitos marginais seja complexa, optou-se por apresentá-los os invés dos tradicionais *log odds-ratios*,  $\ln\left(\frac{P_{ij}}{P_{ik}}\right) = x'_i(\beta_j - \beta_k) = x'_i\beta_j$  se  $\beta_k = 0$ , ou seja, as chances de escolha de uma alternativa em relação à alternativa base. A vantagem de apresentar os efeitos marginais é que eles podem ser calculados para cada uma das alternativas, inclusive para a alternativa base. No caso do modelo MNL, o efeito de uma mudança de uma unidade de um regressor sobre a probabilidade de escolha da variável  $j$  é dado pela seguinte função:

$$\frac{\partial P_{ij}}{\partial x_i} = P_{ij}(\beta_j - \bar{\beta}) \quad (2)$$

em que  $\bar{\beta} = \sum_k P_{ik}\beta_k$ . No caso da alternativa base, o efeito marginal é simplesmente  $-p_{i0}\bar{\beta}$ .

Esta equação deixa claro o porquê da dificuldade de interpretação dos coeficientes: dado o sinal de  $\beta_j$ , seu efeito marginal poderá ser positivo ou negativo, a depender de  $\bar{\beta}$ . Além disso, os efeitos marginais dependem dos valores assumidos pelas variáveis em  $x'_i$ , uma vez que a probabilidade  $P_{ik}$  depende desses valores. Neste trabalho, são apresentados no corpo do texto os efeitos marginais utilizando o valor médio de cada uma das variáveis. Nos anexos, são incluídos os resultados em *log-odds ratio*.

Finalmente, para o cômputo dos erros padrões, as observações foram agrupadas a nível do município de residência em 2000 ou 2010, permitindo correlações entre os erros dos indivíduos cujo destino da migração foi o mesmo.

## 4. Análise dos resultados

### 4.1 Análise descritivas

Inicialmente, é interessante analisar as características amostrais dos dados. A Tabela 3 apresenta a distribuição dos migrantes de acordo com sua hierarquia de destino nos anos de 2000 e 2010, bem como o percentual da população *total* residente em cada uma dessas hierarquias. A partir desses dados, é possível avaliar o peso dos migrantes nos diferentes níveis da rede urbana. Pela tabela, não é possível notar uma sobre representação de chefes migrantes em nenhuma das hierarquias em comparação com a população total, de forma que, para as migrações de data fixa no período analisado, não existe uma preponderância das metrópoles como era comum em períodos anteriores da economia brasileira. Além disso, as tabelas reforçam a conclusão de que há uma tendência de (pequena) queda na força de atração das grandes metrópoles em favor dos centros médios, que contam com representação maior na população de migrantes que na população total.

Nas Tabelas 4 e 5, são apresentadas análises de contingência para os Censos 2000 e 2010, respectivamente. Nelas, a linha se refere à residência na origem, ou seja, cinco anos antes, enquanto a coluna diz respeito ao destino, a residência no ano corrente de realização do Censo. Dessa forma, uma pessoa da linha  $i$  emigrou de uma cidade de hierarquia definida pela linha para um destino definido pela coluna  $j$ . Assim, cada um desses quadros nos dá um indicativo dos fluxos migratórios em níveis de hierarquia das cidades. Além disso, tem-se uma descrição dos fluxos migratórios para os períodos 1995-2000 e 2005-2010.

**Tabela 3 - Distribuição dos chefes imigrantes por hierarquias urbanas**

	Participação na População (%)		População Residente (%)		Imigrantes Identificados	
	2000	2010	2000	2010	2000	2010
Grande Metrópole Nacional (GMN)	10,28	8,88	10,73	10,53	588.069	492.863
Metrópole Nacional (MN)	7,14	6,94	8,18	8,18	408.650	385.039
Metrópole (M)	14,86	13,96	14,15	14,41	849.855	774.511
Capital Regional A (CRA)	7,98	8,13	6,54	6,93	456.638	451.176
Capital Regional B (CRB)	5,25	5,73	4,48	4,66	300.351	318.082
Capital Regional C (CRC)	8,03	8,17	7,51	7,74	459.493	453.269
Centro Sub-Regional A (CSRA)	5,67	5,90	5,06	5,09	324.210	327.332
Centro Sub-Regional B (CSR B)	3,05	3,32	3,26	3,33	174.653	184.396
Centro de Zona A (CZA)	5,39	5,76	5,57	5,55	308.430	319.704
Centro de Zona B (CZB)	5,15	5,40	5,54	5,40	294.305	299.407
Centros Locais (CL)	27,19	27,81	28,97	28,19	1.555.466	1.542.794
Total	100,00	100,00	100,00	100,00	5.720.120	5.548.573

Fonte: Resultado de pesquisa.

Antes de proceder com análises mais acuradas, duas observações devem ser feitas com relação à diagonal principal – GMN-GMN, MN-MN, M-M, etc. – das Tabelas 4 e 5 a fim de que esses dados sejam analisados com o devido cuidado. Primeiramente, pelas tabelas, nos resultados, ficam mescladas informações de migração pendular com dados que dizem respeito a migrações dentro da mesma hierarquia, pois, como refere-se a uma informação de migração de data fixa, nada impede que dentro do período a pessoa tenha saído e voltado a viver na mesma cidade de antes de sua migração<sup>6</sup>.

Em segundo lugar, parte do peso existente do fluxo na mesma hierarquia verificado até Capitais Regionais A (CRA), carrega um efeito das regiões metropolitanas, pois todas as capitais de estado e a grande maioria das aglomerações estão concentradas até essa classe, de forma que existe dentro dessas diagonais um efeito de migração que, apesar de ser dentro da mesma hierarquia, exibe uma especificidade maior que essa, pois é dentro da própria Área de Concentração de População (ACP).

Para o primeiro período, que compreende 1995-2000, representado na Tabela 4, percebe-se que existe uma polarização dos movimentos de saída nos dois extremos da rede urbana, especialmente se as metrópoles forem consideradas simultaneamente. Além disso, depois dos Centros Locais, as Metrópoles aparecem como principal destino dos migrantes. Uma possível causa desse movimento diz respeito ao fato de os nove centros que compõem esses lugares (todos capitais de estado) estarem bastante dispersos ao longo do território nacional e, dessa forma, fora de sua ACP, estão rodeadas por muitos Centros Locais, de modo que as pessoas no entorno enxergam essas capitais como uma boa fonte de oportunidades.

No período 2005-2010, representado pela Tabela 5, os padrões de migração entre as hierarquias urbanas permaneceram muito similar aos do período anterior, com modificações marginais. As principais mudanças são percebidas na redução da proporção de migrantes que buscaram as metrópoles, assim como uma redução dessa proporção com relação à origem dos migrantes. Também se percebe uma redução de migrantes que saíram de Centros Locais, embora a proporção daqueles que tiveram essa hierarquia como destino permaneceu quase inalterada. Assim, combinados esses efeitos, tem-se como resultado um fortalecimento dos fluxos migratórios entre hierarquias intermediárias.

<sup>6</sup> Como mencionado anteriormente, os migrantes chamados nesse trabalho de pendulares foram mantidos aqui, sendo excluídos apenas da amostra utilizada nas regressões a seguir. Isso acontece porque esse tipo de migração determina completamente o movimento entre hierarquias urbanas, isto é, um migrante pendular necessariamente será da categoria que não alterou hierarquia.

**Tabela 4 - Hierarquia de origens e destinos dos migrantes 1995-2000 (%)**

	GMN	MN	M	CRA	CRB	CRC	CSRA	CSRB	CZA	CZB	CL	Total
GMN	6,35	0,16	0,46	0,57	0,35	1,01	0,44	0,22	0,37	0,34	1,65	11,90
MN	0,15	4,39	0,35	0,27	0,11	0,22	0,24	0,13	0,11	0,17	0,64	6,77
M	0,37	0,36	7,90	0,36	0,38	0,65	0,33	0,22	0,38	0,32	1,63	12,88
CRA	0,26	0,25	0,30	3,52	0,15	0,26	0,22	0,10	0,20	0,17	0,90	6,33
CRB	0,23	0,12	0,43	0,22	1,59	0,17	0,22	0,12	0,19	0,15	0,80	4,25
CRC	0,39	0,22	0,72	0,34	0,22	2,76	0,27	0,17	0,24	0,22	1,37	6,93
CSRA	0,26	0,21	0,53	0,34	0,31	0,27	1,55	0,15	0,24	0,22	1,30	5,38
CSRB	0,15	0,10	0,36	0,17	0,17	0,25	0,14	0,76	0,13	0,13	0,89	3,24
CZA	0,26	0,16	0,65	0,37	0,34	0,37	0,30	0,13	1,51	0,29	1,56	5,92
CZB	0,24	0,21	0,57	0,31	0,29	0,31	0,32	0,16	0,32	1,45	1,69	5,87
CL	1,64	0,97	2,59	1,51	1,36	1,77	1,64	0,90	1,70	1,68	14,78	30,53
Total	10,28	7,14	14,86	7,98	5,25	8,03	5,67	3,05	5,39	5,15	27,19	100

Fonte: Resultado de pesquisa.

**Tabela 5 - Hierarquia de origens e destinos dos migrantes 2005-2010 (%)**

	GMN	MN	MN	CRA	CRB	CRC	CSRA	CSRB	CZA	CZB	CL	Total
GMN	5,46	0,18	0,38	0,48	0,29	0,97	0,35	0,19	0,29	0,24	1,32	10,14
MN	0,19	4,36	0,37	0,29	0,13	0,24	0,29	0,13	0,12	0,19	0,77	7,08
MN	0,38	0,43	7,75	0,48	0,47	0,74	0,39	0,27	0,46	0,40	2,08	13,84
CRA	0,26	0,28	0,32	3,69	0,18	0,29	0,25	0,12	0,25	0,22	1,08	6,94
CRB	0,22	0,13	0,43	0,27	1,75	0,21	0,26	0,16	0,24	0,18	0,96	4,79
CRC	0,39	0,21	0,68	0,35	0,24	2,81	0,29	0,17	0,29	0,24	1,46	7,13
CSRA	0,25	0,19	0,48	0,37	0,39	0,31	1,67	0,19	0,30	0,27	1,44	5,87
CSRB	0,14	0,11	0,37	0,19	0,22	0,26	0,19	0,86	0,15	0,14	0,94	3,57
CZA	0,23	0,17	0,60	0,37	0,41	0,36	0,32	0,15	1,62	0,32	1,62	6,17
CZB	0,24	0,20	0,47	0,33	0,34	0,33	0,31	0,17	0,35	1,58	1,68	6,00
CL	1,13	0,68	2,11	1,31	1,32	1,66	1,58	0,93	1,68	1,61	14,46	28,46
Total	8,88	6,94	13,96	8,13	5,73	8,17	5,90	3,32	5,76	5,40	27,81	100

Fonte: Resultado de pesquisa.

Na Tabela 6, são analisados os deslocamentos intra e interclasses, em que os chefes de família que migraram para uma cidade de mesma hierarquia não alteraram sua posição, enquanto aqueles que foram para uma hierarquia superior subiram de classe e os que foram para uma menor desceram. Adicionalmente a isso, são mostrados os saldos migratórios – dados pela diferença entre imigrantes e emigrantes – em cada um dos períodos para cada uma das hierarquias consideradas.

Além disso, os saldos migratórios mostram que houve uma redução da magnitude dos fluxos nas hierarquias. Isso, combinado com o fato de que houve pouca mudança na direção das migrações, embora ainda se percebam fluxos negativos nos extremos da hierarquia e positivos nas intermediárias, apontando para um fortalecimento das cidades de centralidade média. Cabe agora uma exploração questionando se existe uma diferenciação do perfil migratório daqueles cujas cidades de destino representaram uma mudança de padrão hierárquico em termos da rede de cidades.

**Tabela 6 - Fluxos migratórios entre hierarquias nos intervalos de cinco anos**

Hierarquia de destino	Não se alterou (%)		Desceu (%)		Subiu (%)		Saldo	
	2000	2010	2000	2010	2000	2010	2000	2010
Grande Metrópole Nacional (GMN)	62	61	-	-	38	39	-92.678	-69.742
Metrópole Nacional (MN)	61	63	2	3	36	35	21.116	-7.894
Metrópole (M)	53	56	5	5	41	39	113.115	6.609
Capital Regional A (CRA)	44	45	15	15	41	39	94.705	66.219
Capital Regional B (CRB)	30	30	20	19	51	51	57.508	52.150
Capital Regional C (CRC)	34	34	29	30	37	36	63.344	57.428
Centro Sub-Regional A (CSRA)	27	28	31	31	42	41	16.568	1.728
Centro Sub-Regional B (CSR B)	25	26	39	37	39	38	-10.944	-13.874
Centro de Zona A (CZA)	28	28	38	37	37	35	-30.382	-22.628
Centro de Zona B (CZB)	28	29	42	41	33	30	-41.476	-33.429
Centros Locais (CL)	54	52	46	48	-	-	-190.876	-36.567
Total	47	46	26	26	29	28	-	-

Fonte: Resultado de pesquisa.

#### 4.2 Análise econométrica

Como já destacado, além da análise de movimentos gerais entre classes de cidades, neste trabalho, procura-se avaliar se existe uma diferenciação de perfis associados à migração para diferentes classes hierárquicas da rede, em especial no que diz respeito ao nível de escolaridade dessas pessoas e diferenças regionais nos padrões migratórios. Assim, aqui são apresentados os resultados do modelo Logit Multinomial estimado para os anos de 2000 e 2010.

Nesta análise econométrica, as variáveis de interesse são as seguintes: i) região em que vivia na data de referência do Censo; ii) região em que vivia cinco anos antes do Censo; iii) variável categórica de nível de instrução formal (1 – Sem instrução e fundamental incompleto, 2 – Fundamental completo e médio incompleto, 3 – Médio completo e superior incompleto e 4 – Superior completo); iv) variável categórica se não estuda (1), se é estudante não universitário (2) ou se é universitário (3).

Além disso, foram adicionados controles relativos às condições de localidade, demografia, família e trabalho, as seguintes variáveis: i) variável binária zona urbana (0 se não vive, 1 caso contrário); ii) variável binária indicando se retornou para a cidade de nascimento; iii) variável binária indicando se migrou dentro da mesma UF; iv) variável binária de sexo (0 se for mulher, 1 se é homem); v) idade; vi) variável categórica cor da pele (1 – Branca, 2 - Preta, 3 Parda); vii) se vive com cônjuge (1), se não vive mais (2) e se nunca viveu (3); viii) número de filhos (enteados inclusive); ix) logaritmo do seu rendimento mensal total; x) variáveis *dummies* indicando se trabalha até 30 horas ou mais, sendo zero em ambas caso não trabalhe.

A Tabela 7 traz uma rápida análise descritiva das variáveis independentes utilizadas na análise econométrica, levando em conta apenas os chefes de família migrantes que não fizeram migração pendular, ou seja, considera-se apenas a amostra analítica final. De modo geral, a média das variáveis utilizadas pouco se alterou ao longo do período. Em especial, houve uma redução na proporção dos migrantes do sexo masculino, assim como um aumento na proporção de migrantes com mais estudos e daqueles que ainda estudam. Ainda, aumentou a proporção daqueles migrantes que não vivem com o cônjuge e redução no número médio de filhos.

**Tabela 7 - Descrição da amostra analítica – 2000 e 2010**

	2000		2010			2000		2010	
	Média	Desvio Padrão	Média	Desvio Padrão		Média	Desvio Padrão	Média	Desvio Padrão
Norte (destino)	0,0809	0,2727	0,0835	0,2767	Fund. Incomp.	0,592	0,4914	0,3949	0,4888
Nordeste (destino)	0,2107	0,4078	0,2108	0,4078	Médio Incomp.	0,142	0,3498	0,1642	0,3705
Sudeste (destino)	0,4189	0,4934	0,4007	0,4900	Superior incomp.	0,185	0,3888	0,2876	0,4526
Sul (destino)	0,1788	0,3832	0,1894	0,3918	Superior completo	0,079	0,2699	0,1533	0,3602
C-O (destino)	0,1107	0,3138	0,1156	0,3198	Não estuda	0,938	0,2405	0,8954	0,3060
Norte (origem)	0,0759	0,2649	0,0793	0,2702	Est. não universit.	0,038	0,1918	0,0396	0,1950
Nordeste (origem)	0,2434	0,4291	0,2539	0,4353	Universit.	0,023	0,1512	0,0650	0,2465
Sudeste (origem)	0,4061	0,4911	0,3869	0,4870	Vive com cônjuge	0,726	0,4460	0,6590	0,4740
Sul (origem)	0,1792	0,3835	0,1838	0,3873	Já viveu c/ cônjuge	0,186	0,3895	0,2095	0,4070
C-O (origem)	0,0953	0,2937	0,0961	0,2947	Nunca c/ cônjuge	0,087	0,2825	0,1315	0,3379
Urbana	0,8430	0,3638	0,8742	0,3317	Renda mensal	851	2.339	1.738	5.843
Homem	0,7988	0,4009	0,6536	0,4758	Trabalha até 30 hrs	0,068	0,2531	0,1111	0,3142
Idade	38,919	13,636	38,433	14,305	Trabalha mais 30 hrs	0,697	0,4595	0,6280	0,4833
Branca	0,5639	0,4959	0,5118	0,4999	Número de filhos	0,944	1,2875	0,5939	0,9727
Preta	0,0635	0,2439	0,0834	0,2764	Retornante	0,088	0,2841	0,1119	0,3153
Parda	0,3726	0,4835	0,4048	0,4909	Mesma UF	0,675	0,4682	0,6559	0,4751

Fonte: Resultado da pesquisa.

Na Tabela 8 são apresentados resultados do modelo Logit Multinomial estimado para os anos de 2000 e 2010 utilizando a amostra e as variáveis conforme descritos anteriormente. Aqui, é dado destaque para a análise dos resultados a partir dos efeitos marginais calculados utilizando os valores médios das covariadas, conforme a Equação (2). Vale observar que os efeitos marginais reportam a mudança na probabilidade de ocorrência de cada uma das escolhas, dada uma variação na variável dependente. Assim, entre as escolhas, os efeitos marginais devem somar zero. As razões de risco relativo estimadas para este e outros dois modelos são apresentadas no apêndice deste texto, nas Tabelas A1 e A2, juntamente com o resultado de testes de ajuste desses modelos referentes às estatísticas Pseudo-R<sup>2</sup>, AIC e BIC, as quais indicam que o modelo apresentado no texto, que inclui além das variáveis de interesse outros controles relativos à demografia, características da família e do trabalho, possui o melhor ajuste aos dados. O Pseudo-R<sup>2</sup> deste modelo é o maior entre as três alternativas, enquanto os critérios AIC e BIC são os menores.

**Tabela 8 - Efeitos marginais na probabilidade de migração entre hierarquias: chefes de família migrantes (2000 e 2010)**

	2000			2010		
	Mesma	Ascendente	Descendente	Mesma	Ascendente	Descendente
Norte	-0,030 (0,0394)	-0,160*** (0,0419)	0,1893*** (0,0597)	-0,0795** (0,0397)	-0,0955** (0,0409)	0,1750*** (0,0598)
Nordeste	0,005 (0,0327)	-0,2304*** (0,0357)	0,2251*** (0,0521)	-0,0481* (0,0286)	-0,1823*** (0,0357)	0,2305*** (0,0495)
Sul	-0,047 (0,0301)	-0,05455 (0,0348)	0,1012* (0,0521)	-0,0698** (0,0339)	-0,0213 (0,0332)	0,0911 (0,0561)
Centro-Oeste	-0,008 (0,0296)	0,007901 (0,0504)	-0,0002 (0,0564)	-0,0340 (0,0356)	0,0424 (0,0474)	-0,0083 (0,0655)
Origem Norte	-0,091** (0,0384)	0,2725*** (0,0291)	-0,1818*** (0,0241)	-0,0534 (0,0381)	0,2230*** (0,0266)	-0,1695*** (0,0268)
Origem Nordeste	-0,106*** (0,0258)	0,3685*** (0,0237)	-0,2620*** (0,0237)	-0,0602*** (0,0230)	0,2996*** (0,0250)	-0,2393*** (0,0230)
Origem Sul	-0,045 (0,0324)	0,1881*** (0,0152)	-0,1433*** (0,0292)	-0,0344 (0,0341)	0,1263*** (0,0160)	-0,0918*** (0,0343)
Origem Centro-Oeste	-0,026 (0,0352)	0,0819*** (0,0286)	-0,05548 (0,0388)	-0,0097 (0,0342)	0,0256 (0,0263)	-0,0159 (0,0449)
Zona Urbana	-0,086*** (0,0107)	0,2269*** (0,0154)	-0,1413*** (0,0112)	-0,0762*** (0,0106)	0,2171*** (0,0157)	-0,1408*** (0,0117)
Fundamental	0,005 (0,0059)	-0,0460*** (0,0043)	0,0406*** (0,0053)	-0,0101** (0,0044)	-0,0217*** (0,0044)	0,0319*** (0,0047)
Médio	-0,003 (0,0084)	-0,0517*** (0,0067)	0,0542*** (0,0094)	-0,0059 (0,0075)	-0,0255*** (0,0063)	0,0315*** (0,0077)
Superior	-0,014 (0,0190)	-0,0640*** (0,0106)	0,0781*** (0,0202)	0,01027 (0,0174)	-0,0571*** (0,0092)	0,0468** (0,0193)
Estudante (até o médio)	-0,016*** (0,0048)	0,0351*** (0,0057)	-0,0194*** (0,0054)	0,0115** (0,0057)	0,0202*** (0,0063)	-0,0318*** (0,0068)
Estudante (universitário)	-0,040*** (0,0103)	0,0908*** (0,0089)	-0,0507*** (0,0116)	-0,0362*** (0,0063)	0,0927*** (0,0060)	-0,0564*** (0,0070)
Retornante	-0,081*** (0,0109)	-0,0841*** (0,0081)	0,1648*** (0,0113)	-0,0566*** (0,0094)	-0,0298*** (0,0088)	0,0864*** (0,0135)
Mesma UF	0,145*** (0,0208)	-0,0547*** (0,0150)	-0,0897*** (0,0190)	0,1321*** (0,0191)	-0,0525*** (0,0123)	-0,0796*** (0,0195)
Observações	511.978	511.978	511.978	475.635	475.635	475.635

Nota: As estimativas dos efeitos marginais foram obtidas a partir do modelo logístico multinomial, com valores calculados para a média das observações. Estão incluídas como controles as seguintes variáveis: *dummy* indicando se vive em zona urbana ou rural, se a migração foi para cidade de nascimento, se a migração foi dentro da mesma UF, sexo, se trabalha mais de 30 horas semanais, variável categórica para cor da pele (branca, preta e parda), variável categórica para situação familiar (se vive com cônjuge, se já viveu, ou se nunca viveu), número de filhos (inclusive enteados) e logaritmo do rendimento mensal total. Erros padrões robustos entre parênteses (\*\*\*p<0.01; \*\*p<0.05; \*p<0.10).

Fonte: Resultado da pesquisa.

Dos resultados estimados, a partir dos efeitos marginais, percebe-se que não há uma associação entre a região de destino do migrante e a escolha de migração para cidades de igual hierarquia no ano 2000, uma vez que a probabilidade desse tipo de migração não se altera entre nenhuma região, com todos os efeitos marginais estimados não sendo estatisticamente diferente de zero. Já para 2010, os migrantes cujo destino foi a região Norte, Nordeste e Sul apresentaram menor tendência a permanecer na mesma hierarquia comparativamente aos que escolheram a região Sudeste.



Considerando a relação entre mudanças de hierarquia – tanto ascendente como descendente – e a região de destino, percebe-se que os efeitos marginais são em sua maioria significativos, com exceção da região Centro-Oeste. Assim, quando comparados com os migrantes cujo destino foi a região Sudeste, aqueles que escolheram Norte e Nordeste tiveram uma probabilidade maior de descer de hierarquia, tanto em 2000 como 2010. O efeito marginal para 2010 sobre a probabilidade de movimento descendente é muito similar ao de 2000.

Com respeito à região de origem, observa-se justamente o efeito contrário do para região de destino. Quando comparados a migrantes cuja origem é a região Sudeste, migrantes oriundos do Norte, Nordeste e Sul apresentam maiores chances de buscar cidades de maior hierarquia, sendo esse efeito arrefecido entre 2000 e 2010.

A partir destes resultados, é possível chegar a algumas conclusões a respeito das mudanças no padrão regional de migração no período. Em primeiro lugar, as probabilidades de permanecer na mesma hierarquia, descer de hierarquia ou subir de hierarquia para migrantes cuja origem e destino são cidades da região Sudeste é de respectivamente 56%, 20% e 24%, um pouco diferentes das calculadas para toda a amostra analítica, de 46%, 26% e 28% para o ano de 2010<sup>7</sup>. Esse dado é relevante, pois a interpretação dos movimentos migratórios para as regiões tem como base as migrações que ocorrem dentro do Sudeste, categoria base para a região de origem e destino.

Assim, percebe-se que migrações cujo destino são as regiões Norte, Nordeste e Sul tendem a ser para cidades de menor hierarquia, ou seja, têm em sua maioria um caráter desconcentrador quando comparadas com migrações para cidades do Sudeste – e, também, Centro-Oeste, dado que os efeitos marginais para essa região não são significativos. Porém, quando a migração tem origem nessas três regiões, ela tende a ser concentradora, uma vez que migrantes oriundos dessa região têm maior probabilidade de ascender de hierarquia.

Para facilitar a interpretação, segue o exemplo. Um migrante cuja origem e o destino são cidades no Nordeste tem uma chance de ascender de região 13p.p. maior do que de uma pessoa que faz uma migração dentro do Sudeste (-23p.p. por ter seu destino no Nordeste, e +36p.p. por sua origem ser o Nordeste); ao mesmo tempo, ele tem uma chance 4% menor de descer de hierarquia (22p.p. - 26p.p.). Já um imigrante que sai do Nordeste com destino à região Sudeste tem chance 36p.p. maior de ir para uma cidade de maior hierarquia do que um migrante dentro do Sudeste e chance 26p.p. menor de descer de hierarquia. Finalmente, um migrante que sai do Sudeste com destino ao Nordeste tem chance 23p.p. menor de migrar para cidade de menor hierarquia do que migrantes com destino ao Sudeste e uma chance 22p.p. maior de descer de hierarquia.

A partir desses resultados, conclui-se que: i) o Sudeste continua atuando como polo concentrador da hierarquia urbana no Brasil, pois migrações cujo destino são essa região têm maior probabilidade de representar um movimento ascendente; ii) movimentos de saída do Sudeste tendem a ser desconcentradores, na medida em que há uma maior probabilidade de o destino ser uma cidade de menor hierarquia; iii) movimentos intrarregião (eg. Nordeste-Nordeste) tendem a ser concentradores, uma vez que a soma dos efeitos marginais de origem e destino apontam para uma maior chance de aumento na hierarquia; iv) há uma pequena mudança no período de 2010, com uma leve queda na magnitude dos efeitos, havendo uma convergência dessas regiões para o comportamento observado no Sudeste, onde há uma forte tendência de manutenção da hierarquia, seguida de uma fraca força de desconcentração.

Finalmente, migrantes dentro de uma mesma UF tendem a fazê-lo para cidades da mesma hierarquia. Embora não se possa concluir diretamente, esse resultado pode ser interpretado como evidência indireta da migração intrarregião metropolitana, uma vez que todas as cidades da região metropolitana recebem o mesmo nível hierárquico – pois estão na mesma ACP. Porém, não são observadas mudanças significativas nos efeitos marginais dessa variável entre os anos 2000 e 2010.

A segunda parte exploratória diz respeito à vinculação da escolha de migração entre hierarquias e o nível de instrução dos indivíduos. Para isso, considera-se a hipótese de que cidades com diferentes

<sup>7</sup> Surpreendentemente, esses números permanecem muito similares em 2000, em 57%, 20% e 23%, enquanto para todo o Brasil eles são de 27%, 26% e 29%.

hierarquias exercem atração diferenciada sobre pessoas com graus de qualificação (em termos de instrução formal) distintos. Contrariamente aos estudos que mostram uma relação de reforço entre migração de pessoas mais qualificadas e cidades/regiões onde já existe forte presença dessas pessoas (TAVEIRA; ALMEIDA, 2014; FREGUGLIA; GONÇALVES; SILVA, 2014; NASCIMENTO; SILVA; LUCAS, 2017; BARUFI, 2018), os resultados encontrados para hierarquias não revelam essa tendência, pelo contrário. Pessoas com ensino médio completo e superior completo têm chances maiores de descer de hierarquia, sendo essas chances quase 8p.p. a mais para aqueles com diploma de ensino superior.

Para entender esse resultado, é importante compreender a diferença que existe entre a abordagem aqui proposta e a dos demais trabalhos citados. No geral, os trabalhos anteriores consideram características como renda *per capita* para caracterizar cidades mais ou menos desenvolvidas, assim encontrando que as cidades mais desenvolvidas tendem a atrair migrantes mais qualificados. Neste trabalho, porém, utiliza-se o REGIC (2008), construído com base na centralidade dos municípios. Assim, um movimento descendente de hierarquia, não necessariamente é um movimento de uma cidade com renda *per capita* maior para uma de renda menor. Pode ser, por exemplo, um movimento de saída de uma capital regional de volta para uma cidade polarizada por esse local. Ou então de uma capital de estado para alguma dessas capitais regionais. Assim, não necessariamente há uma contradição entre a correlação encontrada na literatura entre migração de pessoas com maior qualificação e cidades com maior qualidade de vida, ou salários médios maiores, pois as abordagens não iguais.

Entretanto, o que pode estar por trás dessa preponderância de movimentos descendentes para migrantes com maior qualificação? Uma possível explicação para esse resultado é que, dada a concentração da oferta de ensino médio e, principalmente, superior em centros urbanos com maior hierarquia, jovens buscam esses centros para sua formação, para em um segundo momento retornar para cidades de menor nível hierárquico onde seu trabalho qualificado é também demandado. Essa hipótese vê respaldo nos resultados acima, uma vez que chefes de família que ainda são estudantes, principalmente os matriculados no ensino superior, apresentam maior probabilidade de fazer migrações para municípios de maior hierarquia.

Por fim, como a migração é um dos fatores relacionados à reorientação regional de um lugar, os resultados para migração aqui encontrados dão destaque à reorientação da rede de cidades, com ressignificação do papel das cidades de como um todo, em especial das que não estão na ponta superior da hierarquia dentro da rede (IPEA/IGBE/NESUR, 2001; SIMÕES; AMARAL, 2011). Os resultados ligados às probabilidades de migração dos mais qualificados, com uma tendência à mudança em direção a hierarquias menores, reforçam isso.

## 5. Considerações finais

Neste trabalho, procurou-se identificar características pessoais e regionais sobre as escolhas migratórias feitas por chefes de família entre as diferentes classes de cidades. Este exercício é importante por auxiliar na compreensão sobre eventuais preponderâncias hierárquicas exercidas pelas cidades na atratividade de imigrantes e, ao mesmo tempo, se esses imigrantes, em função de suas características pessoais, estão optando por lugares hierarquicamente distintos dentro da rede de cidades para migrar. Caso exista uma heterogeneidade do tipo, há implicações de médio e longo prazo com repercussão importante sobre a dinâmica regional entre os diferentes lugares. Até onde pode ser verificado, a exploração aqui trazida, combinando hierarquias urbanas e perfis dos migrantes, é pioneira na literatura nacional.

Para isso, foram combinadas informações censitárias para migrações de data fixa nos anos 1995-2000 e 2005-2010, com a classificação das cidades brasileiras em grupos hierárquico/funcionais proposta pelo REGIC (2008) e, disso, buscou-se caracterizar os fluxos migratórios baseados em uma orientação de hierarquias das cidades. Isso é feito com o objetivo de explorar as características dos migrantes por grupos hierárquicos da rede urbana, bem como de avaliar em que medida é possível identificar características pessoais e regionais que estão correlacionadas com as escolhas migratórias

entre as classes de cidades. Para atender a esse objetivo, foram combinadas análises exploratórias dos dados com um modelo logit multinomial para chefes de famílias.

A partir dos resultados descritivos, foi possível verificar uma estabilização dos fluxos migratórios entre as hierarquias, reduzindo-se assim o peso histórico exercido especialmente por São Paulo e Rio de Janeiro como destino de migrantes provenientes de outras partes, fato esse já observado por Lima (2013).

Além disso, parece haver uma intensificação de fluxos migratórios dentro da mesma hierarquia, reforçando assim a ideia de crescimento das migrações de curta distância. Isso porque, nos extremos das hierarquias, e em especial nas quatro classes superiores, migrações intraclasse geralmente estão associadas a movimentos entre cidades vizinhas que, no caso metropolitano, estão dentro da mesma ACP. Nota-se ainda o forte aumento de migrações de retorno para cidade natal e para a cidade da última etapa. Como uma das possíveis causas disso está o desalento na nova cidade de moradia ou então o aumento da atratividade de sua cidade anterior. Como nesse trabalho são avaliados agregados, não é possível inferir qual deles é preponderante.

Os resultados da análise econométrica buscaram avaliar como algumas características pessoais e regionais influem na decisão de escolha entre migrar para municípios de hierarquia igual, superior ou inferior à do município de origem. Nesse sentido, foi possível notar que um processo de concentração ainda se mantém na região Sudeste, uma vez que migrantes com destino a essa região tendem a buscar cidades de maior hierarquia, enquanto aqueles que migram para outras regiões tendem a buscar cidades de menor hierarquia. Em especial, esse processo é mais destacado para o Nordeste e, em menor magnitude, para o Norte e Sul. A exceção é a região Centro-Oeste, que apresenta padrões similares ao do Sudeste. Nota-se, porém, que essas diferenças entre as regiões se reduzem no período analisado.

Além disso, um maior nível de escolaridade está associado a uma maior tendência de migração para cidade de menor hierarquia. Esse resultado é complementado com o fato de que estudantes e universitários tendem a migrar para cidades de maior hierarquia, corroborando a hipótese de que jovens buscam cidades maiores para sua formação, retornando para cidades menores para exercer sua profissão. Esse processo, somado à rápida expansão do ensino superior ocorrida no período, aponta que, concomitante a uma maior concorrência nas cidades maiores, as cidades menores podem estar passando por dois processos: melhor oferta de serviços educacionais e/ou com demanda por profissionais mais qualificados sendo mais bem satisfeita.

Como desdobramentos futuros, seria interessante investigar o que tem ocorrido com as migrações intraclasses, se os resultados representam mudanças para cidades similares ou se, na verdade, esses movimentos estão representando migrações de retorno para suas cidades da última etapa. Da mesma forma, investigar o que tem ocorrido com os migrantes em função de sua escolaridade coloca-se como pertinente, pois, como ficou evidenciado pelos resultados econométricos, nos períodos analisados, essas pessoas apresentaram maiores probabilidades de migrar para hierarquias inferiores, deixando assim em aberto a questão de se elas estão se dirigindo para esses locais em função de não encontrarem ofertas condizentes onde estavam ou, ao contrário, os locais hierarquicamente inferiores estão se tornando mais atrativos para elas.

## **Referências**

- AMARAL, E. F. L. Brazil: internal migration. In: NESS, I (org.). *The Encyclopedia of Global Human Migration*, Wiley-Blackwell, 2013.
- ANDRADE, T. A.; SANTOS, A. M. S. P.; SERRA, R. V. *Fluxos migratórios nas cidades médias e regiões metropolitanas brasileiras: a experiência do período 1980/96*. Instituto de Pesquisa Econômica Aplicada, 2000. (Textos para Discussão, n. 747).

- BAENINGER, R. Migrações internas no Brasil século 21: evidências empíricas e desafios conceituais. In: CUNHA, M. P. J. (org.). *Mobilidade espacial da população: desafios teóricos e metodológicos para o seu estudo*, 2011. p. 108-121.
- BAENINGER, R. Rotatividade migratória: um novo olhar para as migrações internas no Brasil. *Revista Interdisciplinar da Mobilidade Humana*, v. 20, n. 39, p. 77-100, 2012.
- BAPTISTA, E. A. Detecção de clusters migratórios no Brasil. *Revista da ANPEGE*, v. 13, n. 22, p. 40-56, 2017.
- BARUFI, A. M. B. Labour Market Dynamics in Brazil. In: STOUGH, R. R.; KOURTIT, K.; NIJKAMP, P.; BLIEN, U. (org.). *Modelling Aging and Migration Effects on Spatial Labor Markets*. Suíça: Springer, 2018. p. 89-106.
- BELL, M.; EDWARDS, E.; UEFFING, P.; STILLWELL, J.; KUPISZEWSKI, M.; KUPISZEWSKA, D. Internal migration and development: comparing migration intensities around the world. *Population and Development Review*, v. 41, n. 1, p. 33-58, 2015.
- BERNARD, A.; ROWE, F.; BELL, M.; UEFFING, P.; EDWARDS, E. Comparing internal migration across the countries of Latin America: a multidimensional approach. *PLoS One*, v. 12, n. 3, p. 1-24, 2017.
- BORJAS, G. J. *Friends or strangers: the impact of immigrants on the US economy*. New York: Basic Books, 1990.
- BRITO, F. R. A.; CARVALHO, J. A. M. As migrações internas no Brasil: as novidades sugeridas pelos Censos demográficos de 1991 e 2000 e pelas PNADs recentes. *Parcerias Estratégicas*, v. 11, n. 22, p. 441-455, 2010.
- BUCH, T.; HAMANN, S.; NIEBUHR, A.; ROSSEN, A. What makes cities attractive? The determinants of urban labour migration in Germany. *Urban Studies*, v. 51, n. 9, p. 1960-1978, 2014.
- CHRISTALLER, W. *Central places in southern Germany*, Prentice-Hall, 1966.
- CUSHING, B.; POOT, J. Crossing boundaries and borders: regional science advances in migration modelling. *Papers in Regional Science*, v. 83, n. 1, p. 317-338, 2003.
- FRANKLIN R. *Census 2000 special reports – migration of the young, single, and college educated: 1995 to 2000*. US Government Printing Office, 2003.
- FREGUGLIA, R. S.; GONÇALVES, E.; SILVA, E. R. Composition and determinants of the skilled out-migration in the Brazilian formal labor market: a panel data analysis from 1995 to 2006. *Economia*, v. 15, n. 1, p. 100-117, 2014.
- GOLGHER, A. *Fundamentos da migração*. Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, 2004. (Textos para Discussão, n. 231).
- GREEN, A. Understanding the drivers of internal migration. In: CHAMPIONS, T.; COOKE, T.; SHUTTLEWORTH, I. (eds.). *Internal Migration in the Developed World*. Routledge, 2017.
- GREENE, W. H. *Econometric analysis*. Pearson Education India, 2007.
- GREENWOOD, M. J. Internal migration in developed countries. In: ROSEMBERG, M. R.; STARK, O. (org.). *Handbook of Population and Family Economics*. Holanda: North Holland, 1997. p. 647-720.
- HAAPANEN, M.; TERVO, H. Migration of the highly educated: evidence from residence spells of university graduates. *Journal of Regional Science*, v.52, n.4, p. 587-605, 2012.


- HARRIS, J. R.; TODARO, M. P. Migration, unemployment and development: a two sectors analysis. *American Economic Review*, v. 60, n. 1, p. 126-142, 1970.
- IPEA, IBGE e UNICAMP. *Configuração atual e tendências da rede urbana*. Série caracterização e tendências da rede urbana do Brasil, IPEA, IBGE, NESUR. Brasília, 2001.
- LIMA, A. C. C. Desenvolvimento regional e fluxos migratórios no Brasil: uma análise para o período 1980-2010. Tese (Doutorado em Economia). Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR). Universidade Federal de Minas Gerais (UFMG), 2013.
- LIMA, A. C. C.; SIMÕES, R.; HERMETO, A. M. Desenvolvimento regional, hierarquia urbana e condição de migração individual no Brasil entre 1980 e 2010. *EURE (Santiago)*, v. 42, n. 127, p. 29-54, 2016.
- LÖSCH, A. *The economics of location*, Yale U. P., 1954.
- LUCAS, R. *Internal migration in developing economies: an overview*. Global Knowledge Partnership on Migration and Development (KNOMAD), 2015. (Working Paper, n. 6).
- MASSEY, D. S.; ARANGO, J.; HUGO, G.; KOUAOUCCI, A.; PELLEGRINO, A.; TAYLOR, J. E. Theories of international migration: a review. *Population and Development Review*, v. 19, n. 3, p. 431-466, 1993.
- MASSEY, D.; ARANGO, J.; HUGO, G.; KOUAOUCCI, A.; PELLEGRINO, A. TAYLOR, J. *Worlds in Motion*. Understanding International Migration at the end of the Millenium. Oxford University Press, 1998.
- MATA, D. D., OLIVEIRA, C. W. D. A.; PIN, C.; RESENDE, G. *Quais características das cidades determinam a atração de migrantes qualificados?* Instituto de Pesquisa Econômica Aplicada, 2007. (Texto para Discussão, n. 1305).
- MATOS, R. A contribuição dos imigrantes em áreas de desconcentração demográfica do Brasil contemporâneo. *Revista Brasileira de Estudos de População*, v. 19, n. 1, p. 49-74, 2013.
- MCCANN, P. *Modern urban and regional economics*. Oxford University Press, 2013.
- MORETTI, E. *The new geography of jobs*. Houghton Mifflin Harcourt, 2012.
- NASCIMENTO, T. C. L.; SILVA, R. P.; LUCAS, L. A. P. Tendência das migrações brasileiras: diferenciais de sexo, idade, distância e volume dos migrantes intermunicipais para 1995-2000 e 2005-2010. In: Encontro Nacional de Estudos Populacionais, 20, 2017. *Anais...* Brasília: ABEP, 2017.
- PARIKH, A.; VAN LEUVENSTEIJN, M. Internal migration in regions of Germany: a panel data analysis. *Applied Economics Quarterly*, v. 49, n. 2, p. 173-192, 2003.
- PARR, J. B. The location of economic activity: central place theory and the wider urban system. In: McCANN, P. (org.). *Industrial Location Economics*, 2002, p. 32-82.
- REGIC. *Região de influência das cidades - 2007*. Rio de Janeiro: IBGE, 2008.
- RIGOTTI, J. R. Dados censitários e técnicas de análise das migrações no Brasil: avanços e lacunas. In: CUNHA, M. P. J. (org.). *Mobilidade espacial da população: desafios teóricos e metodológicos para o seu estudo*, 2011.
- RODRÍGUEZ-POSE, A.; KETTERER, T.; CASTELLS-QUINTANA, D. Do we follow the money? The drivers of migration across regions in the EU. *REGION*, v. 2, n. 2, p. 27-45, 2015.

- SILVEIRA NETO, R. M. Do Public income transfer to the poorest affect internal inter-regional migration? evidence for the case of Brazilian Bolsa Família Program. In: Encontro Nacional de Economia, 36, 2008. *Anais...* Salvador: ANPEC, 2008.
- SIMÕES, R.; AMARAL, P. Interiorização e novas centralidades urbanas: uma visão prospectiva para o Brasil. *Economia*, v.12, n.3, p.553-579, 2011.
- TAVEIRA, J. G.; ALMEIDA, E. S. Os determinantes regionais da atração do migrante qualificado. *Análise Econômica*, v. 32, n. 62, p. 199-224, 2014.
- TODARO, M. P. A model of labor migration and urban unemployment in less developed countries. *The American Economic Review*, v. 59, n. 1, p. 138-148, 1969.
- United Nations Population Division. *Methods of measuring internal migration: manuals on methods of estimating population*. Manual VI. New York, 1970.
- WINTERS, J. V. Why are smart cities growing? Who moves and who stays. *Journal of Regional Science*, v. 51, n. 2, p. 253-270, 2011.

## ORCID

Clauber Eduardo Marchezan Scherer  <https://orcid.org/0000-0003-0596-7053>

Alysson Lorenzon Portella  <https://orcid.org/0000-0001-7008-6784>

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.

## Apêndice

Tabela A1 - Modelo multinomial - 2000: razões de risco relativo e testes de ajuste

(continua)

	Modelo 1		Modelo 2		Modelo 3	
	Ascendente	Descendente	Ascendente	Descendente	Ascendente	Descendente
Norte	0.619*	2.459***	0.666*	1.925*	0.694*	2.003*
	(0.132)	(0.564)	(0.122)	(0.544)	(0.126)	(0.564)
Nordeste	0.352***	2.590***	0.436***	1.952**	0.462***	1.951**
	(0.0652)	(0.493)	(0.0618)	(0.455)	(0.0638)	(0.455)
Sul	0.962	1.902**	1.019	1.728*	1.015	1.674
	(0.139)	(0.418)	(0.113)	(0.460)	(0.112)	(0.444)
Centro-Oeste	1.093	1.183	1.025	0.996	1.042	1.023
	(0.217)	(0.299)	(0.157)	(0.299)	(0.157)	(0.305)
Origem	3.119***	0.627**	3.009***	0.728	3.001***	0.754
Norte	(0.500)	(0.100)	(0.572)	(0.144)	(0.567)	(0.149)
Origem	4.578***	0.428***	3.886***	0.509***	3.809***	0.530***
Nordeste	(0.626)	(0.0533)	(0.458)	(0.0793)	(0.444)	(0.0826)
Origem	2.169***	0.681**	2.188***	0.751	2.127***	0.750
Sul	(0.222)	(0.0997)	(0.268)	(0.137)	(0.261)	(0.137)
Origem	1.454*	0.889	1.490*	0.923	1.485*	0.936
Centro-Oeste	(0.244)	(0.142)	(0.242)	(0.176)	(0.241)	(0.178)
Fundamental	1.006	1.067	0.901***	1.086**	0.865***	1.125***
	(0.0315)	(0.0379)	(0.0235)	(0.0331)	(0.0220)	(0.0321)
Médio	1.082	1.135*	0.934	1.160**	0.871***	1.200***
	(0.0464)	(0.0649)	(0.0362)	(0.0630)	(0.0306)	(0.0551)
Superior	1.143	1.332*	0.942	1.304*	0.868*	1.330**
	(0.0819)	(0.166)	(0.0722)	(0.166)	(0.0605)	(0.141)
Estudante	1.360***	0.940*	1.290***	0.930**	1.153***	0.983
(até o médio)	(0.0385)	(0.0236)	(0.0370)	(0.0241)	(0.0296)	(0.0269)
Estudante	1.717***	0.954	1.653***	0.939	1.425***	0.948
(universitário)	(0.0783)	(0.0766)	(0.0800)	(0.0773)	(0.0668)	(0.0739)
Zona Urbana			2.450***	0.812***	2.429***	0.813***
			(0.164)	(0.0501)	(0.157)	(0.0476)
Retornante			1.010	2.200***	1.007	2.182***
			(0.0403)	(0.127)	(0.0399)	(0.125)
Mesma UF			0.551***	0.489***	0.558***	0.484***
			(0.0487)	(0.0538)	(0.0494)	(0.0529)

**Tabela A1 - Modelo multinomial - 2000: razões de risco relativo e testes de ajuste**

(continuação)

	Modelo 1		Modelo 2		Modelo 3	
	Ascendente	Descendente	Ascendente	Descendente	Ascendente	Descendente
Masculino					0.994 (0.0209)	1.204*** (0.0314)
Preta					0.786*** (0.0297)	0.723*** (0.0278)
Parda					0.862*** (0.0190)	0.810*** (0.0191)
Idade					0.995*** (0.001)	1.006*** (0.001)
Já viveu com cônjuge					1.036 (0.018)	1.092*** (0.018)
Nunca viveu com cônjuge					1.624*** (0.052)	1.094** (0.036)
Nº de Filhos					1.031*** (0.005)	0.974*** (0.006)
Log da renda total mensal					0.996 (0.010)	0.964** (0.013)
Trabalha até 30h/s					0.856** (0.051)	1.099 (0.085)
Trabalha mais de 30h/s					0.962 (0.055)	1.003 (0.074)
Observações	511.978		511.978		511.978	
Pseudo R <sup>2</sup>	0.041		0.065		0.070	
AIC	9095824		8868610		8816373	
BIC	9096136		8868989		8816975	

Erros-padrões robustos entre parênteses (\*\*\*p<0.01; \*\*p<0.05; \*p<0.10).

Fonte: Resultado da pesquisa.



Tabela A2 - Modelo multinomial - 2010: razões de risco relativo e testes de ajuste

(continua)

	Modelo 1		Modelo 2		Modelo 3	
	Ascendente	Descendente	Ascendente	Descendente	Ascendente	Descendente
Norte	0.842 (0.180)	2.393*** (0.565)	0.944 (0.175)	2.120** (0.610)	0.978 (0.181)	2.160** (0.622)
Nordeste	0.455*** (0.085)	2.412*** (0.421)	0.576*** (0.085)	2.201*** (0.466)	0.617*** (0.088)	2.179*** (0.464)
Sul	1.200 (0.166)	1.833** (0.416)	1.202 (0.137)	1.693 (0.475)	1.183 (0.133)	1.685 (0.477)
Centro-Oeste	1.306 (0.225)	1.202 (0.344)	1.216 (0.170)	1.055 (0.359)	1.229 (0.170)	1.073 (0.367)
Origem	2.478*** (0.366)	0.660** (0.103)	2.357*** (0.429)	0.700 (0.139)	2.343*** (0.422)	0.725 (0.145)
Norte	3.649*** (0.488)	0.514*** (0.050)	2.944*** (0.335)	0.523*** (0.064)	2.808*** (0.312)	0.548*** (0.068)
Origem	1.707*** (0.176)	0.816 (0.123)	1.762*** (0.230)	0.883 (0.174)	1.727*** (0.224)	0.879 (0.173)
Nordeste	1.119 (0.168)	0.964 (0.156)	1.145 (0.163)	0.974 (0.194)	1.144 (0.163)	0.992 (0.199)
Centro-Oeste	1.155*** (0.028)	1.074** (0.026)	1.056* (0.022)	1.077*** (0.023)	0.971 (0.022)	1.137*** (0.027)
Fundamental	1.246*** (0.047)	1.048 (0.047)	1.086* (0.036)	1.065 (0.042)	0.947 (0.033)	1.120** (0.045)
Médio	1.154* (0.068)	1.129 (0.126)	0.948 (0.058)	1.102 (0.122)	0.816*** (0.048)	1.112 (0.110)
Superior	1.294*** (0.041)	0.874*** (0.031)	1.227*** (0.037)	0.851*** (0.032)	1.024 (0.029)	0.876*** (0.031)
Estudante (até o médio)	1.891*** (0.067)	0.946 (0.052)	1.846*** (0.068)	0.952 (0.054)	1.439*** (0.048)	0.942 (0.045)
Estudante (universitário)			2.484*** (0.159)	0.857** (0.049)	2.397*** (0.152)	0.837** (0.047)
Zona Urbana			1.097** (0.039)	1.544*** (0.092)	1.094** (0.038)	1.541*** (0.092)
Retornante			0.560*** (0.045)	0.525*** (0.057)	0.567*** (0.045)	0.523*** (0.057)
Mesma UF						

**Tabela A2 - Modelo multinomial - 2010: razões de risco relativo e testes de ajuste**

(continuação)

	Modelo 1		Modelo 2		Modelo 3	
	Ascendente	Descendente	Ascendente	Descendente	Ascendente	Descendente
Masculino					0.988 (0.014)	1.106*** (0.016)
Preta					0.800*** (0.031)	0.841*** (0.033)
Parda					0.866*** (0.020)	0.904*** (0.025)
Idade					0.990*** (0.001)	1.005*** (0.001)
Já viveu com cônjuge					1.132*** (0.020)	1.074*** (0.016)
Nunca viveu com cônjuge					1.774*** (0.055)	1.149*** (0.044)
Nº de Filhos					0.991 (0.007)	0.984** (0.006)
Log da renda total mensal					1.017*** (0.004)	1.011* (0.005)
Trabalha até 30h/s					0.790*** (0.017)	0.846*** (0.018)
Trabalha mais de 30h/s					0.822*** (0.016)	0.793*** (0.017)
Observações	475.635		475.635		475.635	
Pseudo R-sq	0.036		0.053		0.061	
AIC	8761750		8607433		8534919	
BIC	8762060		8607809		8535517	

Erros-padrões robustos entre parênteses (\*\*\*p<0.01; \*\*p<0.05; \*p<0.10).

Fonte: Resultado da pesquisa.

**ESTRUTURA FAMILIAR E PADRÕES ESPACIAIS DA ESCOLHA RESIDENCIAL\***

**Tatiana Kolodin Ferrari**

Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR) da Universidade Federal de Minas Gerais (UFMG)

E-mail: [tatianak.ferrari@gmail.com](mailto:tatianak.ferrari@gmail.com)

**Antônio Miguel Vieira Monteiro**

Laboratório de investigação em Sistemas Socioambientais (LiSS). Coordenação Geral de Observação da Terra (CGOBT). Instituto Nacional de Pesquisas Espaciais (INPE)

E-mail: [miguel.monteiro@inpe.br](mailto:miguel.monteiro@inpe.br)

**Pedro Vasconcelos Maia do Amaral**

Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR) da Universidade Federal de Minas Gerais (UFMG)

E-mail: [pedroamaral@cedeplar.ufmg.br](mailto:pedroamaral@cedeplar.ufmg.br)

**RESUMO:** A ocupação residencial das áreas urbanas ocorre de maneira heterogênea, em que se observam aglomerações de grupos similares em determinadas localizações. A estrutura familiar é apontada na literatura como um potencial fator para determinar a localização residencial. A hipótese é que as distintas estruturas familiares têm necessidades específicas de consumo de espaço e de acessibilidade, além do que a renda deve ser considerada de forma *per capita* na função de restrição. Para verificar essa hipótese, o presente trabalho busca explorar a localização espacial de diferentes estruturas familiares no município de São Paulo e observar uma possível relação entre a constituição familiar e os atributos construtivos. A metodologia deste estudo é baseada em análise exploratória de dados espaciais e método de análise de *clusters*. Os principais resultados encontrados evidenciam a existência de padrões de localização com base na estrutura familiar e a renda domiciliar *per capita* como um importante indutor da localização.

**Palavras-chave:** Localização residencial; Padrões espaciais; Estrutura familiar.

**Classificação JEL:** R14; R20; J12.

**FAMILY STRUCTURE AND SPATIAL PATTERNS OF RESIDENTIAL CHOICE**

**ABSTRACT:** The residential land use of urban areas presents spatial heterogeneity, in which agglomerations of similar groups are observed in certain locations. Household structure is pointed out in the literature as a potential factor in determining residential location. The hypothesis is that different household structures have specific needs of consumption of space and accessibility. In addition to that the income restriction should consider the number of members in the household. To verify this hypothesis, the present work investigates the distribution of household structures in the city of São Paulo and observes a possible relationship between household composition and constructive attributes. The methodology of this study is based on exploratory spatial data analysis and cluster analysis method. The main results show the existence of location patterns based on family structure and per capita household income as an important inducer of location.

**Keywords:** Residential location; Spatial patterns; Household structure.

**JEL Codes:** R14; R20; J12.

## 1. Introdução

De acordo com o Censo Demográfico de 2010, 84,4% da população brasileira se encontra nas áreas urbanas (IBGE, 2010), indicando um alto grau de concentração. Essa aglomeração espacial faz com que a atividade residencial tenha grande importância na alocação de usos do solo urbano e se torne responsável por boa parte da configuração espacial das cidades. Em vista disso, o entendimento do comportamento de escolha do local de moradia tem se tornado primordial para o planejamento urbano e para a formulação de políticas públicas.

Cada parcela do solo de uso residencial possui características próprias, assim como também sofre influência dos atributos locais e de vizinhança, fazendo com que cada indivíduo tenha que realizar uma escolha, ponderando as diferentes características. De acordo com os modelos tradicionais da Economia Urbana, os indivíduos realizam um *trade-off* entre acessibilidade, espaço e amenidades locais no seu processo de escolha residencial (FUJITA, 1989).

O trabalho de Alonso (1964) foi pioneiro na modelagem do uso do solo urbano. Tendo como base o sistema de lances, o estudo mostra que os indivíduos escolhem seu local de residência maximizando sua função de utilidade, função essa que depende do tamanho do lote, da distância ao centro da cidade e da restrição orçamentária.

Ao se analisar a configuração dos espaços urbanos, nota-se que, em geral, esses são heterogêneos e são organizados por aglomerações de determinados grupos similares em determinadas localizações. O trabalho de Cecchini et al. (2019) mostra que a atividade imobiliária nas áreas metropolitanas tem se tornado complexa, de forma que está sendo menos influenciada pelas forças econômicas tradicionais (forças de aglomeração) e se tornando mais associada à dinâmica demográfica, de migração, de gentrificação, de acessibilidade e de atrativos culturais/naturais.

Nos grandes centros urbanos, é possível observar um fenômeno de diferenciação espacial com base nas características sociodemográficas. O presente trabalho tem especial interesse em explorar os diferenciais de localização residencial que ocorrem com base na estrutura familiar constituída. Nesse aspecto, o trabalho de Beckmann (1973) foi pioneiro ao estender o modelo de Alonso (1964) considerando o problema da estrutura familiar na escolha residencial. Mas, apesar de vários estudos terem apontado uma relação entre a estrutura familiar e o espaço para habitação, poucos trabalhos têm procurado entender como essa estrutura familiar está associada à localização residencial (JUNG; YANG, 2016).

Dessa forma, é central para ampliar a compreensão da estruturação do espaço intraurbano entender por que determinadas famílias e grupos populacionais escolhem certas localizações de residência e não outras. São essas escolhas que produzem a paisagem urbana. Neste estudo, buscam-se uma exploração e uma caracterização desses padrões de localizações e uma discussão das consequências desses na estruturação do espaço intraurbano de São Paulo. Assim, o objetivo deste trabalho está em responder: Em que medida a estrutura familiar se diferencia espacialmente na cidade? Se isso ocorre, quais características podem desempenhar um papel para essa diferenciação na escolha residencial?

Para a análise, busca-se primeiramente identificar a existência de um padrão de localização, dado o tipo de estrutura domiciliar, através da utilização de análises estatísticas e espaciais. A metodologia deste estudo é baseada em análise multivariada e exploratória de dados. Os resultados encontrados evidenciam os diferentes padrões de localização. Procura-se ainda verificar a existência de características comuns das construções entre as áreas através da análise de *clusters*, a fim de inferir características que podem contribuir para o padrão de localização das famílias.

O trabalho contribui para a literatura ao mostrar que a estrutura familiar é um fator determinante do local de moradia das famílias. Por exemplo, a presença de filhos, ou seja, de dependentes na estrutura familiar, apresentou-se como um fator importante para a localização dessas famílias na borda da cidade, como consequência temos uma demanda diferenciada no local por serviços e tipo de habitação. Essa constatação é relevante, pois, do ponto de vista das políticas públicas, esses padrões geram diferentes demandas locais que devem ser analisadas nas formulações de políticas habitacionais e de planejamento urbano.

Além desta introdução, o trabalho traz uma revisão da literatura na seção 2; apresenta a metodologia utilizada, descrevendo a área de estudo, a classificação da estrutura domiciliar utilizada, a base de dados e o método de análise aplicado, na seção 3; mostra os resultados na seção 4; e, por fim, a seção 5 aborda as principais conclusões.

## 2. Teoria urbana e localização residencial

O principal objetivo de estudo da economia urbana está no entendimento da estrutura interna das cidades, que analisa o mercado de terras, a fim de entender como estão alocados os diferentes agentes e as atividades econômicas (THISE, 2011).

Dentre os diferentes usos do solo urbano, o uso residencial engloba uma parcela importante de consumo do espaço urbano, sendo que, em geral, mais da metade do uso do solo urbano é alocado para essa atividade. Assim, o entendimento da estrutura espacial e das características da alocação residencial é importante para o planejamento urbano das cidades.

O bem terra possui características distintas dos demais bens econômicos. A terra é um bem indestrutível, imóvel e sua oferta é fixa, sendo a função de oferta inelástica no curto prazo, para determinado espaço considerado. Além disso, cada parcela do solo possui características únicas, que refletem suas características físicas e naturais, assim como propriedades institucionais. A ação do homem e o uso destinado a cada parcela do solo afeta o uso e o valor das terras vizinhas, ou seja, cada unidade do bem terra possui externalidades sobre as demais unidades (FUJITA, 1989). Dadas essas características, a pergunta que surge é: como os indivíduos e as famílias decidem onde irão residir?

Desde a década de 1960, um número crescente de estudos tem se concentrado em investigar os determinantes das escolhas de localização residencial urbana e sua influência no surgimento de padrões espaciais capazes de afetar o cotidiano dos habitantes urbanos. A base teórica da abordagem dominante atual para a localização residencial urbana tem suas raízes no modelo de Alonso (1964). Seguindo os princípios defendidos por esses modelos, uma ordem única e eficiente é alcançada por meio de escolhas residenciais que equilibram um *trade-off* entre consumo de habitação e custos de deslocamento para o trabalho.

A ideia é que cada família, ao decidir o local de sua residência, busca maximizar sua utilidade, pondera a acessibilidade ao CBD em relação à sua necessidade de espaço da habitação, estando sujeita a restrições orçamentárias e de tempo. Partindo-se de uma cidade monocêntrica, a ideia é que, quanto maior for a acessibilidade ao CBD, maior será o preço pelo espaço de habitação. As famílias possuem uma renda fixa ( $Y$ ) e a cada período gastam no consumo de bens ( $z$ ) e no consumo da sua unidade residencial ( $s$ ). Assim, dada uma certa distância  $r$  ao CBD, a escolha residencial pode ser representada como:

$$\begin{aligned} & \max_{r,z,s} U(z, s), \\ & \text{sujeito a:} \\ & z + R(r)s = Y - T(r) \end{aligned} \tag{1}$$

em que  $R(r)$  se refere ao preço por unidade de terra na distância  $r$  e  $T(r)$  ao custo de transporte em  $r$ , considerando que  $r > 0$ ,  $z > 0$ ,  $s > 0$ .

Como resultado, os modelos de cidade monocêntrica apresentam uma estrutura em círculos, em que cada classe econômica ocupa determinado anel ao redor da cidade, sendo a dispersão da população uma função unicamente da distância ao CBD.

Esse é o modelo básico que trata a escolha residencial de forma simplificada, deixando elementos importantes de fora da análise. Os espaços urbanos são heterogêneos e possuem uma diferenciação espacial com base em atributos e características dos indivíduos, apontando que existem outras características que condicionam a escolha locacional além desse *trade-off* apontado.

Beckmann (1973) foi o primeiro a trazer o problema da estrutura familiar para a escolha residencial. O grupo familiar possui membros que trabalham ( $n$ ) e membros que são dependentes

desses ( $d$ ). Quanto maior o tamanho da família, maior será a necessidade de consumo de espaço de habitação. Em particular, quanto maior o número de dependentes, os que não trabalham, em relação aos membros que trabalham, menor será o lance de aluguel que uma família conseguirá pagar por um determinado espaço a uma distância  $r$ . Assim, o que Beckmann (1973) procurou responder é como as famílias de diferentes tamanhos e renda irão se localizar e o quanto de espaço vão demandar.

Fujita (1989) apresenta a formalização do modelo de Beckmann (1973) sobre o problema de maximização da utilidade da família. Considerando o tempo dos membros que trabalham na família ( $t$ ), que precisa ser dividido entre o tempo de trabalho ( $t_w$ ), o tempo de lazer ( $t_l$ ) e o deslocamento ( $b_r$ ), temos o problema de maximização da escolha residencial da família como:

$$\begin{aligned} & \max_{r,z,s,t_l,t_w} U(z, s, t_l; d, n), \\ & \text{sujeito a:} \\ & z + R(r)s + nar = Y_N + nW_{tw} \\ & e \\ & t_l + t_w + b_r = t \end{aligned} \quad (2)$$

em que a primeira restrição se refere à renda, em que  $z$  é o consumo de bens,  $R(r)s$  o aluguel por unidade de terra na distância  $r$ ,  $n$  o número de pessoas que trabalham e  $ar$  o custo de deslocamento, que deve ser igual à renda ( $Y_N$ ) mais o salário por hora de trabalho ( $nW_{tw}$ ). A segunda restrição se refere ao tempo dos membros que trabalham na família.

Como resultado, o modelo mostra que: i) quanto maior o número de dependentes na família, maior será sua distância ao CBD, devido à necessidade de maior número de unidades de terra; ii) para uma família que não possui dependentes, a localização é independente do tamanho da família.

Estudos recentes têm enfatizado o padrão espacial que emerge de características culturais, demográficas e de consumo em nível das famílias. Jung e Yang (2016) discutem o processo de suburbanização e o papel da estrutura familiar nesse processo. Os autores partem do pressuposto de que a estrutura familiar está associada com o local de residência, uma vez que a composição e o tamanho da família geram diferentes necessidades. Realizando uma análise para 29 áreas metropolitanas nos Estados Unidos, Jung e Yang (2016) encontraram evidências que suportam a hipótese de que há diferença na estrutura familiar entre a área central e o subúrbio. Particularmente, observaram que indivíduos sozinhos são mais comuns nas áreas centrais e que famílias nucleares estão mais presentes nos subúrbios do que nas áreas centrais.

Dessa forma, tem-se que:

*“one-person households tend to live in central cities. Young adults are known to reside in central city locations due to the neighborhood liveliness and temporary housing consumption patterns. Neighborhood environments such as “proximity to restaurants, mixed-used land uses, and night life” attract young householders. Also, “temporary housing needs” make young one-person householders remain in central city [...] In contrast, married-couple households tend to live in suburbia. Historically, suburbia has been portrayed as an ideal place for married-couple households that have children. The favorable living conditions of suburbia, the provision of safety, a peaceful and ample space, and physical comfort, meet parents’ expectation of raising children”.* (JUNG; YANG, 2016, p. 4).

Buzar et al. (2005) apresentam uma discussão sobre o tema, enfatizando o potencial de estudo abordando diferentes aspectos da estrutura familiar na geografia humana. Dentro dessa discussão, mostram que a ocorrência de concentração de alguns tipos de estruturas familiares em determinadas partes da cidade gera diferenças territoriais em termos de consumo de bens e serviços e na produção do espaço urbano.

Existe uma relação entre o local construído e as características sociodemográficas de sua população. Assim, se existe um padrão de localização com base na estrutura familiar, há demandas

construtivas específicas com base nas suas características, de forma que tal padrão de localização das famílias pode ser impulsionado ou impulsionar determinados padrões construtivos modelando o espaço urbano.

A discussão sobre a suburbanização apresentada por Jung e Yang (2016) pode representar a preferência das famílias nucleares por maior espaço de habitação, de forma que o mercado imobiliário para atender essas famílias cria áreas que atendem aos seus desejos nos subúrbios da cidade. Consequentemente, surge também toda uma infraestrutura, como escolas e creches, para atender a esse tipo de família em específico. Por outro lado, o valor do metro quadrado mais elevado no centro das cidades induz o mercado imobiliário a uma verticalização no padrão de construção, sendo um motivo para afastar certas famílias dessas áreas.

Apesar dessa discussão, pouca importância e estudos empíricos levam em conta a estrutura familiar para explicar o padrão residencial nas cidades. Mais do que isso, Cecchini et al. (2019) chamam atenção que a pesquisa ligando mecanismos locais de expansão urbana (e econômica) e transformações sociodemográficas, como um caminho de desenvolvimento imobiliário e segmentação em mercados de construção, é cada vez mais necessária para fornecer estruturas interpretativas sobre as mudanças espaciais nas metrópoles.

### 3. Metodologia

#### 3.1 Área de estudo

A área de estudo deste trabalho é a cidade de São Paulo, capital do estado de São Paulo. Cidade mais populosa do Brasil, com pouco mais de 11 milhões de habitantes, distribuídos sobre um território de 1.523,3 Km<sup>2</sup>.<sup>1</sup>

A cidade é dividida em nove zonas administrativas (Figura 1), sendo a zona do centro, centro-sul e o começo da zona oeste<sup>2</sup> consideradas as principais áreas de negócio e comércio da cidade, caracterizando-se como o CBD da região.

#### 3.2 Base de dados e algumas especificações

O trabalho parte da hipótese de que existem padrões de localização com base na estrutura familiar. Assim, primeiramente, faz-se necessário classificar esses diferentes grupos.

Considerando a ideia de arranjos familiares<sup>3</sup>, criou-se uma primeira tipologia a partir do chefe do domicílio<sup>4</sup>. Da identificação do chefe do domicílio, são definidas as demais posições dos moradores, sendo: cônjuge ou companheiro, filho ou enteado e demais parentes. Do exposto, foram criadas quatro categorias:

- Arranjo Matrimonial: casal, independente do sexo, unidos matrimonialmente ou por união estável, sem filhos.
- Arranjo Nuclear: casal, independente do sexo, unidos matrimonialmente ou por união estável, com a presença de filhos.
- Arranjo Monoparental: domicílios que tem a presença de apenas um dos pais na criação dos filhos.

<sup>1</sup> IBGE. Dados do Censo Demográfico, 2010. Disponível em: <https://censo2010.ibge.gov.br/sinopse>.

<sup>2</sup> Cabe ressaltar que essa área é similar aos CBDs identificados por Campos (2017).

<sup>3</sup> Segundo Medeiros e Osorio (2001), por arranjo entende-se a combinação de pessoas classificadas segundo diferentes categorias de parentesco (inclusive não parentes) em grupos e/ou subgrupos de residentes em uma determinada unidade domiciliar”.

<sup>4</sup> Segundo o IBGE (2010), o chefe do domicílio se define como a pessoa (homem ou mulher), de 10 anos ou mais de idade, que é o responsável pelo domicílio, ou é assim considerada pelos demais moradores.

- Arranjo Composto: o domicílio composto traz a noção de que o domicílio não abrange apenas o casal e filhos. Pessoas agregadas e com outros vínculos familiares podem estar integradas ao domicílio.

Cabe ressaltar que foi também considerado o arranjo unipessoal. Apesar de os indivíduos que moram sozinhos não serem considerados como arranjos familiares, mas sim como um arranjo “não família”, esses mesmos foram incluídos neste estudo. Sua inclusão é importante devido ao elevado crescimento dessa forma de domicílio nas últimas décadas (ALVES; CAVENAGHI, 2012).

Além de analisar a distribuição espacial, procuramos identificar áreas homogêneas com relação ao padrão construtivo de forma que poderia ser um efeito e/ou afetar a ocorrência de determinado padrão de localização desses grupos. Para esse fim, foram selecionadas as seguintes dimensões do espaço construído: tipo de construção, número de cômodos, renda domiciliar *per capita*.

O tipo de construção indica a preferência dos indivíduos por determinado perfil de densidade local. Locais centrais possuem maior densidade por área, em que predominam as construções verticais, ou seja, prédios em detrimento de casas. Sendo o oposto nos subúrbios.

Já o número de cômodos nos remete à questão do tamanho da família e da preferência por espaço. Como o valor do metro quadrado é mais caro nas regiões centrais, em geral, residências com um número maior de cômodos tendem a ser construídas em áreas periféricas. Exceção ocorre em locais de assentamento informal, os quais são alguns dos principais mecanismos de espraiamento urbano, que se caracterizam por casas pequenas de pouco planejamento.

Para captar essa peculiaridade, considera-se também na análise a renda domiciliar *per capita*. Apesar de se tratar de uma característica do domicílio, a variável é uma *proxy* para o valor do imóvel, mostrando a capacidade de as famílias de determinada renda residirem no local e indicando também valor do aparato construtivo do local.

Acredita-se que essas variáveis afetem a decisão locacional das famílias, assim, a identificação de áreas homogêneas com determinadas características desses parâmetros pode mostrar o porquê da preferência de certos arranjos por determinadas áreas e apontar tendências de localização.

Os dados foram obtidos pelo Censo Demográfico (IBGE, 2010). Neste artigo, foram utilizadas as áreas de ponderação como recorte populacional. No Censo Demográfico de 2010, foram usados métodos e sistemas automáticos de formação de áreas de ponderação que conjugam critérios como tamanho, contiguidade e homogeneidade em relação a um conjunto de características populacionais e de infraestrutura conhecidas. A fração amostral dos domicílios variou conforme o tamanho da população residente em cada município. O município de São Paulo contou com 310 áreas de ponderação, com fração amostral efetiva de 8,81%.

Foram contabilizadas pouco mais de 3,5 milhões de famílias no município de São Paulo. Essas foram então classificadas de acordo com os arranjos descritos anteriormente e, para a análise, transformadas em taxa por mil famílias, de forma que as diferentes áreas de ponderação pudessem ser comparadas. Assim, o cálculo da taxa de arranjo para cada área de ponderação é dado por:

$$\frac{\text{n}^{\circ} \text{ de famílias de determinado arranjo na área de ponderação } i}{\text{total de famílias residentes na área de ponderação } i} \cdot 1000 \quad (3)$$

com  $i = 1, 2, \dots, 310$ .



Figura 1 - Zoneamento administrativo da cidade de São Paulo



Fonte: Prefeitura Municipal de São Paulo (PMSP)<sup>5</sup>.

<sup>5</sup> Disponível em: <https://www.prefeitura.sp.gov.br/cidade/secretarias/subprefeituras/subprefeituras/mapa/index.php?p=250449>. Acesso em: 28/01/2019.

### 3.3 Método de análise

A configuração de localização dos arranjos, dada sua estrutura, será explorada com a utilização de técnicas de análise espacial de áreas. Segundo Bailey e Gatrell (1995), o principal interesse nesse tipo de análise está em inferir se há alguma tendência ou algum padrão espacial nos valores dos atributos de um dado conjunto de áreas.

O índice de autocorrelação espacial  $I$  de Moran e os valores de estatística descritiva são utilizados para medir a posição e dispersão dos arranjos sobre as áreas de ponderação da cidade de São Paulo.

O índice de Moran ( $I$ ) mede a autocorrelação espacial através do produto dos desvios em relação à média ( $y^* - y$ ). A sua estimativa é dada por:

$$I = n \frac{\sum \sum W_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum (y_i - \bar{y})^2} \quad (4)$$

em que  $n$  é o número de observações.  $W_j$  é uma matriz de vizinhança para o par  $i$  e  $j$ , e  $y$  se refere à média. Como o interesse está em observar se há comportamento similar entre as áreas de fronteira direta, não importando a direção, adotou-se uma matriz de vizinhança do tipo *Queen* com um grau de contiguidade.

De uma forma geral, o índice de Moran presta-se a um teste cuja hipótese nula é de independência espacial; nesse caso, seu valor seria zero. Valores positivos (entre 0 e +1) indicam correlação direta e negativos, (entre 0 e -1) correlação inversa.

O índice  $I$  de Moran é um indicador global que retorna apenas um único valor de associação espacial para o conjunto todo de dados. Entretanto, é de interesse obter valores específicos para cada área do conjunto de dados.

Para analisar de forma individual, os dados de cada arranjo foram espacializados, utilizando-se a divisão em três classes geradas por quebra natural de Jenks, a qual determina a melhor divisão de valores entre as classes. Em seguida, utilizou-se o Indicador de Associação Espacial Local (LISA), que produz um valor específico para cada área, permitindo a identificação de agrupamentos de áreas com atributos semelhantes (*clusters*) e objetos anômalos (*outliers*). Seu cálculo é derivado do índice de  $I$  de Moran, só que utilizando os desvios em relação à média como uma medida de covariância. De maneira formal, temos:

$$I_i = \frac{(y_i - \bar{y}) \cdot W_{zi}}{\sigma^2} \quad (5)$$

sendo,  $W_{zi}$  o valor médio dos desvios dos objetos vizinhos de  $i$  e  $\sigma^2$  a variância da distribuição dos valores dos desvios.

Além disso, tem-se o interesse em verificar se há características espaciais similares entre as áreas de ponderação que condicionam ou são condicionadas pela presença de determinadas estruturas de arranjos domiciliares. Nesse aspecto, foi utilizada a análise de *clusters*, cujo objetivo está em agrupar um determinado conjunto de dados de acordo com algum critério de similaridade. A ideia básica é criar conjuntos que tenham alta similaridade entre si, mas que também apresentem alta dissimilaridade com os objetos que não pertencem ao *cluster*.

A metodologia empregada foi de *cluster* hierárquico pelo método aglomerativo. De acordo com Kaufman e Rousseeuw (2009), esse método parte de todos os elementos separados entre si, designada de etapa 0, em que o número de *clusters* é igual ao número de observações. No caso deste trabalho, em que as observações são as áreas de ponderação do município de São Paulo, partimos de 310 *clusters*. As áreas de ponderação são então agrupadas sequencialmente de acordo com a sua similaridade, dado critério de distância euclidiana no campo das variáveis e método de agrupamento baseado em dissimilaridade máxima.

Os resultados dos *clusters* hierárquicos são apresentados graficamente em um dendograma, cujos nós representam cada etapa do processo de agrupamento. Entretanto, esse método não retorna um

número ótimo de *clusters* a ser considerado. A sua determinação se deu pela análise de demais indicadores através do algoritmo *Nb-Clust*. Esse algoritmo foi desenvolvido por Charrad et al. (2012) para o programa R, justamente para determinar a quantidade ótima de *clusters*, com base na avaliação simultânea de 30 índices.

## 4. Resultados

### 4.1 Análise da distribuição espacial

Apresentamos nesta seção os resultados da análise dos dados e sua distribuição espacial. Como mostra a Tabela 1, das 3,5 milhões de famílias do município de São Paulo, a maioria, 1,4 milhões, são do tipo nuclear. Assim, é de se esperar um maior percentual médio de sua distribuição sobre as diferentes áreas de ponderação. Também é o que possui a maior dispersão dos dados, com desvio padrão de 7,66. O arranjo do tipo nuclear é o mais frequentemente encontrado em todas as áreas de ponderação, possuindo um percentual médio de 40,15%, bastante superior aos demais grupos. O índice global de Moran mostra que existe autocorrelação espacial significativa na localização dos arranjos sobre o território de São Paulo, sendo bastante elevado principalmente entre os unipessoais (0,851) e o arranjo nuclear (0,825) (Tabela 1).

**Tabela 1 - Estatística descritiva da distribuição dos arranjos domiciliares segundo área de ponderação, São Paulo, 2010**

Estatísticas	Unipessoal	Matrimonial	Nuclear	Monoparental	Composto
Total de famílias	512 mil	541 mil	1.422 mil	472 mil	596 mil
Mínimo	5,87	8,93	13,96	5,76	8,01
Máximo	45,27	22,96	54,29	19,53	23,45
Média	14,45	15,23	40,15	13,33	16,81
Desvio-Padrão	7,52	3,27	7,66	2,45	2,83
Mediana	11,89	14,62	41,50	13,46	17,01
Variância	56,64	10,75	58,80	6,02	8,03
<i>I</i> Moran	0,851	0,675	0,825	0,499	0,397

Fonte: IBGE, 2010. Elaboração própria.

O padrão de localização dos arranjos unipessoais fica bastante claro no Figura 2. Grande parte das áreas de ponderação de São Paulo apresentam baixos percentuais de indivíduos morando sozinhos, com exceção de alguns poucos com taxas superiores a 30%, todos localizados na zona central da cidade. O índice local LISA apresenta o padrão alto-alto na região central e o padrão baixo-baixo nas bordas norte, leste e sul da cidade de São Paulo. Apesar dos altos percentuais no centro da cidade, a média da distribuição dos arranjos unipessoais é de apenas 14,45%, decorrência dos baixos percentuais nos arredores da cidade. O desvio-padrão desse grupo foi de 7,52, valor elevado que indica a existência de uma maior dispersão dos dados (Tabela 1).

Analogamente, os domicílios matrimoniais também apresentam concentração no centro da cidade (Figura 3), no entanto, com menor intensidade e maior similaridade entre as áreas de ponderação, tendo um desvio padrão de apenas 3,27 (Tabela 1). Vale ressaltar que o arranjo do tipo matrimonial tende a ser um tipo de arranjo instável, no sentido de que muitos casais sem filhos classificados nessa categoria, no momento, podem estar em transição para um arranjo do tipo nuclear e com isso já procurem locais mais semelhantes a esse outro<sup>6</sup>.

<sup>6</sup> Como apresentado por Jung e Yang (2016), casais recém-casados tendem a morar próximos a centros de emprego, mas sua localização tende a mudar a partir do nascimento de filhos, sendo, portanto, um arranjo localmente instável.

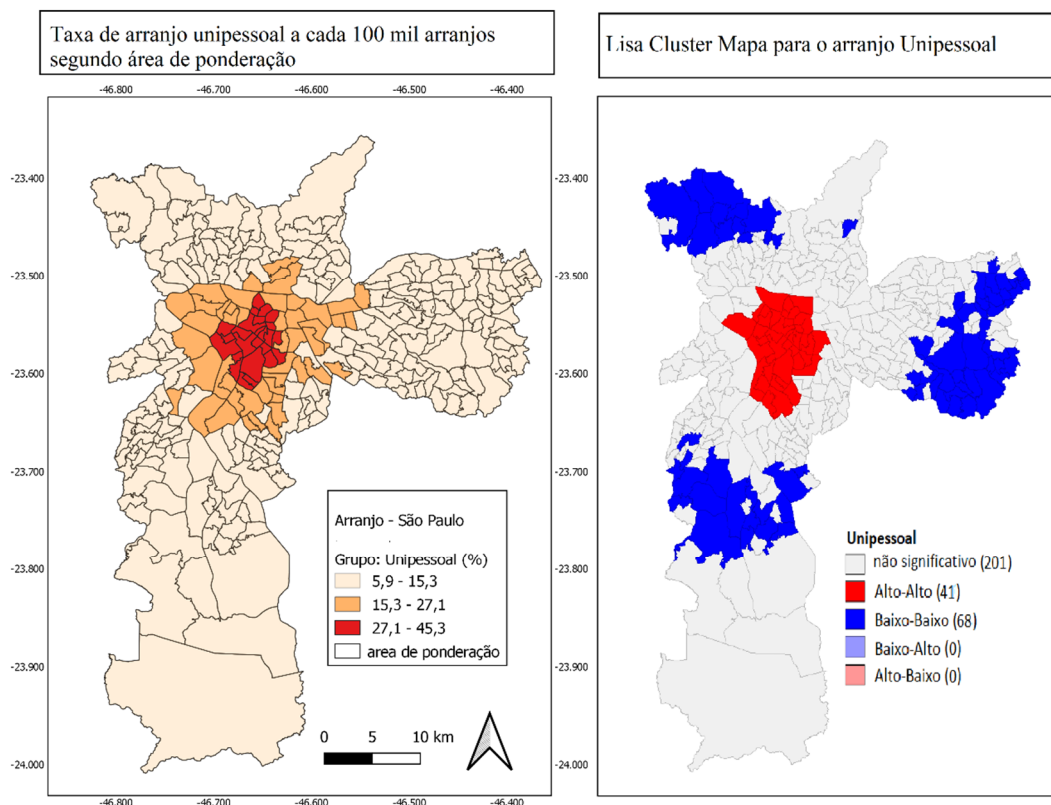
As famílias do tipo nucleares (Figura 4) têm padrão de localização nos subúrbios, com padrão alto-alto nas zonas sul, noroeste e leste.

Com relação aos arranjos monoparentais, esses possuem uma menor representatividade. A sua participação nas áreas de ponderação não passa de 19,53%, com média de 13,33% e mediana no mesmo patamar, 13,46%. O índice global de Moran foi de 0,499. Conforme observado no Figura 5, esses possuem uma tendência de se localizar nas bordas da cidade, afastados do centro, mas não têm um padrão tão homogêneo. Quatro *clusters* de padrão alto-alto foram identificados pelo índice local LISA: 1) no extremo da zona norte, abrangendo as subprefeituras de Perus, Pirituba, Brasilândia e Cachoeirinha; 2) no extremo da zona leste, principalmente nas subprefeituras de Cidade Tiradentes, Guaianazes e Itaim Paulista; 3) na zona sul, nas subprefeituras de Campo Limpo e M'Boi Mirim; e 4) e também zona sul, mais a leste, na subprefeitura de Cidade Ademar.

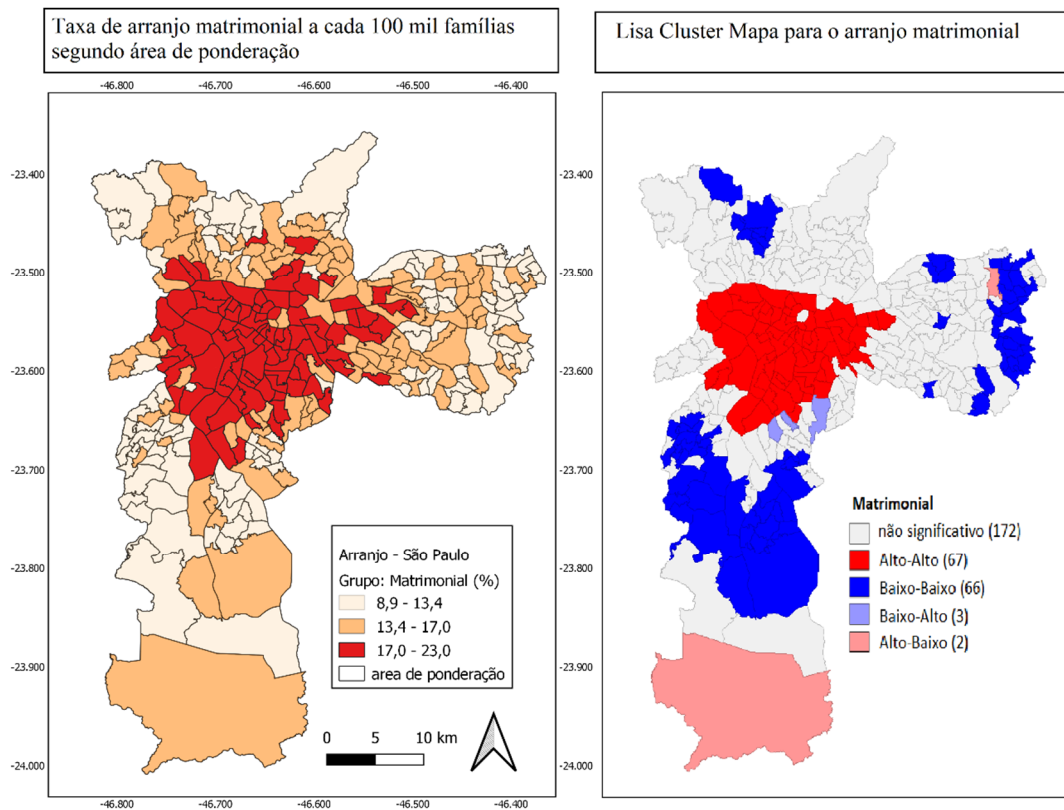
Por fim, os arranjos compostos apresentaram o menor índice de autocorrelação global *I* de Moran, sendo de 0,397. A distribuição do arranjos compostos (Figura 6) também mostra uma tendência a se localizar fora do centro da cidade. Nesse caso, as áreas de ponderação que apresentam as maiores participações, de 18,9% a 23,4%, não são contíguas. O índice local LISA identificou um *cluster* de padrão alto-alto principalmente na região das subprefeituras de Vila Maria e da Mooca.

O fato aqui verificado de que os unipessoais e matrimônios têm uma maior concentração no centro, enquanto os demais arranjos tendem a se localizar afastados desse pode estar ligado a características do local e das construções. O centro da cidade, em geral, é caracterizado por um elevado preço do metro quadrado, construções mais antigas, com menores amenidades ambientais e menor segurança. A preferência de famílias maiores e com filhos em se localizar nos extremos da cidade pode estar ligada a um menor preço do metro quadrado, o que permite adquirir residências maiores, mais condizentes com esse tipo de arranjo familiar. A próxima seção investiga essas hipóteses.

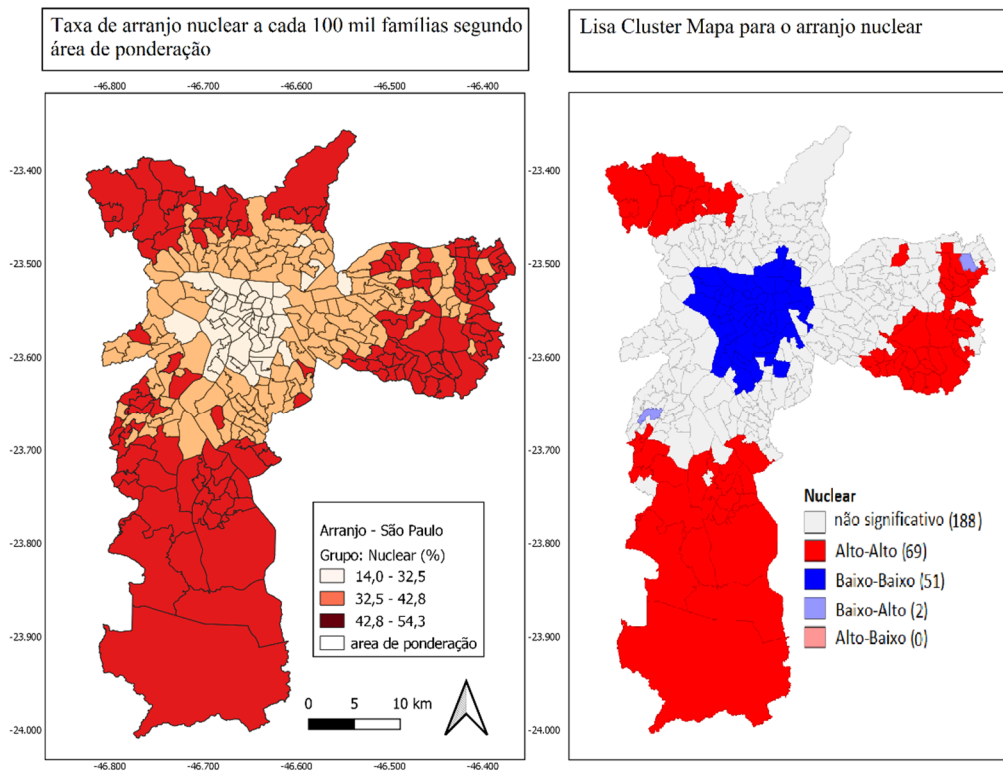
**Figura 2 - Distribuição do arranjo unipessoal e *cluster* espacial por área de ponderação, São Paulo, 2010**



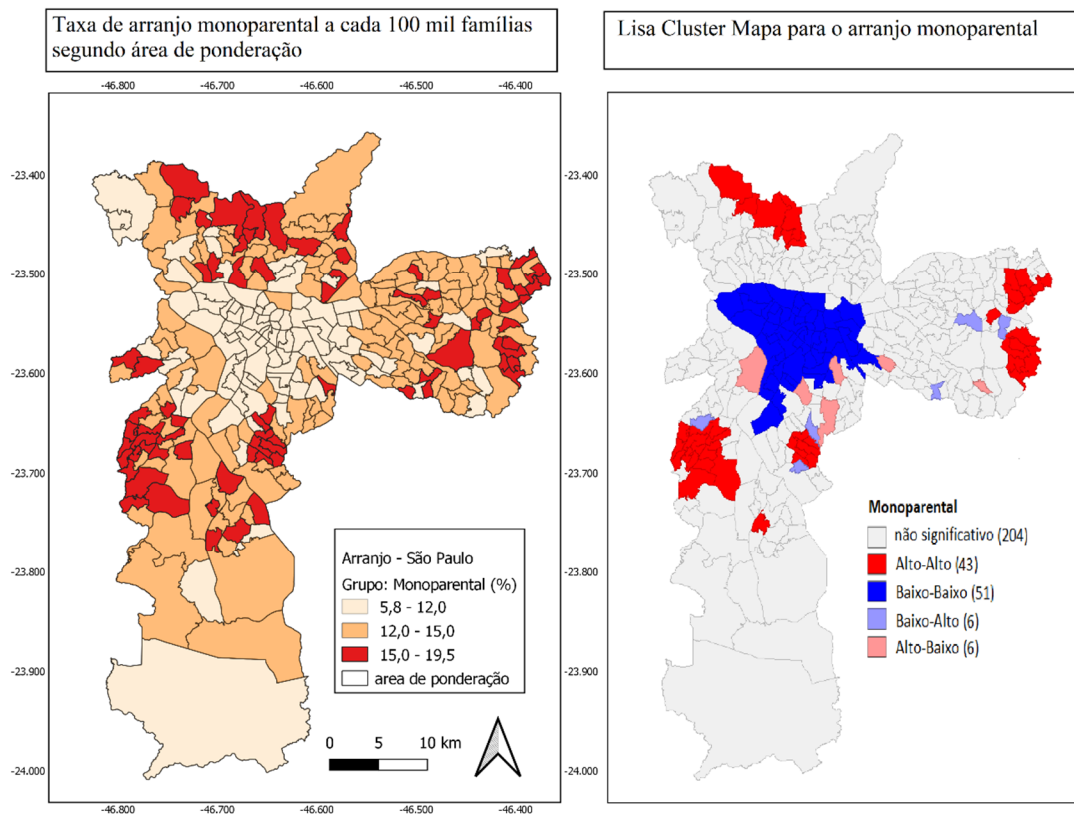
**Figura 3 - Distribuição do arranjo matrimonial e *cluster* espacial por área de ponderação, São Paulo, 2010**



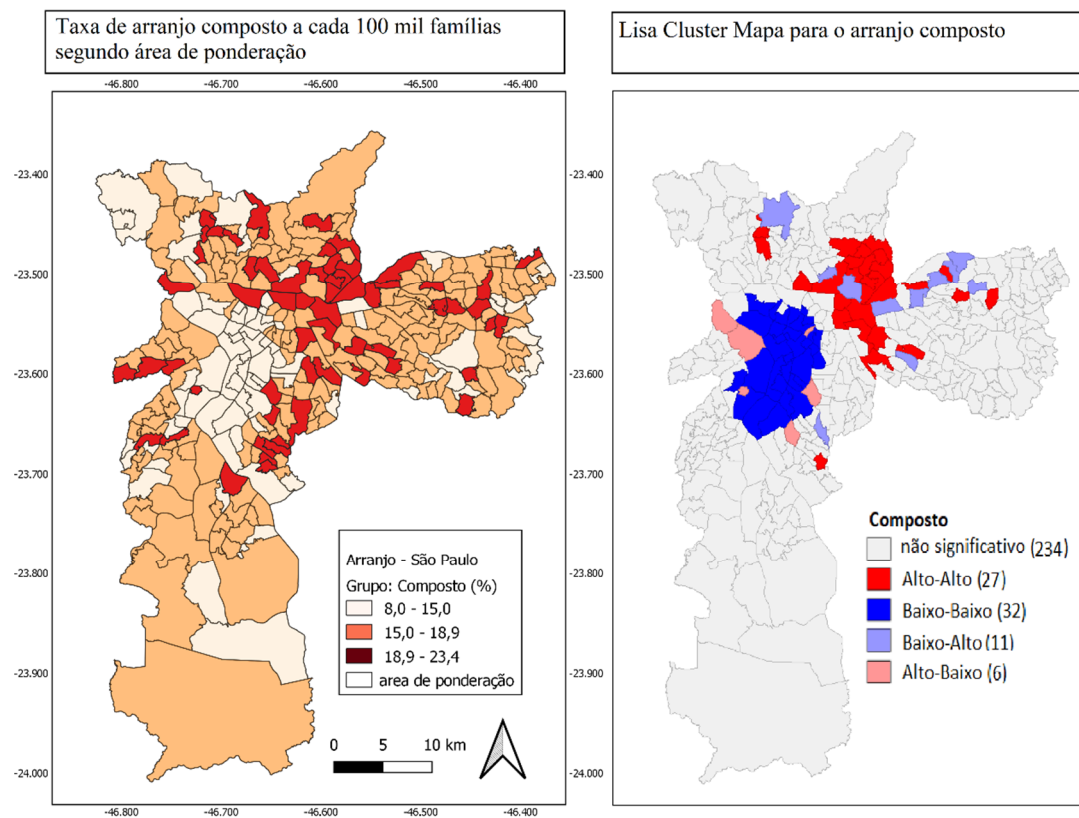
**Figura 4 - Distribuição do arranjo nuclear e *cluster* espacial por área de ponderação, São Paulo, 2010**



**Figura 5 - Distribuição do arranjo monoparental e *cluster* espacial por área de ponderação, São Paulo, 2010**



**Figura 6 - Distribuição do arranjo composto e *cluster* espacial por área de ponderação, São Paulo, 2010**



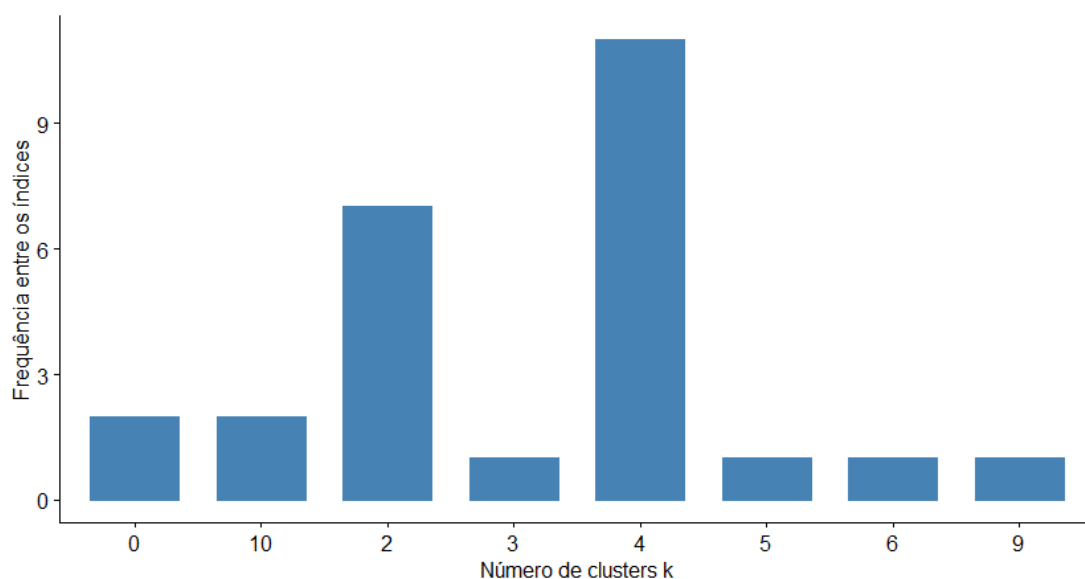
#### 4.1 Análise de clusters

A análise de *cluster* hierárquico visa mostrar as similaridades e divergências no padrão espacial construtivo, observando a média de número de cômodos, o perfil vertical ou horizontal das construções e a renda dos domicílios nas áreas de ponderação da cidade de São Paulo.

Os dados foram padronizados e a similaridade calculada utilizando a distância Euclidiana. Existem diferentes métodos para vincular as variáveis a uma árvore de *cluster* hierárquica. O trabalho realizou essa ligação utilizando cinco algoritmos: máximo, mínimo, pela média, centroide e variância de Ward. Em seguida, foi calculada a correlação entre as distâncias geradas por cada método e a distância dos dados originais, com o objetivo de determinar qual desses métodos reflete com maior precisão os dados originais. O método de máximos, ou completo, retornou o maior coeficiente de correlação, sendo de 0,876. Nesse algoritmo, a distância entre dois *clusters* é definida como o valor máximo de todas as distâncias emparelhadas entre os elementos no *cluster* 1 e os elementos no *cluster* 2 (KASSAMBARA, 2017). Os resultados do dendograma gerado são apresentados na Figura 8.

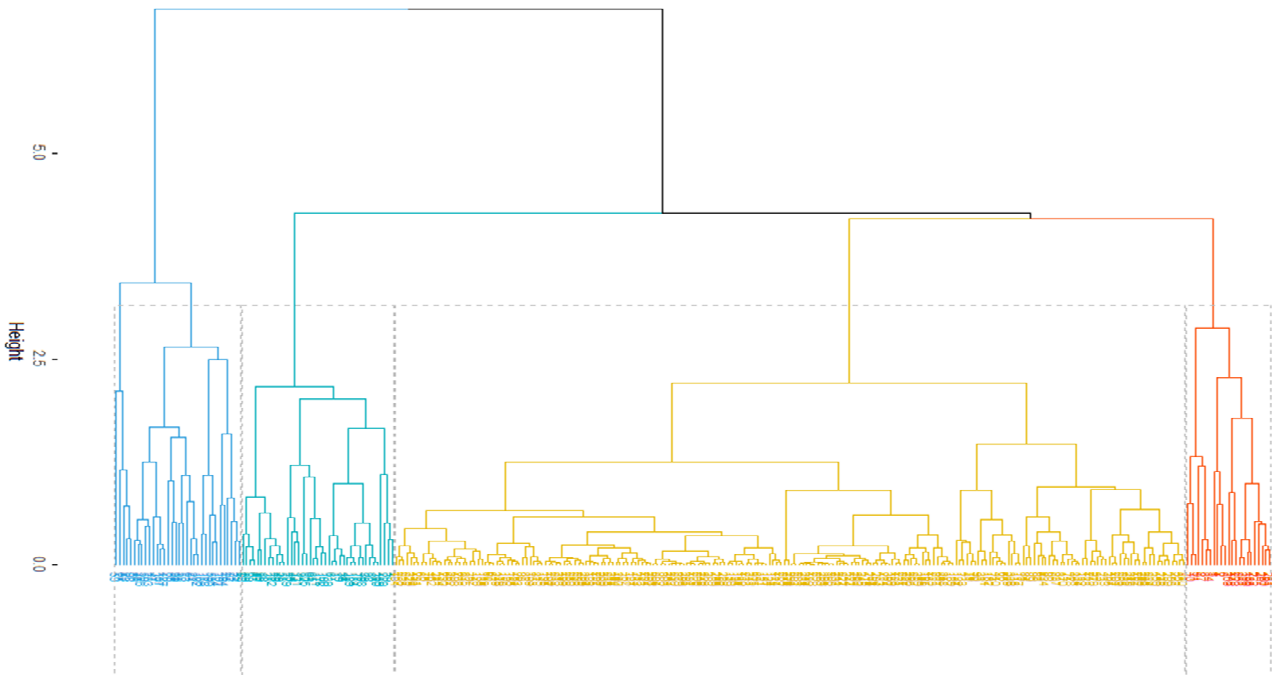
Outra questão na análise de *clusters* é determinar o número ótimo de agrupamentos. Os resultados da aplicação do algoritmo desenvolvido por Charrad et al. (2012) estão resumidos na Figura 7, o qual mostra que 11 índices determinaram que quatro *clusters* seria o ótimo, em seguida, tivemos 7 índices determinando uma quantidade de dois *clusters*.

**Figura 7 - NbClust análise para o melhor número de clusters**



Dessa forma, quatro *clusters* foram estabelecidos com base na média do número de cômodos, padrão vertical das construções e a renda domiciliar *per capita*. As cores apresentadas no dendograma da Figura 8 estabelecem a divisão desses grupos.

A maior parte das áreas de ponderação (212) fazem parte do *cluster* em amarelo, grupo caracterizado por baixa renda, padrão horizontal das construções e baixo número de cômodos na residência. Por outro lado, o agrupamento em vermelho possui a menor quantidade de elementos no grupo, apenas 23 áreas de ponderação. Esse agrupamento, assim como o amarelo, tem baixa média de renda domiciliar *per capita* e residências com poucos cômodos, mas se distingue do *agrupamento amarelo* pelo padrão de construções do local, sendo que possuem média de verticalização de 73%. A Tabela 2 apresenta os valores médios de número de cômodos, renda e padrão de construção, medido pelo percentual de construções verticais.

**Figura 8 - Cluster Hierárquico dos padrões de construção por áreas de ponderação, São Paulo, 2010****Tabela 2 - Características médias dos clusters**

Cluster	Cor de representação	Áreas de Ponderação	Cômodos	Renda (R\$)	Construção Vertical (%)
1	Vermelho	23	5,0	1758,82	73,25
2	Verde	41	6,6	2822,70	47,30
3	Azul	34	7,9	5694,26	74,62
4	Amarelo	212	4,8	888,96	11,71

Já os outros dois *clusters*, representados pelas cores verde e azul no dendrograma, possuem elevado número de cômodos na residência. O que os diferencia é que o *cluster* em azul, com um total de 34 áreas de ponderação, é caracterizado por áreas de alta renda e padrão bastante verticalizado das construções, enquanto o *cluster* em verde, com 41 áreas de ponderação, remete a locais de renda média e um padrão de construção tendendo a casas, visto que o grau de verticalização é em média de 47%. Para entender como esses padrões construtivos se relacionam com os arranjos domiciliares, os resultados foram especializados, conforme apresenta a Figura 9.

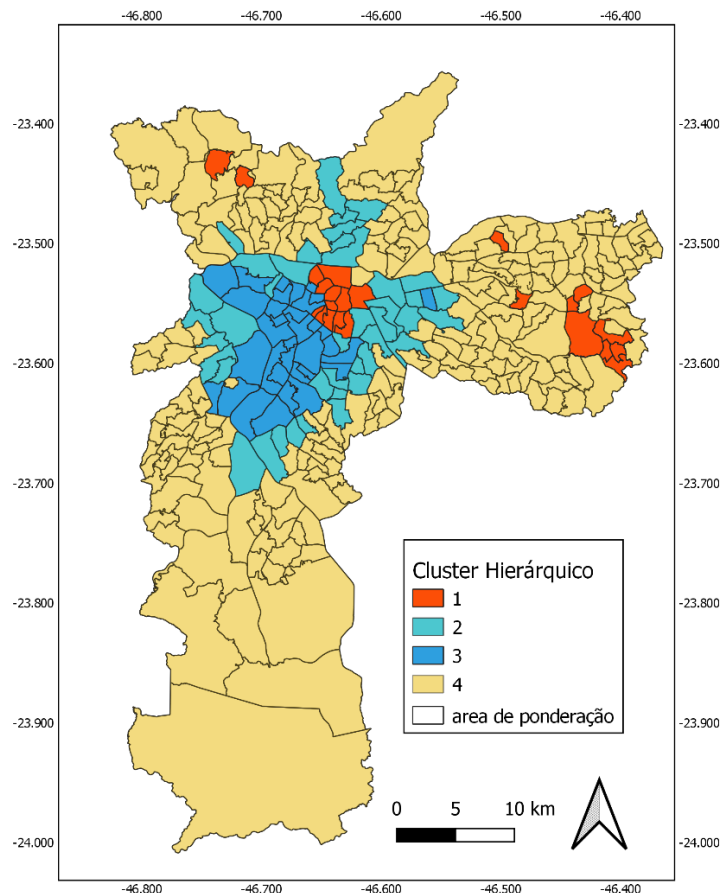
Como se esperava, a região central é mais valorizada, de forma que concentra os domicílios com maior poder aquisitivo. Nota-se que as aglomerações geradas estão fortemente ligadas à distribuição de renda no município, assimilando-se com o padrão de anéis derivado do modelo básico de Alonso (1964).

O padrão espacial das famílias nucleares, monoparentais e compostas nas bordas pode estar ligado justamente a uma menor renda domiciliar *per capita*, visto que esses arranjos possuem maior número de dependentes. Uma extensão para a análise é estratificar os domicílios segundo o número de membros e a proporção dos dependentes no meio familiar, visto que seu padrão locacional indica estar fortemente ligado à renda *per capita*, conforme aponta o modelo de Beckmann (1973). Tal fato aponta a tendência dos arranjos unipessoais de residirem no centro, os quais, possuindo uma maior renda, conseguem se estabelecer nos locais mais valorizados. Importante notar uma peculiaridade na cidade de São Paulo de que residências com maior número de cômodos estão localizadas em áreas de



maior densidade, isto é, com padrão construtivo vertical. Esperava-se que a busca por mais espaço levassem as famílias a residirem em locais com maior presença de casas afastadas do centro, mas a análise apontou que locais de baixa verticalidade tendem a possuir uma média menor de número de cômodos nas residências. O número de cômodos parece estar correlacionado à renda dos domicílios. As áreas de ponderação do *cluster* em azul possuem uma renda domiciliar *per capita* média acima de R\$ 5 mil e uma média de 8 cômodos na residência. Já o grupo do *cluster* em amarelo de menor renda (média de R\$ 888,96) possui em média cinco cômodos em suas residências.

**Figura 9 - Espacialização dos *clusters* hierárquicos, São Paulo**



Uma hipótese é de que as famílias de melhor renda, que buscam a suburbanização como modo de vida, podem estar indo para outros municípios da RMSP, uma vez que essa é altamente interligada, de forma que esse movimento somente no município não seja percebido. No caso de São Paulo, o deslocamento das famílias para as bordas do município parece estar mais ligado a um padrão de renda e de verticalização das construções.

## 5. Considerações finais

A literatura de economia urbana mostra que existe uma relação entre a estrutura familiar constituída e a escolha residencial sobre o espaço urbano. Assim, o primeiro objetivo foi identificar o padrão de localização desses arranjos sobre o território do município de São Paulo.

Em uma análise preliminar, ficou clara a existência de padrões de localização. Em geral, as famílias unipessoais e matrimoniais tendem a se localizar no centro da cidade, sendo essa última mais dispersa. Já as famílias nucleares e monoparentais localizam-se mais afastadas no centro. Famílias

com mais membros, denominadas aqui de compostas, também estão afastadas do centro, localizando-se na região norte e leste da cidade.

A análise de *cluster* do padrão construtivo local estratificou o conjunto de dados em quatro grupos: 1) uma área de alta renda, padrão vertical e elevado número de cômodos na área central do município de São Paulo; 2) circundando esse grupo, temos um *cluster* de renda média, elevado número de cômodos e uma área proporcional em relação a casas e apartamentos; 3) circundando essa, ocupando a maior parte do território, temos um *cluster* de áreas de renda domiciliar *per capita* baixa, padrão horizontal e residências menores; por fim, 4) sem um padrão espacial claro, observa-se um *cluster* de áreas de renda domiciliar *per capita* baixa, residências menores, mas altamente verticalizado.

As análises sugerem a renda domiciliar *per capita* como principal indutor para a localização das estruturas familiares. Existe uma alta correlação entre a renda domiciliar *per capita* e o número de dependentes no domicílio, de forma que o tamanho da família e o número de dependentes passam a ser importantes além do seu arranjo familiar típico. A hipótese é de que os arranjos unipessoais, por não possuírem dependentes, conseguem arcar com um custo maior de moradia na região central do município, com relação aos demais arranjos. Dessa forma, o modelo de Beckmann (1973) parece ser um caminho correto para se tratar a questão.

O padrão construtivo também ficou claro, sendo altamente verticalizado para os arranjos unipessoais e de padrão mais horizontal para os demais arranjos. Mas isso não indica uma preferência por espaço, visto que o tamanho da residência não é um atributo das áreas mais afastadas. Assim, o *trade-off* entre espaço e acessibilidade ao CDB não foi claro na configuração espacial da cidade de São Paulo.

Assim, apesar do mesmo padrão de configuração espacial das estruturas familiares ser encontrado entre este trabalho e de Jung e Yang (2016) nas cidades americanas, as explicações que levam a essa configuração podem ser bem diferentes. Vale ressaltar a diferença nos padrões de periferização entre os dois países. Nos EUA, temos um padrão de suburbanização voltado para o alto consumo do espaço de habitação, caracterizado por amenidades naturais e elevados custos de residência e transporte. O que explica que Jung e Yang (2016) também encontrem forte relação com a renda, mas no padrão americano o aumento de renda é associado com a preferência de residência no subúrbio. No Brasil, temos um duplo movimento de periferização, um tal como o padrão americano, mas na sua maioria esse processo é marcado pela exclusão, sem controle urbano ou social nas periferias, o que resulta em baixo consumo de espaço residencial nas periferias e maior valorização das áreas perto ao centro da cidade.

Os resultados encontrados neste trabalho mostram que a estrutura residencial é um fator a determinar o local de moradia das famílias. Aqui, analisamos apenas a estratificação de acordo com a estrutura familiar, mas a análise de outras estratificações seria interessante em trabalhos futuros a fim de trazer mais evidências do padrão de localização. Do ponto de vista das políticas públicas, esses padrões geram diferentes demandas locais que devem ser analisadas nas formulações de políticas habitacionais e de planejamento urbano. A presença de filhos, ou seja, de dependentes na estrutura familiar, apresentou-se como um fator importante para a localização dessas famílias na borda da cidade, como consequência temos uma demanda diferenciada no local por serviços e tipo de habitação. Este trabalho possui uma série de limitações tanto com relação aos dados como metodológicos. O objetivo foi trazer um primeiro olhar para esse fator que ainda é pouco explorado na literatura. Trabalhos futuros seriam interessantes no sentido de verificar estatisticamente fatores como habitacionais, de vizinhança, cultural, entre outros, a influenciar nesse padrão de localização. Além disso, diferentes estratificações podem ser consideradas, assim como uma análise temporal para observar mudanças e se há uma intensificação de aglomeração espacial das estruturas familiares ao longo do tempo.

## Referências

- ALONSO, W. *Location and land use*. Harvard University Press Cambridge, 1964.
- ALVES, J. E. D.; CAVENAGHI, S. *Tendências demográficas, dos domicílios e das famílias no Brasil*. Aparte: Inclusão Social em Debate, Instituto de Economia, Universidade Federal do Rio de Janeiro, 2012.
- BAILEY, T. C.; GATRELL, A. C. *Interactive spatial data analysis*. Longman Scientific & Technical Essex, 1995.
- BECKMANN, M. J. Equilibrium models of residential land use. *Regional and Urban Economics*, v. 3, n. 4, p. 361-368, 1973.
- BUZAR, S., OGDEN, P. E.; HALL, R. Households matter: the quiet demography of urban transformation. *Progress in Human Geography*, v. 29, n. 4, p. 413-436, 2005.
- CAMPOS, R. The housing market in the municipality of São Paulo: a spatial hedonic prices approach. *Nova Economia*, v. 27, n. 1, p. 303-337, 2017.
- CECCHINI, M.; ZAMBON, I.; SALVATI, L. Housing and the city: A spatial analysis of residential building activity and the socio-demographic background in a mediterranean city, 1990-2017. *Sustainability*, v.11, n.2, p.375-402, 2019.
- CHARRAD, M.; GHAZZALI, N.; BOITEAU, V.; NIKNAFS, A. Nbclust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, v.61, n.6, p.1-36, 2014.
- FUJITA, M. *Urban economic theory: land use and city size*. Cambridge University Press, 1989.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Censo Demográfico*. 2010. Acesso em: julho de 2018. Disponível em: <<http://www.ibge.gov.br/home/>>.
- JUNG, G.; YANG, T. C. Household structure and suburbia residence in us metropolitan areas: Evidence from the american housing survey. *Social Sciences*, v.5, n.4, p. 74-92, 2016.
- KASSAMBARA, A. *Practical guide to cluster analysis in R: unsupervised machine learning*. STHDA, Alboukadel Kassambara, 2017.
- KAUFMAN, L.; ROUSSEEUW, P. J. *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009.
- MEDEIROS, M.; OSORIO, R. *Arranjos domiciliares e arranjos nucleares no Brasil: classificação e evolução de 1977 a 1998*. Instituto de Pesquisa Econômica Aplicada, 2001. (Texto para Discussão, n. 788).
- THISE, J. F. Geografia econômica. In: CRUZ, B. O.; FURTADO, B. A.; MONASTERIO, L. R. JÚNIOR (eds.). *Economia regional e urbana: teorias e métodos com ênfase no Brasil*. Instituto de Pesquisa Econômica Aplicada, 2011.

## ORCID

Tatiana Kolodin Ferrari  <https://orcid.org/0000-0003-1234-1235>

Antônio Miguel Vieira Monteiro  <https://orcid.org/0000-0003-1477-1749>

Pedro Vasconcelos Maia do Amaral  <https://orcid.org/0000-0002-2505-035X>

**ANÁLISE ESPACIAL DAS AGLOMERAÇÕES DA INDÚSTRIA DE TRANSFORMAÇÃO  
NOS SEGMENTOS MODERNO E TRADICIONAL NO ESTADO DO PARANÁ\***

**Andréia Ferreira Prestes**

Mestre em Gestão e Desenvolvimento Regional pela Universidade Estadual do Oeste do Paraná  
(UNIOESTE) – Campus de Francisco Beltrão  
E-mail: andreiafprestes@hotmail.com

**Fernanda Mendes Bezerra**

Professora do curso de Ciências Econômicas da Universidade Estadual do Oeste do Paraná  
(UNIOESTE) – Campus de Francisco Beltrão  
E-mail: ferpompeia@gmail.com

**Talita Egevardt de Castro**

Professora do curso de Ciências Econômicas da Universidade Estadual do Oeste do Paraná  
(UNIOESTE) – Campus de Francisco Beltrão  
E-mail: talita.egevardt@gmail.com

**RESUMO:** A instalação de uma indústria e a formação de aglomerados são fatores importantes para a dinâmica econômica e a geração de riquezas, contribuindo para o desenvolvimento da região e do seu envoltório. Cabe destacar que o nível tecnológico das indústrias pode influenciar nos resultados gerados. Por esse motivo, esta pesquisa tem como objetivo analisar espacialmente as causas da aglomeração da indústria de transformação (IT), em especial entre os segmentos moderno e tradicional, para os municípios do estado do Paraná no ano de 2017. Para isso, foram divididos os setores da IT em dois segmentos: tradicional (baixa e média baixa tecnologia) e moderno (média alta e alta tecnologia). Posteriormente, mensurou-se o Índice de Concentração Normalizado (ICN) e estimou-se o modelo de regressão mais adequado, isto é, com ou sem efeito espacial. Os resultados demonstraram que o segmento tradicional possui impacto espacial influenciando no município e nos vizinhos contíguos, diferentemente do segmento moderno que não apresentou esse efeito. Além disso, os fatores que influenciam na decisão locacional são, na grande maioria, diferentes para cada segmento.

**Palavras-chave:** Concentração; Economia regional; Econometria espacial.

**Classificação JEL:** C01; C31; O1; R1.

**SPATIAL ANALYSIS OF THE GATHERING OF THE TRANSFORMATION INDUSTRY  
IN THE MODERN AND TRADITIONAL SEGMENTS IN THE PARANA STATE**

**ABSTRACT:** The installation of an industry and the formation of agglomerates are important factors for the economic dynamics and the generation of wealth, contributing to the development of the region and its wrapping. It should be noted that the technological level of the industry can influence the results generated. For this reason, this research aims to analyze spatially the causes of the agglomeration of the transformation industry (IT), especially between the modern and traditional segments, for the municipalities of the state of Paraná in the year 2017. For this reason, the IT sectors in two segments: traditional (low technology) and modern (high technology). Subsequently, the Normalized Concentration Index (NCI) was measured and spatial regression was estimated. The results showed that the traditional segment has a spatial impact influencing the municipality and neighboring neighbors, unlike the modern segment that did not reach significance. In addition, the factors influencing the locational decision are, in the great majority, different for each segment.

**Keywords:** Concentration; Regional economy; Spatial econometrics.

**JEL Codes:** C01; C31; O1; R1.

## 1. Introdução

A aglomeração espacial das atividades produtivas pode representar o sucesso de uma região quanto ao seu desenvolvimento e crescimento econômico. Conforme aponta Krugman (1991), no espaço regional, existem forças que induzem a formação de aglomerados produtivos, privilegiando esses locais com a maior capacidade de geração de riquezas e de oportunidades para ampliação das economias de escala e da eficiência produtiva. Assim, quando uma empresa decide se estabelecer em uma determinada localidade, gera padrões determinantes de distribuição espacial da indústria, propiciando o surgimento de aglomerações, que se originam dos retornos positivos da região, intensificando as externalidades e atraindo novos investimentos.

Segundo Rosa e Raiher (2017), a localização da indústria resulta na vinda de novas unidades produtivas para a região e seus arredores. Dessa forma, a aglomeração pode ser benéfica no espaço em que atua, e em seu entorno, com a absorção de mão de obra desocupada da região menos desenvolvida, formando e fortalecendo cadeias produtivas, elevando a produtividade, o emprego, a renda e os níveis de consumo, beneficiando a sociedade economicamente e socialmente.

Para Kaldor (1970), a evolução industrial possui relação com o progresso tecnológico, destacando que a acumulação de aprendizado tecnológico contribui para o desenvolvimento. Além disso, considera a indústria como motor da economia, por apresentar retornos crescentes de escala, fortes *linkages*, além de receber e difundir tecnologias e gerar renda dos produtos exportados.

Por esse motivo, o foco desta pesquisa é a indústria de transformação (IT), a qual foi dividida por intensidade tecnológica entre os segmentos moderno e tradicional. O segmento moderno corresponde à utilização intensiva de capital, abrangendo, assim, os setores que utilizam o fator de produção tecnologia de maneira mais intensiva. Já no segmento tradicional, estão os setores que utilizam mais o fator mão de obra do que a tecnologia em seu processo produtivo. Essa divisão foi realizada, pois, conforme Rocha et al. (2013), os fatores que levam à concentração variam conforme a especificidade de cada indústria. Além disso, Rosa e Raiher (2017) defendem que as indústrias mais intensivas em tecnologia proporcionam maior efeito *spillover* para as regiões vizinhas. Desse modo, a aglomeração pode impactar de forma diferenciada conforme o teor tecnológico presente na região.

Ademais, a identificação dos motivos que levam à escolha locacional industrial e conseqüentemente à formação de aglomerados com diferentes teores tecnológicos é fundamental para o entendimento e a promoção do desenvolvimento econômico regional. Apesar da relevância, ainda são poucas as contribuições empíricas que demonstram as causas da aglomeração, ainda mais quando comparados setores intensivos em capital e em mão de obra e para o estado do Paraná.

Desta maneira, esta pesquisa visou responder o seguinte problema: Quais os principais fatores responsáveis pela aglomeração industrial no Paraná? Assim, o objetivo consistiu em analisar espacialmente as causas da aglomeração da IT, em especial, entre os segmentos moderno e tradicional, para os municípios do estado do Paraná no ano de 2017.

Parte-se da hipótese de que a aglomeração industrial é causada por diferentes fatores, conforme as especificidades de cada município paranaense. Entretanto, de qualquer maneira, entende-se que a aglomeração pode ser favorável para o desenvolvimento da economia local, como também para as economias vizinhas, por meio da interação entre as firmas a qual possibilita o compartilhamento de conhecimentos e técnicas de trabalho, a proximidade com os *inputs* e *outputs* do processo produtivo, o acesso à tecnologia, a redução de custo, entre outros. Além disso, aglomerados produtivos com maior teor tecnológico tendem a gerar maiores transbordamentos para os municípios.

Para o atendimento do objetivo proposto, foi construído o Índice de Concentração Normalizado (ICN) para o segmento moderno e tradicional, utilizando os dados dos empregos formais, extraídos do Ministério do Trabalho (MTE). Estimaram-se dois modelos de regressão espacial e utilizou-se como variável dependente o ICN correspondente a cada segmento e um total de 9 variáveis explicativas mencionadas nas teorias e literatura como os principais fatores responsáveis pela

aglomeração industrial. As variáveis foram extraídas de diversas fontes de dados e estão expostas detalhadamente na metodologia.

Esta pesquisa se encontra dividida em seis seções, incluindo esta introdução. Na segunda, consta o referencial teórico com a Teoria da Segmentação do Mercado de Trabalho (TSMT), as teorias regionais de localização, as ideias *marshallianas*, os teóricos da aglomeração e a visão da Nova Geografia Econômica (NGE). Na terceira, são apresentados os principais trabalhos nacionais e internacionais que abordam os temas voltados para a concentração industrial. Na quarta seção, estão apresentados os métodos e os dados que visam responder os objetivos da pesquisa. Na sequência, tem-se os resultados e, finalizando, as considerações finais.

## 2. Fundamentação teórica

Utilizam-se como fundamentação teórica da concentração espacial as principais teorias de localização, que incluem as contribuições de Marshall juntamente com alguns teóricos que também analisam a concentração regional e a NGE. Mas, antes de adentrar nesse tema, para ficar mais claro ao leitor, faz-se necessário explicar a divisão do mercado de trabalho em segmentos e, para isto, utilizou-se a TSMT abordada a seguir.

### 2.1 Teoria da segmentação do mercado de trabalho

A TSMT ou modelo dual defende a existência de uma descontinuidade do mercado de trabalho entre dois segmentos: o primário e o secundário. Esses, por sua vez, são restritos por barreiras sociais e tecnológicas, que podem ser explicadas por fatores institucionais, culturais e pela concentração de mercado (LEMOS, 1995).

No segmento primário, caracterizam-se os trabalhadores que possuem um trabalho estável, cujas habilidades são adquiridas dentro da própria empresa, a remuneração é mais elevada e as chances de subir de cargo são maiores (REICH et al., 1973). Para Souza (1978), esse segmento atende a necessidade do empregador quanto à aplicação de recursos para o treinamento e à minimização da rotatividade. E também satisfaz os próprios trabalhadores, no que se refere à estabilidade, às vantagens referentes às habilidades, aos conhecimentos adquiridos e à probabilidade de autopromoção. Além disso, para Lemos (1995), o segmento primário apresenta maior progresso tecnológico, o que acarreta um aumento da produtividade, por isso é chamado na economia como setor moderno.

Já no segmento secundário, os trabalhadores possuem condições de trabalho inferiores, pois apresentam menor estabilidade, não recebem qualificação dentro da empresa, os salários são menores, as chances de promoção são mínimas, há ausência de regras, de relação personalizada de emprego, de plano de carreira e apresenta baixa produtividade (REICH et al., 1973; LIMA, 1980; LEMOS, 1995). Nesse mercado, os empregadores não investem e não incentivam treinamentos para seus colaboradores, exigindo apenas o mínimo de qualificação necessário para desempenhar os trabalhos (SOUZA, 1978; LIMA, 1980; LEMOS, 1995). Com isso, esse segmento apresenta alta rotatividade de funcionários, além do baixo nível tecnológico e de produtividade (LIMA, 1980), por isso é chamado na economia de setor tradicional.

Em suma, a TSMT defende que a existência do mercado de trabalho segmentado, por meio do segmento primário (moderno) e do secundário (tradicional), resulta em distinções do mercado de trabalho não justificada pela capacidade produtiva. Essa segmentação pode ser analisada de diversas maneiras, nesta pesquisa será analisada de forma regional, a fim de averiguar espacialmente a concentração desses segmentos dentro do estado do Paraná. Por esse motivo, na sequência, constam as principais teorias regionais que contribuem para esta análise.

## 2.2 Teorias regionais de localização e aglomeração

Partindo da conjectura que o crescimento econômico, a renda e o desenvolvimento entre as regiões não se distribuem igualmente no espaço, economistas e geógrafos econômicos passaram a estudar as características específicas das regiões e suas influências na economia, visando compreender a distinção entre os locais. Nesse contexto, surgem inúmeras teorias com a finalidade de investigar os motivos que levam as firmas a escolherem sua localização geograficamente, em muitos casos gerando aglomerados e conseqüentemente resultando na concentração espacial (BARROS; PRATES, 2014). Uma das primeiras correntes teóricas que abordaram esse tema foi a teoria da localização da firma e da distribuição das atividades produtivas, na qual se encontram os modelos de Von Thünen e Alfred Weber.

Em 1826, a teoria de Von Thünen tratou da organização da atividade econômica, levando em consideração a utilização do solo e a razão pela qual a produção de mercadorias se aloca ao redor do mercado, ou seja, qual o padrão de ocupação no espaço (CAVALCANTE, 2008; CUNHA, 2008; BARROS; PRATES, 2014).

Não obstante, o autor defende que o custo de transporte está diretamente ligado a distância, isto é, quando a produção estiver mais próxima do mercado, menor será a distância e conseqüentemente o custo de transporte. Dessa forma, a renda é máxima quando a distância for zero gerando vantagens locacionais; e mínima quando a distância for grande o suficiente para que o custo de transporte consuma todo o lucro gerado (CUNHA, 2008; BARROS; PRATES, 2014).

Por sua vez, em 1909, Alfred Weber defendeu três fatores importantes para a decisão locacional: custo de transporte, custo de mão de obra e um fator local resultante da aglomeração e desaglomeração (CAVALCANTE, 2008). Para Weber, o custo de transporte é medido desde a matéria-prima até o mercado consumidor, de forma que quanto menor for o custo com transporte maior será o lucro do produtor. Para isso, a firma deve buscar se localizar no ponto ótimo, a fim de minimizar os custos e maximizar os lucros (CUNHA, 2008; BARROS; PRATES, 2014). Em outras palavras, se o custo de mão de obra for maior do que o custo de transporte, as empresas tendem a se concentrar próximas das regiões onde os trabalhadores se localizam. E, se o custo da matéria-prima for superior ao de transporte do produto final, a firma se localizaria mais próxima do processo inicial, caso contrário seria mais próxima do consumidor (CAVALCANTE, 2008).

As teorias mencionadas até o momento, Thünen e Weber, estudaram especificamente o motivo que leva as firmas a escolherem determinada localização para se instalarem. No entanto, segundo Amarante e Silva (2016), a questão da concentração espacial das atividades é discutida desde o século XIX, por meio dos escritos de Alfred Marshall, um dos precursores dessa análise.

Em seus princípios econômicos, Marshall relata os benefícios da concentração de empresas e empregos do mesmo ramo na mesma localidade. Essas vantagens advêm de duas fontes de economias de escala que explicam a expansão produtiva industrial: as economias internas e externas. A primeira depende do tamanho de cada firma, dos recursos, da organização e da eficiência individual da administração, ou seja, é a geração do desenvolvimento da própria capacidade produtiva da firma. A segunda é o desenvolvimento da indústria concentrada, ou seja, do progresso geral do setor. Por isso, o conceito de economias externas surgiu do pressuposto de que a concentração espacial de firmas gera vantagens competitivas para o conjunto dos produtores, que não seria possível se estivessem atuando distantes uns dos outros (MARSHALL, 1985).

De acordo com Marshall (1985), são inúmeros os motivos que interferem na concentração geográfica das indústrias, dentre eles, três se destacam: i) o agrupamento de trabalhadores com habilidades especializadas produzindo novas ideias, fornecendo um mercado de trabalho constante, beneficiando o empregado e também o empregador pela disponibilidade de mão de obra qualificada; ii) as características físicas regionais, como o clima e o solo, a acessibilidade e a variedade de insumos com baixos custos reforçando a eficiência industrial; iii) o transbordamento de conhecimento e a troca de informações técnicas entre a mão de obra especializada, os quais fluem mais facilmente com a proximidade das firmas do que com a distância. Para o autor, as

vantagens dos aglomerados produtivos são inúmeras, contribuindo para explicar a concentração geográfica das atividades econômicas juntamente com o desenvolvimento do setor industrial.

No entanto, após a Segunda Guerra Mundial, surgiram novas contribuições a fim de estudar a dinâmica regional e encontrar explicações para o subdesenvolvimento local. Dentre os principais teóricos estão Gunnar Myrdal e Albert O. Hirschman.

Myrdal (1957) desenvolveu a teoria da Causação Circular Cumulativa (CCC) que parte de uma economia instável e em desequilíbrio. O efeito cumulativo pode ocorrer de forma positiva ou negativa, em que um fator negativo gera outro fator negativo, o mesmo ocorre com o fator positivo, levando ao maior ou menor desenvolvimento regional. Por isso, o autor defende a importância da intervenção pública, uma vez que visa minimizar os efeitos cumulativos negativos e influenciar os efeitos de propulsão do desenvolvimento.

Além disso, defende que, quando uma região está se desenvolvendo, os vizinhos que a circundam também receberão efeitos de impulsão, causando um processo cumulativo positivo de expansão na região e nos vizinhos devido à complementariedade entre si. Esse fato advém principalmente da mão de obra e da matéria-prima absorvidas pela região em desenvolvimento das suas circunvizinhas. Com isso, as regiões vizinhas também irão expandir, pois diminuirá o desemprego e aumentarão a renda, o consumo, a demanda e a produção (MYRDAL, 1957).

Hirschman (1958), apesar de concordar com Myrdal em relação aos efeitos positivos e negativos da aglomeração, considera a desigualdade e a sequência de desequilíbrio regional necessárias, pois funcionam como oportunidade de ajustes para desenvolver as regiões periféricas. Contudo, seu foco está em analisar como o desenvolvimento pode ser transferido entre as regiões.

No entanto, Hirschman (1958) relata que, no processo de desenvolvimento econômico de uma região, desigualdades inter-regionais são inevitáveis devido às pressões causadas pela complementariedade regional. Por isso, a região em expansão também se depara com repercussões positivas e negativas sob a estagnada. Os efeitos favoráveis se encontram na compra, na absorção de mão de obra, no aumento da produtividade e do Produto Interno Bruto (PIB) *per capita* da região menos desenvolvida. Já entre os efeitos desfavoráveis, pode-se destacar a redução da competitividade das atividades manufatureiras e de exportação perante a região mais desenvolvida. Para que o efeito positivo prevaleça e desenvolva o conjunto das regiões, deve-se aumentar a complementariedade entre elas.

Inicialmente, o processo de desenvolvimento se dá em algumas indústrias e regiões, mas em um segundo momento o desenvolvimento tende a se disseminar por meio do encadeamento e da complementariedade setorial e regional, que auxilia para descobrir os setores e as atividades que compõem a dinâmica econômica e social, dentre esses, os que teriam maior especialidade de forma a gerar aglomerados. Assim, a capacidade de investir é um mecanismo de transmissão que atua como indutor do crescimento estimulando o desenvolvimento, o qual depende dos setores mais modernos da economia e também do empreendedorismo local. No caso de regiões subdesenvolvidas, a capacidade de investir em setores modernos é menor por exigir maior quantidade de recursos, por isso é necessário estimular atividades rotineiras, em que o sucesso não depende da inovação (HIRSCHMAN, 1958).

Frente a isto, as indústrias desempenham o papel de propulsão, de indução e de transformação da estrutura da região, incentivando o surgimento de atividades complementares, em que o conjunto articulado dessas atividades impulsiona o crescimento de toda a estrutura econômica, principalmente para os centros urbanos, gerando aglomerações, as quais exercem forças centrípetas relacionadas à compra e venda de produtos e serviços e à propagação da inovação e do progresso tecnológico com todo o grupo. Com isso, ocorre uma atração de capital e de mão de obra na busca de emprego e acúmulo de capital (HIRSCHMAN, 1958).

Por meio das teorias regionais abordadas até o momento, nota-se a importância da decisão das indústrias em escolher um local para se instalar, dado que essa decisão pode influenciar no sucesso da firma, na expansão da produção, na redução dos custos e no aumento dos lucros. Além disso, pode impactar na dinamicidade da região como um todo, mediante o aumento do investimento local



advindo de outras empresas, na geração de emprego, na disponibilidade de bens e serviços, entre outros fatores.

No entanto, cabe destacar que o impacto no desenvolvimento da região pode ser diferente dependendo da intensidade tecnológica da firma que irá se instalar. Como destacado na TSMT, o mercado de trabalho está dividido entre o segmento tradicional, que exige maior quantidade de mão de obra com um preço mais baixo, e o segmento moderno, que usufrui de mão de obra qualificada com salários mais elevados. Dessa forma, os reflexos para as instituições e para com a sociedade em geral podem ser diferentes para cada segmento industrial.

### *2.3 Teorias da nova geografia econômica*

Para Fujita et al. (1999), os microfundamentos dos modelos clássicos das teorias tradicionais da economia regional não foram consistentes e suficientes para explicar como os agentes econômicos se organizam no espaço, abrindo brecha para o surgimento de novas teorias. Para atender a necessidade de suprir esta lacuna, surgiu a NGE. A NGE herda e incorpora conceitos das teorias da localização e das teorias de desenvolvimento regional e reinsere na dimensão ambiental com ênfase na economia de aglomeração.

Com vistas na desigualdade da distribuição espacial, tanto das empresas como das pessoas, e na sua persistência ao longo do tempo, o objetivo da NGE é compreender, por meio de forças econômicas opostas, a aglomeração (centrípeta) e a dispersão (centrífuga), a distribuição no espaço da atividade humana. No caso da força centrífuga, se uma empresa se instalar em uma região que tenha alta concorrência setorial, essa tende a reduzir seus preços e assim os salários, incentivando os trabalhadores a se deslocarem para locais cuja remuneração seja mais elevada. Entretanto, a redução do preço dos produtos e do custo de transporte eleva o salário real da população, fazendo com que aumente a migração para essa região, isso fará com que aumente o mercado consumidor, estimulando a instalação de novas empresas nesse local, caracterizando a força centrípeta. Por isso, para a NGE, a distribuição da atividade econômica depende do resultado dessas forças e, quando ocorre a desigualdade regional, significa que as forças centrípetas estão se sobressaindo às centrífugas (FUJITA et al., 1999).

Além disso, para a NGE, a forma de compreender o processo de concentração de atividades é por meio das externalidades, em que a análise dos retornos crescentes de escala contribui para o entendimento das aglomerações nas regiões (CUNHA, 2008). Os retornos crescentes estão diretamente ligados com a concentração espacial de atividades econômicas, vinculados com o processo de força, chamado por Fujita et al. (1999) de “conexões para trás e para frente associadas aos mercados locais”.

O que se percebe é que a aglomeração industrial se torna um espaço privilegiado, pois tem maior capacidade de gerar externalidades positivas devido à interação entre as firmas, a qual possibilita ampliação das economias de escala, especialização, transbordamento de conhecimento, encadeamento a jusante e a montante do processo, acesso à tecnologia, menor custo de transporte, presença de serviços complementares, disponibilidade de infraestrutura básica, entre outros fatores que seriam minimizados se a indústria estivesse longe das demais.

No entanto, os motivos que incentivam as empresas a se aglomerarem podem se diferenciar conforme os fatores de produção utilizados, ou seja, se a indústria é intensiva em mão de obra ou em tecnologia. Desta forma, cabe nesta pesquisa analisar espacialmente as causas da localização da IT paranaense, para cada segmento. Para isso, foram incluídas como variáveis explicativas da aglomeração industrial os fatores remuneração, custo de transporte e competitividade, extraídas com base nas teorias abordadas até o momento. Ademais, a próxima seção visa abordar algumas contribuições empíricas a respeito desse tema, além de auxiliar na escolha de outras variáveis para explicar a distribuição industrial no território paranaense.

### 3. Síntese de abordagens empíricas sobre a aglomeração industrial

À vista de aprofundar os temas decorridos até o momento, esta seção tem como objetivo reunir estudos da literatura nacional e internacional sobre a concentração industrial.

No nível nacional, Rocha et al. (2013) tiveram como objetivo investigar os fatores responsáveis para explicar a concentração da IT brasileira para os anos de 1996 a 2006. Para isto, utilizaram o modelo La Fountain e as variáveis aluguel médio, crime, densidade demográfica, escolaridade, produtividade, participação da indústria no emprego local e Índice Herfindahl. Os resultados evidenciaram que os setores Alimentícios e Bebidas, Confecção e Metalurgia apresentaram vantagens naturais, enquanto a indústria de Produtos Químicos é influenciada pelas externalidades produtivas e as indústrias de Máquinas e Equipamentos e Veículos Automotores contam com vantagens no acesso ao mercado.

Para os anos de 2006 e 2013, Almeida et al. (2017) averiguaram a coaglomeração e os fatores de aglomeração da IT brasileira, por meio do Índice de coaglomeração e da regressão em painel, com os dados do número de trabalhadores por localidade e setores, informações referentes ao perfil dos trabalhadores (ocupação e escolaridade) e matriz insumo produto. Concluíram que os setores se coaglomeram de forma aleatória, porém em diversos setores as forças *marshallianas* têm papel fundamental nesse padrão.

Com o intuito de investigar os fatores que influenciaram a concentração industrial nos estados brasileiros para os anos 2003 a 2014, Santos (2017) utilizou como metodologia o Índice de Concentração Ellison e Glaeser (EG) e vetores autorregressivos em painel. As variáveis utilizadas foram: participação do governo, transbordamento de conhecimento, parcela das empresas por estado, competitividade entre firmas e custo de transporte. Os resultados evidenciaram que a influência do Governo e as externalidades possuem maior significância na concentração. Além disso, choques relacionados ao transbordamento de conhecimento impactam positivamente e arrecadação do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) negativamente na concentração industrial.

Em termos estaduais e regionais, Souza e Muniz (2010) analisaram os fatores determinantes da localização das indústrias nos municípios de Goiás. Os autores fizeram uso do Método dos Mínimos Quadrados Ordinários (MQO) com as variáveis: número de indústrias, distância da capital, população urbana, PIB *per capita*, agências bancárias e postos bancários. Concluíram que a distância da capital, o mercado consumidor, a existência de infraestrutura e a facilidade de adquirir crédito são fatores que explicam grande parte da decisão locacional das indústrias goianas nos anos de 2009/2010.

Com a finalidade de identificar e analisar espacialmente a evolução da aglomeração do setor de confecções na região sul do Brasil, Rodrigues et al. (2012) fizeram uma comparação entre os anos de 1995, 1999, 2003 e 2007, utilizando como métodos o ICN e a Análise Exploratória de Dados Espaciais (AEDE), com número de empregos e de estabelecimentos. Evidenciaram que houve uma evolução dos aglomerados e do transbordamento espacial beneficiando as regiões vizinhas.

Bonini (2014) também buscou identificar os determinantes da aglomeração industrial, porém para o estado de São Paulo. O autor utilizou a regressão linear de corte transversal e as variáveis: PIB Industrial, população residente, população com nível superior, número de IT, distância da capital, número de empregos na IT, fluxo de comércio, PIB *per capita*, especialização da região e salário médio da IT. Os resultados evidenciaram que o número de estabelecimentos da IT e a população com nível superior influenciam positivamente e a distância da capital afeta negativamente a aglomeração. Em termos municipais, concluiu que um simples aglomerado de indústrias é suficiente para atrair novas empresas. E no nível microrregional a mão de obra qualificada influencia na localização industrial.

Rosa e Raiher (2017) analisaram espacialmente as aglomerações produtivas por intensidade tecnológica e o desenvolvimento econômico dos municípios paranaenses. Para isso, utilizaram o Índice de Desenvolvimento Municipal, o ICN, a AEDE e a econometria espacial. Como variáveis, a taxa de crescimento do desenvolvimento econômico, o índice de desenvolvimento municipal, a taxa

de crescimento da indústria de alta, média alta, média baixa e baixa tecnologia. As autoras evidenciaram que as aglomerações têm efeito positivo no desenvolvimento econômico paranaense, principalmente na indústria de baixa tecnologia. Entretanto, a indústria de alta tecnologia não foi estatisticamente significativa para impulsionar o desenvolvimento.

Os trabalhos supracitados tiveram como principais fontes de dados o MTE, o Instituto Brasileiro de Geografia e Estatística (IBGE), o Atlas de Desenvolvimento Humano no Brasil (ATLAS), o Banco Central do Brasil (BACEN), o IpeaData, o Ministério do Desenvolvimento, Indústria e Comércio (MDIC), o AliceWeb e algumas fontes estaduais, como o portal das cidades paulistas e o Instituto Paranaense de Desenvolvimento Econômico e Social (IPARDES).

Em nível internacional, Gonzalez (2013) verificou os determinantes da aglomeração industrial no México. A metodologia utilizada consistiu em uma estimação de dados em painel com as variáveis: produtividade, dotação de fatores (remuneração/número de indústrias), economia de escala (tamanho médio da indústria/número de empresas) e intensidade de bens intermediários (produção total da indústria/valor adicionado da indústria). Os resultados indicaram que as economias de escala e a intensidade da utilização de insumos intermediários são os principais fatores que explicam a concentração industrial absoluta no México.

Também analisando a aglomeração do México, em especial, a fronteira Norte do país, Cruz et al. (2014) aplicaram questionários e realizaram uma análise descritiva. Os resultados indicaram que a concentração industrial se dá devido à localização geográfica dos municípios próximos aos EUA, local de onde vem os insumos e para o qual é enviada a produção, com redução do tempo e do custo de transporte. Além disso, há disponibilidade de mão de obra não qualificada e barata e de infraestrutura urbana. No entanto, não há evidências de fatores de externalidades pecuniárias e tecnológicas que expliquem a aglomeração de indústrias de alta tecnologia, devido à não existência de cadeia produtiva, a falta de transbordamento de conhecimento e a não qualificação da mão de obra.

Barchet (2016) analisou comparativamente as aglomerações industriais entre a região Sul do Brasil e a Província de Québec/CA. Os métodos utilizados pela autora foram: dados em painel, Quociente Locacional (QL), Coeficiente de assimetria de Pearson e Coeficiente de Concentração de Gini. Os dados para o Brasil foram: pessoas ocupadas na educação e nas atividades financeiras, população urbana e consumo de energia elétrica; já para Québec: o número de pessoas ocupadas no setor de transporte e das atividades financeiras, número populacional, número de cooperativas e sindicatos e escolaridade média. A autora verificou que o capital social, físico, financeiro e o potencial de mercado são fatores determinantes para a formação de aglomerados produtivos industriais, de forma que a elevação do capital humano e financeiro pode inibir a concentração, principalmente em espaços com dinamismo socioeconômico mais estável, como é o caso de Québec. Santa Catarina teve 56 aglomerações produtivas em 2013, aumentou 30% desde 2006. O Rio Grande do Sul apresentou 65 agrupamentos em 16 setores e no Paraná aumento de 20%. Em Québec, 6 agrupamentos em 4 subsetores para 2012, representando redução de 33% de 2005-2012.

Como se pode verificar, existem diversos estudos que envolvem o tema de aglomerações produtivas, uma vez que o entendimento dos fatores que atraem as indústrias para um determinado local tende a beneficiar a região por meio do seu desenvolvimento. No entanto, nota-se que poucos consideram o fator espacial como forma de explicação das aglomerações. Além disso, tais estudos, em sua grande maioria, estão voltados para os níveis estaduais ou para países como um todo, enquanto que as pesquisas a nível municipal ainda são minoria. Contudo, são importantes para o acompanhamento mais detalhado de cada município, assim podendo auxiliar na elaboração de políticas locais de acordo com a necessidade de cada região.

Este estudo contribui no sentido de conciliar a segmentação do mercado de trabalho em nível tecnológico com a concentração industrial dos municípios paranaenses, a fim de identificar os fatores de aglomeração de cada segmento em termos espaciais. Para dar sequência, na seção a seguir, são apresentados os métodos utilizados neste estudo para atender a este propósito.

## 4. Procedimentos metodológicos

### 4.1 Descrição das variáveis

Os setores da IT foram obtidos com base nas informações determinadas pelo IBGE, segundo a Comissão Nacional de Classificação (CONCLA, 2019), a partir da divisão da Classificação Nacional de Atividades Econômicas 2.0 (CNAE 2.0), seção C, que representa um total de 24 setores. Esses foram separados em dois segmentos de acordo com a intensidade tecnológica de cada ramo: o segmento primário, conhecido como moderno, que abrange os setores que utilizam alta e média alta tecnologia; e o secundário representado pelo setor tradicional, composto pelos segmentos com baixa e média baixa intensidade tecnológica. Para realizar tal divisão, utilizou-se a classificação determinada pelo IBGE por meio da Pesquisa Industrial Anual (PIA), elencados no Quadro 1.

**Quadro 1 - Classificação dos setores da indústria de transformação por intensidade tecnológica**

Setor – CNAE 2.0	Intensidade Tecnológica	Segmento
10-Fabricação produtos alimentícios	Baixa	Tradicional
11-Fabricação de bebidas	Baixa	Tradicional
12-Fabricação de produtos do fumo	Média alta	Moderno
13-Fabricação de produtos têxteis	Baixa	Tradicional
14-Confeção de artigos do vestuário e acessórios	Baixa	Tradicional
15-Preparação e fabricação de artefatos de couro, artigos p/ viagem e calçados	Média baixa	Tradicional
16-Fabricação de produtos de madeira	Baixa	Tradicional
17-Fabricação de celulose, papel e produtos de papel	Média alta	Moderno
18-Impressão e reprodução de gravações	Baixa	Tradicional
19-Fabricação de coque, de produtos derivados do petróleo e de biocombustível	Baixa	Tradicional
20-Fabricação de produtos químicos	Média alta	Moderno
21-Fabricação de produtos farmoquímicos e farmacêuticos	Média alta	Moderno
22-Fabricação de produtos de borracha e de material plástico	Média baixa	Tradicional
23-Fabricação de produtos de minerais não metálicos	Média baixa	Tradicional
24-Metalurgia	Média baixa	Tradicional
25-Fabricação de produtos de metal, exceto máquinas e equipamentos	Média baixa	Tradicional
26-Fabricação de equipamentos de informática, produtos eletrônicos e ópticos	Alta	Moderno
27-Fabricação de máquinas, aparelhos e materiais elétricos	Alta	Moderno
28-Fabricação de máquinas e equipamentos	Alta	Moderno
29-Fabricação de veículos automotores, reboques e carrocerias	Alta	Moderno
30-Fabricação de outros equipamentos de transporte, exceto veículos automotores	Alta	Moderno
31-Fabricação de móveis	Baixa	Tradicional
32-Fabricação de produtos diversos	Média alta	Moderno
33-Manutenção, reparação e instalação de máquinas e equipamentos	Alta	Moderno

Fonte: Elaborado pelos autores com base no IBGE-CONCLA (2019).

Conforme a classificação, dos 24 setores da IT, 13 são do segmento tradicional e 11 do moderno. A IT foi escolhida por permitir essa divisão entre os segmentos, além de ser um dos

setores com menor possibilidade de mão de obra informal no mercado de trabalho e possuir a maior participação do Valor Adicionado Fiscal (VAF), com 33% conforme IPARDES (2019).

Cabe destacar que, apesar dos segmentos serem compostos pela soma de diversos setores, não estão presentes em todos os municípios. Dos 399 municípios paranaenses, o segmento tradicional está presente em 389. O segmento moderno, por exigir maior volume de recursos e mão de obra qualificada, está presente em 280 municípios. Dentre os dois segmentos, o tradicional é o maior empregador, com 72,91% da mão de obra da IT e o moderno com 27,08% (BRASIL, 2019a).

No segmento tradicional, o maior empregador é a indústria de Fabricação de produtos alimentícios, que possuía um total de 4.153 estabelecimentos e 186.650 empregados no ano de 2017, esse número é mais expressivo do que a soma dos empregos dos 11 setores do segmento moderno. Já o ramo que mais emprega na indústria alimentícia é o abate de suínos, aves e outros pequenos animais com 92.000 empregos representando 49,29% do setor 10 (BRASIL, 2019a).

No segmento moderno, o setor que mais emprega é Fabricação de veículos automotores, reboques e carrocerias e Fabricação de máquinas e equipamentos, consecutivamente. De acordo com o IPARDES (2007), o mercado automobilístico cresceu a uma taxa acumulada de 179,3% de 1999 a 2007, fato que é fruto da melhoria nas condições de crédito, com taxas reduzidas, aumento do prazo de financiamento e da recuperação do emprego e renda. Já o aquecimento do segundo setor está atrelado com a expansão do segmento agrícola do Brasil.

Tais informações confirmam a importância da IT na geração de empregos e para a economia do estado como um todo, já que segundo Brasil (2019a) este setor é responsável por empregar formalmente mais de 20% dos trabalhadores do estado. Não obstante, essa importância pode ser intensificada quando as indústrias fazem parte de aglomerados, fortalecendo a estrutura produtiva e aumentando ainda mais os benefícios para as empresas e para a população.

#### 4.2 Índice de Concentração Normalizado (ICN)

Para mensurar a concentração industrial, utilizou-se o ICN por ser o mais comum na literatura selecionada, também por ser um método completo, composto pelo conjunto de índices que indicam a concentração no espaço, são eles: QL, Hirschman Herfindal (HH) e Participação Relativa (PR). O  $QL_{ij}$  representa se o município ( $j$ ) possui especialização em determinado setor ( $i$ ). O  $HH_{ij}$  avalia o grau de concentração do setor ( $i$ ) no município ( $j$ ). O  $PR_{ij}$  capta a importância do setor ( $i$ ) no município ( $j$ ). De acordo com Crocco et al. (2006), os referidos índices podem ser calculados da seguinte forma:

$$QL_{ij} = \frac{E_j^i / E_j}{E_{PR}^i / E_{PR}} \quad HH_{ij} = \left( \frac{E_j^i}{E_j} \right) - \left( \frac{E_{PR}^i}{E_{PR}} \right) \quad PR_{ij} = \frac{E_j^i}{E_{PR}^i} \quad (1)$$

em que,  $E_j^i$  trata-se do emprego do setor ( $i$ ) no município ( $j$ );  $E_j$  é o total de emprego no município ( $j$ );  $E_{PR}^i$  é o total de emprego no setor ( $i$ ) no estado do Paraná (PR); e  $E_{PR}$  é o total de empregos no Paraná (PR).

Conforme Crocco et al. (2006), o ICN permite obter as seguintes características de aglomeração: 1) a especificidade de um setor em uma região; 2) sua importância sobre a estrutura industrial da região; 3) a relevância do setor em nível nacional; 4) sua escala na estrutura produtiva industrial local. Para atender essas características, o ICN é composto pela soma dos índices  $(QL_{ij}, HH_{ij} \text{ e } PR_{ij})^1$  todos normalizados ( $n$ ), conforme Equação 2. A normalização dos índices é

<sup>1</sup> Para mais detalhes sobre o cálculo, ver Crocco et al. (2006).

necessária, pois cada um apresenta diferentes unidades de medidas, para isso, foi utilizado o método de máximo e mínimo, o qual irá normalizar os valores do IC entre zero e um.

$$ICn_{ij} = \theta_1 QLn_{ij} + \theta_2 HHn_{ij} + \theta_3 PRn_{ij} \quad (2)$$

em que,  $\theta^2$  são os pesos atribuídos a cada indicador, de forma que a soma deve ser igual a um ( $\theta_1 + \theta_2 + \theta_3 = 1$ ).

Para calcular o ICN, utilizaram-se os números de empregos formais da IT, extraídos do MTE por meio da Relação Anual de Informações Sociais (RAIS) para os 399 municípios paranaenses, por retratar a situação do mercado de trabalho, o desempenho da indústria e a renda da economia. Em seguida, o ICN foi dividido pelos segmentos, obtendo-se um índice para o segmento tradicional (ICN-T) e outro para o moderno (ICN-M).

Cabe destacar que Rodrigues et al. (2012) e Rosa e Raiher (2017) também estimaram o ICN por meio dos dados do emprego e para identificar se os municípios são aglomerados ou não, seja do segmento moderno ou tradicional, o critério utilizado foi a média geral do ICN em cada segmento. Se o município estiver com o valor do ICN abaixo da média geral, o segmento não está concentrado naquele município, mas, se o ICN estiver acima da média geral, o município concentra os setores que compõem o segmento analisado.

Quanto às variáveis explicativas, essas foram selecionadas com base nas teorias e na revisão de literatura supramencionadas. A competitividade é um fator que compromete o desempenho da indústria. Essa variável teve como base a TSMT, a NGE e os estudos de Santos (2017) e Rosenthal e Strange (2003). Os autores mensuraram o nível de competitividade industrial dividindo a quantidade de estabelecimentos pela quantidade de empregados. Nesta pesquisa, realizou-se da mesma maneira, mas foram utilizados dados para cada município e segmento.

A produtividade consta na TSMT como um dos fatores que difere o segmento moderno e o tradicional. Rocha, Bezerra e Mesquita (2013) calculam dividindo o Valor Adicionado Bruto (VAB) pelo pessoal ocupado. Como o VAB não é calculado para nível municipal, utilizou-se como *proxy* o Valor Adicionado Fiscal (VAF) da IT, o que não compromete o resultado da pesquisa, pois as variáveis captam o mesmo efeito. Já o pessoal ocupado foi separado por segmento. Outra variável bastante discutida nas teorias regionais e na NGE é o custo de transporte, para seu cálculo, utilizou-se a distância de cada município até a capital conforme Souza e Muniz (2010) e Bonini (2014). A disponibilidade de infraestrutura também é um fator bastante relatado pelas teorias regionais e pela NGE, para mensurá-la, procedeu-se conforme Barchet (2016), através do consumo de energia elétrica. Nesta pesquisa, considerou-se o consumo de energia elétrica *per capita*.

Segundo as teorias regionais e a NGE, parte-se do pressuposto que quanto maior o território utilizado para a produção agrícola, menor será a concentração industrial. Por isso, para representar os municípios que usam de forma mais intensiva os setores primários, considerou-se a área plantada. A cobrança de impostos e a participação do governo podem ser outros fatores que influenciam na concentração industrial. Assim, utilizou-se o ICMS conforme Santos (2017), mas em termos *per capita*. O autor utiliza o ICMS sobre o VAB industrial.

Para verificar o desenvolvimento local municipal, Rosa e Raiher (2017) fizeram o cálculo do índice de desenvolvimento do município. No presente estudo, utilizou-se o Índice Firjan de Desenvolvimento Municipal (IFDM) (FIRJAN, 2019), que também considera em seu cálculo informações sobre educação, saúde e emprego. Conforme a TSMT, um fator que diferencia o segmento moderno e o tradicional é a remuneração. Tal como Bonini (2014), considerou-se o salário médio da IT discriminado por segmentos.

Para verificar a disponibilidade de mão de obra, Souza e Muniz (2010) e Barchet utilizaram a quantidade de população urbana, já Rocha, Bezerra e Mesquita (2013), a densidade demográfica e Bonini (2014), a população residente. Neste estudo, utilizou-se a taxa de crescimento geométrico

<sup>2</sup> Para a sua obtenção, foi utilizado o método multivariado Análise de Componentes Principais (ACP) e calculado no *software Statistical Package for the Social Sciences (SPSS)*.

populacional, pois parte-se do pressuposto de que as indústrias também podem se localizar na área rural, é o caso dos laticínios, por exemplo. Além disto, residentes nessas áreas também podem trabalhar nas indústrias na área urbana.

O Quadro 2 apresenta um breve resumo das variáveis utilizadas na abordagem empírica, das siglas, da descrição, do ano de obtenção, bem como da fonte dos dados. Nota-se que, para ambos os modelos, foram consideradas as variáveis: custo de transporte, consumo de energia elétrica, área plantada, ICMS, IFDM e crescimento populacional, já competitividade, produtividade e remuneração foram discriminadas por segmento.

**Quadro 2 - Resumo das variáveis utilizadas no modelo empírico**

Variáveis	Sigla	Descrição	Ano	Fonte
Variável Dependente				
Índice de Concentração Normalizado Tradicional	ICN-T	Concentração de cada segmento	2017	MTE
Índice de Concentração Normalizado Moderno	ICN-M	Concentração de cada segmento	2017	MTE
Variáveis Explicativas				
Competitividade Tradicional	COMPET-T	Nº de estabelecimentos do segmento tradicional/Nº de empregados do mesmo segmento	2015	MTE
Competitividade Moderno	COMPET-M	Nº de estabelecimentos do segmento moderno/Nº de empregados do mesmo segmento	2015	MTE
Produtividade Tradicional	PRODUT-T	Valor Adicional Fiscal da IT/Número de trabalhadores do segmento tradicional	2015	IPARDES e MTE
Produtividade Moderno	PRODUT-M	Valor Adicional Fiscal da IT/Número de trabalhadores do segmento moderno	2015	IPARDES e MTE
Custo de Transporte	CT	Distância do município à capital	2015	IPARDES
Consumo de Energia Elétrica do Setor Industrial	CONS-ENERG	Consumo de Energia Elétrica/Nº população	2015	IPARDES
Área plantada	AREA-PLAN	Produção agrícola	2015	IPARDES
Imposto Sobre Circulação de Mercadorias e Serviços	ICMS	ICMS <i>per capita</i>	2015	IPARDES
Índice Firjan de Desenvolvimento Municipal	IFDM	IFDM (Educação + saúde + emprego)	2015	FIRJAN
Remuneração Tradicional	RENDA-T	Remuneração média do segmento tradicional	2015	MTE
Remuneração Moderno	RENDA-M	Remuneração média do segmento moderno	2015	MTE
Crescimento Populacional	CRESC-POPUL	Taxa de Crescimento Geométrico Populacional	2010	IPARDES

Fonte: Elaborado pelos autores

Como citado até o momento, para analisar a concentração industrial, as variáveis dependentes, ICN-T e ICN-M, tiveram como base o ano de 2017, o qual foi escolhido por ser o ano mais recente com disponibilidade de dados para os 12 meses do ano. Para minimizar o problema de

endogeneidade<sup>3</sup>, considerou-se a defasagem no tempo das variáveis explicativas. Assim, buscou-se extrair para o ano de 2015 ou para o último ano disponível.

#### 4.3 Econometria espacial<sup>4</sup>

Para identificar os fatores que explicam a aglomeração, utilizou-se a regressão espacial, pois, além de identificar as variáveis que influenciam as empresas a se aglomerarem no espaço, é um método pouco explorado nas teorias de aglomeração. Conforme Almeida (2012), a regressão espacial é uma extensão da econometria clássica, a qual deve ser incorporada na estrutura espacial, devido à dependência entre as observações afetar na explicação.

Entretanto, fez-se necessário seguir algumas etapas para averiguar qual modelo é mais adequado. De acordo com Baumont (2004) e Almeida (2012), inicialmente, deve-se estimar uma regressão não espacial por MQO. Em seguida, testa-se a autocorrelação espacial nos resíduos via *I* de Moran. O teste é realizado com distintas matrizes de pesos espaciais (*W*)<sup>5</sup>. De acordo com Baumont (2004), a matriz mais adequada é aquela cujo *I* de Moran apresenta o maior valor com significância estatística. Segundo Almeida (2012), a escolha da matriz é um processo muito importante, dado que a escolha inadequada da matriz pode acarretar resultados viesados e inconsistentes. Caso não haja autocorrelação espacial, o modelo adequado é o não espacial, mas, se houver, devem ser estimados os modelos espaciais: *Spatial Auto Regressive* (SAR), *Spatial Error Model* (SEM), *Spatial Durbin Model* (SDM) *Spatial Durbin Error Model* (SDEM). Dentre os modelos espaciais estimados, o mais adequado é aquele que não apresentar autocorrelação espacial nos resíduos e obtiver o menor critério de informação<sup>6</sup>.

Estimaram-se os modelos espaciais SAR, SEM, SDM, SDEM para cada segmento. Para o segmento tradicional, considerou-se como variável dependente o ICN-T e como variáveis explicativas: COMPET\_T, PRODUT\_T, CT, CONS\_ENERG, AREA\_PLAN, ICMS, IFDM, RENDA\_T e CRESC\_POPUL. Para o segmento moderno, considerou-se como variável dependente o ICN-M e como variáveis explicativas: COMPET\_M, PRODUT\_M, CT, CONS\_ENERG, AREA\_PLAN, ICMS, IFDM, RENDA\_M e CRESC\_POPUL. As variáveis explicativas foram resumidas na matriz X conforme o segmento.

O modelo SAR indica que uma alteração na variável dependente em algum município afeta diretamente o próprio município e indiretamente os valores das variáveis independentes dos municípios vizinhos (ALMEIDA, 2012). Conforme o autor, o modelo pode ser expresso por:

$$ICN_j = \alpha + \beta X'_j + \rho WICN_j + \varepsilon_j \quad (3)$$

em que *ICN* é a variável dependente;  $\beta$  é o vetor de parâmetros das variáveis explicativas;  $\rho$  é o coeficiente autorregressivo espacial; *WICN<sub>j</sub>* é matriz de defasagem espacial da variável dependente; e  $\varepsilon$  é o termo erro ( $\varepsilon \sim N(0, \sigma^2 I_n)$ ).

Diferentemente do SAR, no modelo SEM, a dependência espacial é por meio do termo de erro e não da variável dependente. Segundo Almeida (2012), no SEM, os erros associados com alguma observação são uma média dos erros das regiões vizinhas mais um componente de erro aleatório. A formulação geral é dada conforme a Equação (4), em que  $\xi = \lambda W\xi + \varepsilon$ .

<sup>3</sup> Segundo Roberts e Whited (2013), endogeneidade ocorre quando a variável explicativa se correlaciona com o termo de erro, levando a estimadores inconsistentes, viesados e ineficientes dos parâmetros da regressão. No entanto, conforme Wooldridge (2010 apud SANTOS, 2016), a correção do problema de endogeneidade pode ser resolvida estimando o modelo por Máxima Verossimilhança (MV).

<sup>4</sup> As estimações dos modelos espaciais foram realizadas por MV.

<sup>5</sup> Os pesos espaciais da matriz *W* indicam o quão dependentes são os municípios. As matrizes podem ser, por exemplo, de contiguidade (rainha ou torre) e/ou de distância (*k*-vizinhos mais próximos). Conforme o modelo, a matriz pode ser utilizada para obter a variável dependente defasada espacialmente (*Wy*), as variáveis explicativas defasadas espacialmente (*WX*) ou o termo de erro defasado espacialmente (*Wε*). Para mais detalhes, ver Almeida (2012).

<sup>6</sup> Critério de Informação de Akaike (AIC) e/ou critério Bayesiano de Schwarz (BIC) que garantem a qualidade do modelo.



$$ICN_j = \alpha + \beta X'_j + \xi_j \quad (4)$$

Além dos modelos globais<sup>7</sup> SAR e SEM, que analisam a propagação no espaço, neste estudo estimou-se o transbordamento espacial das variáveis, também para ambos os segmentos. Para averiguar tal efeito, estimaram-se os modelos com efeitos locais, isto é, SDM e SDEM.

Segundo Almeida (2012), o SDM incorpora o efeito transbordamento mediante a defasagem das variáveis explicativas ( $WX$ ), o qual supõe o impacto de algum fator que justifique a inclusão das variáveis endógenas defasadas espacialmente. Conforme o autor, o SDM pode ser expresso conforme a Equação 5, em que  $\tau$  é o vetor de parâmetros espaciais.

$$ICN_j = \alpha + \beta X'_j + WX_j\tau + \varepsilon_j \quad (5)$$

Por sua vez, no SDEM, os componentes de alcance local se manifestam nas variáveis explicativas exógenas e, no termo de erro, segundo Almeida (2012), esse modelo é expresso de acordo com a Equação 6.

$$ICN_j = \alpha + \beta X'_j + WX_j\tau + \xi_j \quad (6)$$

Na próxima seção, encontram-se expostos os resultados dos modelos.

## 5. Resultados e discussões

No segmento tradicional, primeiramente, estimou-se o modelo não espacial por MQO. Testaram-se os resíduos do modelo para confirmar a existência ou não de autocorrelação espacial. O critério de Baumont (2004) apontou a presença de dependência espacial na matriz de 20 vizinhos mais próximos, indicando que há interação espacial entre os municípios paranaenses. Dessa forma, o MQO é inadequado para analisar a concentração industrial do segmento tradicional. Por isso, seguiu-se com a estimação dos modelos espaciais, os quais estão descritos na Tabela 1.

O modelo escolhido foi o SAR, pois o componente espacial apresentou significância estatística e o modelo os menores critérios de informação AIC e BIC. O modelo SAR permite fazer a decomposição dos efeitos, isto é, estimar os efeitos da concentração do segmento tradicional dentro do próprio município (direto), nos municípios vizinhos por meio dos transbordamentos espaciais (indireto) e do efeito total (direto + indireto), conforme Tabela 2.

Os resultados do efeito direto indicam que a aglomeração do segmento tradicional possui uma relação negativa dentro do próprio município com a competitividade, a produtividade e a área plantada e um impacto positivo no que diz respeito ao custo de transporte, consumo de energia elétrica, desenvolvimento e crescimento populacional.

A competitividade apresentou valor negativo tanto no município quanto nos vizinhos, indicando que para cada aumento da competitividade reduz a concentração do segmento tradicional em 0,27. Santos (2017) obteve resultado diferente ao analisar a concentração dos estados brasileiros, concluiu que a competitividade causa um aumento da concentração em 0,50. No entanto, entende-se que enquanto tiver mercado consumidor as indústrias continuam atuando nesse local, mas, a partir do momento em que a competitividade fica mais acirrada, as empresas tendem a se realocarem em outras regiões. Para Fujita, apesar da decisão de localização depender da própria empresa, o autor defende que a dispersão das atividades advém do aumento da competição. Krugman contribui afirmando que um dos fatores que influenciam na aglomeração é a competição local (FUJITA et al., 1999).

<sup>7</sup> A autocorrelação espacial global permite concluir se a(s) variável(eis) segue(m) padrões de associações espaciais por meio da visualização no mapa. No âmbito local, trata-se da análise individual de cada observação e evita o ocultamento de padrões de autocorrelação, além de possibilitar a visualização dos *clusters*, da significância e do *I* de Moran (ALMEIDA, 2012).

**Tabela 1 - Resultado das estimativas dos modelos espaciais do segmento tradicional**

Variáveis	MQO	SAR	SEM	SDM	SDEM
COMPET_T	-0.2854*	-0.2673*	-0.2635*	-0.2510*	-0.2554*
PRODUT_T	-0.0000*	-0.0000*	-0.0000*	-0.0000*	-0.0000*
CT	0.0002*	0.0001*	0.0003*	0.0004	0.0004
CONS_ENERG	0.0857*	0.0841*	0.0797*	0.0822*	0.0839*
AREA_PLAN	-0.0000*	-0.0000**	-0.0000**	-0.0000**	-0.0000**
ICMS	0.0000**	0.0000	0.0000	0.0000	0.0000
IFDM	0.0000*	0.0000**	0.0000*	0.0000	0.0000
RENDA_T	-0.0000	-0.0000	0.0000	0.0000	-0.0000
CRESC_POPUL	0.0334*	0.0288*	0.0348*	0.0351*	0.0354*
WCOMPET_T				-0.1202	-0.2008
WPRODUT				0.0000	-0.0000
WCT				-0.0004	-0.0003
WCONS_ENERG				0.0630***	0.0901**
WAREA_PLAN				-0.0000	-0.0000
WICMS				0.0000**	0.0000**
WIFDM				0.0000**	0.0000*
WRENDA_T				-0.0002*	-0.0003*
WCRESC_POPUL				-0.0108	-0.0038
$\rho$		0.3393*		0.2145**	
$\lambda$			0.3844*		0.2342**
R <sup>2</sup>	0.4092	0.4239	0.4066	0.4593	0.4599
AIC	-375.0565	-392.4794	-386.575	-397.1268	-397.6707
BIC	-335.1669	-348.6009	-342.697	-321.3365	-321.8804

Fonte: Elaborado pelos autores.

Notas: Os erros padrão foram obtidos pela correção de White. Siglas e definições das variáveis estão presentes no Quadro 02. \*Coeficiente significativo a 1%; \*\* a 5% e \*\*\* a 10%.

No que diz respeito à produtividade, essa impacta de forma negativa na concentração. Segundo a TSMT (REICH et al., 1973; LIMA, 1980; LEMOS, 1995), o segmento secundário apresenta baixa produtividade devido à menor qualificação e remuneração. Porém, conforme Marshall (1985), uma das vantagens da aglomeração industrial é a economia de escala, isto é, a expansão produtiva das indústrias, a qual gera vantagens competitivas para o complexo como um todo, que não seria possível se estiver atuando no mercado de forma isolada. Gonzalez (2013) evidenciou que um dos principais fatores para as indústrias se localizarem próximas umas das outras é o aumento da produtividade.

A área plantada, a qual representa a presença da agricultura nas regiões, obteve relação negativa com a concentração industrial do segmento tradicional. Segundo a teoria “Anéis de Von Thünen”, a utilização do solo se dá ao redor do mercado (CAVALCANTE, 2008; CUNHA, 2008; BARROS; PRATES, 2014).

As teorias da localização, da concentração e da NGE abordaram que a concentração industrial está diretamente vinculada com o custo de transporte, de forma que, quanto menor a distância dos grandes centros, maior a tendência de gerar aglomerados. Em outras palavras, o custo de transporte é menor quando está próximo da matéria-prima ou do consumidor, cujos lugares tendem a apresentar maior acesso aos serviços vinculados com o processo produtivo e com as vias de comercialização. O modelo identificou que o segmento tradicional não tem uma relação positiva

com a distância do município com a capital do estado. De acordo com a teoria “Triângulo locacional de Weber”, a indústria busca o máximo de produção com o mínimo de custo para auferir maiores lucros, por isso buscam se localizar próximos do mercado consumidor ou da matéria-prima (CAVALCANTE, 2008; CUNHA, 2008; BARROS; PRATES, 2014). Por isso, entende-se que o segmento tradicional, cujo setor principal é a agropecuária, isto é, a produção de alimentos, bebidas, entre outros, não está localizado na capital, mas sim no interior do estado, próximo da matéria-prima.

**Tabela 2 - Efeitos espaciais do segmento tradicional do modelo SAR**

Variáveis	Direto	Indireto	Total
COMPET_T	-0.2706*	-0.1340*	-0.4046*
PRODUT_T	-0.0000*	-0.0000**	-0.0000*
CT	0.0002*	0.0001*	0.0003*
CONS_ENERG	0.0852*	0.0422*	0.1274*
AREA_PLAN	-0.0000**	-0.0000**	-0.0000**
ICMS	0.0000	0.0000	0.0000
IFDM	0.0000**	0.0000**	0.0000**
RENDA_T	-0.0000	-0.0000	-0.0000
CRESC_POPUL	0.0292*	0.0144*	0.0436*

Fonte: Elaborado pelos autores.

Notas: \*Coeficiente significativo a 1%; \*\* a 5% e \*\*\* a 10%.

Segundo Bizinela e Furlanetti (2011), um dos fatores que leva uma região a crescer acima da média do estado, tornando-se um polo de desenvolvimento, é a disponibilidade de infraestrutura. Por isso, para estimar a disponibilidade de infraestrutura, a variável utilizada foi consumo de energia elétrica, a qual obteve sinal positivo e estatisticamente significativo, indicando um efeito positivo no próprio município e também nos vizinhos ao redor. Souza e Muniz (2010) e Cruz et al. (2014) também evidenciaram que a disponibilidade de infraestrutura adequada é fundamental para a aglomeração industrial.

Sob o prisma do desenvolvimento, verificou-se uma relação positiva com a concentração industrial do segmento tradicional, tanto para o município como para seus vizinhos. Segundo Marshall (1985), a aglomeração possibilita à população maior acesso a bens e serviços com menor custo e maior agilidade, dessa maneira, favorecendo a população e contribuindo para o seu desenvolvimento. Rosa e Raiher (2017) também evidenciaram o efeito de transbordamento espacial entre os aglomerados produtivos com o desenvolvimento, principalmente para as indústrias com baixo teor tecnológico e justificam que essas têm maior facilidade de se instalar, uma vez que possuem maiores vantagens comparativas do que as indústrias de produtos primários.

Como o segmento tradicional demanda maior número de mão de obra, o crescimento populacional influencia positivamente na concentração desse segmento, de forma que o impacto é maior na região da concentração e menor nos vizinhos, mas, mesmo assim, evidencia-se que a aglomeração não é benéfica apenas para a região, mas também para as regiões vizinhas mediante a absorção de mão de obra desocupada, contribuindo, assim, para a geração de renda nas regiões menos desenvolvidas.

No que tange ao segmento moderno, os passos para as estimações foram os mesmos. Por isso, calculou-se o MQO não espacial (Tabela 3). O diagnóstico do teste *I* de Moran indicou autocorrelação espacial na matriz rainha de ordem de sete vizinhos. Ao se tratar da escolha do melhor modelo, verificou-se que no segmento moderno nenhum coeficiente da dependência espacial foi estatisticamente significativo, indicando que a dependência espacial do ICN-M não tem efeito. Cabe destacar que esta pesquisa partiu do pressuposto de que o aglomerado de indústrias

intensivas em capital gera maior efeito *spillover* nas regiões ao seu redor, no entanto, não foi isso que os resultados demonstraram. Conforme dados do MTE, no ano de 2017, o segmento moderno obteve uma participação de apenas 27,08% da mão de obra da IT. Além disso, enquanto o segmento tradicional estava presente em 389 dos 399 municípios paranaenses, o moderno estava em apenas 280, ou seja, sua participação é pequena e ainda possui elevado encadeamento externo (BRASIL, 2019a).

**Tabela 3 - Resultado das estimativas dos modelos espaciais do segmento moderno**

Variáveis	MQO	SAR	SEM	SDM	SDEM
COMPET_T	-0.1047*	-0.1024*	-0.1027*	-0.0993*	-0.0991*
PRODUT_T	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
CT	-0.0001*	-0.0001*	-0.0001*	-0.0002**	-0.0002**
CONS_ENERG	0.0208	0.0199*	0.0206*	0.0193*	0.0192*
AREA_PLAN	-0.0000	-0.0000	-0.0000	-0.0000**	-0.0000**
ICMS	0.0000	0.0000	0.0000	0.0000	0.0000
IFDM	0.0000	0.0000**	0.0000**	0.0000**	0.0000**
RENDA_T	0.0001*	0.0001*	0.0001*	0.0001*	0.0001*
CRESC_POPUL	-0.0018	-0.0025	-0.0022	-0.0001	-0.0001
WCOMPET_T				-0.0922	-0.0934
WPRODUT				0.0000	0.0000
WCT				0.0000	0.0000
WCONS_ENERG				-0.0497	-0.0488
WAREA_PLAN				0.0000*	0.0000*
WICMS				-0.0000	-0.0000
WIFDM				0.0000	0.0000
WRENDA_T				-0.0000	-0.0000
WCRESC_POPUL				0.0024	0.0027
$\rho$		-0.0825		0.0380	
$\lambda$			0.0994		0.0125
R <sup>2</sup>	0.3672	0.3685	0.3667	0.3786	0.3785
AIC	-594.7061	-0.3271	-596.7707	-587.9460	-587.8952
BIC	-554.8165	0.1622	-560.8700	-520.1336	-520.0828

Fonte: Elaborado pelos autores.

Notas: \*Coeficiente significativo a 1%; \*\* a 5% e \*\*\* a 10%.

Rosa e Raiher (2017), apesar de se basearem no mesmo pressuposto, também identificaram transbordamento apenas nas indústrias com baixo teor tecnológico. No entanto, as autoras identificaram que os aglomerados das indústrias de alta tecnologia obtiveram elevação no número de aglomerados no período de 1998 a 2014, as quais são importantes na dinamização da atividade econômica e na geração de renda para o estado do Paraná.

Dada a insignificância estatística dos modelos espaciais, analisaram-se apenas os resultados do MQO. A competitividade, da mesma forma que no segmento tradicional, impacta na concentração do segmento moderno de forma negativa, no entanto de forma mais amena (0,10). No caso do segmento tradicional, por exemplo, um frigorífico instalado em uma região que utiliza grande parte da mão de obra, dificilmente outra empresa do mesmo ramo se instalará nesse local, em vista de que terá que concorrer pela mão de obra. No caso do segmento moderno, uma empresa de *software*, por exemplo, além de utilizar menor número de mão de obra, em muitos casos, essas vêm de outros locais para atender a demanda de qualificação profissional.

Para o segmento moderno, o custo de transporte obteve o sinal negativo de acordo com as teorias e os trabalhos de Souza e Muniz (2010) e Bonini (2014). Indicando que, quanto maior a distância da capital, menor é o número de indústrias dos setores intensivos em tecnologia, devido ao custo com a logística. Outros fatores que influenciam são a disponibilidade dos fatores produtivos e o escoamento da produção, os quais impactam nos custos dos bens e serviços.

Uma das principais diferenças entre o segmento moderno e o tradicional é o valor da mão de obra, sendo que o segmento moderno é melhor remunerado devido à exigência de trabalhadores capacitados para desempenhar suas funções (REICH et al., 1973). Essa diferença pode ser evidenciada nos modelos, uma vez que a concentração responde de forma positiva com a renda do segmento moderno e negativa no segmento tradicional, apesar deste último não ser estatisticamente significativo. Dessa forma, entende-se que as indústrias que compõem o segmento moderno tendem a se localizar onde há maior qualificação da mão de obra.

Em geral, são vários os fatores que contribuem para a formação de aglomerados industriais, além de serem diferentes de acordo com a intensidade tecnológica. Rocha, Bezerra e Mesquita (2013) também chegaram à conclusão de que não se pode generalizar os resultados para explicar as decisões locacionais, visto que chegaram a resultados diferentes para explicar a distribuição geográfica dos setores da IT. As indústrias de alimentos e bebidas, confecção e metalurgia são influenciados pelas vantagens naturais; a indústria de produtos químicos, pelas externalidades produtivas; e as indústrias de máquinas e equipamentos e veículos automotores, pelo acesso ao mercado. Por isso, as autoras defendem que a aglomeração varia conforme suas especificidades e por diferirem na intensidade de cada fator de produção.

Já Almeida, Rocha e Justo (2017) evidenciaram que a aglomeração para a indústria brasileira é influenciada pela disponibilidade de mão de obra e dos *inputs* do processo de produção. Barchet (2016), por sua vez, defende que o fator responsável pela aglomeração industrial é o capital social, físico e financeiro.

Apesar dos fatores serem diferentes em cada setor, podendo gerar efeitos distintos entre as regiões, Hirschman (1958) defende que é necessário desenvolver uma relação de complementariedade entre as regiões, a fim de fortalecer os efeitos positivos da aglomeração industrial também nos circunvizinhos. Para Fujita et al. (1999), uma vez que as vantagens econômicas se sobressaíam nos aglomerados industriais, as firmas se beneficiam pelo maior acesso ao mercado e à população devido à redução do custo de vida. Além disso, a tendência é a intensificação da aglomeração com o tempo, pois os custos do processo produtivo tendem a se difundir cada vez mais.

## 6. Considerações finais

Esta pesquisa teve como objetivo analisar espacialmente as causas da aglomeração da IT, em especial entre os segmentos moderno e tradicional, para os municípios do estado do Paraná no ano de 2017. Por meio da econometria espacial, no segmento tradicional, foi possível detectar relação espacial dentro do município e também nos seus vizinhos por meio do transbordamento entre as regiões. Por sua vez, o modelo do segmento moderno não evidenciou relação no espaço, dessa forma não podendo inferir que os fatores selecionados afetam espacialmente a concentração das indústrias intensivas em tecnologia no município, nem nos seus vizinhos na forma de transbordamento.

Os modelos possibilitaram identificar que os fatores que influenciam na geração de aglomerados do segmento tradicional no território paranaenses são: competitividade, produtividade, área plantada, custo de transporte, consumo de energia elétrica, desenvolvimento e crescimento populacional. Enquanto que no segmento moderno os fatores são: competitividade, custo de transporte e renda. Dessa maneira, é possível verificar que os fatores que determinam a localização industrial variam de acordo com a intensidade tecnológica. Evidenciou-se que apenas a competitividade é um fator comum entre os segmentos. Por sua vez, o custo de transporte impacta

de maneira diferenciada em cada segmento. Além disso, foi possível detectar que a produtividade, a área plantada, o consumo de energia elétrica, o IFDM e o crescimento populacional são significativos apenas para o segmento tradicional; e a renda apenas para o segmento moderno.

Dessa maneira, a primeira hipótese norteadora da pesquisa – de que a aglomeração industrial é causada por diferentes fatores, conforme as especificidades de cada município paranaense – não foi rejeitada. No entanto, a segunda hipótese – de que aglomerados produtivos com maior teor tecnológico, tendem a gerar maiores transbordamentos para os municípios – foi rejeitada. No entanto, o não atendimento desta última não compromete a contribuição do estudo, pelo contrário, contribui evidenciando que aglomerados de indústrias com alto teor tecnológico também podem contribuir para a região, mas ainda detém pequena participação no território paranaense.

Por isso, entende-se que os setores com maior teor tecnológico almejam crescer e no território paranaense tem espaço, visto que estão presentes em apenas 280 dos 399 municípios. Além disso, segundo IPARDES (2019), no terceiro trimestre de 2018, a taxa de desocupação está em 8,6%, apesar de estar abaixo da média nacional, representa que tem muita mão de obra necessitando trabalhar e, como visto, o segmento moderno é o principal gerador de renda e pode influenciar na redução da pobreza e no desenvolvimento regional.

Para tanto, a intervenção governamental, em todas as suas esferas (federal, estadual e municipal), no apoio à instalação de uma indústria com alto ou baixo teor tecnológico, influencia a região e as suas circunvizinhas, não apenas no curto prazo, mas também no seu futuro. Por isso, este estudo contribui com a análise dos fatores geradores de aglomerados produtivos e se diferencia das demais pesquisas por utilizar o fator espacial e analisar o mercado de trabalho da IT em segmentos, os quais são influenciados por diferentes fatores, contribuindo para auxiliar nas tomadas de decisões e na elaboração de políticas capazes de promover complementariedade entre os municípios, gerando externalidade positivas além da fronteira.

Ademais, esta pesquisa deixou de incorporar à análise algumas variáveis importantes, porém difíceis de mensurar, como: disponibilidade de recursos naturais; subsídios e incentivos do governo em determinadas indústrias; parcela de mercado das indústrias, pois empresas monopolistas e oligopolista podem impactar de maneira diferente do que na concorrência imperfeita; capacidade de empreender pelo fato de ter nascido na região e estar próximo da família; custo de vida; e mobilidade de mão de obra que pode ser mensurado pelo acréscimo no modelo da variável dependente defasada. A obtenção dessas informações é possível por meio de pesquisa de campo, por isso, fica como sugestões para pesquisas futuras o acréscimo desses fatores que também podem contribuir para explicar os motivos que levam as empresas a se aglomerarem espacialmente.

## Referências

- ALMEIDA, E. *Econometria espacial aplicada*. Campinas, SP: Editora Alínea, p. 498, 2012.
- ALMEIDA, J. V. P.; ROCHA, R. M.; JUSTO, W. R. Por que as Indústrias se Coaglomeram? Evidências para o Brasil. In: Congress Latin American and Caribbean Regional Science Association International. Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos, 15, 2017. *Anais...* São Paulo: ABER, 2017.
- AMARANTE, P. A.; SILVA, M. V. B. Economias de aglomeração nas atividades econômicas dos municípios brasileiros nos anos de 2000 e 2009: evidências a partir de equações salariais, *Revista Brasileira de Economia de Empresas*. v 16, n. 1, p. 25-51, 2016.
- ATLAS DO DESENVOLVIMENTO HUMANO DO BRASIL. IDH Paraná. Disponível em: [http://www.atlasbrasil.org.br/2013/pt/perfil\\_uf/parana/](http://www.atlasbrasil.org.br/2013/pt/perfil_uf/parana/). Acesso em: jan/2019.
- BARCHET, I. *Agglomerações industriais e polos econômicos regionais: uma análise comparativa entre a região Sul do Brasil e a Província de Québec/CA*. Tese (Doutorado em Desenvolvimento Regional e Agronegócio), Universidade Estadual do Oeste do Paraná. UNIOESTE. Toledo 2016.

- BARROS, F. R. P. M.; PRATES, T. M. A nova geografia econômica e a explicação do diferencial de produtividade entre cidades: estudo de caso sobre Maceió e Recife. *Registro Contábil – UFA I*, v. 5, n. 1, p. 128-147, 2014.
- BAUMONT, C. *Spatial effects in housing price models: do house prices capitalize urban development policies in the agglomeration of Dijon (1999)?* Laboratoire d'économie et de gestion. 2004. (Research Report, n. 5118).
- BIZINELA, A.; FURLANETTI, G. I. O anel de interligação do Paraná no contexto dos polos de crescimento e da economia local. In: Seminário do Centro de Ciências Sociais Aplicadas de Cascavel, 10, 2011. *Anais...* Cascavel: CCSAC. 2011.
- BONINI, A. L. S. F. *Determinantes da aglomeração espacial da atividade industrial no estado de São Paulo*. Dissertação (Mestrado profissional em Economia da Mundialização e do Desenvolvimento), Pontifícia Universidade Católica de São Paulo. São Paulo, 2014.
- BRASIL. Ministério do Trabalho e Emprego. *RAIS: Relação Anual de Informações Sociais*. Brasília, DF, 2019a. Disponível em: <http://bi.mte.gov.br/bgproger/login.php>. Acesso em: mar/2019.
- CAVALCANTE, L. R. M. T. Produção teórica em economia regional: uma proposta de sistematização. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 2, n. 1, p. 9-32, 2008.
- CONCLA. Comissão Nacional de Classificação. Pesquisa CNAE 2.0. Rio de Janeiro. Disponível em: <http://www.cnae.ibge.gov.br>. Acesso em: 01 mai. 2019.
- CROCCO, M. A.; GALINARI, R.; SANTOS, F.; LEMOS, M. B.; SIMÕES, R. Metodologia de identificação de aglomerações produtivas locais. *Nova Economia*, v. 16, n. 2, p. 211-241, 2006.
- CRUZ, J. A. P.; ÁLVAREZ, G. I. C.; CALDERÓN, A. R. C. Los factores que explican la mayor aglomeración de la industria de alta tecnología en la frontera norte de México: el caso de Matamoros y Reynosa. *Estudios Fronterizos*, v. 15, n. 29, p. 173-206, 2014.
- CUNHA, J. C. *Nova geografia econômica: um ensaio para o Brasil*. Dissertação (Economia Aplicada), Faculdade de Ciências Econômicas da Universidade Federal do Rio Grande do Sul. Porto Alegre, 2008.
- FIRJAN. Índice Firjan de Desenvolvimento Municipal. Federação das Indústrias do Rio de Janeiro. Disponível em: <http://www.firjan.com.br/ifdm>. Acesso em 01.fev. 2019.
- FUJITA, M.; KRUGMAN, P.; VENABLES A. *The spatial economy: cities, regions and international trade*. Cambridge: MIT Press, 367 p. 1999.
- GONZALEZ, I. D. H. *Determinantes de la concentración industrial en México*. CIDE - Centro de Investigación y Docencia Económicas A. C., DF, México, 2013.
- HIRSCHMAN, A. O. *The strategy of economic development*. New Haven: Yale University Press, 1958.
- IPARDES. Instituto Paranaense de Desenvolvimento Econômico Social. *Dinâmica recente da indústria paranaense: estrutura e emprego*. Curitiba: IPARDES, 87p., 2007.
- IPARDES. Instituto Paranaense de Desenvolvimento Econômico Social. Base de Dados do Estado – BDEweb. Disponível em: <http://www.ipardes.pr.gov.br/imp/index.php>. Acesso em: 14 abr. 2019.
- KALDOR, N. The case for regional policies. *Scottish Journal of Political Economy*, v. 17, n. 3, p. 337-348, 1970.
- KRUGMAN, P. Increasing returns and economy geography. *Journal of Political Economy*, v. 99, n. 3, p. 483-499, 1991.


- LEMOS, M. L. F. Origins and development of the theory of labor Market segmentation. *Revista Universitária Rural*, v, 17, n. 2, p. 13-26, 1995.
- LIMA, R. Mercado de Trabalho: o capital humano e a teoria da segmentação. *Pesquisa e Planejamento Econômico*, v. 10, n. 1, p. 217-272, 1980.
- MARSHALL, A. *Princípios de economia*. Coleção dos Economistas. São Paulo: Nova Cultura, v.1, 272 p., 1985.
- MYRDAL, G. *Economic theory and under-developed regions*. Gerald Duckworth & CO. LTD: London, 1957.
- REICH, M. G., GORDON, D. M., EDWARDS, R. C. A theory of labor market segmentation. *The American Economic Review*, v. 63, n. 2, p. 359-365, 1973.
- ROBERTS, M. R.; WHITED, T. M. Endogeneity in empirical corporate financel. In: CONSTANTINIDES, G.; HARRIS, M.; STULZ, R. (eds.). *Handbook of the Economics of Finance*, v.2, Parte A, p. 493-572, 2013.
- ROCHA, R. M.; BEZERRA, F. M.; MESQUITA, C. S. Uma análise dos fatores de aglomeração da indústria de transformação brasileira. *Revista Economia*, v. 14, n. 1A, p.61-90, 2013.
- RODRIGUES, M. A.; MONTEIRO, W. F.; CAMPOS, A. C.; PARRÉ, J. L. Identificação e análise espacial das aglomerações produtivas do setor de confecções na região Sul. *Economia Aplicada*. v. 16, n. 2, p. 311-338, 2012.
- ROSA, J. R.; RAIHER, A. Aglomerações produtivas por intensidade tecnológica e o desenvolvimento econômico dos municípios paranaenses: uma análise espacial. *Desenvolvimento Regional em Debate*, v. 7, n. 1, p. 115-143, 2017.
- ROSENTHAL, S. S.; STRANGE, W. C. Geography, industrial organization, and agglomeration. *The Review of Economics and Statistics*, v. 85, n. 2, p. 377-303, 2003.
- SANTOS, J. C. *Determinantes das concentrações industriais entre os estados brasileiros: uma análise PVAR no período de 2003 a 2014*. Dissertação (Mestrado em Economia), Universidade Federal do Rio Grande do Norte, Natal, 2017.
- SANTOS, S. A. *Os impactos do desmatamento na incidência de malária na Amazônia: uma análise espacial*. Dissertação (Mestrado em Economia), Universidade Federal do Rio Grande do Norte, Natal, 2016.
- SOUZA, L. A.; MUNIZ, A. L. P. Os fatores determinantes da localização das indústrias Goianas. *Revista CESUC*, Ano XIII, n. 23, p. 161-175, 2010.
- SOUZA, M. C. C. Mercados de trabalho: abordagens duais. *Revista de Administração de Empresas*, v. 18, n. 1, p. 59-69, 1978.

## ORCID

Andréia Ferreira Prestes  <https://orcid.org/0000-0002-3906-5608>

Fernanda Mendes Bezerra  <https://orcid.org/0000-0002-3307-0107>

Talita Egevardt de Castro  <https://orcid.org/0000-0003-4213-4069>

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.



**A EVOLUÇÃO DA INTERAÇÃO UNIVERSIDADE-EMPRESA NO RIO GRANDE DO SUL: O CASO DAS INDÚSTRIAS EXTRATIVAS E DE TRANSFORMAÇÃO NO PERÍODO 2008-2014\***

**Rodrigo Morem da Costa**

Secretaria de Planejamento, Orçamento e Gestão do Rio Grande do Sul

E-mail: [rodrigo-costa@planejamento.rs.gov.br](mailto:rodrigo-costa@planejamento.rs.gov.br)

**RESUMO:** Este estudo investiga a evolução da interação universidade-empresa nas indústrias extrativas e de transformação do Rio Grande do Sul no período 2008-2014, comparando esse processo com o observado em estados brasileiros e em nações industrializadas selecionadas. Os resultados encontrados mostram que houve uma evolução positiva das capacitações tecnológicas de empresas e de universidades gaúchas, com desempenho acima da média nacional, mas em patamar ainda inferior ao observado em nações desenvolvidas. Essa deficiência aparenta ser relativamente mais crítica para as empresas industriais do que para as universidades. Quanto à interação universidade-empresa no Rio Grande do Sul, mesmo com intensificação no período considerado, essa ainda é pouco frequente e coloca-se em nível inferior ao verificado em países avançados. Além disso, mesmo existindo elementos facilitadores para a interação universidade-empresa, outros fatores atuam como limitantes importantes, como as relativamente baixas capacitações tecnológicas das empresas industriais, a defasagem em alguns setores de alta e de média-alta intensidade tecnológica e a estrutura da indústria gaúcha, com predominância de setores industriais *low-tech* e pequena participação de firmas de portes médio e grande no total de empresas, bem como a relativa baixa proporção do quadro de docentes em exercício nas instituições de ensino superior que possui atuação em pesquisa.

**Palavras-chave:** Interação universidade-empresa; Inovação; Rio Grande do Sul.

**Classificação JEL:** O31; O43; O54.

**THE EVOLUTION OF UNIVERSITY-ENTERPRISE INTERACTION IN RIO GRANDE DO SUL: THE CASE OF MINING AND QUARRYING AND MANUFACTURING INDUSTRIES IN THE PERIOD 2008-2014**

**ABSTRACT:** This study investigates the evolution of university-enterprise interaction in the mining and quarrying and manufacturing industries of Rio Grande do Sul in the period 2008-2014, with comparisons between this process and that one observed in selected Brazilian states and industrialized nations. The results points that has been occurring a positive evolution in the technological capabilities of the firms and universities of Rio Grande do Sul, with a performance above of the national average, but still at an inferior level when compared to that of developed nations. This limitation appears to be relatively more critical to the industrial enterprises than to the universities. About the university-enterprise interaction in Rio Grande do Sul, even with its intensification in the period considered, it is still infrequent and is at a lower level than that seen in advanced countries. In addition, even with the existence of elements that facilitate the university-enterprise interaction, other factors acts as important constraints, such as the relatively still low technological capabilities of the industrial enterprises, the technological gap in some sectors of high and medium-high technology intensities and the industry structure, with a predominance of low-tech industries and small participation of medium and large sizes enterprises as a share of the total, as well as the relative low proportion of professors in exercise in higher education institutions that have engagement in research activities.

**Keywords:** University-enterprise interaction; Innovation; Rio Grande do Sul.

**JEL Codes:** O31; O43; O54.

## 1. Introdução

O objetivo deste estudo é analisar a evolução do padrão de interação universidade-empresa (U-E) nas indústrias extrativas e de transformação do Rio Grande do Sul entre 2008 e 2014. É bem sabido que a interação entre universidades e empresas é um nexos fundamental em sistemas de inovação, cuja intensificação contribui para o surgimento e a difusão de inovações, com efeitos positivos para a competitividade das empresas e para o desenvolvimento econômico. O período 2008-2014 testemunhou aspectos positivos e negativos no processo de inovação, em geral, e para a interação universidade-empresa em específico.

Pelo lado positivo, no Brasil, as políticas industriais e de ciência e tecnologia com foco no apoio à inovação nas empresas, que até então eram esparsas, ganharam centralidade como estratégia de desenvolvimento econômico a partir de 2003, sendo intensificadas até 2014 (MAZZUCATO; PENNA, 2016; RAPINI et al., 2017). A título de ilustração, algumas das políticas industriais e de ciência e tecnologia implantadas no Brasil pelo Governo Federal, com destaque no apoio à inovação após 2003, foram: Política Industrial, Tecnológica e de Comércio Exterior (PITCE), Fundo Nacional de Desenvolvimento Científico e Tecnológico, Lei da Inovação, Plano de Desenvolvimento Produtivo (PDP), Plano Brasil Maior.

Similarmente, esse período também exibiu ações de fomento à inovação sendo intensificadas pelos governos de alguns estados, incluindo-se aí o Rio Grande do Sul desde 1999, sobretudo, nos últimos anos. No Rio Grande do Sul, dentre as ações elaboradas e intensificadas pelo Governo do estado de 1999 a 2014 e que tiveram relativo maior vulto, podem-se elencar as seguintes: Programa de Apoio aos Sistemas Locais de Produção (SLPs), Fundo Operação Empresa do Estado do Rio Grande do Sul (FUNDOPEM/RS); Programa Estadual de Fortalecimento das Cadeias e Arranjos Produtivos Locais (APLs); Programa Gaúcho de Apoio aos Parques Tecnológicos (PGTEC). Um maior detalhamento das ações pode ser encontrado em Antunes Júnior et al. (2017).

Pelo lado negativo, releva assinalar que, entre 2008 e 2014, o Brasil passou por momentos conjunturais adversos, de incertezas e de crise econômica. Essas adversidades afetaram negativamente as decisões empresariais de investimento, inclusive em P&D, sobretudo, em 2009 e 2013-2014 (IBGE, 2013; 2016).

Nesse contexto, a pergunta de pesquisa que orientou o trabalho foi: como o padrão de interações U-E das indústrias extrativas e de transformação do Rio Grande do Sul evoluiu no período 2008-2014, em relação a outros estados brasileiros e a nações industrializadas selecionadas? O estudo orienta-se pelo conceito de Sistemas de Inovação da teoria evolucionária, com ênfase nos relacionamentos existentes entre empresas e universidades. A metodologia empregada no estudo consiste de revisão da literatura e de análise de dados coletados em diferentes bases: Pesquisa de Inovação (PINTEC) do IBGE; diretório de grupos de pesquisa do CNPq; CAPES; INEP; *Community Innovation Survey* (CIS), dentre outras. Um conjunto de indicadores que refletem as capacitações tecnológicas de empresas industriais e de instituições de ensino superior e de pesquisa no Rio Grande do Sul foram definidos e analisados, bem como o padrão de interação universidade-empresa no estado, de 2008 a 2014. O foco do estudo centra-se nas empresas das indústrias extrativas e de transformação e nas universidades<sup>1</sup> e instituições de ensino superior do Rio Grande do Sul.

Afora esta Introdução, o texto está organizado em mais quatro seções abordando: (1) a revisão teórica – em que é discutido o conceito de sistema de inovação, o papel de seus atores e os benefícios e condicionantes da interação universidade-empresa para o desenvolvimento de inovações; (2) a caracterização das capacitações tecnológicas das empresas industriais gaúchas, em termos de número e qualificação dos pesquisadores e de esforço tecnológico; (3) caracterização e capacitações em ensino e pesquisa das universidades do Rio Grande do Sul, relativas à qualidade dos cursos e à qualificação do corpo docente; e (4) o padrão de interação universidade-empresa no estado e sua evolução no período considerado, através de dados da PINTEC e do Diretório de Grupos de Pesquisa do CNPq. As considerações finais encerram o texto.

---

<sup>1</sup> Ao longo deste trabalho, o termo “universidade” será utilizado em seu sentido mais amplo, abarcando as próprias universidades e também as demais instituições de ensino superior.

## 2. A Interação universidade-empresa em sistemas de inovação

Sob a ótica da teoria econômica evolucionária, é aceito que inovações, sobretudo as tecnológicas, desempenham um papel importante para a competitividade das empresas e para o desenvolvimento econômico. O conceito de Sistemas de Inovação (doravante referido como SI) estabelecido por esta teoria mostra-se útil ao entendimento do processo de desenvolvimento de novas tecnologias pelas empresas. Em uma abordagem ampla, um Sistema de Inovação pode ser definido como sendo o conjunto de firmas e organizações – inseridas em um contexto institucional<sup>2</sup> que estrutura e condiciona seus relacionamentos – que interagem com objetivo de explorar, buscar, produzir, transmitir e aprender novos conhecimentos científicos e tecnológicos, úteis ao processo de desenvolvimento e à difusão de inovações (LUNDVALL, 2010). Em uma abordagem restrita, o SI engloba as organizações e instituições mais diretamente envolvidas nos processos de inovação (FREEMAN, 2010). Ambas as definições podem ser aplicadas a diferentes dimensões de SIs: nacional, regional, setorial e local.

A ideia básica de um SI é que os processos dos quais as inovações emergem são complexos, envolvendo o surgimento e a difusão de conhecimentos codificados e tácitos com possibilidades científicas e tecnológicas, sendo originários de múltiplas fontes. Isto requer a sua “tradução” em novos produtos ou processos produtivos e organizacionais e o seu lançamento no mercado, envolvendo uma série de relações interativas e mecanismos de *feedback* (EDQUIST, 2011). Em outras palavras, as empresas não desenvolvem novos conhecimentos e projetam inovações apenas a partir de suas capacitações tecnológicas (humanas, financeiras, infraestrutura de laboratórios, etc.); de seus esforços formais de pesquisa e desenvolvimento (P&D); de fontes informais de aprendizado; e de relações internas entre seus departamentos. O aprendizado de novos conhecimentos e o acesso a recursos para desenvolver ou implementar inovações também decorrem da interação da empresa com o ambiente produtivo, científico, político e institucional no qual está inserida, assim como da relação que essas esferas mantêm com o resto do mundo (FREEMAN, 2008). Assim, as interações externas da firma envolvem competição, mas também comunicação e coordenação, permitindo o estabelecimento de relações de cooperação com outros atores para desenvolver inovações.

Embora os conhecimentos possam ser gerados por quaisquer dos agentes de um Sistema de Inovação, são as empresas que basicamente os materializam na forma de inovações em produtos, processos produtivos, rotinas organizacionais, etc., o que as torna centrais ao sistema. Além delas, três outros conjuntos de atores exercem papéis relevantes em um SI. O primeiro é formado pelas organizações de ensino e pesquisa – públicas e privadas – como universidades, institutos tecnológicos, laboratórios de metrologia, organizações de treinamento profissional, de consultoria, etc. O segundo consiste na ação do Estado, através de órgãos públicos e empresas estatais, enquanto formulador de política industrial e tecnológica; de fonte de recursos à inovação; de regulação da atividade econômica; e demandante, fomentador e empreendedor em inovações importantes à sociedade, mas que se mostrem onerosas e/ou arriscadas demais para serem desenvolvidas pelo setor privado (MAZZUCATO, 2014). E o terceiro grupo consiste em organizações pertencentes ao sistema financeiro nacional, devido a uma parcela do esforço tecnológico ocorrer através de financiamento, via crédito ou outros mecanismos de alavancagem de recursos no mercado de capitais, como, por exemplo, o *venture capital*.

Quanto às universidades, sua atuação possui relevância particular para o desempenho do sistema de inovação, pois elas geram e transferem conhecimentos à sociedade por diferentes canais. Elas formam recursos humanos, dentre os quais se encontram pesquisadores em nível de mestrado e de doutorado; realizam pesquisas em ciência e tecnologia; prestam serviços tecnológicos; e servem como infraestrutura<sup>3</sup> à realização de P&D. Observe-se, também, que muitas empresas de base tecnológica (*start-ups*) são formadas pelo empreendedorismo de indivíduos ligados às pesquisas em universidades. Devido ao estabelecimento de vínculos com a produção científica internacional, as universidades constituem-se como um dos canais de

<sup>2</sup> O conceito de instituições adotado tem o sentido de hábitos e rotinas de conduta, regras, contratos e organizações, que estruturam e condicionam as interações entre agentes econômicos, gerando ordem e estabilidade para a realização de ações coletivas (HODGSON, 1998).

<sup>3</sup> Além de equipamentos e laboratórios, a infraestrutura também pode compreender incubadoras e parques tecnológicos.

entrada em um SI de conhecimentos gerados externamente. Em alguns casos, os conhecimentos produzidos pelas pesquisas em universidades podem posteriormente dar origem a novas aplicações, inclusive radicais, ou solucionar eventuais gargalos à evolução tecnológica, gerando novas oportunidades de mercado às empresas (FREEMAN, 2010). Ademais, o acesso a pesquisadores e a laboratórios das universidades complementa as habilidades e os recursos físicos, quando a empresa não dispõe de capacitações suficientes<sup>4</sup> à solução de determinados problemas tecnológicos. Este ponto é particularmente relevante quando há descontinuidades na base de conhecimentos (MOWERY; SAMPAT, 2005; SUZIGAN et al., 2011; SMITH, 2011).

Em especial, nos últimos anos, a interação universidade-empresa vem aumentando de importância, devido à elevação da complexidade na base de conhecimento científico e à concomitante necessidade de pesquisa básica para seu avanço, sobretudo em setores *high-techs*. Com isso, vem se tornando muito caro para as empresas investirem na formação integral das capacitações tecnológicas requeridas ao desenvolvimento de muitas das novas tecnologias que estão emergindo (inteligência artificial, *internet* das coisas, novos materiais, etc.). Portanto, ao serem absorvidas ou acessadas pelas empresas, essas funções das universidades contribuem ao seu processo de formação de capacitações tecnológicas e de desenvolvimento de inovações, ampliando suas possibilidades competitivas.

Observa-se que a natureza das interações universidade-empresa é simbiótica, sendo um processo coevolutivo importante ao desenvolvimento de um SI (NELSON; ROSENBERG, 1999). Além das vantagens para as empresas, as universidades se beneficiam de problemas tecnológicos trazidos pelas primeiras, que podem criar novos objetos de pesquisa para os cientistas. Inclusive, inovações radicais que criem novos mercados também podem originar novas áreas de conhecimento e disciplinas na academia. Nessa direção, demandas recorrentes por parte das empresas podem informar a atualização do currículo de disciplinas em cursos universitários. Além disso, avanços tecnológicos em diferentes dispositivos, como computadores, microscópios, telescópios e outros instrumentos, têm contribuído ao avanço científico por melhorarem os processos de observação de fenômenos e de coleta e processamento de dados.

Alguns fatores atuam como facilitadores para a ocorrência da interação universidade-empresa, sobretudo, a proximidade cognitiva, física, social e institucional entre esses atores. A cognitiva favorece a transmissão de conceitos e de conhecimentos complementares. A proximidade física contribui à realização de interações face a face com frequência, inclusive, sendo importante à comunicação de conhecimentos tácitos. Ademais, a existência de capital social entre os partícipes e o contexto institucional compartilhado, contribuindo ao ordenamento desta relação, permitem intensificar a interação universidade-empresa e o fluxo de informações entre esses atores. Inclusive, essa é uma das vantagens atribuídas à organização da produção em aglomerações produtivas especializadas, sobretudo nos chamados Arranjos Produtivos Locais.

Nessa discussão, pesquisadores com mestrado e doutorado nas empresas, devido ao seu treinamento formal e à prática na realização de atividades internas de P&D, constituem-se em facilitadores da interação com universidades. Isto porque esses profissionais fazem a ligação entre os conhecimentos formais e os aplicados; estão familiarizados com o *modus operandi* (rotinas) da pesquisa universitária, que é distinto da empresarial; criam objetivos compartilhados entre essas organizações e, usualmente, dispõem de uma rede de contatos nas universidades (capital social), advindo da passagem prévia quando de sua formação acadêmica, particularmente na pós-graduação. Por outro lado, convém observar que, quando a firma se depara com um problema de elevada complexidade, particularmente em inovações radicais, o elemento orientador da busca pela interação com a universidade constitui-se na excelência em pesquisa acadêmica na área de interesse (COSTA et al., 2011; GARCIA et al., 2014).

Mesmo existindo benefícios mútuos na interação universidade-empresa, nem sempre esses agentes se mostram interessados em buscar tal relacionamento. Assim, cabe indicar alguns dos aspectos mais recorrentes, computados pela literatura, que dificultam sua ocorrência (CAMPOS, 2010; BUCHARTH, 2011; SCHEFFER et al., 2015). Para as empresas, os seus elementos restritivos internos à cooperação com

---

<sup>4</sup> Quando a firma não vislumbra um uso frequente para recursos humanos com determinada habilidade ou para um equipamento específico, a interação com universidades consiste em uma alternativa interessante à solução de gargalos no desenvolvimento de inovações, com menor custo em relação à alternativa de investir na formação interna destas capacitações.

universidades dizem respeito, entre outros, ao estágio do ciclo de vida tecnológico do produto<sup>5</sup>; ao elevado grau de incerteza sobre os resultados de projetos conjuntos; e à amplitude de suas capacitações tecnológicas, quais sejam, nível de qualificação e de conhecimento dos pesquisadores; orçamento para financiar projetos conjuntos de P&D; e infraestrutura à realização dessas atividades internamente. Os limitantes externos para as empresas podem ser atribuídos a fatores como: distanciamento da pesquisa universitária em relação à sua aplicação prática e imediata; dificuldades na regulação relacionada à transferência de tecnologia das universidades para as empresas; complexidade e velocidade da burocracia universitária; e longa duração dos projetos conjuntos.

Quanto às universidades, dentre os seus fatores internos que inibem a interação com empresas podem ser elencados: perda de liberdade de pesquisa; contratos de pesquisa voltados para o curto prazo; elevada burocracia; e dificuldades organizacionais e institucionais no uso de recursos. Dentre os fatores externos limitantes da interação U-E para as universidades, os principais seriam a diferença no nível de conhecimento em relação aos interlocutores da firma (distância cognitiva); o conflito entre o direito de propriedade da patente (empresa) e de divulgação dos resultados da pesquisa (universidade); e dificuldades para obtenção de financiamento para a realização de projetos conjuntos.

### 3. Capacitações tecnológicas das empresas industriais do Rio Grande do Sul

Como pré-requisito à análise da interação das empresas industriais<sup>6</sup> com universidades no Rio Grande do Sul, convém caracterizar ambos os conjuntos de atores deste SI. Em 2014, na indústria gaúcha, as unidades produtivas com mais de cinco pessoas ocupadas eram em número de 21.868, possuíam 735.046 empregados e geraram valor da transformação industrial (VTI) de R\$78,7 bilhões (Tabela 1). Esses números respondem por 8,8% do emprego e por 7,2% do VTI das indústrias extrativas e de transformação do País, o que colocou o estado como o quarto no *ranking* industrial nacional por critério de tamanho do VTI<sup>7</sup>. Em 2014, os oito principais setores industriais do estado responderam juntos por cerca de 70% do VTI e por 66,5% do emprego (Tabela 1). Quando agregadas as atividades industriais por intensidade tecnológica<sup>8</sup>, somando-se os valores do VTI, observa-se que a estrutura do Rio Grande do Sul era formada por 33,7% de setores *high-tech* e 66,3% de *low-tech*, enquanto que para a média do Brasil elas representaram respectivamente 27,6% e 72,4%.

Em termos da estrutura por porte dos estabelecimentos da indústria do Rio Grande do Sul em 2014, 82,9% eram de microempresas, 13,6% pequenas, 2,8% médias e 0,6% grandes, as quais foram responsáveis, respectivamente, por 20,4%; 23,8%; 27,6% e 28,2% do total do emprego<sup>9</sup>. Esses dados revelam que as médias e grandes empresas industriais, que em tese teriam melhores condições para sustentar atividades

<sup>5</sup> Produtos em fases iniciais de seu ciclo de vida tecnológico – radicalmente novos – em geral possuem maior probabilidade de incorrerem em descontinuidades na base de conhecimentos, podendo motivar a interação universidade-empresa.

<sup>6</sup> Nesse texto, o termo “indústria” será empregado para se referir às atividades das indústrias extrativas e de transformação.

<sup>7</sup> Cálculos com base em dados da PIA-Empresa (IBGE, 2017b).

<sup>8</sup> Essa classificação segue a taxonomia de intensidade de esforço tecnológico da OCDE (2011), adaptada para divisões (dois dígitos) da CNAE2.0, a exemplo da empregada pelo Eurostat (2019), devido à indisponibilidade de alguns dados por grupos (três dígitos) da CNAE2.0 que dificultam o uso da classificação mais desagregada. As atividades *high-techs* consistem nas classificadas como de “alta” e “média-alta” intensidade tecnológica e as *low-techs* como de “média-baixa” e “baixa”. Convém alertar que a Taxonomia da OCDE representa a intensidade tecnológica média de cada setor da indústria de transformação, calculada a partir de dados de um conjunto de países membros dessa Organização, com empresas próximas da fronteira tecnológica. Entende-se que cada setor da indústria tenha um paradigma tecnológico e um padrão de concorrência setorial próprios; e de que a pressão competitiva tenda a fazer com que as empresas busquem se aproximar da intensidade de P&D observada na média da sua atividade. Alguns estudos identificaram que, com algumas exceções, os setores da indústria no Brasil, em geral, mostram-se aderentes aos agrupamentos por intensidade tecnológica das Taxonomias da OCDE, o que valida sua adoção (MARQUES et al., 2018; MORCEIRO, 2019). Contudo, o patamar de esforço tecnológico dos setores industriais brasileiros se mostra inferior ao de países da OCDE, devido ao relativo atraso tecnológico nacional. Infere-se que o mesmo ocorra com as atividades da indústria do Rio Grande do Sul.

<sup>9</sup> Cálculos com base no número de estabelecimentos e de empregos da Relação Anual de Informações Sociais (RAIS) (BRASIL, 2017) e delimitação de porte de estabelecimentos industriais por número de empregados, de acordo com a metodologia do Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (SEBRAE, 2017), assim classificados: microempresa, entre 1 e 19 empregados; pequena, entre 20 e 99; média, entre 100 e 499; e grande, com 500 e mais empregados.

contínuas de P&D, representavam 3,4% dos estabelecimentos e 55,8% do emprego da indústria gaúcha. Entre os setores industriais, os com maior participação de estabelecimentos de médio e de grande porte no estado foram: automotivo (10,9%); fabricação de produtos químicos (6,6%); fabricação de produtos de borracha e de material plástico (6,1%); e fabricação de produtos alimentícios (5%).

**Tabela 1 - Dados selecionados das unidades industriais do Rio Grande do Sul com 5 ou mais pessoas ocupadas - 2014**

Classificação Nacional de Atividades Econômicas (CNAE 2.0)	Intensidade Tecnológica (OCDE)	Unid. locais (Nº)	Pessoal ocupado		Valor da transformação industrial	
			(Nº)	(%)	(R\$ Bilhões)	(%)
<b>Total</b>	<b>N/A</b>	<b>21.868</b>	<b>735.046</b>	<b>100,0</b>	<b>78,7</b>	<b>100,0</b>
- (B) Indústrias extrativas	Baixa	372	6.146	0,8	0,6	0,8
- (C) Indústrias de transformação	N/A	21.496	728.900	99,2	78,1	99,2
-- (10) Produtos alimentícios	Baixa	2.758	131.599	17,9	13,6	17,3
-- (28) Máquinas e equipamentos	Média-alta	1.533	70.610	9,6	8,6	11,0
-- (29) Automotivo	Média-alta	512	51.164	7,0	8,0	10,1
-- (20) Produtos químicos	Média-alta	530	17.175	2,3	6,3	8,1
-- (15) Couros e calçados	Baixa	2.802	119.607	16,3	5,3	6,8
-- (25) Produtos de metal	Média-baixa	2.824	62.560	8,5	5,3	6,7
-- (19) Coque, deriv. petróleo e biocombustív.	Média-baixa	17	1.541	0,2	4,9	6,2
-- (22) Produtos de borracha e plástico	Média-baixa	1.014	34.881	4,7	3,4	4,3
-- Outros	N/A	9.506	239.763	32,6	22,7	28,9

Fonte: IBGE (2017b); OCDE (2011).

No que tange à sua evolução no Rio Grande do Sul de 2008 a 2014, a indústria vem apresentando baixo dinamismo, em média, com perda de participação na estrutura do valor adicionado bruto (VAB) e do emprego da economia estadual<sup>10</sup>. No período, o VAB da economia gaúcha cresceu 39,2% e o emprego total elevou-se em 23,3%, enquanto que para a sua indústria esse aumento foi de respectivamente 18,5% e de 10,8%. Desse modo, sua participação na economia do estado declinou de 20% no VAB e de 25,6% no emprego em 2008 para, respectivamente, 17% e 23% em 2014. Deve-se considerar que, além de dificuldades competitivas, essa queda de importância da indústria na economia estadual também está relacionada à conjuntura adversa, devido aos efeitos da crise financeira mundial em 2008-2011 e da desaceleração da economia brasileira no período 2013-2014, que a afetaram mais do que as demais atividades (IBGE, 2013; 2016). Contudo, dado que as quedas de participação no VAB e no emprego foram praticamente contínuas no período, pode-se perceber esse processo como tendo um forte componente de desindustrialização e de mudança estrutural.

Quanto à evolução das capacitações tecnológicas da indústria do Rio Grande do Sul, convém lançar mão da análise de alguns indicadores, enquanto aproximações, para seu potencial para desenvolver ou para assimilar inovações. De 2008 a 2014, observa-se que, em geral, embora o gasto total em atividades de inovação<sup>11</sup> tenha se reduzido no estado e no País, ainda assim, os dispêndios em P&D interno vieram aumentando, com taxas de crescimento<sup>12</sup> de 66,1% para o Rio Grande do Sul e de 22,9% para a média brasileira. Esse movimento se observa em todas as Unidades da Federação analisadas (Tabela 2). Inclusive,

<sup>10</sup> Cálculos realizados com base em dados da FEE (2016) e RAIS (BRASIL, 2017), deflacionando os valores para preços constantes de 2014 pelo IGP-DI/FGV (IPEA, 2017).

<sup>11</sup> As atividades de inovação compreendem: Atividades internas de P&D; Aquisição externa de P&D; Aquisição de outros conhecimentos externos; Aquisição de software; Aquisição de máquinas e equipamentos; Treinamento; Introdução das inovações tecnológicas no mercado; e Projeto industrial e outras preparações técnicas.

<sup>12</sup> Cálculos com base nos dados da Tabela 2.

isso se reflete em maior intensidade de esforço tecnológico interno, mensurado como a fração destes dispêndios em relação à receita líquida de vendas das empresas<sup>13</sup>, que na indústria do Rio Grande do Sul foi de 0,39% em 2008, de 0,60 em 2011 e de 0,65 em 2014, enquanto que na brasileira foi respectivamente de 0,62%; 0,71% e de 0,67%. Ademais, no processo de desenvolvimento de inovações nos triênios 2006-2008, 2009-2011 e 2012-2014, a proporção de empresas que atribuíram relevância alta ou média para as atividades internas de P&D elevou-se de 11,2% para 17,8% e finalmente 21,1% no estado, enquanto que essa evolução no Brasil saiu de 11,5% para 15,9% e para 15,2%<sup>14</sup>.

Esses aspectos indicam que o esforço tecnológico interno voltado ao desenvolvimento próprio de inovações vem ganhando importância na indústria, o que é relevante por resultar em maior absorção de conhecimentos, diferenciação de produtos e formação de infraestrutura à P&D, ampliando as possibilidades futuras das empresas. Trata-se, portanto, de uma mudança qualitativa importante na estrutura dos dispêndios em atividades de inovação, que vem se processando no País nos últimos anos, sendo relativamente mais intensa no Rio Grande do Sul. Entretanto, note-se que a fração dos dispêndios em P&D interno relativo ao total das atividades de inovação no estado encontrava-se abaixo da média nacional em 2014, ainda que o hiato em relação a 2008 tenha se reduzido.

**Tabela 2 - Dispêndios realizados em atividades de inovação por empresas industriais inovadoras no Brasil: 2008-2014**

Valores em R\$ bilhões a preços de 2014

Unidades da Federação	Atividades de Inovação			P&D Interno			P&D Interno/AI - (%)		
	2008	2011	2014	2008	2011	2014	2008	2011	2014
<b>Brasil</b>	<b>60,4</b>	<b>60,3</b>	<b>57,6</b>	<b>14,8</b>	<b>18,0</b>	<b>18,2</b>	<b>24,5</b>	<b>29,8</b>	<b>31,5</b>
Rio Grande do Sul	3,8	3,4	3,5	0,6	0,9	1,1	17,1	27,6	30,3
São Paulo	30,3	27,0	27,4	8,0	7,7	8,8	26,3	28,7	32,2
Rio de Janeiro	5,6	7,3	6,4	2,7	5,0	3,7	47,7	68,3	57,8
Minas Gerais	8,0	7,6	4,5	1,7	1,6	1,2	20,8	20,6	26,3
Paraná	2,9	2,9	3,2	0,5	0,7	0,8	18,2	25,1	25,0
Santa Catarina	2,1	3,5	2,7	0,4	0,9	0,9	16,9	24,5	33,4
Outros	7,8	8,5	9,9	1,0	1,1	1,7	12,5	13,2	17,0

Fonte: IBGE (2010, 2013 e 2016).

Notas: Empresas das indústrias extrativas e de transformação, com 10 ou mais empregados, que implementaram produto e/ou processo novo ou substancialmente aprimorado. Valores deflacionados para preços de 2014 pelo IGP-DI/FGV (IPEA, 2017).

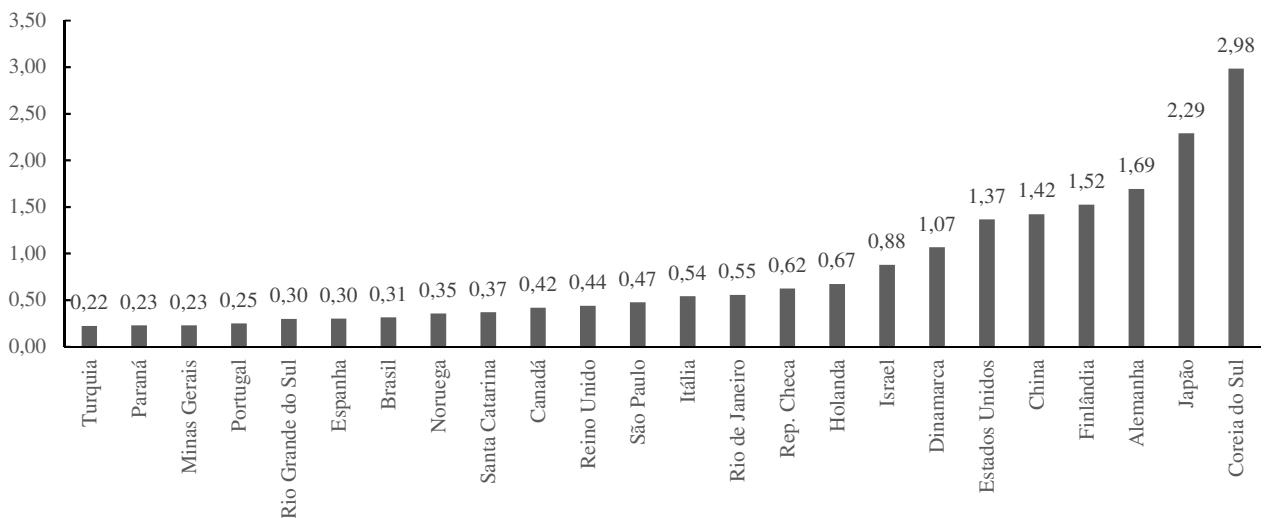
Cabe considerar que o desempenho apresentado nos gastos em atividades de inovação, destacando-se nesse conjunto o P&D interno, por empresas industriais do Rio Grande do Sul e dos demais estados brasileiros, ainda as posiciona em 2014 em um nível intermediário de esforço tecnológico em relação a nações industrializadas. Isso pode ser observado na proporção do esforço tecnológico interno das empresas industriais, computados em termos de gasto em P&D interno em relação ao PIB (Gráfico 1).

Em relação aos outros dois indicadores de capacitações tecnológicas das empresas, ou seja, infraestrutura e recursos humanos, agrupados como *proxy* para os principais tipos de recursos no desenvolvimento de inovações, verifica-se que de 2008 para 2014 houve uma nítida tendência de aumento em algumas das variáveis selecionadas. No que concerne à infraestrutura, uma aproximação para esse tipo de capacitação consiste em avaliar o número daquelas empresas que realizaram atividades internas de P&D, sobretudo de caráter contínuo, pois isso exige alguma dotação, pelo menos mínima, de laboratórios e de equipamentos. Ademais, a continuidade na execução de P&D permite maior domínio das rotinas associadas a essa atividade, gerando maior eficiência em sua realização.

<sup>13</sup> Cálculos com base em dados das PINTECs 2008, 2011 e 2014 (IBGE, 2010, 2013 e 2016).

<sup>14</sup> Cálculos com base em dados das PINTECs 2008, 2011 e 2014 (IBGE, 2010, 2013 e 2016).

**Gráfico 1 - Gastos em P&D interno de empresas industriais em relação ao PIB de estados e de países selecionados em 2014 (%)**



Fonte: IBGE (2016; 2017c); OCDE (2017a e 2017b).

Nota: Empresas das indústrias extrativas e de transformação com 10 ou mais empregados que implementaram inovações.

No Rio Grande do Sul e no Brasil, de um modo geral, dentre as empresas industriais que implementaram inovações de produto e/ou de processo produtivo entre 2008 e 2014, aumentou o número daquelas que inovaram realizando dispêndios em atividades internas contínuas de P&D, ganhando participação. No Rio Grande do Sul, essa proporção passa de 9,2% em 2008 para 18% em 2014, apresentando elevação com continuidade. Na média nacional, a mesma proporção variou de 9,9% em 2008 para 12,4% em 2014, embora com oscilação em seu crescimento. Isso também reflete uma maior intensidade do processo de ganho de capacitações tecnológicas na indústria no Rio Grande do Sul em relação a outros estados brasileiros no período. Inclusive, em 2014, o estado mostrou a maior participação (18%) de empresas com dispêndios contínuos em P&D interno sobre o total que inovou na indústria (Tabela 3). Aponta-se, também, que o número absoluto de empresas industriais gaúchas que inovaram com investimento contínuo em P&D foi de 305 em 2008 e de 728 em 2014, perfazendo um crescimento de 138,4% (IBGE, 2010; 2013; 2016). Inclusive, o Rio Grande do Sul possui o segundo maior contingente da indústria nacional de empresas que inovaram realizando P&D contínuo. Portanto, o ganho em capacitações se deu tanto de modo absoluto quanto relativo. Contudo, convém salientar que ainda é pequena a fração de empresas industriais que realizam atividades internas de P&D com continuidade para inovar, tanto no estado quanto no País.

Quanto aos recursos humanos exercendo atividades de P&D nas empresas industriais do Rio Grande do Sul que implementaram inovações, esse contingente mostrou evolução positiva no período 2008-2014, com crescimento de 151,5%. Na média nacional, essa taxa de crescimento foi de 118,9%.

Uma forma mais acurada de identificar as capacitações tecnológicas em recursos humanos consiste em avaliar a parcela de pós-graduados, em que praticamente sua totalidade está ocupada como pesquisadores. De um lado, normalmente, esses profissionais são detentores de um maior nível de conhecimento e de qualificação para a realização de atividades de P&D. De outro lado, como destacado anteriormente, os pesquisadores pós-graduados são importantes ao estabelecimento de relações com as universidades e demais instituições de pesquisa.

Ao longo do período, o Rio Grande do Sul mostrou o segundo maior contingente brasileiro de recursos humanos em P&D. Contudo, na indústria do Rio Grande do Sul, a fração do pessoal pós-graduado oscilou em torno de uma baixa participação, de 3,9% a 6,5%, do total empregado em atividades de P&D. Nota-se que essa proporção ainda esteve abaixo da média industrial brasileira no período, assim como não recuperou em 2014 o patamar de 2008. No absoluto, o número de pós-graduados em empresas industriais gaúchas se ampliou em 94,9%, de 309 em 2008 para 602 em 2014. No agregado da indústria do Brasil, a proporção de pós-graduados apresentou tendência de decréscimo, de 9,1% em 2008 para 7,4% em 2014 (Tabela 3).



Quanto ao número desses pesquisadores no Brasil, esse cresceu 76,6%, de 4.398 em 2008 para 7.768 em 2014. Observa-se, novamente, que houve ganho relativo de capacitações tecnológicas na dimensão de recursos humanos, sendo mais intensa no Rio Grande do Sul do que na média nacional. Cabe indicar que o ganho de capacitação foi mais pelo aumento absoluto do pessoal em P&D, mas com perda no relativo, dado ter diminuída a proporção de pós-graduados.

**Tabela 3 - Indicadores selecionados de capacitações tecnológicas em infraestrutura e em recursos humanos das empresas industriais brasileiras que implementaram inovações: 2008-2014**

Unidade da Federação	Infraestrutura						Recursos Humanos					
	Empresas com dispêndios em atividades de inovação						Pessoal ocupado em atividades de P&D					
	Número Total			% com P&D Interno Contínuo			Número Total			% Pós-Graduação		
	2008	2011	2014	2008	2011	2014	2008	2011	2014	2008	2011	2014
<b>Brasil</b>	<b>30.645</b>	<b>32.616</b>	<b>34.583</b>	<b>9,9</b>	<b>13,2</b>	<b>12,4</b>	<b>48.096</b>	<b>71.351</b>	<b>105.267</b>	<b>9,1</b>	<b>8,0</b>	<b>7,4</b>
Rio Grande do Sul	3.326	3.831	4.038	9,2	16,2	18,0	4.728	8.208	11.891	6,5	3,9	5,1
São Paulo	10.063	10.477	10.108	13,6	15,9	17,0	24.433	32.909	50.104	8,1	8,2	7,7
Rio de Janeiro	1.127	1.403	1.302	18,1	21,2	15,0	3.939	5.949	6.615	23,7	15,7	16,1
Minas Gerais	4.238	4.582	3.846	5,6	9,8	9,1	3.952	7.942	9.336	12,4	7,6	8,4
Paraná	2.890	2.863	3.331	7,7	14,1	10,8	2.839	4.718	7.810	6,1	7,0	3,1
Santa Catarina	2.710	2.732	3.671	10,5	12,1	7,7	3.562	5.128	8.749	5,8	7,3	3,9
Outras	6.291	6.727	8.285	6,3	7,8	7,9	4.644	6.497	10.762	6,6	7,5	8,2

Fonte: IBGE (2010, 2013, 2016).

Nota: Foram consideradas apenas as empresas das indústrias extrativas e de transformação com 10 ou mais empregados que implementaram produto e/ou processo tecnologicamente novo ou substancialmente aprimorado.

Em conjunto, embora tenha ocorrido um ganho de capacitação tecnológica na indústria do Rio Grande do Sul e do Brasil, nos três principais fatores relevantes ao desenvolvimento ou absorção de inovações – investimentos, recursos humanos e infraestrutura à P&D – essas ainda podem ser consideradas baixas, dada a pequena representação nos respectivos totais.

#### 4. Caracterização e capacitações das universidades do Rio Grande do Sul

No universo das instituições de ensino superior e de pesquisa existentes em 2014, o Rio Grande do Sul contava com uma infraestrutura formada por um total de 120 organizações, divididas em 19 universidades, cinco centros universitários, 93 faculdades e três Institutos Federais (IFs)/Centros Federais de Educação Tecnológica (CEFETs) (INEP, 2017). No estado, havia ainda três parques tecnológicos em operação até 2014: Tecnopuc (PUCRS), Tecnosinos (UNISINOS) e Valetec (FEEVALE) (BERNARDINI et al., 2014). Nesse universo de instituições, em 2014, o Rio Grande do Sul contava com 2.292 cursos presenciais em nível de graduação e 356 programas de pós-graduação (INEP, 2017; CAPES, 2017b).

Em relação aos 356 programas de pós-graduação do Rio Grande do Sul em 2014, 54 (15,2%) estavam no rol daqueles que se pode considerar como de “excelência” acadêmica, ou seja, que atingiram conceitos 7 ou 6 na avaliação da CAPES<sup>15</sup>. Por esse critério, o estado figurou na terceira colocação nacional, sendo

<sup>15</sup> Na avaliação de cursos de pós-graduação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), são atribuídos conceitos, que variam de 1 (mínimo) a 7 (máximo), para diferentes critérios: (a) titulação e experiência em orientação

precedido por São Paulo (20,2%) e Rio de Janeiro (15,6%) (Tabela 4). Nota-se que, embora o Rio Grande do Sul detenha 9,5% dos cursos de pós-graduação do Brasil, as universidades gaúchas respondem por 13% dos cursos de “excelência” acadêmica. Entre as universidades do Rio Grande do Sul, a UFRGS se destaca com 32 programas de pós-graduação considerados como sendo de excelência acadêmica, sendo sucedida pela PUC-RS com 11.

**Tabela 4 - Total de programas de pós-graduação e daqueles com excelência acadêmica no Brasil em 2014**

Unidade da Federação	Total de Programas (TP)		Programas de Excelência Acadêmica (EA)		
	TP	TP - %	EA	EA - %	EA/TP - %
<b>Brasil</b>	<b>3.748</b>	<b>100,0</b>	<b>415</b>	<b>100,0</b>	<b>11,1</b>
Rio Grande do Sul	356	9,5	54	13,0	15,2
São Paulo	840	22,4	170	41,0	20,2
Rio de Janeiro	435	11,6	68	16,4	15,6
Minas Gerais	378	10,1	53	12,8	14,0
Paraná	288	7,7	11	2,7	3,8
Santa Catarina	141	3,8	18	4,3	12,8
Outros	1.310	35,0	41	9,9	3,1

Fonte: GEOCAPES (CAPES, 2017b).

Observação: O número de programas de pós-graduação abarca os cursos de (a) mestrado profissional, (b) mestrado, (c) doutorado e (d) mestrado/doutorado. Considera-se como “excelência acadêmica” o somatório do número de programas com conceitos 6 e 7 na avaliação da CAPES.

Observa-se que, de 2008 a 2014, as instituições de ensino superior (IES) no Rio Grande do Sul e no Brasil vieram em um processo de ganho de capacitações, de intensidade similar, o que se expressa na evolução da proporção de professores com doutorado sobre o efetivo total (Gráfico 2). Inclusive, é de se observar que esse indicador para o estado foi superior ao do País ao longo do período, o que aponta para maiores capacitações para desempenhar suas atividades que a média das instituições de ensino superior estaduais. Em 2014, as IES gaúchas ficaram em segundo lugar no Brasil, com 44,5% de seus professores com doutorado, abaixo apenas das do Rio de Janeiro com 46,7%<sup>16</sup>. Esse diagnóstico também foi constatado para outros indicadores representativos das capacitações das instituições de ensino superior gaúchas no período 2009-2012 por Bernardini et al. (2014). Ademais o mesmo estudo identificou um desempenho similar em indicadores de resultado, como na proporção de jovens matriculados em cursos de graduação e de pós-graduação e de produção científica por autor.

O processo de ganho de capacitações das instituições de ensino superior também pode ser visto por outra ótica, através da parcela dos professores em exercício nas IES que possuem atuação em pesquisa. Essa fração representa os profissionais das IES que estão empenhados em gerar e transferir novos conhecimentos científicos e tecnológicos à sociedade, o que também é uma aproximação para sua qualificação, pois buscam estar na fronteira relativa aos seus objetos de pesquisa. Adicionalmente, uma maior realização de pesquisas nas universidades e IES também aumenta as possibilidades para a ocorrência de interações universidade-empresa. No período 2010-2014, as IES do Rio Grande do Sul e do Brasil indicam estar intensificando suas atividades de pesquisa (Gráfico 3)<sup>17</sup>. O número de docentes em exercício com atuação em pesquisa aumentou 32,5%, passando de 7.100 para 9.410. Esse crescimento no estado ficou 4,7 pontos percentuais

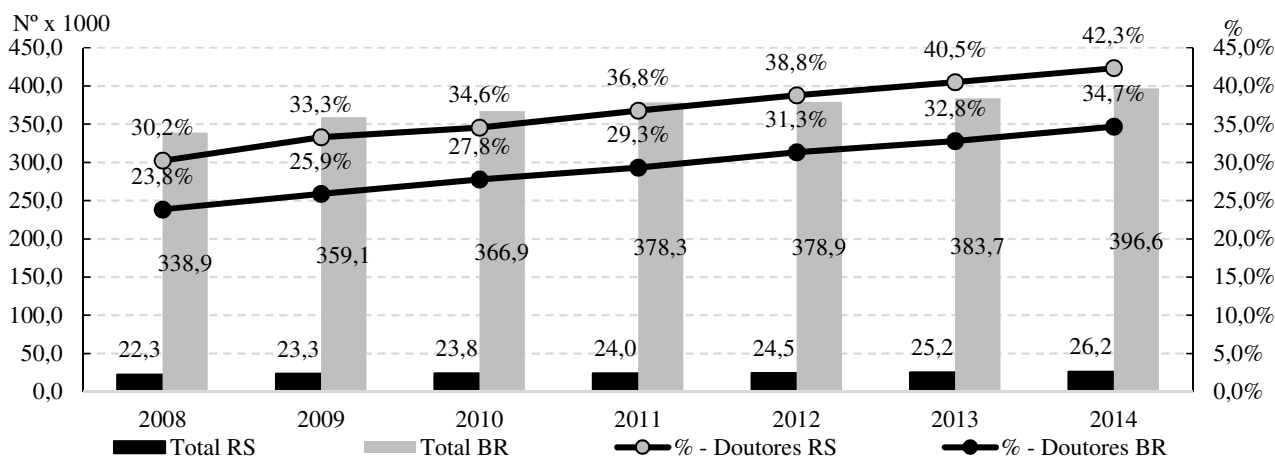
do corpo docente, (b) produção intelectual, (c) infraestrutura de ensino e pesquisa, (d) interação com grupos de pesquisa no exterior e (e) realização de programas de iniciação científica com alunos de graduação (CAPES, 2017a).

<sup>16</sup> Cálculos com base em dados do INEP (2017).

<sup>17</sup> É importante alertar que o total de docentes em exercício, no Gráfico 3, não inclui aqueles que estão afastados, enquanto que o Gráfico 2 trata-se do quadro total, inclusive os afastados. A diferença entre os dois totais é de cerca de 1,3% para o RS e o Brasil.

abaixo do observado para o País. No total do Brasil, a quantidade desses professores variou em 37,2%, de 77.305 para 106.086.

**Gráfico 2 - Total de professores e percentual com doutorado em instituições de ensino superior no Rio Grande do Sul e no Brasil**



Fonte: INEP (2017).

Em razão da elevação apresentada, a fração de professores em exercício que atua em pesquisa também se ampliou no Rio Grande do Sul, passando de 31,2% em 2010 para 37,1% em 2014 (Gráfico 3). Na média nacional, a mesma fração evoluiu de 22,4% para 27,7%. Observe-se que os outros estados mais industrializados, quando individualmente considerados, também apresentaram crescimento positivo no número de professores em exercício atuando em pesquisa<sup>18</sup>. Dentre esses estados, o Rio Grande do Sul se mostra como um dos que possui o maior percentual de professores em exercício em suas IES que possuem atuação em pesquisa, tendo ocupado a primeira posição em 2013 e 2014. Assim, com base na evolução dos indicadores calculados, pode-se entender que as capacitações das IES gaúchas e sua contribuição ao sistema estadual de inovação, enquanto qualificação de recursos humanos e realização de pesquisas, tenha se ampliado. Contudo, em termos dos professores em exercício envolvidos em pesquisa, entende-se que essa proporção de 37,1% em 2014 ainda pode ser considerada baixa. Esse é um limitante relevante à geração de conhecimentos científicos e tecnológicos, assim como para as possibilidades de interações com empresas.

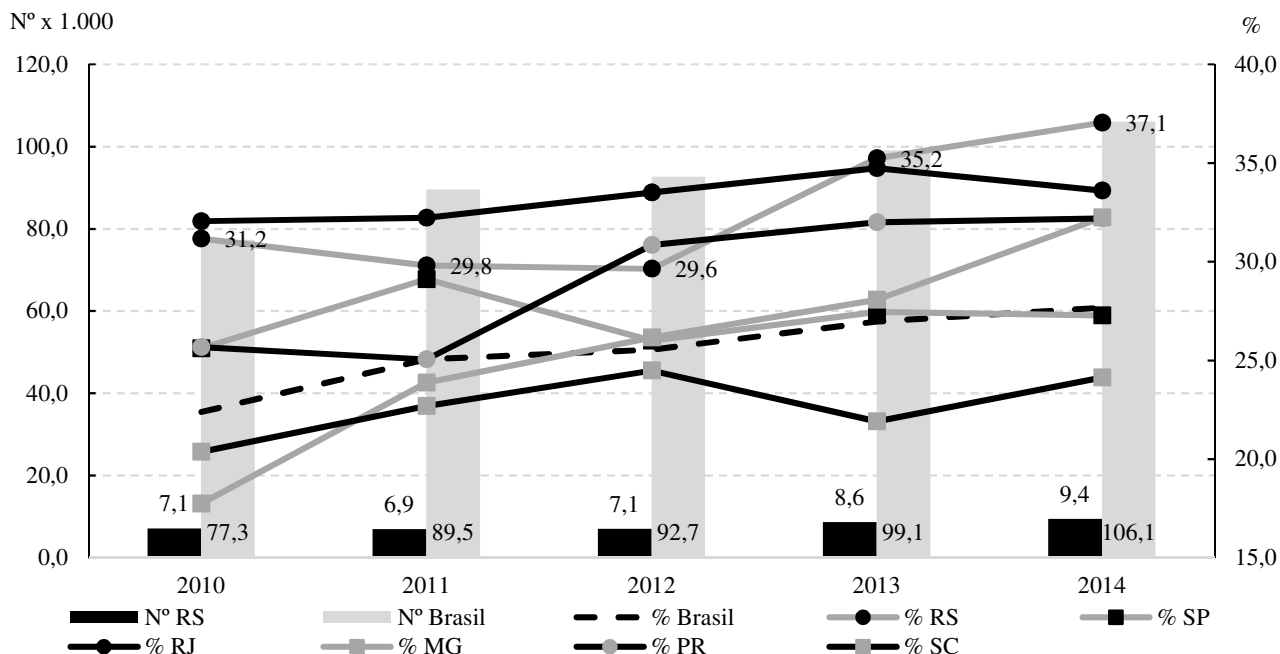
Em uma análise ampla das informações sobre as instituições de ensino superior do Rio Grande do Sul, quando comparadas nacionalmente, pode-se concluir que essas organizações apresentam um desempenho considerado bom e, em alguns casos, excelente, em indicadores de capacitações em geração de novos conhecimentos científicos e tecnológicos, de formação de pessoal qualificado em pesquisa e de infraestrutura para a realização de atividades de P&D. Entretanto, na comparação internacional, apenas sete universidades brasileiras (1,4%) conseguiram figurar entre as 500 melhores no *ranking* internacional de qualidade QS em 2015: USP-SP (143<sup>a</sup>), UNICAMP-SP (195<sup>a</sup>), UFRJ-RJ (323<sup>a</sup>), UFRGS-RS (entre 451<sup>a</sup> e 460<sup>a</sup>), UNESP-SP (entre 481<sup>a</sup> e 490<sup>a</sup>), UNB-DF e UNIFESP-SP (entre 491<sup>a</sup> e 500<sup>a</sup>) (TOPUNIVERSITIES, 2017). Chama a atenção que São Paulo abriga quatro das sete universidades nacionais no Top 500.

No cômputo total, o quadro que se forma no período 2008-2014 é o de que o Rio Grande do Sul possui um dos Sistemas de Inovações mais desenvolvidos, para o padrão nacional, na dimensão de suas empresas industriais e universidades. O estado figurou como um dos cinco mais industrializados, bem como com firmas com capacitações tecnológicas acima da média nacional em alguns indicadores. Dentre as instituições de ensino superior do Brasil, as gaúchas se destacam com um dos maiores potenciais para a formação de recursos humanos qualificados e para a geração de conhecimentos científicos e tecnológicos. Inclusive, observou-se um processo positivo de evolução das capacitações de ambos os atores nos últimos anos. Isso pode ser atribuído ao estabelecimento de um amplo conjunto de políticas industriais, de ciência e tecnologia

<sup>18</sup> Observado nos microdados do INEP (2019) que originaram o Gráfico 3.

e educacionais a partir de 2003, que colocaram a inovação no centro da agenda (MAZZUCATO; PENNA, 2016).

**Gráfico 3 - Número de professores com atuação em pesquisa e percentual relativo ao quadro de docentes em exercício em instituições de ensino superior do Brasil e de estados industrializados selecionados**



Fonte: Microdados do Censo da Educação Superior (INEP, 2019).

Nota: O percentual de professores que atuam em pesquisa foi discriminado apenas para o Rio Grande do Sul.

Contudo, para os padrões internacionais, esses SIs ainda estão em um patamar intermediário de desenvolvimento. Em parte, isso se atribui ao caráter relativamente tardio da criação de universidades e instituições de pesquisa, bem como do processo de industrialização brasileiro, conforme assinalado por Suzigan e Albuquerque (2011, p.17). Embora as capacitações tecnológicas e o desempenho na geração de inovações destes SIs sejam superiores ao de nações pouco industrializadas e ao de outras em patamar similar de industrialização, ainda estão bem aquém das que exercem liderança tecnológica.

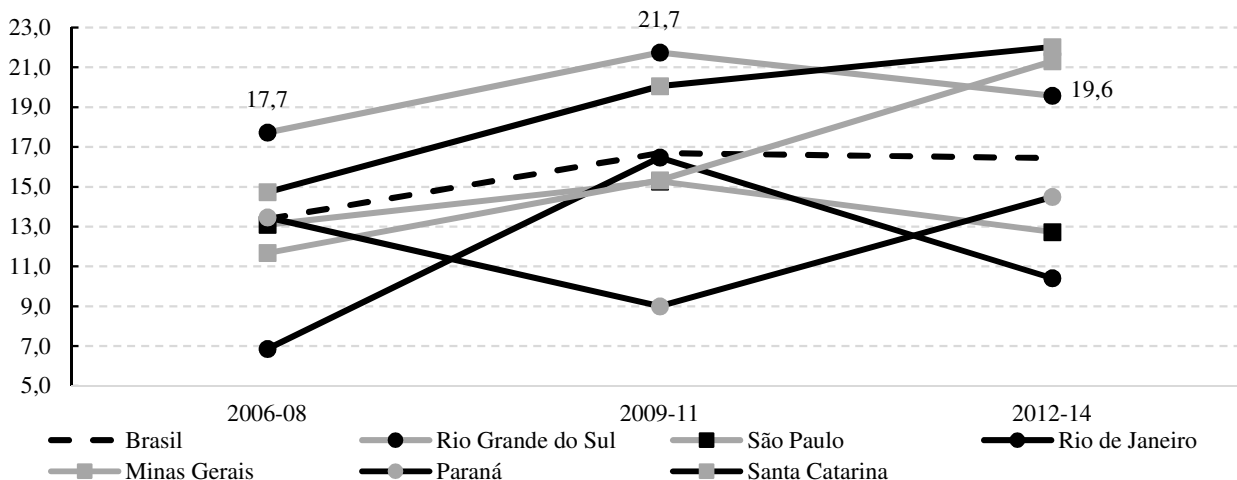
## 5. Padrão de interação universidade-empresa no Rio Grande do Sul

A partir da caracterização das capacitações tecnológicas dos atores, essa seção orienta-se para o padrão de interação universidade-empresa (U-E) no Rio Grande do Sul e sua evolução recente. Nessa direção, uma primeira constatação é de que, em média, uma baixa proporção das empresas industriais gaúchas e brasileiras que inovaram entre 2006-2008 e 2012-2014, inferior a 22%, atribuíram importância alta ou média às informações provenientes de Universidades ou Outros Centros de Ensino Superior (Gráfico 4).

Para efeito de comparação, no triênio 2012-2014, no Rio Grande do Sul, foram considerados mais relevantes para a inovação – com grau alto ou médio – os conhecimentos oriundos de clientes/consumidores (80,3%), de redes de informação informatizadas (76,7%), de fornecedores (70%) e de feiras/exposições (64,9%), sendo que essa distribuição havia sido semelhante nos dois triênios anteriores<sup>19</sup>. Apesar disso, as instituições de ensino superior vieram em um processo de aumento de importância como fonte de informação para as firmas industriais do Rio Grande do Sul do triênio 2006-2008 para o 2009-2011, com queda no período seguinte.

<sup>19</sup> Cálculos com base em dados das PINTECs 2008, 2011 e 2014 (IBGE, 2010, 2013 e 2016).

**Gráfico 4 - Participação das empresas industriais que implementaram inovações e que classificaram como de “alta” ou “média” relevância as informações provenientes de universidades ou outros centros de ensino superior**



Fonte: IBGE (2010, 2013 e 2016).

Notas: Foram consideradas as empresas das indústrias extrativas e de transformação com 10 ou mais empregados que implementaram produto e/ou processo novo ou substancialmente aprimorado. Os percentuais discriminados referem-se ao Rio Grande do Sul.

Dentre as empresas industriais que introduziram inovações nos três triênios, a proporção das que realizaram desenvolvimento em cooperação com algum parceiro considerado de alta ou média relevância apresentou crescimento, tanto no Rio Grande do Sul quanto na média nacional e em outros estados industrializados (Tabela 5). Em geral, constata-se no Brasil que as relações de cooperação relevantes dificilmente ultrapassam a marca de 20% do total das empresas que implementaram inovações. Nesse universo, uma fração ainda menor do total das firmas industriais inovadoras realizou algum tipo de desenvolvimento conjunto com universidades e institutos de pesquisas, avaliado como tendo importância alta ou média. Especificamente sobre a interação universidade-empresa, menos de 6% das empresas industriais inovadoras gaúchas e menos de 5% das brasileiras fizeram codesenvolvimento com universidades, considerado como tendo alta ou média importância nos três triênios. Entre as empresas industriais dos estados brasileiros selecionados, as do Rio Grande do Sul apresentaram as maiores taxas de inovações em parceria com universidades no período 2012-2014, de 5,2%.

**Tabela 5 - Empresas industriais que implementaram inovações em cooperação com outras organizações e que atribuíram grau alto ou médio de importância para a parceria por triênio, no Rio Grande do Sul e no Brasil**

Unidades da Federação	Total com Cooperação			Cooperação com Universidades e Institutos de Pesquisa		
	2006-08	2009-11	2012-14	2006-08	2009-11	2012-14
<b>Brasil</b>	<b>10,1</b>	<b>15,9</b>	<b>14,3</b>	<b>3,0</b>	<b>4,8</b>	<b>3,5</b>
Rio Grande do Sul	12,5	9,3	19,7	4,4	3,5	5,2
São Paulo	11,7	17,0	15,4	3,3	4,7	4,4
Rio de Janeiro	7,0	30,1	17,1	3,0	4,7	2,8
Minas Gerais	11,3	9,0	17,1	2,2	2,3	3,7
Paraná	11,2	20,3	14,1	2,4	1,8	1,4
Santa Catarina	7,3	20,2	12,6	2,7	8,1	1,9

Fonte: IBGE (2010, 2013 e 2016).

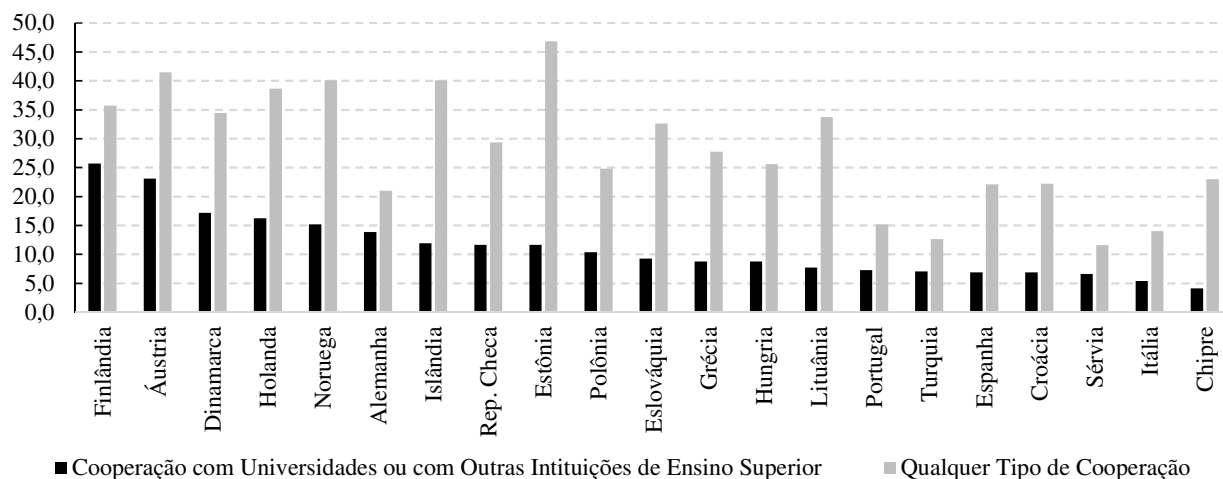
Notas: Foram consideradas as empresas das indústrias extrativas e de transformação com 10 ou mais empregados que implementaram produto e/ou processo novo ou substancialmente aprimorado no triênio analisado.

Para dimensionar o estágio atual das interações U-E no Rio Grande do Sul e no Brasil (Tabela 5), pode-se avaliá-lo em relação ao padrão observado em economias avançadas (Gráfico 5). Para países selecionados da Europa, em que pese diferenças metodológicas em relação à PINTEC<sup>20</sup>, em geral, as empresas industriais que inovaram em cooperação com outras organizações no triênio 2012-2014 representaram mais de 20% do total de inovadoras, acima do apresentado pelo estado (19,7%) e pelo País (14,3%) nesse período.

No que concerne à interação das empresas industriais com universidades e outras instituições de ensino superior, observa-se que, em geral, essa não chega a atingir 25% do total daquelas que inovaram no período nos casos selecionados. Ainda assim, a proporção de empresas industriais europeias que cooperaram com universidades para inovar foi significativamente maior do que aquelas apresentadas em 2012-2014 no Rio Grande do Sul (menos de 6%) e no Brasil (menos de 5%). Mesmo nações em patamar intermediário de industrialização, como Turquia ou Portugal, por exemplo, possuem uma maior proporção de interações universidade-empresa, do que aquelas apresentadas pelo RS e pelo BR. Essa proporção mais baixa no Brasil alinha-se ao diagnóstico de Suzigan e Albuquerque (2011, p.19) de que existam apenas alguns poucos “pontos de interação” universidade-empresa no País.

Além dos aspectos mencionados, em que pesem as variações causadas pelas características institucionais de cada sistema nacional de inovação, nota-se, nos casos europeus, que aquelas nações conhecidas como tendo maiores níveis de desenvolvimento e de capacitações científicas e tecnológicas apresentaram maiores taxas de inovações oriundas de cooperação U-E (Gráfico 5). Esse padrão observado na Europa está de acordo com o preconizado pela literatura, que constata que maiores capacitações científicas e tecnológicas das empresas geram uma maior ocorrência de cooperação U-E para desenvolvimento de inovações.

**Gráfico 5 - Empresas da indústria de transformação de países da Europa que implementaram inovações desenvolvidas com relações de cooperação em geral e com universidades ou outras instituições de ensino superior no triênio 2012-2014**



Fonte: *Community Innovation Survey 2014* (EUROSTAT, 2017).

Notas: Consideraram-se as empresas da indústria de transformação com 10 ou mais empregados dos países selecionados que implementaram inovações de produto ou de processo produtivo.

Conforme a literatura sobre interação universidade-empresa, a proximidade geográfica facilita a interação, o que também se constata no Brasil (GARCIA et al., 2014). No Rio Grande do Sul, 43,5% das interações U-E concentram-se em até 50 km de distância, o que aponta para a relevância de sistemas locais

<sup>20</sup> Na comparação entre a Tabela 5 e o Gráfico 5, o dado da PINTEC de número de empresas que interagiram com cada tipo de organização é computado e apresentado por categorias de importância “alta”, “média” ou “baixa ou não relevante” do parceiro para a inovação, enquanto que o dado da CIS apresenta apenas as empresas inovadoras que interagiram com cada tipo de parceiro. Infere-se que a soma das categorias de “alta” e “média” importância da PINTEC seja uma *aproximação* razoável ao dado disponibilizado na CIS.

de inovação para o desenvolvimento tecnológico de empresas industriais (COSTA et al., 2011). Inclusive, observa-se que as instituições de pesquisa locais, em geral, possuem algumas atividades de P&D e de extensão voltadas às empresas de sua região. Apesar disso, em recente pesquisa realizada pela Fundação de Economia e Estatística – FEE/RS sobre determinantes da competitividade de empresas em onze aglomerações produtivas e arranjos produtivos locais do Rio Grande do Sul, constatou-se uma baixa interação U-E em alguns desses casos (MACADAR; COSTA, 2016). Em geral, nos casos estudados, foram as empresas de portes médio e grande, dotadas de maiores capacitações e próximas à fronteira tecnológica em suas atividades<sup>21</sup>, que interagiram mais com as universidades.

Logo, a baixa proporção de interações U-E em relação ao observado entre nações industrializadas indica que as relativamente baixas capacitações tecnológicas na indústria do Rio Grande do Sul e do Brasil, sobretudo em número de pesquisadores pós-graduados, seja um dos principais fatores explicativos para esse comportamento, por gerar, entre outros, um distanciamento cognitivo entre esses agentes. Nesse sentido, Rapini, Chiarini e Bittencourt (2017) também constata a nível nacional que o *déficit* de pessoal qualificado, tanto na indústria quanto nas universidades, seja um dos impeditivos para a ocorrência de interações. Entretanto, mesmo que a cooperação das empresas industriais para inovar com as universidades e outras instituições de ensino superior seja relativamente baixa, ela vem aumentando, passando de 4,4% em 2006-2008 para 5,2% em 2012-2014 no Rio Grande do Sul; e de 3,0% para 3,5% no Brasil.

Igualmente, constata-se que tanto o número total de grupos de pesquisa quanto daqueles que interagiram com empresas se elevaram no período<sup>22</sup>. Em termos de taxa de crescimento<sup>23</sup> de 2008 para 2014, a do total dos grupos de pesquisa foi de 43,9% no Rio Grande do Sul e de 55,4% no Brasil, enquanto que a daqueles com relacionamento com empresas foi de 171,3% no estado e de 242,9% no País. Logo, verifica-se que a proporção de grupos de pesquisa com interação com empresas ganhou importância no Brasil, sendo que, nesse quesito, o estado ficou com a terceira posição entre os estados mais industrializados em 2010 e 2014 (Tabela 6). Isso indica estar havendo uma maior aproximação entre universidades e empresas nessas áreas geográficas, o que também pode ser creditado à evolução dos instrumentos de política científica e tecnológica, sobretudo na última década (DE NEGRI et al., 2015).

Em 2008 e em 2010, as grandes áreas do conhecimento com maior número de grupos de pesquisa que mantiveram interações com empresas no Rio Grande do Sul foram as engenharias, seguidas das ciências agrárias e as da saúde (Tabela 7). Essas também foram as mais acessadas pelas empresas gaúchas e brasileiras, apresentando maior densidade de interações por grupo de pesquisa, o que já era de se esperar, dada a natureza do conhecimento em que essas áreas estão envolvidas. Como constatado anteriormente, alguns programas de pós-graduação em engenharias, ciências biológicas e da saúde destacam-se como sendo de excelência. Esse padrão indica existir uma associação próxima entre as expertises e linhas de pesquisa das universidades e institutos de pesquisa e a estrutura da matriz produtiva do estado, a qual possui expressiva participação dos setores de base metalomecânica (automotivo, máquinas e equipamentos, produtos de metal e outras) e da cadeia produtiva de fabricação de alimentos (atividades do setor primário e agroindústrias).

Segundo a indicação de líderes de grupos de pesquisa no Rio Grande do Sul, os motivos predominantes para interação com empresas industriais brasileiras foram majoritariamente a “pesquisa científica com considerações de uso imediato dos resultados”, seguida pela “transferência de tecnologia desenvolvida pelo grupo para o parceiro” e pela “pesquisa científica sem considerações de uso imediato dos resultados”, cujo somatório<sup>24</sup> resulta em cerca de 62% do total informado em 2008 e em 2010 (Tabela 8). Em especial, essa distribuição de motivos para a interação U-E observada no Rio Grande do Sul se assemelha ao padrão

<sup>21</sup> Também foi constatado que algumas dessas empresas integravam verticalmente a fabricação de itens de maior sofisticação devido à falta de fornecedores capacitados.

<sup>22</sup> Os dados do Diretório de Grupos de Pesquisa do CNPq se referem às interações com empresas que são declaradas pelo líder do grupo de pesquisa, de forma autodeclaratória. Pode existir algum grau de subestimação no número de interações, sobretudo em determinadas áreas, por essas não terem sido informadas. Contudo, considera-se que o número de interações informadas se constitua em uma *proxy* fidedigna para o comportamento do total de interações entre grupos de pesquisa e empresas.

<sup>23</sup> Cálculos com base nos dados da Tabela 6.

<sup>24</sup> Cálculos com base em dados da Tabela 8.

identificado na literatura. Isto é, em geral, os tipos de relacionamentos buscados pelas empresas industriais com organizações de pesquisa gaúchas objetivam complementar lacunas em suas capacidades tecnológicas ou solucionar problemas no desenvolvimento de inovações, bem como obter resultados no curto prazo.

**Tabela 6 - Número de grupos de pesquisa que relataram pelo menos um relacionamento com empresas, segundo a Unidade da Federação onde o grupo está localizado**

UF do grupo	2008			2010			2014		
	Com relação (R)	Total de grupos (T)	% (R)/(T)	Com relação (R)	Total de grupos (T)	% (R)/(T)	Com relação (R)	Total de grupos (T)	% (R)/(T)
Brasil	2.726	22.797	12,0	3.506	27.523	12,7	9.348	35.424	26,4
Rio Grande do Sul	345	2.304	15,0	404	2.677	15,1	936	3.315	28,2
São Paulo	576	5.938	9,7	747	6.359	11,7	1.922	7.278	26,4
Rio de Janeiro	286	2.779	10,3	358	3.313	10,8	1.216	4.147	29,3
Minas Gerais	300	2.135	14,1	396	2.848	13,9	920	3.482	26,4
Paraná	242	1.915	12,6	302	2.264	13,3	686	3.011	22,8
Santa Catarina	184	1.070	17,2	239	1.263	18,9	489	1.612	30,3
Outros	793	6.656	11,9	1.060	8.799	12,0	3.179	12.579	25,3

Fonte: CNPq (2017b).

Notas: Foram contabilizados os grupos de pesquisa e aqueles com no mínimo um relacionamento com empresas. Um grupo pode manter relacionamento com mais de uma empresa no ano. O CNPq não realizou recenseamento no ano de 2012.

**Tabela 7 - Interações de grupos de pesquisa do Rio Grande do Sul com empresas brasileiras por grande área do conhecimento**

Grande área do conhecimento	2008					2010				
	Grupos		Empresas		Densidade	Grupos		Empresas		Densidade
	Nº	%	Nº	%	(E/G)	Nº	%	Nº	%	(E/G)
<b>TOTAIS</b>	345	100,0	683	100,0	1,98	404	100,0	829	100,0	2,05
- Engenharias	113	32,8	288	42,2	2,55	128	31,7	311	37,5	2,43
- Ciências Agrárias	66	19,1	112	16,4	1,70	82	20,3	175	21,1	2,13
- Ciências Exatas e da Terra	40	11,6	74	10,8	1,85	45	11,1	87	10,5	1,93
- Ciências da Saúde	40	11,6	91	13,3	2,28	53	13,1	113	13,6	2,13
- Ciências Biológicas	34	9,9	47	6,9	1,38	41	10,1	58	7,0	1,41
- Outras	52	15,1	71	10,4	1,37	55	13,6	85	10,3	1,55

Fonte: CNPq (2017a).

Notas: As empresas mencionadas não estão localizadas, necessariamente, no mesmo estado do grupo. O CNPq descontinuou o Plano Tabular do Diretório de Grupos de Pesquisa após 2010.



**Tabela 8 - Tipos de relacionamento predominantes entre grupos de pesquisa do Rio Grande do Sul e empresas industriais brasileiras**

Tipos de relacionamentos	2008		2010	
	Nº	%	Nº	%
Total	560	100,0	556	100,0
- Pesquisa científica com considerações de uso imediato dos resultados	175	31,3	183	32,9
- Transferência de tecnologia desenvolvida pelo grupo para o parceiro	101	18,0	94	16,9
- Pesquisa científica sem considerações de uso imediato dos resultados	74	13,2	72	12,9
- Fornecimento, pelo parceiro, de insumos materiais para as atividades de pesquisa do grupo sem vinculação a um projeto específico de interesse mútuo	28	5,0	53	9,5
- Atividades de engenharia não rotineira inclusive o desenvolvimento de protótipo cabeça de série ou planta-piloto para o parceiro	44	7,9	30	5,4
- Outros	138	24,6	124	22,3

Fonte: CNPq (2017a).

Notas: Foram admitidas até três formas predominantes de relacionamento. As empresas mencionadas não estão localizadas, necessariamente, no mesmo estado do grupo. O CNPq descontinuou o Plano Tabular do Diretório de Grupos de Pesquisa após 2010.

Dentre as empresas gaúchas que cooperaram com grupos de pesquisa no estado ou no País, as industriais representaram 40,5% em 2008 e 40,3% em 2010. Os setores de atividades industriais que individualmente mostraram o maior número de interações em relação ao total da indústria foram o de “máquinas e equipamentos” e o de “produtos químicos” (Tabela 9). No agregado de atividades industriais, o destaque pertence às de base metalomecânica.

**Tabela 9 - Empresas industriais do Rio Grande do Sul que cooperaram com grupos de pesquisa no Brasil por setor de atividade, total de empresas no setor e proporção de interações U-E**

Divisão de Atividade CNAE 2.0	2008				2010			
	Cooperaram		Total	% - (C/T)	Cooperaram		Total	% - (C/T)
	Nº(C)	(%)	Nº(T)		Nº(C)	(%)	Nº(T)	
Total	632	N/A	424.775	0,15	611	N/A	452.891	0,13
-Indústrias	256	100,0	50.845	0,50	246	100,0	52.931	0,46
--Prod. farmoquímicos e farmacêuticos	8	3,1	71	11,27	8	3,3	65	12,31
--Equip. info., prod. eletrônicos e ópticos	26	10,2	457	5,69	14	5,7	473	2,96
--Produtos químicos	33	12,9	989	3,34	27	11,0	974	2,77
--Automotivo	5	2,0	761	0,66	14	5,7	772	1,81
--Metalurgia	14	5,5	690	2,03	11	4,5	649	1,69
--Máq., aparelhos e mater. elétricos	9	3,5	629	1,43	8	3,3	618	1,29
--Máquinas e equipamentos	45	17,6	2.471	1,82	34	13,8	2.655	1,28
--Borracha e plástico	21	8,2	1.957	1,07	18	7,3	1.949	0,92
--Produtos alimentícios	14	5,5	6.023	0,23	30	12,2	6.384	0,47
--Produtos de metal	29	11,3	6.626	0,44	17	6,9	7.190	0,24
--Couros e calçados	12	4,7	6.492	0,18	15	6,1	6.420	0,23
--Outros	40	15,6	23.679	0,17	50	20,3	24.782	0,20

Fonte: CNPq (2017a) e IBGE (2017a).

Notas: A área geográfica da empresa foi informada pelo líder do grupo e pesquisa. Os grupos não estão localizados, necessariamente, no Rio Grande do Sul. O CNPq descontinuou o Plano Tabular do Diretório de Grupos de Pesquisa após 2010.

Um indicador mais representativo da importância setorial das interações U-E consiste na sua proporção em relação ao total de empresas do respectivo setor no estado. Sob essa ótica, aqueles com maiores percentuais de empresas que cooperaram com grupos de pesquisa foram os de “produtos farmacêuticos e químicos”; de “equipamentos de informática, produtos eletrônicos e ópticos”; de “produtos químicos”; de “metalurgia”; e de “máquinas e equipamentos”. Nota-se que essa distribuição se aproxima daquela observada na taxonomia de intensidade de esforço tecnológico da OCDE, pois a maior densidade de interações ocorre nos setores classificados como de “alta” ou de “média-alta”, com exceção da metalurgia (média-baixa). Em outras palavras, os setores que em média possuem maior intensidade tecnológica foram também os que, em geral, apresentaram maior proporção de empresas com interações U-E. Assim, a baixa participação relativa de setores *high-tech* (33,7%), bem como de empresas de médio e de grande portes (3,4%), indica que a estrutura da matriz industrial do estado seja um dos determinantes da baixa ocorrência de interações universidade-empresa. Observa-se que Bastos e Britto (2017) fazem constatação semelhante para o Brasil. Por fim, nota-se que o número de empresas industriais que mantiveram relacionamentos reduziu de 2008 para 2010, em 3,9% para as industriais e em 3,3% para o total, o que pode ser atribuído às incertezas associadas à crise financeira mundial (2008-2009).

## 6. Considerações finais

A interação universidade-empresa é um elemento importante do funcionamento de Sistemas de Inovação e da evolução do conhecimento científico e tecnológico na sociedade, afetando em âmbito mais amplo o processo de desenvolvimento econômico. Os resultados encontrados neste estudo visam fornecer informações para orientar a elaboração de política industrial e de ciência e tecnologia, bem como a tomada de decisão empresarial, visando ao fortalecimento do SI do Rio Grande do Sul e seus efeitos positivos sobre a competitividade das empresas e o crescimento econômico.

Quando avaliadas no período 2008-2014, observa-se que o desempenho das empresas industriais, em termos de suas capacitações tecnológicas (investimentos em inovação, recursos humanos e infraestrutura à P&D), e o das universidades, quanto ao seu potencial para a formação de recursos humanos e de realização de pesquisas, encontram-se posicionadas acima da média nacional. Além disso, observa-se em diversos indicadores das capacitações de empresas industriais e de universidades e outras instituições de pesquisa, em geral, uma tendência de melhoria absoluta e relativa no período 2008-2014, tanto no estado quanto no País. Essa tendência pode ser atribuída às políticas de educação superior, de ciência, tecnologia e industriais e de inovação aplicadas no País desde 2003. Esse processo de melhora tem se mostrado relativamente mais intenso no estado.

Entretanto, em que pese esses destaques positivos, é de se indicar que o sistema de inovação do Rio Grande do Sul ainda se encontrava em 2014 em patamar inferior àquele apresentado por países industrializados, situando-se em nível intermediário de desenvolvimento. Isso mostra que existem espaços para as empresas industriais e as universidades aprimorarem suas capacitações científicas e tecnológicas, o que fortaleceria o SI estadual. Contudo, há uma assimetria entre esses agentes, cuja situação aparenta ser mais crítica para as empresas industriais do que para as universidades. Isso se evidencia em alguns indicadores importantes para o processo de desenvolvimento de inovações, cujos percentuais em 2014 podem ser considerados como relativamente baixos, como no caso da intensidade dos dispêndios em atividades internas de P&D (0,65%), de participação de pesquisadores pós-graduados (5,1%) e de empresas industriais que inovaram com gastos contínuos em P&D interno (18%). Esse fraco desempenho acaba afetando negativamente a competitividade na indústria do estado, sobretudo, em seus segmentos de maior intensidade tecnológica.

Pode-se considerar que a interação universidade-empresa no Rio Grande do Sul ainda se apresenta pouco frequente, uma vez que menos de 22% das empresas industriais inovadoras atribuíram alta ou média relevância às informações provenientes de universidades e demais instituições de pesquisa nos três triênios analisados, assim como menos de 6% desse total declarou ter desenvolvido inovações em parceria com estas

organizações. Essa baixa intensidade de interações universidade-empresa, embora acima da média nacional, ainda está bem aquém daquela constatada em economias desenvolvidas. Como aspecto positivo, porém, nos últimos anos, percebe-se uma tendência de elevação da importância da interação U-E, indicando estar ocorrendo uma maior aproximação das universidades e instituições de pesquisa com empresas industriais no Rio Grande do Sul. Ademais, o número de grupos de pesquisa que interagiu com empresas se elevou a uma taxa superior à do total de grupos, ganhando maior representatividade. Nesse processo, o padrão observado nas interações U-E manteve-se como sendo semelhante àquele apontado na literatura, isto é, em geral, as empresas industriais do estado procuram este tipo de relação para complementar alguma lacuna em suas capacitações ou para solucionar problemas tecnológicos, buscando obter resultados de aplicação imediata. Portanto, ainda existe um baixo acesso pelas empresas aos conhecimentos e recursos gerados pelas universidades, de modo que há espaço para uma maior aproximação desses agentes e para a intensificação de suas relações.

A menor proporção de interações U-E em relação aos países mais industrializados indica que existem alguns fatores atuando como limitadores. Dentre os principais que foram identificados neste estudo, sem ter a pretensão de esgotar essa listagem, entende-se que sejam: (1) as relativamente baixas capacitações tecnológicas das empresas industriais gaúchas, sobretudo em qualificação de recursos humanos, com reduzido número e baixa proporção de pesquisadores pós-graduados; (2) a pequena proporção de empresas que executam atividades de P&D interno em caráter contínuo; (3) a conhecida distância da fronteira tecnológica internacional em alguns segmentos de “alta” e de “média-alta” intensidade tecnológica – tais como na indústria farmacêutica, na produção de tecnologias de informação e comunicação e de componentes eletrônicos; (4) a estrutura da matriz produtiva industrial, com predominância de atividades *low-tech*, somando-se à baixa incidência de médias e grandes empresas nessa estrutura; e (5) a relativa baixa proporção do quadro de docentes em exercício nas instituições de ensino superior que possui atuação em pesquisa.

## Referências

- ANTUNES JÚNIOR, J. A. V.; HORN, C. H.; PELLEGRIN, I.; VAZ, I. E. A. (orgs.). *Remando contra a maré: a política industrial e de desenvolvimento econômico no Rio Grande do Sul (2011-2014)*. Porto Alegre/RS: Bookman, 376p., 2017.
- BASTOS, C. P.; BRITTO, J. Inovação e geração de conhecimento científico e tecnológico no Brasil: uma análise dos dados de cooperação da PINTEC segundo porte e origem de capital. *Revista Brasileira de Inovação*, v. 16, n. 1, p. 35-62, 2017.
- BERNARDINI, R.; CONCEIÇÃO, C. S.; TARTARUGA, I. P.; OLIVEIRA, L. L. S.; COSTA, R. M.; PICHLER, W. A.; ROSSI, A. S. (eds.). *Ciência, tecnologia e inovação no Rio Grande do Sul: indicadores selecionados 2014*. Porto Alegre/RS/BR: FEE, 80p., 2014.
- BRASIL. Ministério do Trabalho e Emprego (MTE). *Relação de Informações Sociais (RAIS)*. Brasília, DF, 2017. Acesso em: 05 abr. 2017.
- BURCHARTH, A. L. L. A. What drives the formation of technological cooperation between university and industry in less-developed innovation systems? evidence from Brazil. *Revista Brasileira de Inovação*, v. 10, n. 1, p. 101-128, 2011.
- CAMPOS, A. L. A review of the influence of long-term patterns in research and technological development (R&D) formalisation on university-industry links. *Revista Brasileira de Inovação*, v. 9, n. 2, p. 379-410, 2010.
- CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. *Sobre a avaliação*. 2017. Disponível em <http://www.capes.gov.br/avaliacao/sobre-a-avaliacao>. Acesso em: 20 abr. 2017a.


- CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. *GEOCAPES: Sistema de Informações Georreferenciadas* – 2015. Disponível em: <http://geocapes.capes.gov.br/geocapes2/> Acesso em: 27 fev. 2017b.
- CNPq. Conselho Nacional de Desenvolvimento Científico e Tecnológico. Diretório dos Grupos de Pesquisa no Brasil. *Plano Tabular*. 2017. Disponível em: <http://plsql1.cnpq.br/planotabular/>. Acesso em: 10 jan. 2017a.
- CNPq. Conselho Nacional de Desenvolvimento Científico e Tecnológico. Diretório dos Grupos de Pesquisa no Brasil. *Súmula Estatística*. 2017. Disponível em: <http://lattes.cnpq.br/web/dgp/grupos1>. Acesso em: 10 jan. 2017b.
- COSTA, A. B.; RUFFONI, J.; PUFFAL, D. Proximidade geográfica e interação universidade-empresa no Rio Grande do Sul. *Revista de Economia*, v. 37, p. 213-238, 2011.
- DE NEGRI, F.; CAVALCANTE, L. R.; ALVES, P. A. *University-enterprise interaction in Brazil: the role of the public research infrastructure*. Instituto de Pesquisa Econômica Aplicada, 2015. (IPEA Discussion Paper, n. 0206).
- EDQUIST, C. Systems of innovation approaches – their emergence and characteristics. In: EDQUIST, C. (ed.). *Systems of innovation: technologies, institutions and organization*. Nova York: Routledge. p.1-35, 2011.
- EUROSTAT. *Community Innovation Survey 2014*. Luxembourg: European Commission, 2017. Disponível em: <https://ec.europa.eu/eurostat/data/database>. Acesso em: 07 mar. 2017.
- EUROSTAT. Glossary: High-tech classification of manufacturing industries. Eurostat Statistics Explained. Disponível em: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:High-tech\\_classification\\_of\\_manufacturing\\_industries](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:High-tech_classification_of_manufacturing_industries). Acesso em: 23 ago. 2019.
- FEE. Fundação de Economia e Estatística Siegfried Emanuel Heuser. *Estrutura do valor adicionado por setores de atividade*. Porto Alegre: FEE, Centro de Indicadores Econômicos e Sociais, Núcleo de Contas Regionais. 2016.
- FREEMAN, C. Formal scientific and technical institutions in the national systems of innovation. In: LUNDEVALL, B.-A. (ed.). *National systems of innovation: toward a theory of innovation and interactive learning*. Londres: Anthem Press. p.173-192, 2010.
- FREEMAN, C. Technological infrastructure and international competitiveness. In: FREEMAN, C. (ed.). *Systems of Innovation: Selected essays in evolutionary economics*. Cheltenham: Edward Elgar Publishing Limited, p.6-37, 2008.
- GARCIA, R.; ARAUJO, V. A.; MASCARINI, S.; SANTOS, E. G. Efeitos da qualidade da pesquisa acadêmica sobre a distância geográfica das interações universidade-empresa. *Estudos Econômicos*, v. 44, n. 1, p. 105-132, 2014.
- HODGSON, G. M. The approach of institutional economics. *Journal of Economic Literature*, v. 36, n. 1, p. 166-192, 1998.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Estatísticas do Cadastro Central de Empresas – CEMPRE*. In: Sistema IBGE de Recuperação Automática - SIDRA, 2017. Disponível em: <https://sidra.ibge.gov.br/pesquisa/cempre/quadros/brasil/2014>. Acesso em: 12 abr. 2017a.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Pesquisa de Inovação (PINTEC)-2008*. Rio de Janeiro, 2010.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Pesquisa de Inovação (PINTEC)-2011*. Rio de Janeiro, 2013.

- IBGE. Instituto Brasileiro de Geografia e Estatística. *Pesquisa de Inovação (PINTEC)-2014*. Rio de Janeiro, 2016.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Pesquisa Industrial Anual Empresa – 2014*. Sistema de Recuperação Automática - SIDRA, 2017. Disponível em: <https://sidra.ibge.gov.br/pesquisa/pia-empresa/quadros/brasil/2014>. Acesso em: 06 fev. 2017b.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Produto Interno Bruto dos Municípios*. Sistema IBGE de Recuperação Automática - SIDRA, 2017. Disponível em: <https://sidra.ibge.gov.br/pesquisa/pib-munic/tabelas>. Acesso em: 12 abr. 2017c.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sinopse estatística da educação superior: 2008 a 2014*. Brasília: INEP, 2017. Disponível em: <http://portal.inep.gov.br/web/guest/sinopses-estatisticas-da-educacao-superior>. Acesso em: 21 fev. 2017.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Censo da Educação Superior: 2010 a 2014*. Brasília: INEP, 2019. Disponível em: <http://portal.inep.gov.br/web/guest/microdados>. Acesso em: 15 ago. 2019.
- IPEA. Instituto de Pesquisa Econômica Aplicada. *IPEADATA: IGP-DI - geral - índice (ago. 1994=100)*. Fundação Getúlio Vargas (FGV). 2017. Disponível em: <http://www.ipeadata.gov.br>. Acesso em: 10 fev. 2017.
- LUNDEVALL, B.-A. Introduction. In: LUNDEVALL, B.-A (ed.). *National systems of innovation: toward a theory of innovation and interactive learning*. Londres: Anthem Press. p.1-19, 2010.
- MACADAR, B. M.; COSTA, R. M. D. Uma avaliação geral das aglomerações produtivas e dos APLs selecionados no Rio Grande do Sul. In: MACADAR, B. M.; COSTA, R. M. (org.). *Agglomerações e arranjos produtivos locais no Rio Grande do Sul*. Porto Alegre: FEE, p.566-584, 2016.
- MARQUES, M. D.; ROSELINO, J. E.; MASCARINI, S. *Análise da aderência das taxonomias industriais à realidade da indústria de transformação brasileira*. São Paulo: Blucher, p.1481-1497, 2018.
- MAZZUCATO, M. *O estado empreendedor: desmascarando o mito do setor público vs. setor privado*. Tradução de Elvira Serapicos. São Paulo: Portfólio-Penguin, 1ª ed., 340p., 2014.
- MAZZUCATO, M.; PENNA, C. *The Brazilian innovation system: a mission-oriented policy proposal. Temas estratégicos para o desenvolvimento do Brasil*. Avaliação de Programas em CT&I. Brasília: Centro de Gestão e Estudos Estratégicos, 119 p., 2016.
- MORCEIRO, P. C. *Nova classificação de intensidade tecnológica da OCDE e a posição do Brasil. Temas de Economia Aplicada*. Disponível em: <http://downloads.fipe.org.br/content/downloads/publicacoes/bif/bif461-8-13.pdf>. Acesso em: 23 ago. 2019.
- MOWERY, D. C.; SAMPAT, B. N. Universities in national innovation systems. In: FAGERBERG, J.; MOWERY, D. C.; NELSON, R. R. (eds.). *The Oxford Handbook of Innovation*. Nova York: Oxford University Press Inc., p.209-239, 2005.
- NELSON, R.; ROSENBERG, N. Science, technological advance and economic growth. In: CHANDLER, A.; HAGSTRÖM, P.; SÖVELL, Ö. *The dynamic firm: the role of technology, strategy, organization and regions*. Oxford: Oxford University Press, p.45-59, 1999.
- OCDE. Organização para a Cooperação e Desenvolvimento Econômico. *ISIC Rev.3 Technology Intensity Definition: classification of manufacturing industries into categories based on R&D intensities*. Paris, 2011.

- OCDE. Organização para a Cooperação e Desenvolvimento Econômico. OECD Stat. *Business Enterprise R-D Expenditure by Industry (ISIC 4)*. Paris, 2017. Disponível em: <http://stats.oecd.org>. Acesso em: 13 jun. 2017a.
- OCDE. Organização para a Cooperação e Desenvolvimento Econômico. OECD Stat. *Gross Domestic Product (GDP)*. Paris, 2017. Disponível em: <http://stats.oecd.org>. Acesso em: 13 jun. 2017b.
- RAPINI, M. S.; CHIARINI, T.; BITTENCOURT, P. F. Obstacles to innovation in Brazil: the lack of qualified individuals to implement innovation and establish university-firm interactions. *Industry and Higher-Education*, v. 31, n. 3, p. 168-183, 2017.
- SCHAEFFER, P. R.; RUFFONI, J.; PUFFAL, D. Razões, benefícios e dificuldades da interação universidade-empresa. *Revista Brasileira de Inovação*, v. 14, n. 1, p. 105-134, 2015.
- SEBRAE. Serviço Brasileiro de Apoio às Micro e Pequenas Empresas. *Critério para classificação de empresas: MEI – ME – EPP*. 2017. Disponível em: <http://www.sebrae-sc.com.br/leis/default.asp?vcdtexto=4154>. Acesso em: 07 mar. 2017.
- SMITH, K. Economic infrastructures and innovation systems. In: EDQUIST, C. (ed.). *Systems of innovation: technologies, institutions and organization*. Nova York: Routledge. p.86-106, 2011.
- SUZIGAN, W.; ALBUQUERQUE, E. M. A interação entre universidades e empresas em perspectiva histórica no Brasil. In: SUZIGAN, W.; ALBUQUERQUE, E. M.; CARIO, S. A. F. (Eds.). *Em busca da inovação: interação universidade-empresa no Brasil*. Belo Horizonte: Autêntica Editora, (Coleção Economia Política e Sociedade, 3), p.17-44, 2011.
- SUZIGAN, W.; ALBUQUERQUE, E. M.; CARIO, S. A. F. Introdução. In: SUZIGAN, W.; ALBUQUERQUE, E. M.; CARIO, S. A. F. (eds.). *Em busca da inovação: interação universidade-empresa no Brasil*. Belo Horizonte: Autêntica Editora, Coleção Economia Política e Sociedade, p.3-16, 2011.
- TOPUNIVERSITIES. *QS World University Rankings® 2015*. 2017. Disponível em: <https://www.topuniversities.com/university-rankings/world-university-rankings/2015>. Acesso em: 03 abr. 2017.

## ORCID

Rodrigo Morem da Costa  <https://orcid.org/0000-0002-9097-644X>

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.