

- DEFORESTATION AND HUMAN DEVELOPMENT IN THE BRAZILIAN AGRICULTURAL FRONTIER: AN ENVIRONMENTAL KUZNETS CURVE FOR MATOPIBA** 161-182
Pedro Henrique Batista de Barros, Alysson Luiz Stege
- ANALYSIS OF THE EFFECT OF TERRITORIAL FACTORS ON REGIONAL INNOVATION IN THE STATE OF SÃO PAULO, BRAZIL** 183-200
Suelene Mascarini, Renato Garcia, José Eduardo Roselino
- DETERMINANTES DA INOVAÇÃO AMBIENTAL: UMA ANÁLISE EMPÍRICA SOBRE A INDÚSTRIA DE TRANSFORMAÇÃO DE MINAS GERAIS** 201-222
Maria Alice Ferreira, Maria Ivanilde Pereira Santos, Emerson Costa dos Santos
- DIFERENCIAIS DE CUSTOS DE VIDA ENTRE OS ESTADOS BRASILEIROS: UMA APLICAÇÃO DE PREÇOS HEDÔNICOS A PARTIR DOS DADOS DA PNAD** 223-242
Marcello Luiz de Souza Junior, Alexandre Nunes de Almeida, João Gomes Martines Filho, Maria Cristina Galvão
- SERVIÇOS DE SANEAMENTO BÁSICO EM MINAS GERAIS E SEUS DETERMINANTES LOCACIONAIS, DEMOGRÁFICOS E SOCIOECONÔMICOS** 243-267
Juliana Brito de Oliveira, Gabriel Teixeira Ervilha
- EFEITOS DE LONGO PRAZO DO AJUSTE FISCAL SOBRE A ECONOMIA SERGIPANA** 268-287
Luiz Carlos de Santana Ribeiro, Kênia Barreiro Souza
- SUICÍDIO NO RN E SUA RELAÇÃO COM DETERMINANTES ESPACIAIS, URBANIZAÇÃO, DESENVOLVIMENTO E OUTROS FATORES SOCIOECONÔMICOS** 288-308
Stefany Silva Amaral

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*.

**DEFORESTATION AND HUMAN DEVELOPMENT IN THE BRAZILIAN
AGRICULTURAL FRONTIER: AN ENVIRONMENTAL KUZNETS CURVE FOR
MATOPIBA***

Pedro Henrique Batista de Barros

Doutorando no Departamento de Economia da Faculdade de Economia e Administração da Universidade de São Paulo (USP/FEA)

E-mail: pedrohbarros@usp.br

Alysson Luiz Stege

Professor Adjunto do Departamento de Economia da Universidade Estadual de Ponta Grossa (UEPG)

E-mail: alyssonstege@gmail.com

ABSTRACT: This paper aims to estimate an Environmental Kuznets Curve (EKC) for the current Brazilian agricultural frontier, located in the region known as MATOPIBA. The question to be answered can be summarized as: how does human development affect the region's environment, captured by deforestation? Specifically, we analyzed the presence of spatial dependence and heterogeneity, as well as the existence of clusters between the 337 municipalities of MATOPIBA in 2010, using exploratory spatial data analysis (ESDA) and spatial econometrics. We identified the presence of spatial dependence for deforestation, which led to the incorporation of this effect into the econometric modeling, which resulted in the SLX as the best spatial model. In addition, we got an inverted-U shape for the EKC; thus, deforestation increases until a certain threshold, as the region develops, from which it begins to fall. The "turning point", where development reaches its maximum impact on the environment, is a Human Development Index of 0.57 and 28.18% of the municipalities are below this value, which highlights environmental concerns, since their development could boost degradation. To worsen this scenario, we identified many variables, especially related to the agricultural frontier expansion, which induces deforestation in MATOPIBA.

Keywords: Environmental Kuznets Curve (EKC); Brazilian agricultural frontier; MATOPIBA; Spatial dependency.

JEL Codes: Q01; Q56.

**DESMATAMENTO E DESENVOLVIMENTO HUMANO NA FRONTEIRA
AGRÍCOLA BRASILEIRA: UMA CURVA AMBIENTAL DE KUZNETS PARA O
MATOPIBA**

RESUMO: Este trabalho busca estimar uma Curva Ambiental de Kuznets (CKA) para a atual fronteira agrícola brasileira, localizada na região conhecida como MATOPIBA. A pergunta a ser respondida pode ser resumida em: como o desenvolvimento humano impacta o meio ambiente da região, captada pelo desmatamento? De forma específica, analisou-se a presença de dependência e heterogeneidade espacial, bem como a existência de clusters entre os 337 municípios do MATOPIBA em 2010, utilizando-se da análise exploratória de dados espaciais (AEDE) e econometria espacial. Identificou-se a presença de dependência espacial, fato que levou à incorporação desse efeito na modelagem econométrica, procedimento que resultou no SLX como melhor modelo. O formato encontrado para a CKA foi de "U" invertido, isto é, o desenvolvimento tende a elevar o desmatamento até certo patamar, a partir do qual a relação se inverte, resultando na queda do impacto ambiental. O "ponto de virada", onde o desenvolvimento atinge seu máximo impacto, é um IDH de 0.57, sendo que 28.18% dos municípios estão abaixo desse valor, fato que levanta preocupações ambientais, pois o desenvolvimento pode acelerar o desmatamento nesses locais. Para piorar esse cenário, identificou-se diversas variáveis, principalmente relacionadas ao avanço da fronteira agrícola, que induzem o desmatamento no MATOPIBA.

Palavras-chave: Curva de Kuznets Ambiental (CKA); Fronteira agrícola brasileira; MATOPIBA; Dependência espacial.

Classificação JEL: Q01; Q56.

1. Introduction

MATOPIBA is the main region of the current Brazilian agricultural frontier. According to Araújo et al. (2019), agricultural frontiers are regions dominated by natural vegetation and which are facing intensive agriculture-related land occupation. The term MATOPIBA refers to the initial syllables of the states that comprise this region: Maranhão, Tocantins, Piauí and Bahia. The agricultural frontier in the area has been expanding due to the implementation of technologies adapted to local conditions, which allows an increase in agricultural productivity. The low price of land and the easy adoption of technologies have attracted investments and intensified their occupation, resulting in a significant growth in the local production, mainly of grains like soybeans and maize (MIRANDA et al., 2014; BRAGANÇA, 2018; ARAÚJO et al., 2019).

Recognizing the region's strategic importance for the Brazilian agribusiness, the country's government created the Decree No. 8447 on May 6, 2015 with the main objective of establishing an Agricultural Development Plan for MATOPIBA seeking to guide federal projects and actions specifically for the region (BRASIL, 2015).

The intensive occupation of MATOPIBA for agricultural production began in the 1980s, and the process is still underway. The existence of underutilized land, where low-productivity techniques are adopted, allows the increase in production with adoption of technology. In addition, there are areas where native forests of the Cerrado biome prevail, and their incorporation into the dynamic areas of MATOPIBA as the agricultural frontier advances is possible (BATISTELLA; VALLADARES, 2009; STUDTE, 2008; BOLFE et al., 2016; ARAÚJO et al., 2019). According to Chagas and Andrade (2017), economic agents located in agricultural frontiers face a considerable opportunity cost by not clearing forest areas for economic use, in light of possible present and future returns.

According to Bolfe et al. (2016) and Bragança (2018), the increase in the use of high-capacity land, combined with the adoption of productivity-enhancing technologies, has enabled the region to present significant increases in its production levels and, consequently, economic growth. Thus, the existence of underutilized and/or not yet occupied land, together with the agricultural frontier expansion and the economic development in MATOPIBA, may result, in the coming years, in a process of deforestation. In fact, according to Borges and Santos (2009), the current deforestation of the Cerrado has been located mainly in these sparsely occupied areas, as MATOPIBA, due to the establishment of new agricultural frontiers.

The factors mentioned have enabled MATOPIBA to present increasing levels of production, especially in the cultivation of soybeans, and local economic growth (BOLFE et al., 2016; ZANIN; BACHA, 2017; BRAGANÇA, 2018; ARAÚJO et al., 2019). To illustrate this, according to Araújo et al. (2019), MATOPIBA had a significant growth in soy production, from 260,624 t in 1990 to 10,758,927 t in 2015, an increase of 4,028% in the period.

In this context, this paper aims to analyze the relationship between economic development and deforestation in MATOPIBA, with a special focus on their connection with the agricultural frontier expansion, in view of its importance for the economic growth of the region. In addition, we investigate the spatial distribution of deforestation, as well as the formation of spatial clusters and the presence of spatial dependence. One of the hypotheses is that there is a spatial effect in deforestation, with centripetal forces acting in the attraction of productive activities, especially agricultural, and consequently generating spatial autocorrelation in this variable. These forces are generated by the existence of different production techniques, climate, topography and soil conditions among the municipalities of MATOPIBA, factors that lead to regional differences.

The basic hypothesis connecting economic development and deforestation comes from the pioneering work of Grossman and Krueger (1991), who state that, at low levels of development, a growth in per capita income leads to an increase in environment degradation. However, from a certain level, this logic would reverse itself, with an increase in per capita income leading to a decrease in environmental degradation. The curve that represents this relationship has an inverted-U shape and is known in the literature as the Environmental Kuznets Curve (EKC).

However, a recurrent problem in the literature, often not considered in the estimations of EKC, is the use of per capita income as a proxy for economic development; it is inadequate for such because it captures only one aspect of development (HILL; MAGNANI, 2002; JHA; MURTHY, 2003; COSTANTINI; MONNIM, 2008; STIGLITZ et al., 2009; COSTANTINI; MARTIGINI, 2010; KUBISZEWSKI et al., 2013; NEVE; HAMAIDE, 2017). Therefore, the present paper proposes to solve this problem by replacing the per capita income with the Human Development Index (HDI) of the MATOPIBA municipalities, following the approach adopted by Hill and Magnani (2002), Jha and Murthy (2003), Costantini and Monni (2008), Costantini and Martini (2010), and Lamb and Rao (2015).

Although there are several papers that have sought to estimate an EKC for Brazil using several indicators of environmental degradation, none has used the HDI as a proxy for economic development. Moreover, deforestation has been used in some studies for the Amazon and the Cerrado biomes, but there are no works specifically developed for the MATOPIBA region. This paper proposes to fill both gaps in the literature.

Finally, the work is structured into four sections, including this introduction. The second section discusses the theoretical framework on the EKC and the relationship between economic development and environmental degradation. In the third section, we detail the methodology and the database used. The results and their analysis are displayed in the fourth section, followed by the final considerations.

2. Theoretical framework

Grossman and Krueger (1991) attempted to decompose the effects that are behind the relationship between economic growth and environmental quality in the EKC, resulting in the identification of three main effects: scale, composition, and technical. The scale effect occurs due to the increase in production, which causes a pressure on the environment due to a greater use of natural resources. The composition effect is the change that occurs in the composition of the goods and services produced. Finally, the technical effect is related to technological advances that increase productivity and/or can make production “cleaner”, generating less waste. The composition and technical effects can be large enough to minimize the scale effect. The descending part of the EKC, therefore, occurs because of this overlay of effects (GROSSMAN; KRUEGER, 1991).

Several authors have attempted to broaden the understanding of why the relationship between economic development and environmental degradation takes on an inverted-U shape. Among them, we can mention Shafik and Bandyopadhyay (1992), Selden and Song (1994), Stern et al. (1996), De Bruyn et al. (1998), and Culas (2007). Shafik and Bandyopadhyay (1992) conducted an empirical study and most of the indicators showed results similar to those found by Grossman and Krueger (1991). Selden and Song (1994), in turn, argue that the environmental pressure is due to increases in income and consumption, which lead to greater use of natural resources. However, the authors claim that there are some damping factors that could mitigate the negative effects on the environment, and even reverse it in the long run. The reasons and processes that lead EKC to have an inverted-U shape, according to Selden and Song (1994), mainly stem from positive income elasticity for environmental quality, changes in the composition of production and consumption and technological innovations that increase productivity, induced by market competition and/or adjustments to imposed legislation.

Many researchers, however, argue that the descendant part of the EKC occurs because polluting industries tend to move from developed to underdeveloped countries, in a move encouraged by restrictions imposed by legislation (SURI; CHAPMAN, 1998; STERN et al., 1996). This theory is known in the literature as the Pollution Haven Hypothesis. In the other hand, Institutional changes, according to Culas (2007), are an important element to explain the inverted-U relationship between environmental degradation and per capita income. As income rises, factors such as increased utility from underdeveloped areas, population awareness of the importance of environmental sustainability, and strengthening government capacity for environmental protection help to explain the decline in environmental impact.

Despite the evidence for the existence of an EKC, some authors, such as De Bruyn et al. (1998) argue that this relationship is not sustained in the long run, since there is another turning point in which per capita growth leads once again to environmental degradation. Therefore, an N-shaped curve, rather than an inverted-U shape, would better represent the relationship between development and environment. In addition, according to the author, there is the possibility of EKC assuming other shapes beyond the usual, with a need for verification for each specific case.

According to Neve and Hamaide (2017), per capita income is not the most adequate proxy for economic development to verify the relationship with environmental degradation, given that it only partially captures the development of a region. In this context, Hill and Magnani (2002), Jha and Murthy (2003), Stiglitz et al. (2009), Kubiszewski et al. (2013) and Neve and Hamaide (2017) recommend using a variable that is more related to well-being in general than just economic performance. This occurs because several indicators can lead to increased social welfare, resulting in less environmental damage, without income growth necessarily occurring.

Hill and Magnani (2002) and Jha and Murthy (2003) sought to avoid this problem by using the HDI as a proxy for economic development, replacing the per capita income in the EKC model. The authors also verified that, when using the HDI, there is an improvement in the environmental degradation prediction in relation to the per capita income, suggesting that the HDI is able to capture the relationship with the environment better. From the pioneering works by Hill and Magnani (2002) and Jha and Murthy (2003), many authors have sought to replace per capita income by the HDI, or similar indicators, in their estimates. In general, they have verified an improvement in model adjustment (COSTANTINI; MONNI, 2008; COSTANTINI; MARTINI, 2010; LAMB; RAO, 2015; NEVE; HAMAIDE, 2017).

There is no consensus about the existence of a traditional inverted-U EKC for deforestation in the literature (SHAFIK; BANDYOPADHYAY, 1992; SHAFIK, 1994; CROPPER; GRIFFITHS, 1994; BHATTARAI; HAMMING, 2001; KOYUNCU; YILMAZ, 2009). Chiu (2012) states that the empirical results are controversial and argues that an analysis must be carried out for each locality, since it is not possible to infer causality from studies on other regions.

Shafik e Bandyopadhyay (1992) and Shafik (1994) have not found statistically significant relationships between deforestation and economic growth. On the other hand, analyzing this relationship for three continents, Africa, Latin America and Asia, Cropper and Griffiths (1994) found statistically significant results for the first two. Bhattarai and Hammig (2001), conducting a similar study for the three continents, found statistically significant results for all between growth and forest cover, with an inverted-U relationship. According to the authors, at low levels of development, the structure of demand causes deforestation, but as economic growth occurs, such demand structure tends to change, moving to goods that affect the environment less. In addition, income growth induces an increase in replanting efforts, which ends up reversing the deforestation process in the long run. Koyuncu and Yilmaz (2009) found that the increase in demand for arable land also has a significant impact on deforestation alongside economic growth.

For Brazil, there have been studies seeking to identify the existence of the EKC using deforestation. However, practically all the papers focus on the Legal Amazon, with only one on the Cerrado and no studies specifically for MATOPIBA. For Amazon, we have controversial empirical evidence, which varies according to the year analyzed or method adopted. Gomes and Braga (2008), Prates (2008), Santos et al. (2008), Polomé and Trotignon (2016), Tritsch and Arvor (2016) found evidence of an inverted-U relationship, while Araújo et al. (2009) and Jusys (2016) captured a U-shaped EKC, and Oliveira et al. (2011) and Oliveira and Almeida (2011) identified an N-shaped relationship. For the Cerrado, the paper by Colusso et al. (2012) is the only one which estimates an EKC for this biome. The authors projected several spatial models, which corroborated significant results for an N-shaped curve for the Cerrado, indicating that, in the long run, economic growth is not sufficient to prevent the deforestation of the biome.

However, among the papers for Brazil, there is none that specifically investigates the existence of an EKC for the MATOPIBA region. Therefore, based on Chiu (2012), it is clear that we need a

study focused on the Brazilian agricultural frontier in the Cerrado. The purpose of the next section is to describe the occupation dynamics in this biome, as well as the expansion of its agricultural frontier.

2.1. Occupation of the Cerrado and the expansion of its agricultural frontier

The Cerrado is located in the central region of Brazil, occupying about 25% of the national territory, with an area of approximately 2,039,243 km², covering 1,389 Brazilian municipalities. The biome is the richest savanna in the world and is of much importance to the balance of the global ecosystem. However, its intensive occupation, especially after the 1970s, with the advancement of the Brazilian agricultural frontier, has caused serious damage to the biome, with many irreparable environmental losses (MYERS et al., 2000).

The Brazilian government has played an active role in the occupation and expansion of the agricultural frontier in the Cerrado. This role began in the 1970s with the military governments, especially after the Second National Development Plan (II PND). In practice, the incentives for the occupation of the Cerrado, especially in the Central-West region, were through the Agricultural Frontier expansion. The basic instrument used was a subsidized rural credit offer, combined with the implementation of an infrastructure that enabled the occupation of the territory. Such incentives and measures resulted in rapid changes in land coverage and use in the region. There is a close relationship between agricultural frontier expansion and the opening of roads, for they allow the creation of access corridors to the region and possibly to deforestation of native vegetation because the road network expansion allows access to previously isolated areas, affecting its environmental degradation rhythm (ASSUNÇÃO; BRAGANÇA, 2015; BRAGANÇA, 2018). Nevertheless, there are no papers in the literature that have sought to analyze this impact of roads on the deforestation of MATOPIBA, or even the Cerrado

According to Chagas and Andrade (2017), human presence in forest areas itself represents a deforestation vector, since the population demands local resources for their subsistence, income growth and material well-being. The agricultural frontier expansion is, in turn, an inductive factor of the occupation of the Cerrado, which increases the pressure to open new areas. This scenario has led to a progressive depletion of the natural resources of the region, making this biome the second that has suffered the most changes due to anthropogenic actions in Brazil, after the Atlantic forest. Despite this, conservation units protect only 7.44% of the territory, which has served to aggravate the intensive use of its natural resources (SANTOS et al., 2009; BORGES; SANTOS, 2009; IBAMA, 2010).

In historical terms, the state that has presented greater deforestation in its Cerrado areas is São Paulo, with 90% of the total, followed by Mato Grosso do Sul (75.87%), the Federal District (70.63%), Paraná (70.00%), Goiás (65.11%), Minas Gerais (56.84%), Mato Grosso (42.83%), Bahia (36.45%), Tocantins (26.40%), Maranhão (22.85%), Piauí (15.10%), and Rondônia (2.88%) (IBAMA, 2010). Although the MATOPIBA states are those with the smallest deforested area, except Rondônia, the current deforestation in the Cerrado has been located mainly in Piauí, Bahia, Tocantins and Maranhão, states that suffered considerable land use changes after the 2000s (BORGES; SANTOS, 2009). Table 1 shows the ten municipalities with the largest deforestation in 2010; we can note that all are located in the states belonging to the MATOPIBA region.

In this context, the Brazilian Government adopted measures to combat and inhibit deforestation in the Cerrado. In 2009, for example, the Brazilian government released the Action Plan for the Prevention and Control of Deforestation and Forest Fires in the Cerrado (PPCerrado) which aimed to reduce continuously and permanently the rate of deforestation, as well as forest fires and wildfires in the Cerrado. In 2014, it launched the second phase of the plan in order to guide the actions, in addition to ratifying the importance of the conservation of the natural resources of the region. An essential element of the plan is the National Policy on Climate Change, Law N° 12.187/2009, which seeks the reduction of greenhouse gas emissions in the atmosphere. It also established a goal of 40% reduction in the deforestation rate in the biome.

Table 1 – Municipalities of the Cerrado that presented greater deforestation in 2010

Municipality	State	Suppression (km ²)	Area (%)
Baixa Grande do Ribeiro	PI	394.29	5.05%
Uruçuí	PI	203.48	2.41%
Formosa do Rio Preto	BA	143.92	0.89%
São Desidério	BA	119.85	0.81%
Mateiros	TO	93.06	0.97%
Barreiras	BA	88.39	1.12%
Balsas	MA	85.24	0.65%
Santa Quitéria	MA	73.88	3.85%
Codó	MA	69.91	1.60%
Riachão das Neves	BA	68.81	1.18%

Source: IBAMA (2010).

The intensive occupation of MATOPIBA for agricultural production began in the 1980s, and this process is not yet completed. This is due to the existence of much underutilized land, where low-productivity techniques are adopted. The very conditions of the region facilitate this process of occupation, such as: good climate for agriculture, flat land that enables the adoption of machinery to enhance land productivity, cheap labor and low price of land. In addition, many spaces are not occupied, where native forests of the Cerrado prevail. Therefore, this availability makes it possible to incorporate these regions into the most dynamic areas of MATOPIBA, as the agricultural frontier advances. (BATISTELLA; VALLADARES, 2009; STUDTE, 2008; BOLFE et al., 2016; BRAGANÇA, 2018; ARAÚJO et al., 2019).

The low price of land and the easy adoption of mechanized and large-scale agriculture have attracted labor along with investments in capital, which have intensified their occupation, resulting in a significant growth in local production, mainly of grains such as soybeans and maize (MIRANDA et al., 2014; ASSUNÇÃO; BRAGANÇA, 2015; BRAGANÇA, 2018; ARAÚJO et al., 2019). However, according to Garcia and Vieira Filho (2018), approximately 68% of the agricultural expansion in the region between 2002 and 2014 was due to the conversion of native areas.

In 2017, for example, the region accounted for approximately 11% of the national soy production, a figure that may increase in the future as the agricultural frontier in the region expands (ZANIN; BACHA, 2017; ARAÚJO et al., 2019). Therefore, the existence of underutilized and/or not yet occupied land, together with the agricultural frontier expansion and the economic development of MATOPIBA, may maintain the deforestation process in the region in the following years. To make matters worse, Garcia and Vieira Filho (2018) point out that inadequate soil management, which causes their progressive degradation, resulted in approximately 9 million hectares of area with moderate degree of desertification and 591,000 hectares with high degree.

The agricultural frontier expansion in MATOPIBA, however, faces some natural challenges, especially in areas transitioning into the Caatinga biome. A transition area normally presents diverse ecosystems and climatic conditions and lower natural fertility. The soybean, for example, is not suitable in regions with annual average rainfall below 1000 mm, which occur in Cerrado areas near the semi-arid. In other words, the annual average rainfall acts as a natural barrier for the agricultural frontier expansion, a scenario that could be reversed with the development of new varieties of soybeans that support rainfall between 800 and 1000 mm. However, this technological innovation could boost deforestation along with the agricultural production in MATOPIBA (ARAÚJO et al., 2019).

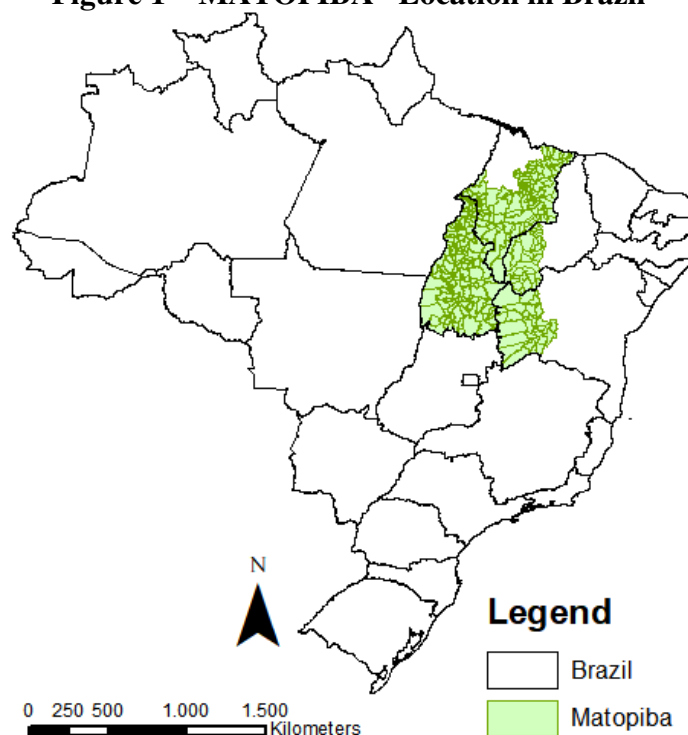
Considering the expansion of the agricultural frontier and the economic development of MATOPIBA, the present paper aims to investigate their relationship with deforestation.

3 Methodology

3.1. Database and region covered

The Brazilian Agricultural Research Corporation (Embrapa) elaborated the MATOPIBA boundary delimitation, having as its main criterion the presence or not of the Cerrado in the four states, as well as other socioeconomic factors, which resulted in 337 municipalities comprising 31 microregions with an area of 73 million hectares. Among the MATOPIBA states, the one with the largest area is Tocantins, with 37.95% of the total (139 municipalities), followed by Maranhão, with 32.77% (135 municipalities), Bahia, with 18.06% (30 municipalities) and Piauí, with 11.21% (33 municipalities) (EMBRAPA, 2017). Figure 1 brings the MATOPIBA location within the Brazilian territory.

Figure 1 – MATOPIBA* Location in Brazil



Source: Research data.

Note: * The shaded area refers to the 337 municipalities of MATOPIBA.

According to the Ministério da Agricultura (2017), The MATOPIBA region reached an 11% share of the total produced by Brazilian agribusiness in 2015. Recognizing the region's strategic importance for the future of the country's agribusiness, the Brazilian government created the Decree No. 8447 in 2015 with the main objective of establishing an agricultural development plan for MATOPIBA aiming to guide Federal projects specifically for the region (BRASIL, 2015)

According to the population census held by IBGE (2010), MATOPIBA has about 6 million inhabitants, and 35% reside in the rural area, considerably above the Brazilian average of 15.3%. Among the states in the region, the most populous is Maranhão, with 57.6% of the total, followed by Tocantins (25.30%), Bahia (12.72%), and Piauí (4.75%). In relation to the income, the region had a per capita income of only 40% of the Brazilian average in 2010 – R\$8,000 in MATOPIBA against R\$19,878.00 for Brazil. However, if considering only Tocantins and Bahia, the percentage would go up to approximately 60% of the national average (BOLFE et al., 2016).

The data used concerning the deforestation of the Cerrado in the MATOPIBA region were obtained in the 2010 Technical Report on the Monitoring of Deforestation in the Cerrado Biome from

IBAMA, the Brazilian Institute of the Environment and Renewable Natural Resources. The deforestation variable (DEFOREST) is used as a proxy for environmental degradation, the dependent variable on the EKC model for the 337 municipalities of MATOPIBA. We consider 2010 as a reference year for deforestation, as well as for all the explanatory variables. We used this timeframe due to the limitation of disaggregated data for the municipalities of the region, which made it impossible to have a more extensive temporal analysis. For the economic development proxy, we used the HDI variable, in line with the methodological advances in the estimation of EKC in the literature (HILL; MAGNANI, 2002; JHA; MURTHY, 2003; COSTANTINI; MONNIM, 2008; STIGLITZ et al., 2009; COSTANTINI; MARTIGINI, 2010; KUBISZEWSKI et al., 2013; NEVE; HAMAIDE, 2017).

We included the HDI in a square and cube format in the estimations to investigate the existence of other formats for EKC, i.e., a quadratic or cubic function. The variables described above, as well as the other explanatory variables used in this paper, are displayed in Table 2. The inclusion of the variables aimed to improve the specification of the econometric model, as well as to better structurally represent the region and identify possible relationships that they may have with deforestation. In addition to the variables directly linked to the agricultural frontier expansion, we also consider some geographic and structural features for control purposes. Among them, we used some vector data to construct the variables specifically to this empirical design: ROADS, RAINFALL, SOIL, FEDERAL.RES and INDIGN.RES. We construct the measures using the spatial joint tool in the GIS software (ArcMap 10.3). Some explanations, however, are worth mentioning.

Table 2 – Description of variables, all for 2010

Variables	Description	Unit	Source
DEFOREST	Deforested area	Ha	IBAMA
HDI	Human Development Index (HDI)	cent	IPEA
HDI ²	HDI Squared	-	-
HDI ³	HDI Cubed	-	-
RURAL CREDIT	Total rural credit	R\$ (BRL)	BACEN
DEM.DENSITY	Demographic density (inhabitants/km ²)	km ²	SIDRA/IBGE
AGRIC.GDP	Agricultural participation in GDP	%	SIDRA/IBGE
CATTLE	Cattle herd size	count	SIDRA/IBGE
CROP	Total crop area	Ha	SIDRA/IBGE
SUGARCANE	Sugarcane Productivity	kg/ha	SIDRA/IBGE
MAIZE	Maize Productivity	kg/ha	SIDRA/IBGE
SOYBEAN	Soy Productivity	kg/ha	SIDRA/IBGE
ROADS	Length of roads	Km	MAPBIOMAS
RAINFALL	Average annual precipitation	Mm	CPRM
SOIL	Good or regular suitability of soil for farming	binary	MMA/IBGE
FEDERAL.RES	Federal Reserve	binary	CSR
STATE.RES	State Reserve	binary	CSR
INDIGEN.RES	Indigenous Reserve	binary	CSR
FOREST.COVER	Remaining forest cover	%	MAPBIOMAS

Source: Research data.

The SOIL variable was constructed using the Map of Brazilian Agricultural Potential compiled by the Brazilian Institute of Geography and Statistics (IBGE) and made available by the Ministry of the Environment. The Brazilian territory was classified according to the agricultural potential of its soils, considering: fertility, physical and morphological characteristics, main limitations and topography. The effort resulted in five basic classifications: i) good; (ii) regular; (iii) restricted; (iv) unfavorable; and (v) inadvisable. Merging the agricultural potential map with the MATOPIBA map,

we identified the predominant type of soil that exists in the municipalities. Finally, we created a binary variable, in which the number 1 was assigned to municipalities with i) good or regular soil and 0 for others. The basic purpose of this procedure is to verify if municipalities with greater agricultural potential soils have higher rates of deforestation. In an indirect way, it will be possible to identify if the Brazilian agricultural frontier expansion in MATOPIBA, caused by the conversion of forests into arable areas, is occurring in municipalities with greater agricultural potential.

The RAINFALL variable is composed of average annual precipitation data (1977 to 2006), from the national hydrometeorological network, compiled by the *Companhia de Pesquisa de Recursos Minerais* (CPRM, 2018) and made available by the Pluviometric Atlas of Brazil. The ROADS variable refer to the length of state and federal highways in a given municipality in kilometers. The data vector are made available by the Mapbiomas project, using data provided by the Brazilian government. We obtained information on protected areas, which generated the binary variables FEDERAL.RES, STATE.RES and INDIGN.RES, from the Center for Remote Sensing of the Federal University of Minas Gerais (CSR-UFMG). Joining the MATOPIBA municipalities with the protected area shape files, it was possible to obtain the presence or not of these areas for each municipality, considering only those created until 2010.

3.2. Descriptive statistics

In order to investigate the characteristics of MATOPIBA municipalities and the changes in the period, Table 3 reports the descriptive statistics for the variables used in the EKC model. In terms of deforestation, there was an average of 15,91 ha of cleared area in 2010 and 5,360 ha in all the MATOPIBA region. Regarding forest cover, in turn, 60% of the total area was comprised of forest remnants, with some municipalities having 98% of their area composed of native forests. The variables related with economic development, as HDI and per capita GDP, the averages were 0.61 and R\$7,359.64. In addition, there was a considerable difference in the characteristics of the municipalities when considering the maximum and minimum values, which may reflect variances in the occupation stage.

Table 3 - Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max	Total
DEFOREST	15.91	26.25	0.00	227.34	5,360.80
HDI	0.61	0.05	0.44	0.79	-
GDP	7,359.64	5,817.02	2,292.04	52,736.02	2,480,199.95
RURAL CREDIT	11,500,000.00	37,700,000.00	5,045.52	485,000,000.00	3,872,740,512.53
DEM.DENSITY	13.43	18.56	0.23	180.79	-
AGRIC.GDP	30.07	14.78	0.62	74.86	-
CATTLE	44,540.95	46,843.11	1,300.00	423,650.00	15,010,299.00
CROP	12,573.65	37,616.84	68.00	441,164.00	4,237,320.00
SUGARCANE	22,048.66	21,816.03	0.00	100,000.00	-
MAIZE	2,048.35	1,716.39	85.00	8,617.00	-
SOYBEAN	1,020.65	1,317.37	0.00	3,449.67	-
ROADS	149.77	101.40	7.50	623.70	50,472.79
RAINFALL	1,456.08	277.71	800.00	2,100.00	-
FOREST.COVER	0.60	0.20	0.04	0.98	-

Source: Research data.

We can also highlight some total values for MATOPIBA, which are related with the region of agricultural frontier expansion. The offer of rural credit reached a value of approximately R\$3.8

billion. The size of the cattle herd and crop area, in turn, reached 15 million heads and 4.2 ha, respectively. Finally, the length of roads totaled 50,472 km². A recurrent problem in the EKC model is multicollinearity, which can invalidate statistical inferences. The Appendix A shows the correlation between the variables used. From them, we can notice no extremely high correlations that could compromise the estimation of the EKC model, with the exception of the CROP and RURAL.CREDIT variables.¹ Therefore, we included only crop area in the econometric model.

3.3. Exploratory Spatial Data Analysis (ESDA) and spatial econometrics

The Exploratory Spatial Data Analysis (ESDA) is a technique used to identify spatial effects, specifically those of spatial dependence and heterogeneity. Both, if identified, should be treated to avoid problems in econometric models, such as bias and inconsistency (ALMEIDA, 2012). Moran's I statistic seeks to capture the spatial autocorrelation between a variable across the regions. The expected value of this statistic is $E(I) = -1/(n-1)$ and statistically larger (minor) values, relative to expected, indicate positive spatial autocorrelation (negative). Mathematically, we have

$$I_t = \left(\frac{n}{S_0} \right) \left(\frac{z_t' W z_t}{z_t' z_t} \right) \quad t = 1, \dots, n \quad (1)$$

where n is the number of regions, S_0 is a value equal to the sum of all elements of W , z is the value of the standardized variable analyzed, and Wz is the average value of the standardized variable in the neighbors according to a weighting matrix W . The local Moran's I (LISA), in turn, is

$$I_i = z_i \sum_{j=1}^J w_{ij} z_j \quad (2)$$

where z_i represents the region's i standardized variable, w_{ij} is the element of the spatial weighting matrix (W) and z_j is the standardized variable of region j .

The spatial component is incorporated into the econometric model with spatially lagged variables. It is possible to propose a general spatial model that, by imposing restrictions on parameters, allows to obtain the desired specifications. Such model is

$$\begin{aligned} y &= \rho W y + X \beta + W X \tau + \xi \\ \xi &= \lambda W \xi + \varepsilon \end{aligned} \quad (3)$$

where X is the matrix of explanatory variables; β is the vector $k \times 1$ of regression coefficients; ε is the error term with average zero and constant variance.

The Spatial Autoregressive Model (SAR) is obtained by imposing the following restrictions on the model (3): $\rho \neq 0$, $\tau = 0$ and $\lambda = 0$. The SAR model seeks to capture the spatial autocorrelation effects of the dependent variable between neighboring units. Therefore, it is included as an explanatory variable in the econometric model the spatially lagged dependent variable ($\rho W y$), which can be interpreted as the average value of that variable in neighboring units. In this paper, the SAR model seeks to identify if the deforestation rate of a given municipality is influenced by the value of that variable of its neighbors, determined according to a spatial weight matrix. If we get a significant $\rho > 0$, we have a positive spatial autocorrelation, while $\rho < 0$ is a negative spatial dependence. The model will suffer from the endogeneity problem of the lagged variable; therefore, it should be estimated with instrumental variables, which are the lagged explanatory variables ($W X$).

¹ In other words, there is a strong relation between the amount of rural credit destined for the region and the size of the cultivated area.

The Spatial Error Model (SEM) emerges if $\rho = 0$, $\tau = 0$ and $\lambda \neq 0$ when spatial dependence manifests itself in the error term. The closer to one the parameter λ is, greater is the effect of this shock on the neighborhood. The estimation by OLS is not adequate, because the error bias makes the model parameters inefficient. According to Kelejian and Prucha (1999), we should estimate the model with maximum likelihood estimation (MLE) or by the generalized method of moments (GMM). The Spatial Lag of X Model (SLX) occurs when $\rho = 0$, $\tau \neq 0$ and $\lambda = 0$. The model seeks to capture the spatial spillovers from the independent variables, using a matrix of spatial weights W as a spatial lag operator. This lag is exogenous because the variables are determined outside the model. For this reason, the model does not present an endogeneity problem, thus it is possible to estimate by OLS.

The Spatial Durbin Model (SDM) and the Spatial Durbin Error Model (SDEM) are a combination of the previous ones. The SDM occurs when $\rho \neq 0$, $\tau \neq 0$ and $\lambda = 0$, with the spatial autocorrelation in the dependent and the explanatory variables. The SDEM, in turn, means $\rho = 0$, $\tau \neq 0$ and $\lambda \neq 0$, when spatial dependence manifests itself in the explanatory variables and the error term. The model choice, however, is not arbitrary, because the spatial effects can manifest in just one of the forms, in some combination of them or even in all. The chosen model that is able to minimize the spatial autocorrelation in the residuals of the models, following Almeida (2012), will be identified ahead.

3.4. Empirical strategy

In this paper, the deforested MATOPIBA area is included as a dependent variable in the EKC model. The municipalities' HDI levels are used as a proxy for the level of economic development. In addition, to better verify the relationship between environmental degradation and economic development, we estimate EKC with quadratic and cubic forms. Therefore, the general equation is

$$DEFOREST_i = \beta_0 + \beta_1HDI_i + \beta_2HDI^2_i + \beta_3HDI^3_i + \varepsilon_i \quad (4)$$

where *DEFOREST* is the percentage of deforested area of the Cerrado in municipality *i*; *HDI* is the Human Development Index. The incorporation of additional explanatory variables is important to avoid the problem of omission of relevant variable. Therefore, geographical, structural and agricultural variables are included, as in Table 2. Hence, in a general form, the model is

$$DEFOREST_i = \beta_0 + \beta_1HDI_i + \beta_2HDI^2_i + \beta_3HDI^3_i + \beta_kZ_i + \varepsilon_i \quad (5)$$

where *Z* is a matrix with the *k* of additional explanatory variables included in the model.

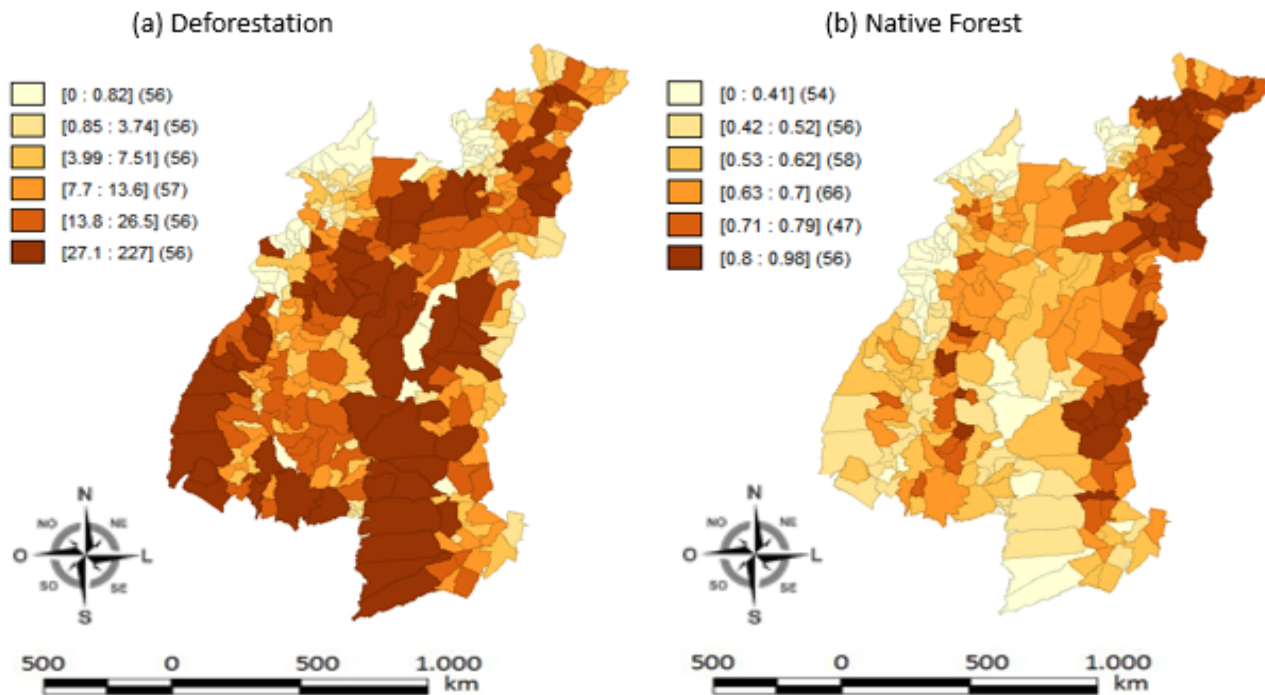
The EKC format is related to the signs and significance presented by the coefficients β_1 , β_2 and β_3 in the model (6). It is a sufficient condition for the curve to present a linear shape when we get a significant $\beta_1 > 0$ or $\beta_1 < 0$, while β_2 and β_3 are not. In this configuration, an increase in the HDI is linearly related to deforestation. For the inverted-U shape, it is sufficient that $\beta_1 > 0$, $\beta_2 < 0$, while for the U shape, $\beta_1 < 0$, $\beta_2 > 0$; both significant while β_3 is not. Finally, the cases in which $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$ or $\beta_1 < 0$, $\beta_2 > 0$ and $\beta_3 < 0$, being all statistically significant, display necessary and sufficient conditions for a N shape or inverted-N shape, respectively.

4. Results and discussion

To verify the relationship between deforestation and forest cover, Figure 2 shows the spatial distribution of deforestation (Figure 2a) and native forest remnants (Figure 2b). The municipalities with high cleared area (Figure 2a) are concentrated especially in western Bahia, the central area of MATOPIBA, southwest of Tocantins and, finally, the northern part of Maranhão. Araújo et al. (2019) argue that these regions, especially the first two, have undergone an intense modernization of their agricultural activity, especially of soybeans, resulting in significant increases in their production and

yield after the 2000s. According to the authors, this phenomenon may be one of the explanations for the recent deforestation in the region.

Figure 2 - Distribution of deforestation (a) and native forest (b) in MATOPIBA.



Source: Research data.

Regarding the native forests (Figure 2b), we can see that most of the municipalities with more than 80% of their territory covered by forests are located east, in transition areas with the Caatinga biome. According to Araújo et al. (2019), a transition area normally presents diverse ecosystems and climatic conditions and lower natural fertility. In addition, the authors argue that soybean cultivation, the main driver of occupation in MATOPIBA, is suitable only in areas with annual average rainfall above 1000 mm, which does not occur in Cerrado areas near the semi-arid region. Therefore, the future agricultural expansion in this region depends on the development of new varieties of soybeans that support rainfall between 800 and 1000 mm and, if this occurs, we can expect profound land use changes with reductions in forest cover.

The spatial concentration of deforestation is visible in Figure 2, indicating the existence of patterns, which may result in spatial dependence and heterogeneity. To verify this hypothesis, Table 4 presents Moran's I statistic, according to several spatial matrix conventions. We confirm the existence of spatial dependence for deforestation, regardless of the convention adopted, indicating that deforestation tends to be spatially concentrated. Theoretically, this may result from spatial spillovers, which stem from productive links and concentration of human and physical capital.

Table 4 - Moran's I for deforestation in MATOPIBA

	Weights Matrix					
	Queen	Rook	Three neigh.	Five neigh.	Seven neigh.	Ten neigh.
Deforestation	0.29*	0.30*	0.31*	0.22*	0.22*	0.15*

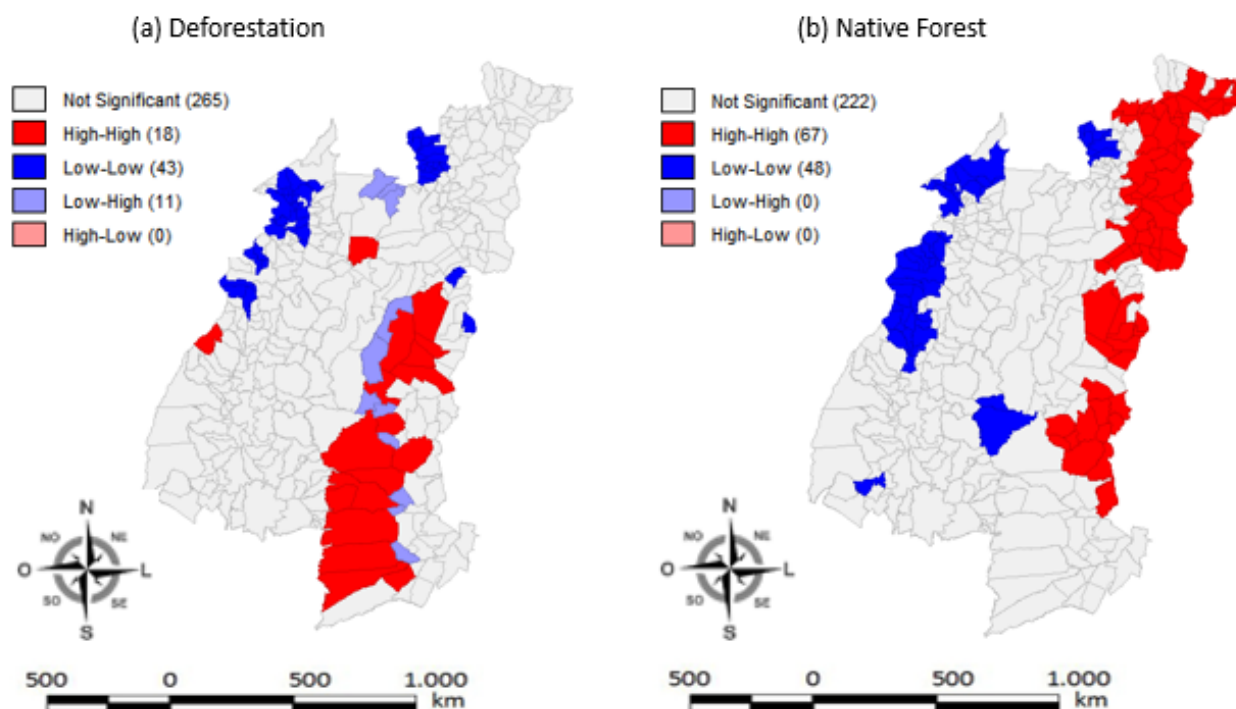
Note: * Level of significance of 1%.

Source: Research data.

Figure 3 shows the LISA maps of spatial clusters for deforestation. It presented two significant High-High clusters for deforestation (Figure 3a): one in western Bahia and the other in the Southwest of Piauí. For the native forest (Figure 3b), we have spatial clusters located especially in the east

transitional areas with the semi-arid biome. Therefore, the clusters presented a similar spatial pattern to Figure 2, corroborating the arguments of Araújo et al. (2019).

Figure 3 – LISA maps for deforestation (a) and native forest (b) in the MATOPIBA



Source: Research data.

Regarding the Environmental Kuznets Curve, we estimated the model in its quadratic and cubic versions using the classical linear regression model estimated by the OLS method. Table 5 presents the estimations.² This procedure is necessary to verify the existence of spatial dependence, captured by the Moran's I of the residuals, which are significant for all estimated models in Table 5.

The best model estimated, according to the Akaike information criterion, is one that incorporates a quadratic relationship between the HDI and deforestation, i.e., the HDI (2). Therefore, this model will be the base for the following analyses and we confirm that the HDI is a better proxy for economic development in line with Hill and Magnani (2002), Jha and Murthy (2003), Costantini and Monni (2008), Costantini and Martini (2010), and Lamb and Rao (2015). Since deforestation, agricultural and geographic variables usually suffer from spatial interactions, Moran's I statistic presented a statistical significance of 1% as expected, which indicates the presence of spatial autocorrelation in the model residuals. In this context, the estimates may not be consistent, which requires the adoption of specific econometric methods to address the presence of spatial effects (ALMEIDA, 2012). Therefore, the next step is to estimate HDI (2) by incorporating spatially lagged variables, aiming to control spatial dependence in the residuals.

Furthermore, from the Jarque-Bera test it was possible to reject the null hypothesis of normality in the residuals with a significance level at 1%. Regarding the variance, the Koenker-Bassett test rejected the homoscedasticity hypothesis, indicating the presence of a non-constant variance in the residuals. The spatial models in Table 6, due to the non-normality in the residuals, are estimated with the Generalized Method of Moments of Kelejian and Prucha (1999). In addition, White's robust error (WHITE, 1980) is employed in the SAR, SLX and SDM models, and the robust error of Kelejian and Prucha (1999) for models SEM and SDEM, both aiming to control the presence of heteroscedasticity.

² The software used to estimate the models was GeoDaSpace, made available by the Center for Spatial Data Science – University of Chicago.

In addition, we chose the spatial lag matrix that generated the largest Moran's I coefficient for the HDI (2) residues (Appendix B) to estimate the spatial models, opting for the rock matrix.

Table 5 – Econometric results for the EKC estimated with OLS

	HDI (2)	HDI (3)	PIB(2)	PIB (3)
CONSTANT	-180.1141***	-180.5291***	-11.6019**	-9.7553*
HDI	587.5692***	601.7303**	-0.0001	-0.0008
HDI ²	-511.8558***	-558.5114**	0.0000	0.0000
HDI ³		39.5611		0.0000
DEM.DENSITY	-0.0815	-0.0840	-0.1180**	-0.1153**
AGRIC.GDP	-0.0562	-0.0548	-0.0061	0.0001
CATTLE	0.0001***	0.0001***	0.0001***	0.0001***
CROP	0.0004***	0.0004***	0.0004***	0.0004***
SUGARCANE	0.0001**	0.0001**	0.0001*	0.0001*
MAIZE	0.0014*	0.0014*	0.0013	0.0015*
SOYBEAN	0.0006**	0.0006**	0.0006**	0.0006**
ROADS	0.0273***	0.0273***	0.0223**	0.0232**
RAINFALL	0.0020	0.0020	0.0016	0.0016
SOIL	-1.1945	-1.1981	-1.5379	-1.4392
FEDERAL.RES	6.5379**	6.5600**	7.1074***	7.0777***
ESTATE.RES	-0.4191	-0.5988	-1.2974	-1.4553
INDIGEN.RES	-1.0752	-1.1256	-1.5783	-1.7170
FOREST.COVER	14.3776***	14.4532***	15.9287***	15.8401***
Akaike	2818.953	2820.911	2826.123	2827.512
Jarque-Bera	4267.002***	4248.649***	4442.388***	4552.182***
Koenker-Bassett	60.185***	60.765***	56.194***	56.521***
Moran's I	0.135***	0.134***	0.124***	0.122***

Source: Research results. Note: *** Significant at 1%; ** Significant at 5%. * Significant at 10%.

According to Almeida (2012) and Raiher and Candido (2018), the best spatial model is the one that minimizes the spatial autocorrelation in the residuals. Therefore, considering Moran's I in the EKC spatial model residuals (Appendix B), the SLX and SDEM models presented the lowest coefficients for this statistic, indicating that they are the models that best controlled the spatial dependence problem. In other words, the approach, by incorporation relevant spatial spillovers related to deforestation and the agricultural frontier expansion, was able to minimize spatial dependence.

However, in the SDEM model in Table 6, the spatial lag of the error term ($W\xi$) did not presented statistical significance, indicating that only the spatial spillovers from the explanatory variables are important to explain deforestation in MATOPIBA. In any case, all the statistical significant variables in the SLX and SDEM models are the same and the coefficients are similar, indicating the robustness of the results. In addition, according to Almeida (2012), structural stability in the model parameters is a sign that the spatial heterogeneity is not present in an extent that invalidates the results. In this context, we considered the SLX model in the following analyses.

Since the quadratic models have statistically significant economic development proxies with coefficients, $\beta_1 > 0, \beta_2 < 0$, we have an inverted-U relationship between deforestation and development in MATOPIBA. This fact demonstrates that deforestation will increase until a certain threshold as the region develops, from which it shall begin to fall, in line with Grossman and Krueger (1991; 1995). In addition, we can cite the robustness of the results, because all the models that included the HDI in their linear and quadratic forms had significant coefficients. This empirical evidence differs from Colusso et al. (2012), who identified an N-shaped relationship when analyzing

the whole biome, which highlights that MATOPIBA may present different characteristics regarding the impact of economic development on the environment.

Table 6 – EKC spatial models for MATOPIBA

Variables	SAR	SEM	SLX	SDM	SDEM
CONSTANT	-121.2634**	-151.3787***	-145.5197***	-164.9267***	-150.2584***
HDI	388.1051**	472.5044***	489.7202***	544.7760***	501.1248***
HDI ²	-337.4297**	-396.9236***	-427.8252***	-458.1134***	-434.3899***
DEM.DENSITY	-0.1046***	-0.0647***	-0.0574*	-0.0609**	-0.0554*
AGRIC.GDP	-0.0213	0.0032	-0.0758	-0.0572	-0.0652
CATTLE	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**
CROP	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***
SUGARCANE	0.0001**	0.0000	4.16E-05	2.79E-05	3.71E-05
MAIZE	0.0013	0.0012	0.0022**	0.0021**	0.0021**
SOYBEAN	0.0004**	0.0003	0.0003	0.0002	0.0003
ROADS	0.0210**	0.0234***	0.0236***	0.0265***	0.0246***
RAINFALL	0.0015*	0.0009	0.0005	0.0003	0.0005
SOIL	-1.0668	-0.8967	-1.0048	-1.0166	-0.9441
FEDERAL.RES	7.4203**	6.9657**	8.1435***	7.5392***	8.1312***
ESTATE.RES	5.3336	4.5345	5.6590	5.9516	5.7087
INDIGEN.RES	-4.6199	-4.3685	-5.7509	-6.0728*	-5.7100
FOREST.COVER	18.5311***	16.4407***	17.2549***	17.1812***	16.8820***
D_OUTLIER	54.9167***	50.9545***	54.6023***	52.9212***	53.3676***
ρ	-0.1581	-	-	0.2890**	
λ	-	0.2748**	-		0.1838
W_DEM.DENSITY			-0.1242**	-0.0651	-0.1150**
W_AGRIC.GDP			0.0583	0.0456	0.0483
W_CATTLE			3.90E-06	-8.90E-06	5.60E-06
W_CROP			-0.0001**	-0.0002***	-0.0001**
W_SUGARCANE			0.0001**	0.0001	0.0001**
W_MAIZE			-0.0018	-0.0018	-0.0016
W_SOYBEAN			0.0003	0.0001	0.0003
W_ROADS			-0.0167	-0.0240*	-0.0184
W_RAINFALL			0.0022	0.0013	0.0020
W_FORES.COVER			-4.4538	-8.4427	-4.3892

Source: Research results. Note: *** Significant at 1%; ** Significant at 5%. * Significant at 10%.

D.OUTLIER is a dichotomous variable for the municipalities comprised of leverage points detected in the ESDA, that is, they reinforce the deforestation pattern observed.³

According to Grossman and Krueger (1991; 1995), at low development levels, growth initially causes a scale effect by increasing the use of natural resources, which leads to deforestation. However, after a “turning point”⁴, the composition and technical effects become large enough to mitigate the scale effect, reducing environmental degradation. Considering the results for the SLX model in Table 6, the turning point for MATOPIBA is an HDI level of 0.57, slightly below the region’s average of

³ Oliveira et al. (2011) proposed the procedure for deforestation in the EKC model for Legal Amazon, which improved econometric estimates.

⁴ A region where the curve reaches its maximum value. We can obtain the “turning point” with: $\tau = -\beta_1/2\beta_2$.

0.61. Therefore, 71.82% of the municipalities in the current Brazilian agriculture frontier have an HDI above the turning point, which indicates that deforestation will decrease its pace as development increases. On the other hand, we still have a trade-off between development and forest conservation for the remaining 28.18% that are below the turning point, since their development will boost deforestation.

In addition, we highlight the statistical significance in the SLX model for the following additional variables: demographic density, cattle herd, crop area, maize productivity, length of roads, Federal Reserve and forest area. Regarding the spatial spillovers from the agricultural frontier expansion, we have demographic density, crop area and sugarcane productivity with statistical significance. With the exception of demographic density and spillovers from crop area, all the variables presented a positive relationship with deforestation, with its increase leading to deforestation in MATOPIBA.

The length of roads has a positive impact on deforestation in MATOPIBA. One possible explanation is that the expansion of the road network allows access to previously isolated areas by creating corridors to the region, reducing transportation costs and pushing the agricultural frontier further by intensifying the migration and occupation of the territory, which causes deforestation (ASSUNÇÃO; BRAGANÇA, 2015; BRAGANÇA, 2018; ARAÚJO et al., 2019). This empirical evidence is an important contribution to the literature on deforestation in Brazil, since there are no papers that have address this issue for this region directly.

Cattle herd is also an important deforestation inductor in MATOPIBA. According to Bragança (2018), this phenomenon is explained mainly by changes in land use due to the advancement in the cultivation of soybeans and sugarcane in recent periods in Brazil, which has adopted more technologically advanced inputs, with greater potential to generate profits. This has led to a displacement of cattle ranching to agricultural frontier regions with lower land prices, causing deforestation.

Crop area presented a positive statistical significance and a negative spillover for Matopiba. Garcia and Vieira Filho (2018) argue that its expansion – for soybeans in particular – is occurring in a considerable part due to forest area reduction, especially at the agricultural frontier. According to the authors, 68% of the agricultural expansion in Matopiba between 2002 and 2014 was due to the conversion of native areas. On the other hand, empirical evidence suggests that crop area expansion in a municipality diminishes its neighbor's deforestation in a negative spatial spillover effect. Productivity spillovers from maize and sugarcane presented significant positive impact on deforestation in MATOPIBA. One possible reason is that both crops recently gained market value due to the increase of national and international demand for animal feed and biodiesel, which resulted in high profitability and environmental degradation.

Demographic density and its spatial spillovers, in turn, presented a statistically significant negative sign, indicating that less densely populated municipalities tend to be deforested more than those with large demographic density. This result contradicts those found by Grossman and Krueger (1995) and especially those by Cropper and Griffiths (1994), who used deforestation as an indicator of environmental degradation. On the other hand, in the Brazilian context, in Oliveira et al. (2011) and Colusso et al. (2012), the demographic density variable did not present statistical significance, indicating that it may not be relevant to explain deforestation in the country. Despite this, the results found in this paper highlight a different characteristic for the MATOPIBA region.

The remaining forest cover in the municipalities is statistically significant, indicating that higher deforestation is associated with greater proportion of native forests. This fact makes logical sense, since some municipalities may deforest less because they do not have much remaining forest area to do it. In addition, a higher proportion of forests is related to regions where agricultural activities, basic infrastructure and migratory attraction have not reach their full potential yet (BOLFE et al., 2016; ZANIN; BACHA, 2017; ARAÚJO et al, 2019). In other words, the growth of these factors translates into the agricultural frontier expansion, which is the main environmental degrader in the Cerrado (GARCIA and VIEIRA, 2018). In addition, according to IBAMA (2010), only 7.44% of the Cerrado biome territory is comprised of conservation units, which has served to aggravate deforestation. However, in the MATOPIBA regime, the Federal Reserve variable presented a significant positive

impact, contradicting the idea that conservation units act as inhibitors of deforestation. This demonstrates a need for an expansion of government supervision over conservation areas in MATOPIBA, since the status granted to these localities does not serve to hold back deforestation.

It is worth mentioning that the characteristics of the soil in the MATOPIBA region do not affect land use changes. In other words, deforestation occurs regardless of the soil being suitable or not for agricultural production, corroborating Bolfe et al. (2016), who point out that a considerable part of the forest conversion in MATOPIBA did not occur on soil with agricultural suitability.

5. Final considerations

This paper aimed to investigate the relationship between human development and environmental degradation at the current Brazilian agricultural frontier, a region known as MATOPIBA. The hypothesis used is the Environmental Kuznets Curve, which states that the level of environment degradation increases with economic development initially, but, after a certain level, the relationship reverses, with an increase in development leading to a reduction of degradation.

The municipalities with high cleared area in MATOPIBA are concentrated especially in western Bahia, the central area of the region, southwest of Tocantins and the northern part of Maranhão. These regions have undergone an intense modernization of their agricultural activity, especially related to soybeans, resulting in significant increases in their production and yield after the 2000s. In addition, we confirm the presence of spatial dependence for deforestation, indicating that it tends to be spatially concentrated, which leads to the adoption of spatial econometrics.

The dependent variable used as a proxy for environmental degradation is the deforestation of the Cerrado in the MATOPIBA region for the year 2010. In addition, we used the Human Development index (HDI) as a proxy for economic development, which is considered more appropriate to represent development when compared to per capita income, usually adopted by the literature, according to methodological advances in the estimation of EKC. The explanatory variables included in the EKC model, in addition to the HDI in its linear, quadratic and cubed form, are: cattle herd, demographic density, productivity of maize, sugarcane and soybeans, crop area, agricultural participation in the GDP, length of roads, average annual precipitation, soil suitability, and presence of federal, state and indigenous reserve, all set at the municipal level.

Initially, we estimated the models using conventional econometric techniques to identify the presence of spatial effects in the residuals. We also estimated several spatial models in order to verify the robustness of the results: SAR, SEM, SLX, SDM and SDEM. The model that best captured the EKC relationship, according to Moran's *I* in the spatial model residuals, are the Spatial Lag of *X* Model (SLX), since it is the one that minimizes spatial dependence. In the estimations, we found an inverted-U shape for the EKC, corroborating the initial hypothesis on economic development and environmental degradation for the current Brazilian agricultural frontier. Therefore, human development, although leading to deforestation initially, induces a sustainable development after a certain level. The turning point of the EKC curve, where economic development reaches its maximum impact on the environment, is an HDI level of 0.57, which is slightly below the region's average of 0.61. Therefore, 71.82% of the municipalities in the current Brazilian agriculture frontier have an HDI level above the turning point, which indicates that deforestation will decrease its pace as these municipalities develop. On the other hand, there is still a trade-off between development and forest conservation for the remaining 28.18% of municipalities that are below the turning point. These highlights environmental concerns for the region, since its development could boost degradation in these underdeveloped municipalities.

To worsen this scenario, we identified many variables, especially related to the expansion of the agricultural frontier in MATOPIBA, which affect the environment negatively. Among the main influences, we have the expansion of roads, which attracts migratory waves and agricultural activities due to its cost reduction, cattle herd, crop area, maize productivity, Federal Reserve and forest area. Regarding the spatial spillovers from the agricultural frontier expansion, demographic density, crop

area and sugarcane productivity are statistically significant. With the exception of demographic density and spillovers from crop area, all these variables presented a positive relationship with deforestation, with their increase leading to deforestation in MATOPIBA.

The empirical evidence from this paper can help to identify the determinants and possible outcomes of deforestation in MATOPIBA and to construct specific agricultural and environmental policies that consider idiosyncratic characteristics of the region along with spatial effects. However, some possible limitations of the present paper are worth mentioning. For example, the IBAMA database adopted here has a limited timeframe, which makes the deforestation phenomenon in MATOPIBA difficult to analyze more broadly. In addition, we recommend the adoption of alternative methodologies, such as the Geographically Weighted Regression (GWR), in order to map the local effects of the agricultural frontier expansion in the region.

References

- ALMEIDA, E. *Econometria espacial aplicada*. Campinas: Alínea, 2012.
- ARAUJO, C.; BONJEAN, C.; COMBES, J. P.; REIS, E. Property rights and deforestation the Brazilian Amazon. *Ecological Economics*, v. 68, n. 8-9, p. 2461-2468, 2009.
- ARAÚJO, M. L. S.; SANO, E. E.; BOLFE, E. L.; SANTOS J. R. N.; SANTOS, J. S.; SILVA, F. B. Spatiotemporal dynamics of soybean crop in the Matopiba region, Brazil (1990–2015). *Land Use Policy*, v. 80, n. 1, p. 57-67, 2019.
- ASSUNÇÃO, J.; BRAGANÇA, A. *Does technical change in agriculture increase deforestation? Evidence from the Brazilian Soybean Revolution*. INPUT, 2015. (Working Paper, n. 8).
- BATISTELLA, M.; VALLADARES, G. S. Farming expansion and land degradation in Western Bahia, Brazil. *Biota Neotropica*, v. 9, n. 3, p. 61-76, 2009.
- BHATTARAI, M.; HAMMING, M. Institutions and the EKC for deforestation: a cross-country analysis for Latin America, Africa, and Asia. *World Development*, v. 29, n. 6, p. 995-1010, 2001.
- BOLFE, E. L.; VICTÓRIA, D. C.; CONTINI, E.; BAYMA-SILVA, G.; SPINELLI-ARAÚJO, L. GOMES, D. MATOPIBA em crescimento agrícola aspectos territoriais e socioeconômicos. *Revista de Política Agrícola*, n. 4, p. 38-62, 2016.
- BORGES, K. M. R.; SANTOS, P. M. C. Modelo Linear de Mistura Espectral – MLME aplicado ao monitoramento do Cerrado. In: Simpósio Brasileiro de Sensoriamento Remoto, 14, 2009. *Anais...* INPE, 2009.
- BRAGANÇA, A. A. The economic consequences of the agricultural expansion in Matopiba. *Revista Brasileira de Economia*, v. 72, n. 2, p. 161-185, 2018.
- BRASIL. Decreto nº 8.447 de 6 de maio de 2015. Link: <http://www.planalto.gov.br/ccivil03/ato2015-018/2015/decreto/d8447.html>. Accessed on: Oct 23, 2017.
- CHAGAS, A. L. S.; ANDRADE, L. C. Custo de Oportunidade da preservação ambiental: o caso da Amazônia Legal Brasileira. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 11, n. 3, p. 384-405, 2017.
- CHIU, Y. Deforestation and the EKC in developing countries: a panel smooth transition regression approach. *Canadian Journal of Agricultural Economics*, v. 60, n. 2, p. 177-194, 2012.
- COLUSSO, M. V. S.; PARRÉ, J. L.; ALMEIDA, E.; Degradação ambiental e crescimento econômico: a curva de Kuznets ambiental para o Cerrado. *Revista de Economia e Agronegócio*. v.10, n. 3, p. 335-358, 2012.

- COSTANTINI, V.; MARTINI, C. A Modified environmental Kuznets Curve for sustainable development assessment using panel data. *International Journal of Global Environmental Issues*, v. 10, n. 1/2, p. 84-122, 2010.
- COSTANTINI, V.; MONNI, S. Environment, human development and economic growth. *Ecological Economics*, v. 64, n. 4, p. 867-880, 2008.
- CPRM – Serviço Geológico do Brasil. *Atlas Pluviométrico (2018)*. Link: <http://www.cprm.gov.br/publique/Hidrologia/Mapas-e-Publicacoes/Atlas-pluviometrico-do-Brasil-1351.html>. Accessed: Oct 03, 2018.
- CROPPER, M.; GRIFFITHS G. The interaction of population, growth and environmental quality. *American Economic Review*, v. 84, n. 2, p.250-254, 1994.
- CULAS, R. J. Deforestation and environmental Kuznets Curve: an institutional perspective. *Ecological Economics*, v. 61, n. 2-3, p. 429-437, 2007.
- DE BRUYN, S. M.; VAN DEN BERGH, J. C. J. M.; OPSCHOOR, J. B. Economic growth and emissions. *Ecological Economics*, v. 25, n. 2, p. 161-175, 1998.
- EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. Proposta de Delimitação Territorial do MATOPIBA. Link: [em: https://www.embrapa.br/gite/publicacoes/NT1_DelimitacaoMatopiba.pdf](https://www.embrapa.br/gite/publicacoes/NT1_DelimitacaoMatopiba.pdf). Accessed: Nov 03, 2017.
- GARCIA, J. R.; VEIRA FILHO, J. E. O papel da dimensão ambiental na ocupação do MATOPIBA. *Confins*, v. 35, n. 2, p. 1-20, 2018.
- GOMES, S. C.; BRAGA, M. J. Desenvolvimento econômico e desmatamento na Amazônia Legal: uma análise econométrica. In: Congresso da Sociedade Brasileira de Economia Administração e Sociologia Rural, 46, 2008. *Anais...* Acre: SOBER, 2008.
- GROSSMAN, G.; KRUEGER A. Economic growth and the environment. *Quarterly Journal of Economics*, v. 110, n. 2, p. 353-377, 1995.
- GROSSMAN, G.; KRUEGER, A. *Environmental impacts of a North American free trade agreement*. Cambridge, MA: National Bureau of Economic Research, 1991. (Working Paper NBER, n. 3914)
- HILL, R.; MAGNANI, E. An exploration of the conceptual and empirical basis of the environmental Kuznets Curve. *Australian Economic Papers*, v. 41, n. 2, p. 239-54, 2002.
- IBAMA – Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. 2010. Link: <https://www.mma.gov.br/>. Accessed: Nov 10, 2017.
- JHA, R.; MURTHY, K. V. B. An inverse global environmental Kuznets curve. *Journal of Comparative Economics*, v. 31, n. 2, p. 352–368, 2003.
- JUSYS, T. Fundamental causes and spatial heterogeneity of deforestation in Legal Amazon. *Applied Geography*, v. 35, n. 10, p. 188-199, 2016.
- KELEJIAN, H. H.; PRUCHA, I. R. A. Generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, v. 40, n. 2, p. 509-533, 1999.
- KOYUNCU, C. YILMAZ, R. The impact of corruption on deforestation: a cross-country evidence. *Journal of Developing Ideas*, v. 42, n. 2, p. 213-22, 2009.
- KUBISZEWSKI, I; CONSTANZA, R.; FRANCO, C.; LAWN, P.; TALBERTH, J.; JACKSON, T.; AYLNER, C. Beyond GDP. *Ecological Economics*, v. 93, p. 57-68, 2013.

- LAMB, W. F.; RAO, N. D. Human development in a climate-constrained world: what the past says about the future. *Global Environmental Change*, v. 33, n. 7, p. 14-22, 2015.
- MIRANDA, E. E.; MAGALHÃES, L. A.; CARVALHO, C. A. *Proposta de Delimitação Territorial do MATOPIBA*. Nota Técnica, n. 1. Brasília: EMBRAPA, 2014.
- MYERS, N.; MITTERMEIER, R.; MITTERMEIER, C.; FONSECA, G.; KENT, J. Biodiversity hotspots for conservation priorities. *Nature*, v. 403, n. 2, p. 853-858, 2000.
- NEVE, M.; HAMAIDE, B. EKC with adjusted net savings as a trade-off between environment and development. *Australian Economic Papers*, v. 56, n. 1, p. 39-58, 2017.
- OLIVEIRA, R. C.; ALMEIDA, E. *Deforestation in the Brazilian Amazonia and spatial heterogeneity: a local environmental Kuznets Curve approach*. In: 57th Regional Science Association International, 2011.
- OLIVEIRA, R.C.; ALMEIDA, E.; FREGUGLIA, R. S.; BARRETO, R. C. S. Desmatamento e Crescimento Econômico no Brasil: uma análise da Curva de Kuznets Ambiental para a Amazônia Legal. *Revista Economia e Sociologia Rural*, v. 49, n. 3, p. 709-739, 2011.
- PALOMÉ, P.; TROTIGNON, J. *Amazonian deforestation, environmental Kuznets Curve and deforestation policy: a cointegration approach*. Groupe d'Analyse et de Théorie Economique (GATE), 2006. (Working Paper GATE, n. 1608).
- PRATES, R. C. *O desmatamento desigual na Amazônia brasileira: sua evolução, suas causas e conseqüências para o bem-estar*. Tese (Doutorado em Economia). Programa de Pós-Graduação do Departamento de Economia da Universidade de São Paulo. São Paulo, 2008.
- RAIHER, A. P.; CANDIDO, M. J. Aglomerações produtivas da região sul do Brasil e sua relação com a produtividade industrial. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 12, n. 1, p. 17-34, 2018.
- SANTOS, P. M. C.; ALVES M. S.; SILVA D. A.; CARNEIRO, C. L.; FERNANDES, M. M. Monitoramento do desmatamento no cerrado, porção oeste da bacia do Rio São Francisco. In: Simpósio Brasileiro de Sensoriamento Remoto, 14, 2009. *Anais...* Natal, 2009.
- SANTOS, R.B.N.; DINIZ, M. B.; DINIZ, M. J. T.; RIVERO, S. L. M.; OLIVEIRA JÚNIOR, J. N. Estimativa da curva de Kuznets ambiental para a Amazônia Legal. In: Congresso da Sociedade Brasileira de Economia Administração e Sociologia Rural, 46, 2008. *Anais...* Acre: SOBER, 2008.
- SELDEN, T. M.; SONG, D. Environmental quality and development. *Journal of Environmental Economics and Management*, v. 27, n. 2, p. 147-162, 1994.
- SHAFIK, N. Economic development and environmental quality: an econometric analysis. *Oxford Economic Papers*, v. 46, n. 46, 1994.
- SHAFIK, N.; BANDYOPADHYAY, S. Economic growth and environmental quality. *Journal of Environmental Economics and Management*, v. 4, n. 10, p.1-24, 1992.
- STERN, D. I.; COMMON, M. S.; BARBIER, E. B. Economic growth and environmental degradation: the EKC and sustainable development. *World Development*, v. 24, n. 7, p. 1151-1160, 1996.
- STIGLITZ, J.; SEN, A.; FITOUSSI, F. *Report of the Commission on the Measurement of Economic Performance and Social Progress*. Paris, 2009. Link: <https://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report>. Accessed: Oct 14, 2019.

STUDTE, M. *A expansão da fronteira agrícola e a agricultura familiar no Cerrado Maranhense*. Tese (Mestrado em Economia). Programa de Pós-Graduação do Departamento de Economia da Universidade de São Paulo. São Paulo, 2008.

TRITSCH, I.; ARVOR, D. Transition in the Brazilian Amazon: emergence of a new pattern of socio-economic development and deforestation. *Land Use Policy*, v. 59, n. 31, p. 446-455, 2016.


WHITE, H. Heteroskedasticity-Consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, v. 48 n. 4 p. 817-38, 1980.

ZANIN, V.; BACHA, C. J. C. A importância dos sojicultores sulistas na nova fronteira agrícola brasileira. *Indicadores Econômicos FEE*, v. 45, n. 1, p. 35-52, 2017.

ORCID

Pedro Henrique Batista de Barros  <https://orcid.org/0000-0002-7968-0197>

Alysson Luiz Stege  <https://orcid.org/0000-0001-9266-1890>

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

Appendices

Appendix A – Correlation for the variables

	HDI	GDP	MAIZE	R.CREDIT	SOY	CANE	CROP	AG.GDP	CATTLE	RAINF	ROADS	F.COVER	D.DENS
HDI	1												
GDP	0.4477	1											
MAIZE	0.2722	0.6058	1										
RURAL CREDIT	0.0877	0.4884	0.5465	1									
SOYBEAN	0.1892	0.4443	0.5305	0.2838	1								
SUGARCANE	0.1038	0.1041	0.2467	0.1945	0.1411	1							
CROP	0.0446	0.5382	0.6052	0.922	0.3312	0.2123	1						
AGRIC.GDP	-0.1764	0.1274	0.3115	0.1792	0.2613	-0.0492	0.2155	1					
CATTLE	0.4196	0.2309	0.1474	0.188	0.0895	0.034	0.0458	0.1268	1				
RAINFALL	0.34	0.1321	-0.0166	-0.0732	0.046	-0.2678	-0.1092	0.1401	0.2084	1			
ROADS	0.2639	0.2789	0.3705	0.3543	0.3013	0.2051	0.3495	-0.0494	0.3107	-0.0535	1		
FOREST.COVER	-0.3541	-0.2657	-0.2297	-0.1029	0.0143	0.2162	-0.0529	-0.1635	-0.3938	-0.3668	0.0223	1	
DEM.DENSITY	0.1105	-0.1385	-0.2566	-0.0955	-0.2702	0.0174	-0.0879	-0.3119	-0.0814	0.0805	-0.1257	-0.0058	1

Source: research results.

Appendix B - Moran's I for the EKC Models - convention matrix decision

	Weights Matrix					
	Queen	Rook	Three neigh.	Five neigh.	Seven neigh.	Ten neigh.
HDI (2)	0.126*	0.135*	0.126*	0.046	0.069*	0.035
HDI (3)	0.126*	0.134*	0.126*	0.046	0.069*	0.035
PIB (2)	0.116*	0.124*	0.111*	0.050	0.069*	0.038
PIB (3)	0.113*	0.122*	0.107*	0.051	0.069*	0.037

Note: * Level of significance of 1%. Source: research data.

Appendix C - Moran's I for the ECK Spatial Models residuals

	Weights Matrix					
	Queen	Rook	Three neigh.	Five neigh.	Seven neigh.	Ten neigh.
SAR	0.08	0.08	0.04	0.01	0.07*	0.04
SEM	0.12*	0.13*	0.07*	0.03	0.08*	0.05
SLX	0.06	0.06	0.03	0.06*	0.06	0.04
SDM	0.07	0.08	0.04	0.01	0.07*	0.04
SDEM	0.07	0.08	0.03	0.00	0.06*	0.04

Note: * Level of significance of 1%. Source: research data.



Revista Brasileira de Estudos Regionais e Urbanos (RBERU)

v. 13, n. 2, p. 183-200, 2019

<http://www.revistaaber.org.br>

ANALYSIS OF THE EFFECT OF TERRITORIAL FACTORS ON REGIONAL INNOVATION IN THE STATE OF SÃO PAULO, BRAZIL*

Suelene Mascarini

Department of Science and Technology at University of Campinas

E-mail: smascarini@gmail.com

Renato Garcia

Institute of Economics at University of Campinas

E-mail: renatogarcia@eco.unicamp.br

José Eduardo Roselino

Department of Geography, Tourism and Humanities at Federal University of São Carlos

E-mail: jeroselino@ufscar.br

ABSTRACT: This paper examines the geography of innovation in the São Paulo micro-regions. Using data on innovation inputs and outputs at the regional level, this paper examines how territorial factors affect innovation in the regions of the state of São Paulo, Brazil. Previous studies show that local innovation is linked to the quantity and quality of innovative inputs, and the main characteristics of regions are configured as regional driving forces of innovation. However, most studies on this subject have focused on developed countries; few studies have examined developing countries. In this way, this paper presents new empirical evidence on the role of territorial factors in fostering innovation in Brazilian regions. The main results show that local R&D expenditures and the local externalities generated by diversified industry and social filters are the main drivers of regional innovation.

Keywords: Geography of innovation; Knowledge production function; Territorial factors.

JEL Codes: O31; O18; R12.

ANÁLISE DOS EFEITOS DOS FATORES TERRITORIAIS DA INOVAÇÃO REGIONAL NO ESTADO DE SÃO PAULO

RESUMO: Este artigo avalia a geografia da inovação nas microrregiões paulistas. Ao utilizar dados sobre insumos e resultados da inovação em nível regional, este artigo examina como os fatores do território afetam a inovação regional no estado de São Paulo. Estudos anteriores mostram que a inovação local está relacionada à quantidade e qualidade de insumos inovadores, de modo que as características das regiões se configuram como forças impulsionadoras da inovação regional. Entretanto, a maioria desses estudos é desenvolvida para países desenvolvidos, poucos estudos têm tratado dos países em desenvolvimento. Dessa forma, este artigo apresenta novas evidências empíricas sobre o papel dos fatores territoriais na promoção da inovação nas regiões brasileiras. Os principais resultados mostram que os gastos com P&D e as externalidades locais geradas pela indústria diversificada, e a qualificação da mão de obra local são os principais impulsionadores da inovação regional.

Palavras-chave: Geografia da inovação; Função de produção do conhecimento; Fatores do território.

Classificação JEL: O31; O18; R12.

**Recebido em: 11/07/2018. Aceito em: 18/03/2019.*

1. Introduction

Innovation is unequally distributed among regions in Brazil, as well as in several other countries. This inequality is frequently attributed to differences in the quality and quantity of innovation inputs, such as R&D efforts and human capital. In addition, the literature indicates that fostering innovation is related to location since the agglomeration of innovation inputs, proximity effects, and local knowledge spillover effects affect innovation (AUDRETSCH; FELDMAN, 1996; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, 2012; JAFFE, 1986).

Previous studies have addressed the issue of regional innovation by examining developed countries (AUDRETSCH; FELDMAN, 1996; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007; JAFFE et al., 1993). However, the lack of understanding of the spatial development of innovation in developing countries, such as India and China (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012), Russia (CRESCENZI; JAAX, 2016), and Mexico (RODRÍGUEZ-POSE; VILLARREAL PERALTA, 2015), has become a subject of concern. Despite the increasing concern over the relationship between innovation and location in developing countries, empirical evidence for the main drivers of the innovation gap among these countries' regions, based on systematic, large-scale and quantitative data, is still lacking (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012).

Hence, the main aim of this paper is to investigate how territorial factors affect regional innovation by applying this subject to the state of São Paulo. Some studies have analysed the innovation gap at the country level (ALBUQUERQUE, 2007) and the spatial distribution of the main indicators of innovation (GARCIA; ARAUJO; MASCARINI, 2013; GONÇALVES; ALMEIDA, 2009). However, a gap remains in our understanding of how the spatial distribution of innovative inputs and the role of territorial factors affect innovation at the regional level.

This paper focuses on the geography of innovation in Brazil by presenting new empirical evidence on how territorial factors can shape the spatial distribution of innovation at the regional level. To do this, we analysed the state of São Paulo by using data at the micro-regional level, which corresponds to the EU NUTS-3. The state of São Paulo is located in the south-eastern part of the country and accounts for 35% of the Brazilian GDP and approximately 58% of the Brazilian R&D expenditure (2005 data from the Brazilian Institute of Geography and Statistics – IBGE). This work is innovative in the use of regional R&D expenditures because previous studies in Brazil have used proxies for regional R&D efforts. In addition, our analysis considered a new driving factor of innovation at the regional level: the role of university-industry collaboration. The main assumption is that the level of local innovation is linked to the quantity and quality of the level of local innovation inputs and to specific regional characteristics, such as agglomeration, diversification of local productive structures, knowledge spillovers and relationships among players.

This paper is organized in six sections, excluding this introduction. The second section presents the main conceptual debates regarding the geography of innovation. Section three presents a brief description of the distribution of innovation activity in the micro-regions of São Paulo. Section four presents the empirical model, variables and proxies. Section five discusses the effects of territorial factors. Finally, section six presents the final remarks.

2. Main conceptual remarks

2.1. Relation between innovation and territory

The territorial dynamics of innovation are gaining increasing attention in the debate on the geography of innovation. The main concern in this debate is the understanding on how territorial factors affect innovation and how firms benefit from local knowledge spillovers as well as knowledge-generation centre agglomeration. Since Jaffe's pioneering study, several authors have found positive relations between local knowledge flows and firms' innovation at the local level. The empirical evidence presented in these studies allowed for the identification of the role of local

knowledge spillovers as well as their importance for fostering innovation (AUDRETSCH; FELDMAN, 1996; JAFFE, 1986, 1989). Some of these studies have used the knowledge production function, which can correlate innovative inputs and outputs. The main assumption is that the clustering of the major inputs for innovation, such as R&D expenditures, the level of human capital accumulation, and the structure of local academic research are the main reasons for the innovation gap among different regions (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007; PONDS; VAN OORT; FRENKEN, 2010; SINGH; MARX, 2013; THOMPSON, 2006).

In addition to the usual innovative inputs, territorial factors can also play an important role in fostering innovation at the local level. The spatial organization of main sources of new knowledge is another factor that affects the extent of localised knowledge externalities, which can also affect the level of innovation output. Studies on the geography of innovation have also emphasized the role of innovation input clustering, knowledge spillovers and proximity effects, which includes the agglomeration and specialization patterns of the local productive structure (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007). In general, firms located geographically closer to knowledge-generating centres can perceive a significant competitive advantage of innovation because they benefit from local knowledge externalities and from wider possibilities for interactive learning (AUDRETSCH; FELDMAN, 1996; FRANCO; MARZUCCHI; MONTRESOR, 2014; GARCIA; ARAUJO; MASCARINI, 2013; GERTLER, 2003; JAFFE, 1989; LAURSEN; REICHSTEIN; SALTER, 2011; STORPER; VENABLES, 2003).

Three sets of factors are normally noted as the main drivers of regional innovation. First, innovative performance is a result of the returns on firms' in-house investments in innovative activities, such as R&D expenditures and human capital (AUDRETSCH; FELDMAN, 1996; JAFFE, 1986). The second set of factors is related to main territorial conditions that emerge from agglomeration economies and local knowledge spillovers (ACS; ANSELIN; VARGA, 2002; CARLINO; CHATTERJEE; HUNT, 2007; GLAESER et al., 1992). Externalities that emerge from territorial conditions are attributable to the clustering of firms and skilled workers, which have positive effects on the creation and diffusion of new knowledge (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007; GLAESER et al., 1992; STORPER; VENABLES, 2003). Finally, a third set of variables emphasizes local-level capabilities, particularly within science and technology (S&T) institutions, such as universities and public research institutes, which shape the regional innovation system (ACS; ANSELIN; VARGA, 2002; ASHEIM; GERTLER, 2009; COOKE, 2001; DE FUENTES et al., 2015; JAFFE, 1989).

In addition, agglomeration externalities can occur in places with a diversified or specialized local production structure. Following Marshall, concentrating on more specialized production structures that operate mainly in a specific industry will generate externalities that are commonly called Marshallian externalities (BEAUDRY; SCHIFFAUEROVA, 2009; MARSHALL, 1920). In contrast, according to Jacobs's view, the concentration of diverse industries encourages imitation, knowledge sharing and recombination of ideas and practices among complementary agents (JACOBS, 1969). Therefore, externalities emerge from a diverse local production structure and take the form of so-called Jacobian externalities (BEAUDRY; SCHIFFAUEROVA, 2009; GLAESER et al., 1992).

Empirical studies provide evidence supporting both perspectives. Some studies show that industrial specialization in regions is important for fostering innovation, thus highlighting the role of Marshallian externalities (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012; VAN DER PANNE, 2004). Several other studies support Jacobs's perspective by evidencing the positive effects of diversity on innovation (CARLINO; CHATTERJEE; HUNT, 2007; CO, 2002; FELDMAN, 1999; GLAESER et al., 1992). Other studies offer mixed evidence, because comparing two different institutional contexts shows that, in Europe, externalities spurred by diversity appear to be important drivers of innovation, whereas externalities generated by specialization are unfavourable for the dynamics of innovative regions. In another national context, that of the United States, externalities stemming from both diversity and specialization have been found to play an

important role in fostering regional innovation (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007).

2.2. Innovation and territory in developing countries

Recently, growing concerns related to the geography of innovation in developing countries have been observed (CRESCENZI; JAAX, 2016; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012; DE FUENTES; DUTRENIT, 2013; GONÇALVES; ALMEIDA, 2009; RODRÍGUEZ-POSE; VILLARREAL PERALTA, 2015). In many countries, empirical evidence shows that agglomeration economies are crucial for promoting innovation. In China, the spatial concentration of innovation is related to agglomeration forces driven by urban population and infrastructure endowment. Evidence also shows that higher innovative performance in Chinese regions results from increased industrial specialization. In India, local innovation is linked to advantageous socioeconomic conditions and S&T investments along with local knowledge spillovers among neighbouring regions (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012). In Russia, the main driver of regional innovation is regional R&D expenditure, and external sources of knowledge, such as inter-regional spatially mediated knowledge flows, also provide relevant innovation inputs (CRESCENZI; JAAX, 2016). In Mexico, direct R&D investment in areas with appropriate social filters is an important driver of regional economic growth (RODRÍGUEZ-POSE; VILLARREAL PERALTA, 2015).

In Brazil, the innovation system has been considered immature mainly because of the relatively small amount of industrial R&D expenditure, which results in poor innovation performance at the firm level and a low number of patent applications (ALBUQUERQUE, 2007). At the regional level, as in other developing countries, innovation in Brazil is unequally distributed across the country because a large share of Brazil's innovation, measured by the number of patent applications, occurs in the southern region, particularly because of that region's degree of urbanization (GONÇALVES; ALMEIDA, 2009), thus highlighting the importance of territorial factors.

Previous studies have shown that innovation and innovative inputs are not randomly distributed in Brazil, thus indicating the importance of local knowledge spillovers (ARAUJO, 2014; GARCIA; ARAUJO; MASCARINI, 2013; MONTENEGRO; BETARELLI JUNIOR, 2009; MONTENEGRO; GONÇALVES; ALMEIDA, 2011). Regions with higher university research efforts and a higher degree of agglomeration tend to be more innovative (MONTENEGRO; BETARELLI JUNIOR, 2009). Specialized and diversified local production systems and higher educational level of the industrial workforce also play an important role in innovation, supporting the ability of local firms to transform new knowledge into new products and processes (MONTENEGRO; GONÇALVES; ALMEIDA, 2011). Innovation is also affected by local levels of industrial and university R&D, urban agglomeration and local productive structure (ARAUJO, 2014). These studies have been important for understanding how innovation is spatially shaped across Brazilian regions. However, the scarcity and limited availability of regional-level industrial R&D data represent important gaps in the literature. Thus, an important contribution of this paper is the use of data to fill this gap by using direct measures of R&D expenditures to investigate the main drivers of regional innovation in the state of São Paulo, which allows for the identification of major drivers and local factors that contribute to the innovation gap among regions.

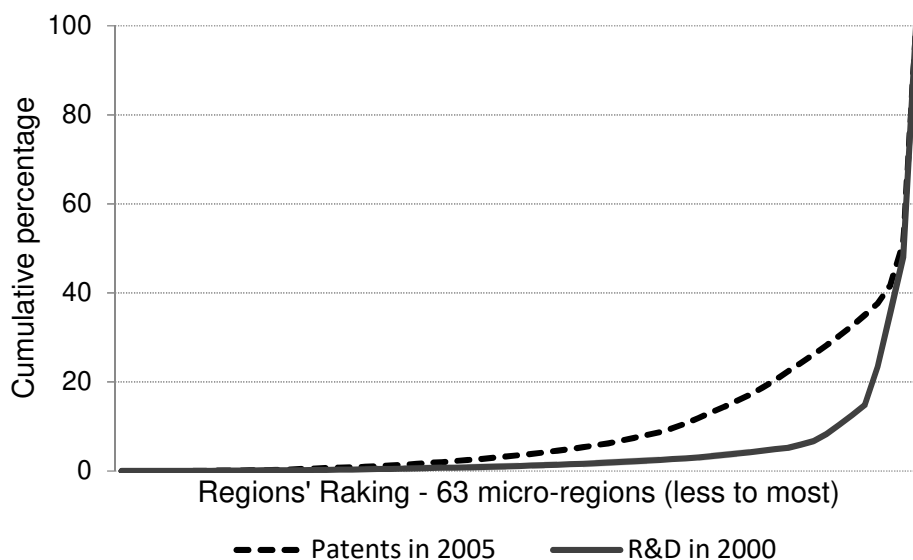
3. Regional distribution of innovation in the state of São Paulo

Brazilian innovation is far from being equally distributed in space. As well as in other developed and developing countries, regional distribution of innovation shows a strong concentration in some regions of Brazil (GARCIA; ARAUJO; MASCARINI, 2013; GONÇALVES; ALMEIDA, 2009). There is a strong concentration in the southern part of the country, especially in the state of São Paulo, which accounts for 35% of the Brazilian GDP; 46.2%

of patent deposits in Brazil; 35.3% of innovative firms; and 58.7% of Brazilian R&D expenditures (data from IBGE, 2005).

The spatial distribution of patents across micro-regions of the state of São Paulo also shows a strong tendency towards an unequal concentration in a few regions. In 2005, approximately 60% of patent applications from the state of São Paulo in the Brazilian Patents Office (INPI – *National Institute of Industrial Property*) were submitted by inventors who resided in three regions of the state: São Paulo (I), Campinas (II) and Osasco (III) (Figures 1 and 2).

Figure 1 – Distribution of total patents and R&D across micro-regions of São Paulo



Source: Authors.

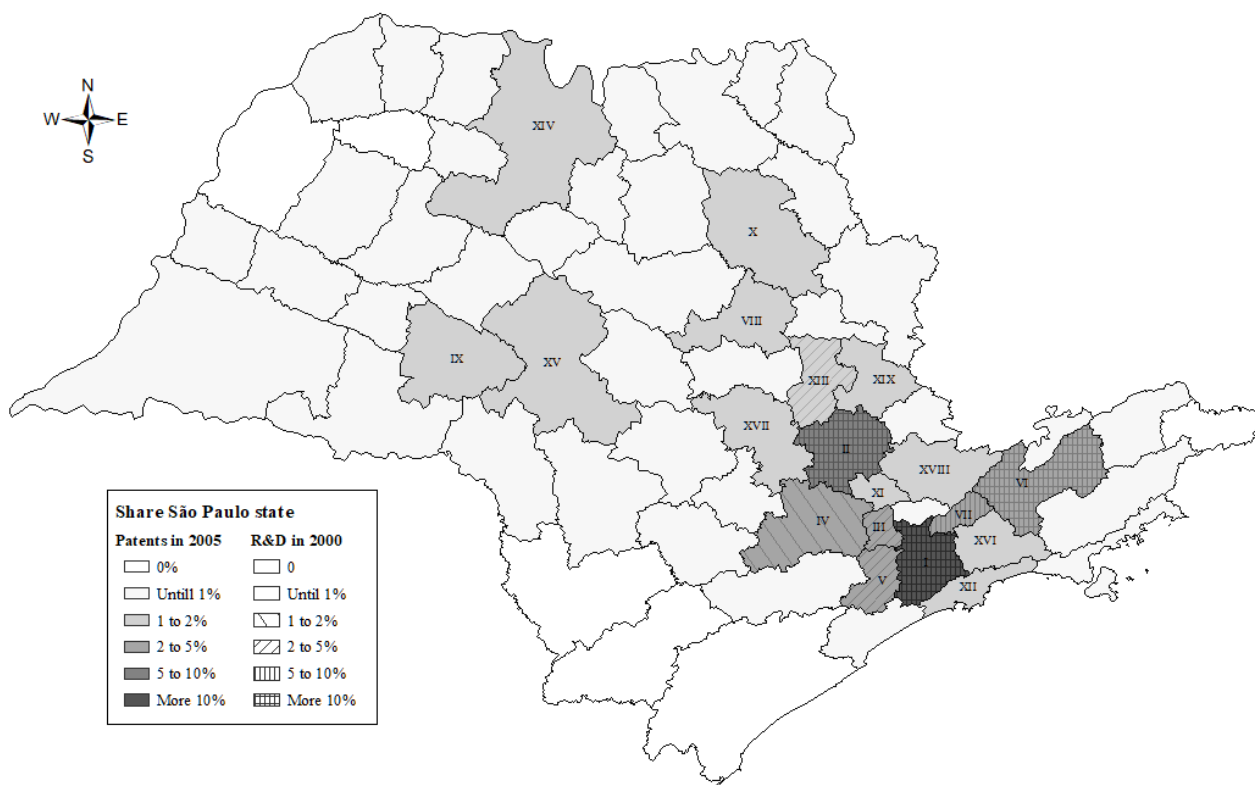
Considering the spatial distribution of innovative inputs, such as total R&D expenditures, the concentration in a few regions remains the same: more than 50% of the total R&D expenditures in 2000 were made in the region of São Paulo (I), followed by Campinas (II), with a share of 12.4%, São José dos Campos (VI), 12.3%, and Guarulhos (VII), with 8.5%. The accumulated share of the total for these four regions is approximately 85%.

The distribution of R&D expenditures in the regions of the state of São Paulo is more concentrated than the distribution of the innovation output, as measured by patents, probably because part of the industrial R&D effort does not generate patents. In fact, the regional concentration of innovation output is partly due to the agglomeration of existing innovative activities in the state, especially in regions whose firms spend more on R&D than the average. These regions tend to show better innovative performance, as seen in the spatial distribution of total patents and R&D expenditures (Figure 2).

A geographic overlap between the most innovative regions of the state can be ensured in terms of the share of total patent applications and the regional distribution of R&D expenditures. In other words, some regions are more innovative than others, probably due to higher R&D expenditures. However, regions such as Marília (IX), Ribeirão Preto (X), São Carlos (VIII), Osasco (III), Sorocaba (IV) and Santos (XIII) have higher innovative performance than other regions, even though they exhibit lower R&D expenditures. For example, in 2005, São Carlos (VIII) and Marília (IX) owned 2% of patents in the state, ranking them 8th and 9th, respectively. However, regarding total R&D expenditures, São Carlos (VIII) and Marília (IX) ranked 18th and 20th, respectively, representing less than 0.3% of R&D in 2000. This suggests that other factors affect regional innovative performance. For example, proximity effects can be an important driver of regional innovation because certain regions, such as Santos (XIII), Sorocaba (IV) and Osasco (III), can benefit from their proximity to the state capital, the city of São Paulo, whose micro-region spends

the largest amount on R&D in the entire state. This indicates the occurrence of regional knowledge spillovers.

Figure 2 – Patents and R&D location



Source: Authors.

Based on this descriptive evidence, it is possible to estimate an empirical model to investigate both the relationship between innovation and space and how territorial factors affect innovation at the regional level.

4. Territorial factors of local innovation

The empirical analysis is based on the knowledge production function formalized by Griliches (1979) and adapted by Jaffe (1989) in the following form:

$$\log(P_{ikt}) = \beta_{1k} \log(I_{ikt}) + \beta_{12} \log(U_{ikt}) + \beta_{31k} [\log(U_{ikt}) + \log(C_{ikt})] + \varepsilon_{ikt} \quad (1)$$

In this equation, P is the patent measure; I is industrial R&D expenditures; U is academic research; C is the measure of the geographic coincidence of university and industrial research; ε is the error determined in the region (states); i , k are the technological areas; and t is time.

Jaffe's variant of the knowledge production function was extremely important for the development of studies related to the geography of innovation because it transferred the unit of observation from the firm to the geographical level. Therefore, it relates to local knowledge spillovers and the geographical proximity of actors involved in innovation (AUDRETSCH; FELDMAN, 1996; CARLINO; CHATTERJEE; HUNT, 2001; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, 2012).

We set up the database using secondary data collected from several sources (INPI, PINTEC, RAIS and IBGE). We restricted the analysis to the state of São Paulo at the micro-regional level for

two main reasons. First, the state accounts for a high share of the country's total innovation efforts. Second, due to the scarcity and limited access to innovation data at the regional level, we were able to gather R&D regional-level data only for the state of São Paulo. The analysis is based on 63 micro-regions covering the entire state of São Paulo. The delimitation of the period is also due to the availability of data. Thus, we estimated the model for 2000-2005, since the results of innovation were measured in 2005, and innovative efforts, such as R&D, were measured in 2000.

To measure innovation, several studies highlight patents as a good proxy. However, we must point out some limitations associated with the use of patent data as a proxy for innovation outputs, especially for developing countries, because innovative efforts of domestic firms do not necessarily generate new patents. Other proxies have also been proposed for innovation inputs. Nevertheless, several studies have used patents as a proxy for innovation. Thus, the specification of innovation should assume a highly similar form as the model presented in previous studies (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, 2012; JAFFE, 1989):

$$\log P_{it} = \alpha + \beta_1 \ln RD_{i,T-t} + \beta_2 \text{Wln} RD_{i,T-t} + \beta_3 \text{Agglom}_{t,T-1} + \beta_4 SF_{i,T-t} + \beta_5 \text{kindex}_{i,T-t} + \beta_6 \ln \text{Coll}_{i,T-t} + \beta_7 \ln P_{i,T-t} + \beta_8 \ln \text{Dist}C_{i,T-t} + \varepsilon \quad (2)$$

where $\ln P_{i,T}$ is the natural logarithm of patent applications per capita granted by the Brazilian Patent Office by micro-regions i at $T=2005$.

Patents, as previously mentioned, are used by several studies as a proxy for innovation (CARLINO; CHATTERJEE; HUNT, 2001; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007; JAFFE, 1989). However, the use of patents as a measure of innovation has several shortcomings and limitations because it is a mechanism for appropriating the economic benefits of knowledge (ACS; AUDRETSCH, 1987). In addition, not all knowledge can be measured by patents, and patents represent only some innovation output. However, the use of patent as an outcome to measure innovation has several advantages: patents enable broad disintegration, both regional and industrial; patents can be a good measure of knowledge spillovers; and patents provide a broad basis for comparison with several studies that use them. Moreover, the reliability of patents as a proxy for innovation was tested, and the results showed that patents may be used as a fairly reliable measure of innovation and no better alternative is not available (ACS; AUDRETSCH, 1987; CARLINO; CHATTERJEE; HUNT, 2001; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, 2012; FELDMAN, 1999).

Regarding the explanatory variables represented in equation 1, $RD_{i,T-t}$ is R&D expenditures; $WRD_{i,T-t}$ is the knowledge spillovers of R&D expenditures in the neighbouring regions; $\text{Agglom}_{i,T-t}$ is agglomeration; and $SF_{i,T-t}$ represents structural characteristics that would make a micro-region more innovative. $\text{Kindex}_{i,T-t}$ is the specialization index of the region, which shows whether it is diversified or specialized, and $\text{Coll}_{i,T-t}$ is the density of collaboration between universities and firms. All measurements were calculated for micro-regions i of the state of São Paulo at time $T-t=2000$. The control variables, represented by $P_{i,T-t}$, are previous innovations, and the distance of a micro-region to the state capital (the São Paulo micro-region, where the most innovative efforts are concentrated) is represented by $\text{Dist}C_{i,T-t}$.

The use of a temporal lag between the dependent variable ($T=2005$) and the explanatory variables ($T-t=2000$) is justified by the fact that innovation does not occur immediately, as it depends on firms' capacity to generate innovations at a given time. Therefore, innovation inputs take time to generate an innovative result. In addition, the use of a temporal lag helps to minimize potential bias of reverse causality.

The R&D expenditure in a region, $\ln RD$, was measured by the natural logarithm of the total R&D expenditure divided by the total number of innovative firms. The source of data was the Brazilian Innovation Survey (PINTEC)¹. This variable represents the most important input of

¹ PINTEC is an Innovation Survey conducted by the Brazilian Institute of Geography and Statistics (IBGE), which is the main provider of data and information about the country. However, due to the low number of firms found in some

innovation. R&D expenditure improves firms' capacity to generate new knowledge and can give them the capacity to internalize knowledge from other sources, such as university R&D and/or other firms.

Thus, knowledge spillovers may represent an important input of regions' innovative activity. The potential of R&D expenditure to extend beyond regional borders is represented by $W\lnRD$ in equation 1 and is calculated by spatially weighting (W_{ij}) the natural logarithm of the R&D expenditures of neighbouring micro-regions. For spatial weighting, a row of a standardized inverse distance matrix is used in which distance, d_{ij} , the Euclidian distance between two micro-regions is, and each element is standardized by dividing it by the number of neighbours, in this case, 62, so the value of each off-diagonal element is $W_{ij} = d_{ij}^{-1}/62$. To calculate the Euclidian distance between two micro-regions, their geographical coordinates are used.

Agglomeration, Agglom, is also included in the model and is measured by urban population density, with data from the IBGE. This explanatory variable was inserted to measure the benefits of regional agglomeration as an important driver of a region's innovative performance.

Another independent variable included in the model is Social Filter, SF, which aims to capture regional structural characteristics that shape the behaviour of the population and the ability to obtain and absorb knowledge, which would make it more prone to innovation. As underlined by the innovation systems approach, the capacity of any given region to generate and use knowledge depends on a complex set of local factors. The social filter variable used in this article is therefore made up of a set of variables available for Brazil that focuses on three main aspects: education capacity, structure of productive resources in S&T and demographic structure of the micro-region.

The educational capacity of the population, or the qualification of the workforce in the region, is measured by the share of people with higher education, Educ, using data from the Brazilian Ministry of Education. This aspect seeks to measure the accumulation of skills at the regional level.

The structure of productive resources in Science and Technology, Tec, is measured by the share of local employees in technological occupations, according to Brazilian Ministry of Labour data². The demographic structure of the micro-region, Age, is measured by the number of people aged between 15 and 24 years in the micro-region, using IBGE data. This variable seeks to identify trends in population dynamics, under the assumption that young people contribute to the renewal of local society that will influence the attitudes toward innovation and social change in general.

Social filter is calculated by principal component analysis, which combines the three measures in a non-correlated order of importance and describes the data variation. It should be noted that it may be necessary to incorporate more than one component, which will depend on the analysis of the component as a combination. The combination of proxies for all these different dimensions into one single composite indicator, social filter, develops a quantitative 'profile' of an innovation-prone regional environment, enabling a comparison of the social filter conditions of different regions across countries (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, 2012). Principal component analysis (PCA) shows that the first principal component alone accounts for 45% of the total variance of the original variables, and the second represents 36.6% (Table 1). From the standardized value of the original variables (Educ, Age and Tec), using the previously listed coefficients, the score that generates the social filter index is estimated. FS1 is the first principal component, putting similar weight on the Educ and Age variables and greater weight on the Tec variable. FS1 is the first principal component that combines the three variables, providing a similar weight on the Educ and Age variables and greater weight on the Tec variable ($SF2 = 0.461Educ + 0.481Age + 0.745Tec$, Table 1). FS2 is the second principal component and is the

of the 63 micro-regions of the state of Sao Paulo, there are problems of confidentiality of information. The availability of data provided by PINTEC was limited to 10 micro-regions and 12 groups of micro-regions. To obtain an analysis of the micro-regions as a unit, a procedure was adopted for weighting groups of micro-regional data (12) from the share of employment in the region. Thus, all variables provided by PINTEC follow this pattern.

² Data were selected from the Brazilian classification of 94 group-level technology-related occupations: 011; 012; 019; 020; 021; 022; 023; 024; 025; 026; 027; 028; 029; 081; 082; 083. Brazilian classification follows the Standard Occupational Classification (SOC) system.

combination of the Educ and Age variables with similar but opposite weights ($SF2 = 0.717Educ - 0.697Age + 0.006Tec$, Table 1).

Table 1 – Results of the principal component analysis

Principal components/correlation				
Component	Eigen value	Difference	Proportion	Cumulative
Comp1	1.351	0.264	0.450	0.450
Comp2	1.087	0.524	0.362	0.813
Comp3	0.563	.	0.188	1.000

Principal components (eigenvectors)				
Variable	Comp1	Comp2	Comp3	Unexplained
Educ	0.462	0.717	0.522	0
Age	0.481	-0.697	0.531	0
Tecn	0.745	0.006	-0.667	0

Source: Authors.

Specialization or diversification of the micro-region is measured by the Krugman index. Represented in equation 1 by $Kindex$, it is calculated using Brazilian Ministry of Labour data on employment in manufacturing in the regions. The Krugman index was calculated as follows:

$$Kindex_{i,T-t} = \sum abs(v_{i,T-t}^k - v_{i,T-t}^{-k}), \text{ with } v_{i,T-t}^k = \frac{\sum_{j \neq i} x_{i,T-t}^k}{\sum_k \sum_{j \neq i} x_{i,T-t}^k} \quad (3)$$

where $v_{i,T-t}^k$ is part of sector k in region i for all businesses in this region, and $v_{i,T-t}^{-k}$ is part of the same enterprise sector from all other regions different from i divided by all firms of other regions different from i . The index has a value close to zero if the micro-region is more diversified and a maximum value of 2 if more specialized.

Finally, in equation 1, $\ln Coll$ represents university-industry collaborations, measured by the natural logarithm of the number of collaborations divided by the number of firms that collaborate in micro-regions. These numbers are based on data obtained from the Brazilian Ministry of Science and Technology. This variable aims at capturing the importance of collaborations between firms and universities because the basis for innovation is to increase the collaboration and flow of knowledge amongst stakeholders and the absorptive capacity of firms (FREEMAN, 1989; GERTLER, 2007; TEECE; PISANO, 1994). Moreover, university-industry collaboration encourages firms to develop the capacity to build different new linkages with more and diverse kinds of partners with whom they could not connect before, thereby continually increasing the flow of essential knowledge for innovation. University-industry collaboration is widely recognized as a factor that promotes innovation; however, few studies on the geography of innovation consider this variable in their empirical applications. Therefore, this variable can contribute decisively for understanding the geography of innovation, especially in Brazil, because firms' R&D expenditure is relatively low in the country.

In addition to the variables presented above, two controls were added to the analysis. The first control variable, the time lag of innovative results, is designated with $P_{i,T-t}$ and measured by the natural logarithm of the initial number of patents per capita when $T-t=2000$ in micro-region i . The main goal of incorporating this variable is to control different initial patterns of technological

capability in the region, which may also reflect differences in a region's propensity for innovative capability.

The second control variable is the natural logarithm of the road distance between micro-region i and the state capital (the São Paulo micro-region) and is designated $\ln\text{DistC}$ in equation 1. This variable attempts to control regional differences, given the proximity of the city of São Paulo, where a significant portion of the assets of the state are concentrated.³

Table 2 – Variables Summary

	Factor	Variable	Proxy(ies)	Source	
Innovation Result	Patent	$\ln P_{i,T}$	Patent application in 2005	PINTEC	
Innovation Inputs	R&D Expenditure	$\ln RD$	Total R&D expenditure divided by the total number of innovative firms	PINTEC	
	Knowledge Spillovers	$W\ln RD$	Weighting the expenditures on R&D	DER and PINTEC	
	Agglomeration	$Agglom$	Urban population density	EMBRAPA	
	Social Filter		$SF1$ and $SF2$	% people with higher education - $Educ$	INEP
				% people aged between 15 and 24 years - Age	IBGE
				% local employees in technological occupations - $Tecn$	RAIS
	Specialization or Diversification	$kindex$	Krugman Index calculated using data on employment in manufacturing in the regions	RAIS	
University-Industry Collaborations	$\ln Coll$	Number of collaborations divided by the number of firms that collaborate	Directory of Research Groups		
Controls		$\ln P_{i,T-t}$	Patents per capita	INPI	
		$\ln DistC$	Distance between a micro-region i and the state capital	DER	

Source: Authors.

5. Results

Table 3 provides the results of analysis based on 22 regressions. Local innovative efforts (i.e., R&D expenditure, $\ln RD$), knowledge spillovers ($W\ln RD$), the initial level of patents ($\ln P$) and the distance from the city of Sao Paulo ($\ln DistC$) are present in all regressions. Territorial factors, such as agglomeration ($Agglom$), social filters ($SF1$ and $SF2$), the Krugman index ($kindex$) and university-industry collaborations ($\ln Coll$), were inserted successively or sequentially until the full model was estimated (regression 19). In regressions 6-8 and 20-22, the individual components of the social filters were included separately.⁴

The results show that R&D expenditures have a positive and significant coefficient (regressions 1-3, 5-8, 10 and 13), indicating that regional innovation increases as industrial R&D expenditures increase. This is an expected result because industrial R&D expenditures represent the most important innovation input, although other activities also foster innovation, and this finding converges with results from previous studies in both developed and developing countries where higher levels of innovation result from greater R&D expenditures (CHARLOT; CRESCENZI;

³ To keep the observations equal to zero, before using the logarithmic transformation, we added 10^{-6} to the value of a variable. This procedure is used for all variables with a logarithmic transformation containing zero: $\ln(P_{i,T} + 10^{-6})$; $\ln(P_{i,T-t} + 10^{-6})$; $\ln(DistC + 10^{-6})$; $\ln(Coll + 10^{-6})$.

⁴ We have carried out a Lagrange multiplier (LM) test and a robust Lagrange multiplier (RLM) test for the spatial error model and the spatial autoregressive model. All test results are not significant, which suggests that the OLS robust regression used in this paper is appropriate.

MUSOLESI, 2012; CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007; JAFFE, 1989; MORENO; PACI; USAI, 2005; STERLACCHINI; VENTURINI, 2014). Nevertheless, previous studies at the regional level did not have this result, which is primarily because we are using industrial R&D expenditures at the regional level to measure innovative inputs, and such a process allows us to show the importance of industrial R&D for innovation at the regional level (MONTENEGRO; GONÇALVES; ALMEIDA, 2011). Thus, we can find a relation between regional R&D expenditures and patents, even considering that we are measuring only a share of both the efforts and the results of innovation.

The introduction of the industrial structure indicator measured by the Krugman Index or the university-industry collaboration indicator removes the significance of industrial R&D expenditures and thus may indicate the importance of territorial factors in fostering regional innovation (regressions 4, 9 11-12, 14-22). This finding confirms that, for developing countries, industrial R&D at the regional level represents only a share of the innovative efforts of local firms. Territorial factors, such as the characteristics of the local industrial structure and collaboration with universities, are also important drivers of local innovation.

Moreover, industrial R&D expenditures tend to extend beyond their own region because they are important for fostering innovation in neighbouring regions. This phenomenon is evidenced by the positive and significant coefficient of the spatial knowledge spillovers variable WRD (regressions 3, 6, 13, 15 and 17). The importance of regional knowledge spillovers is high because high R&D expenditure in a certain region affects innovation in neighbouring regions. This result is similar to findings in developed countries, i.e., in the United States and in European regions, because innovation at the local level is affected by local industrial R&D expenditures and by regional knowledge spillovers related to R&D expenditures in neighbouring regions (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007). Therefore, regions may innovate even with low levels of R&D expenditures because the spatial concentration of R&D expenditures is higher than the geographical concentration of innovation.

The results also show that the coefficient of the agglomeration variable Agglom is not significant, which means that nothing can be inferred about the importance of agglomeration for innovation in the regions of the state of Sao Paulo. This result was also found in similar and previous studies for EU regions (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007), in which the population density did not seem to stimulate regional innovation. This finding can be attributed to two reasons: first, there is a broader set of territorial forces that offset the positive effect of population density; second, the agglomeration in a context of low workforce mobility may encourage the pooling and stratification of inadequate skills (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, p. 698). Both arguments can be applied to our findings. In addition, we are using patents as a proxy for regional innovation, and the positive effects of agglomeration could be translated into forms of innovations that do not involve patents.

The coefficient of one of two components of the social filter variable SF2 is positive and significant, whereas that of SF1 is not significant. This result indicates that the presence of a set of specific local-level structural characteristics of a region is an important driver of the regional system of innovation. Therefore, structural characteristics play an important role in explaining differential innovative performance in the regions of the state of Sao Paulo. More specifically, this result suggests that the combination of a greater availability of highly qualified individuals and a lower amount of young people in the population represents an important innovation input at the regional level because $SF2 = 0.717Educ - 0.697Age + 0.006Tec$ (see Table 1). Therefore, our result reveals that the educational level and demographic structure in São Paulo micro-regions are important factors for promoting regional innovation, while the regional share of technological occupations does not seem to stimulate innovation. The first component of the social filter, which assigns a large weight to the labour force employed in technological occupations ($SF = 0.461Educ + 0.481Age + 0.745Tec$, Table 1), is not significant.

Table 3 – Robust regression - Patents per capita in micro-regions of São Paulo in 2005

		<i>Continues</i>										
Variables		1	2	3	4	5	6	7	8	9	10	11
lnRD	R&D Expenditures	0.283*** (0.097)	0.282*** (0.097)	0.198* (0.103)	0.116 (0.095)	0.187* (0.102)	0.210** (0.088)	0.266** (0.117)	0.294*** (0.096)	0.0933 (0.091)	0.187* (0.103)	0.0222 (0.101)
WlnRD	R&D Spillovers	0.370 (0.361)	0.370 (0.363)	1.212** (0.497)	-0.179 (0.376)	0.374 (0.362)	1.233** (0.532)	0.365 (0.370)	0.527 (0.387)	-0.207 (0.376)	0.374 (0.365)	-0.173 (0.394)
Agglom	Urban Population		9.39e-06 (4.59e-05)							7.67e-05 (6.73e-05)	4.19e-08 (4.01e-05)	
SF1	Social Filter			0.136 (0.164)								
SF2				0.463** (0.185)								
Kindex	Krugman Index				-2.262*** (0.746)					-2.376*** (0.725)		-2.253*** (0.750)
lnColl	Collaboration					0,041 (0.039)					0,041 (0.039)	0,04 (0.035)
Educ	% Population with higher Education						0.008** (0.004)					
Tecn	Employees in C&T							0.001 (0.004)				
Age	% Population with 15-24 age								-0.002 (0.001)			
lnRDT-t	Patent per capita	0.338* (0.169)	0.337* (0.170)	0.238 (0.170)	0.264* (0.152)	0.364** (0.167)	0.188 (0.183)	0.340* (0.171)	0.351** (0.164)	0.253 (0.153)	0.364** (0.167)	0.290* (0.151)
lnDistC	Distance Capital and Micro-region	0.034 (0.026)	0.036 (0.029)	0.070** (0.030)	0.044 (0.028)	0.024 (0.024)	0.072** (0.030)	0.038 (0.030)	0.031 (0.022)	0.057* (0.033)	0.024 (0.026)	0.033 (0.026)
Constant		-12.52*** (3.848)	-12.58*** (3.866)	-20.27*** (5.321)	-5.179 (3.844)	-11.30*** (4.166)	-22.60*** (6.430)	-12.40*** (3.988)	-10.68*** (3.849)	-5.310 (3.801)	-11.30*** (4.173)	-4.018 (4.463)
Observations		63	63	63	63	63	63	63	63	63	63	63
R-squared		0.378	0.378	0.456	0.494	0.397	0.442	0.378	0.402	0.500	0.397	0.512
rank		5	6	7	6	6	6	6	6	7	7	7
ll_0		-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1
ll		-93.16	-93.15	-88.91	-86.67	-92.18	-89.73	-93.14	-91.93	-86.26	-92.18	-85.54
r2_a		0.335	0.324	0.398	0.449	0.344	0.393	0.324	0.349	0.447	0.332	0.459

Table 3 – Robust regression - Patents per capita in micro-regions of São Paulo in 2005

		<i>Continued</i>										
Variables		12	13	14	15	16	17	18	19	20	21	22
lnRD	R&D Expenditures	0.00632 (0.101)	0.193* (0.104)	0.0925 (0.118)	0.0845 (0.127)	0.0704 (0.113)	0.0825 (0.129)	-0.0134 (0.136)	-0.0268 (0.134)	-0.0107 (0.106)	-0.0388 (0.168)	0.0185 (0.099)
WlnRD	R&D Spillovers	-0.198 (0.395)	1.219** (0.507)	0.522 (0.469)	1.222** (0.497)	0.499 (0.477)	1.226** (0.507)	0.541 (0.474)	0.519 (0.486)	0.431 (0.488)	-0.205 (0.392)	-0.045 (0.416)
Agglom	Urban Population	6.73e-05 (6.20e-05)	3.02e-05 (5.61e-05)			8.58e-05 (5.43e-05)	1.95e-05 (5.92e-05)		7.52e-05 (5.40e-05)	5.93e-05 (5.41e-05)	6.75e-05 (6.34e-05)	8.24e-05 (5.59e-05)
SF1	Social Filter		0.135 (0.166)	0.0530 (0.163)	0.177 (0.159)	0.0437 (0.165)	0.176 (0.159)	0.0924 (0.159)	0.0821 (0.161)			
SF2			0.468** (0.190)	0.356** (0.164)	0.459** (0.177)	0.362** (0.168)	0.462** (0.181)	0.353** (0.152)	0.359** (0.156)			
Kindex	Krugman Index	-2.353*** (0.734)		-1.959*** (0.693)		-2.085*** (0.662)		-1.934*** (0.694)	-2.046*** (0.667)	-2.051*** (0.668)	-2.332*** (0.739)	-2.332*** (0.711)
lnColl	Collaboration	0,038 (0.035)			0,042 (0.037)		0,041 (0.037)	0,04 (0.034)	0,037 (0.034)	0,036 (0.035)	0,043 (0.037)	0,037 (0.034)
Educ	% Population with higher Education									0.005* (0.003)		
Tecn	Employees in C&T										0.00237 (0.005)	
Age	% Population with 15-24 age											-0.002 (0.001)
lnRDT-t	Patent per capita	0.279* (0.152)	0.235 (0.172)	0.204 (0.157)	0.260 (0.164)	0.192 (0.160)	0.257 (0.166)	0.225 (0.154)	0.213 (0.157)	0.191 (0.169)	0.285* (0.153)	0.289* (0.147)
lnDistC	Distance Capital and Micro-region	0.046 (0.031)	0.075** (0.034)	0.064** (0.030)	0.063** (0.028)	0.079** (0.033)	0.066** (0.033)	0.058** (0.027)	0.071** (0.031)	0.068** (0.032)	0.053 (0.034)	0.04* (0.027)
Constant		-4.180 (4.453)	-20.54*** (5.493)	-12.02** (4.988)	-19.02*** (5.470)	-12.25** (5.072)	-19.21*** (5.658)	-10.95** (5.313)	-11.21** (5.446)	-11.69* (6.029)	-3.872 (4.603)	-2.640 (4.428)
Observations		63	63	63	63	63	63	63	63	63	63	63
R-squared		0.517	0.457	0.538	0.475	0.546	0.476	0.555	0.561	0.541	0.518	0.538
rank		8	8	8	8	9	9	9	10	9	9	9
ll_0		-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1	-108.1
ll		-85.21	-88.85	-83.81	-87.79	-83.24	-87.76	-82.63	-82.18	-83.56	-85.09	-83.81
r2_a		0.455	0.388	0.479	0.409	0.479	0.398	0.489	0.486	0.473	0.447	0.469

Source: Authors.

This result seems to be reinforced when the individual components of a social filter are assessed separately, because only a positive effect of the variable for the educational level of local population (Educ) is found, implying that high educational levels affect local innovation when considered individually. However, the age of the population (Age) and employees in technological occupations (Tecn) are not important drivers, although their combination functions as a genuine filter, which suggests a relevant synergy that affects innovation at the local level. This result differs from previous studies in developed countries, such as the United States and European regions, in which the combination and individual components were found to contribute to local innovation (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007). In other developing countries, such as India, similar results were found because the joint effect of the social filter index is important, while the individual components do not matter as a driver of innovation at the local level (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012).

Concerning the local industrial structure, the coefficient of the Krugman index (kindex) is negative and significant; therefore, more diversified regions whose indexes are closer to zero have better innovative performance. Innovation results in diversified regions, such as São Paulo (I), Campinas (II), and São José dos Campos (VI, Figure 2), are concentrated in both the state of São Paulo and Brazil. In addition, the introduction of this variable (kindex) removes the significance of the R&D expenditure coefficient. This result may indicate that low R&D expenditures at the local level can be compensated for by the characteristics of a diverse local industry, thereby allowing for the maintenance of high levels of local innovation. Thus, the result supports Jacobs's perspective on the benefits of diversified industrial structures and their positive effects on innovation at the regional level and aligns with findings of other previous studies that indicated that the benefits of Jacobian externalities are important generators of innovation (FELDMAN; AUDRETSCH, 1999; FRITSCH; SLAVTCHEV, 2007). In addition, these findings in Brazil are similar to those in Europe, where regions that are more specialized seem to be persistently disadvantaged in their ability to produce innovation (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007). However, a similar study in China found different results and showed that a greater degree of industrial specialization increases innovative performance at the local level (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2012). Thus, it is possible to ensure that the local industrial structure has a different effect on regional innovation in both developed and developing countries (BEAUDRY; SCHIFFAUEROVA, 2009). The coefficient of university-industry collaborations between $\ln\text{Coll}$ is not significant, indicating that no inference was observed in the importance of collaborative projects with universities in fostering innovation as measured by the number of patent applications at the local level in the state of Sao Paulo. As noted, patents represent only a share of the total innovative outputs and many firms do not patent their innovations. Thus, this finding can be related to the use of patents as a proxy for innovation since many collaboration projects with university can generate other types of innovation, such as new products and improvements to manufacturing processes.

Moreover, the introduction of the university-industry collaboration variable ($\ln\text{Coll}$) removes the significance of the R&D expenditure coefficient. This result may indicate that low industrial R&D expenditures at a local level can be somehow compensated by an interaction with academic research, which is an agent of the regional innovation system. Fostering certain types of relationships among innovation actors across sectors can help increase innovation at the regional level (DE FUENTES et al., 2015). Academic research can act as a substitute for industrial R&D because a higher density of collaborations between universities and firms can result in similar levels of innovative performance (AROCENA; SUTZ, 2010). Nevertheless, the industrial diversity of regions and university-industry collaborations are important drivers of innovation at the local level, and both factors help in the understanding of how regions with low industrial R&D expenditures can generate innovations. This represents an important characteristic of the geography of innovation in Brazil, in which extramural aspects related to territorial factors and the regional innovation system appear to be important drivers in shaping innovation in space.

The coefficient of the initial level of patenting ($\ln P_{i,T-t}$) is positive and significant when the SF variables are excluded (regressions 1-2, 4-5, 7-8, 10-12, and 21-22). Thus, the innovation process is

related to local previous innovations, indicating the importance of the temporal inertia of innovation and knowledge accumulation for innovation. This finding shows that regional innovation is linked to existing innovative activities because its most important input is cumulative knowledge. Accordingly, the influence of past events, the learning process and thus innovation generated in previous periods all positively affect the future of innovation. This result converges with the findings of previous studies in both developed and developing countries (CRESCENZI; RODRÍGUEZ-POSE; STORPER, 2007, 2012).

Finally, the coefficient of the variable that measures the geographical distance from the city of São Paulo DistC is positive and significant whenever the SF variables are included. This result indicates that regions which are geographically distant from the capital may have a higher level of patent per capita than regions located closer to the capital. Although the innovation activities in the latter may be shaped to a greater extent by exposure to interregional knowledge spillovers, proximity to the city of São Paulo appears to hamper innovation. This result is contrary to the expected outcome because São Paulo is a micro-region that fosters many innovations. Nonetheless, this result may be explained by the fact that São Paulo has a group of facilities for patent applicants that does not exist in closer regions, which causes applicants to choose São Paulo instead of them, whereas this choice does not arise in more distant regions. Ultimately, compared with proximity to the city of São Paulo, the greater importance of other regional factors such as industrial structure and density of collaborations may explain this pattern of innovation distribution.

6. Final remarks and policy implications

The existing debate about innovation in Brazil has paid little attention to the role of territorial factors. However, previous studies in both developed and developing countries have found empirical evidence of the importance of territorial determinants for shaping innovation in space. Most studies have applied this subject to developed countries, although there is increasing concern in developing countries regarding the role of territorial factors in fostering innovation at the regional level. Nevertheless, the main drivers of the innovation gap among developing countries' regions are poorly understood based on systematic and large-scale data.

With this gap in mind, further research on this topic is warranted. Using systematic, large-scale and quantitative data, we present new evidence on how territorial factors affect innovation in Brazilian regions by applying this subject to the state of São Paulo at the regional level. The main results demonstrate the links between industrial R&D at the regional level and local innovation as measured by patenting. In addition, industrial R&D has positive effects on innovation in neighbouring regions, which makes it possible to recognize the presence of regional knowledge spillovers. Alongside these, other territorial factors, such as the diversification of the local productive structure and human capital, are also important drivers of innovation at the regional level, showing the role of the Jacobian externalities in regional innovation.

These results can be associated with policy implications, especially related to technology policy for developing countries. First, our results show that industrial R&D is an important driver for regional innovation, both for the region itself and for neighbouring regions. In this way, regional R&D can play an important role in local innovation, even for developing countries, when a significant share of innovation comes from other sources and industrial R&D expenditures tend to be less important for innovation. Policies designed to foster the rise of industrial R&D expenditures can strengthen their main positive effects on innovation.

Another important result is the role of territorial factors, which include not only the industrial R&D expenditures in the neighbouring regions but also local territorial drivers, such as diversification of the local productive structure and local human capital. Linked to this finding, policies should promote measures that can strengthen the role of local factors, thereby stimulating the geographical concentration of innovative inputs, especially those of highly qualified individuals. This implication is very important for technology policies in developing countries, where private innovative efforts are weaker and scarcer than in developed countries. The geographical

concentration of innovative inputs can be an important tool for fostering innovation at the regional level. For example, deficiencies in the qualification of human capital could be an important barrier to fostering innovation in regions of developing countries, not only in terms of industrial researchers and engineers but also in terms of the overall skills of the population as a way to enhance the capacity of local agents to absorb new knowledge and to benefit from local spillovers that have been generated by the system as a whole.

Finally, it is important to mention that this analysis has limitations. We used patents as a proxy for innovation outputs. However, patents are a measure for only a share of the results of innovation. In addition, this shortcoming is especially important for developing countries, where local firms are more distant from state-of-the-art technological developments and tend to use other methods to protect their innovations. Accordingly, it is necessary to analyse the main drivers of local innovation and the role of territorial factors using other proxies for innovation output, such as innovative firms at the regional level.

References

- ACS, Z. J.; ANSELIN, L.; VARGA, A. Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, v. 31, n. 7, p. 1069-1085, 2002.
- ACS, Z. J.; AUDRETSCH, D. B. Innovation, market structure, and firm size. *The Review of Economics and Statistics*, v. 69, n. 4, p. 567-574, 1987.
- ALBUQUERQUE, E. DA M. Inadequacy of technology and innovation systems at the periphery. *Cambridge Journal of Economics*, v. 31, n. 5, p. 669-690, 2007.
- ARAÚJO, V. *Dimensão local da inovação no Brasil: determinantes e efeitos de proximidade*. Tese (Doutorado em Engenharia). Escola Politécnica da Universidade de São Paulo. São Paulo, 2014.
- AROCENA, R.; SUTZ, J. Weak knowledge demand in the South: learning divides and innovation policies. *Science and Public Policy*, v. 37, n. 8, p. 571-582, 2010.
- ASHEIM, B. T.; GERTLER, M. S. The geography of innovation: regional innovation systems. In: FAGERBERG, J.; MOWERY, D. C.; NELSON, R. R. (Eds.). *The Oxford Handbook of Innovation*. 2009. p. 291-317.
- AUDRETSCH, D. B.; FELDMAN, M. P. R&D spillovers and the geography of innovation and production. *The American Economic*, v. 86, n. 3, p. 630-640, 1996.
- BEAUDRY, C.; SCHIFFAUEROVA, A. Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, v. 38, n. 2, p. 318-337, 2009.
- CARLINO, G.; CHATTERJEE, S.; HUNT, R. *Knowledge Spillovers and the new economy of cities*. Federal Reserve Bank of Philadelphia. 2001.
- CARLINO, G.; CHATTERJEE, S.; HUNT, R. M. Urban density and the rate of invention. *Journal of Urban Economics*, v. 61, n. 3, p. 389-419, 2007.
- CHARLOT, S.; CRESCENZI, R.; MUSOLESI, A. *An 'extended' Knowledge Production Function approach to the genesis of innovation in the European regions*. Grenoble Applied Economics Laboratory (GAEL), 2012. (Working Papers 2012-06).
- CO, C. Evolution of the geography of innovation: evidence from patent data. *Growth and Change*, v. 33, n. 4, p. 393-423, 2002.
- COOKE, P. Regional innovation systems, clusters, and the knowledge economy. *Industrial and Corporate Change*, v. 10, n. 4, p. 945-974, 2001.
- CRESCENZI, R.; JAAX, A. Innovation in Russia: the territorial dimension. *Economic Geography*, v. 93, n. 1, p. 66-88, 2016.

- CRESCENZI, R.; RODRÍGUEZ-POSE, A.; STORPER, M. The territorial dynamics of innovation: A Europe-United States comparative analysis. *Journal of Economic Geography*, v. 7, n. 6, p. 673-709, 2007.
- CRESCENZI, R.; RODRÍGUEZ-POSE, A.; STORPER, M. The territorial dynamics of innovation in China and India. *Journal of Economic Geography*, v. 12, n. 5, p. 1055-1085, 2012.
- DE FUENTES, C. et al. Determinants of Innovation and Productivity in the Service Sector in Mexico. *Emerging Markets Finance & Trade*, v. 51, n. 3, p. 578-592, 2015.
- DE FUENTES, C.; DUTRÉNIT, G. SMEs' absorptive capacities and large firms' knowledge spillovers: Micro evidence from the machining industry in Mexico. *Institutions and Economies*, v. 5, n. 1, p. 1-30, 2013.
- FELDMAN, M. P. The new economics of innovation, spillovers and agglomeration: a review of empirical studies. *Economics of Innovation and New Technology*, v. 8, n. 1-2, p. 5-25, 1999.
- FELDMAN, M. P.; AUDRETSCH, D. B. Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review*, v. 43, n. 2, p. 409-429, 1999.
- FRANCO, C.; MARZUCCHI, A.; MONTRESOR, S. Absorptive Capacity, Proximity in Cooperation and Integration Mechanisms. Empirical Evidence from CIS Data. *Industry and Innovation*, v. 21, n. 4, p. 332-357, 2014.
- FREEMAN, C. *Technology policy and economic performance*. Pinter Publishers Great Britain, 1989.
- FRITSCH, M.; SLAVTCHEV, V. Universities and innovation in space. *Industry and Innovation*, v. 14, n. 2, p. 201-218, 2007.
- GARCIA, R.; ARAUJO, V.; MASCARINI, S. The Role of Geographic Proximity for University-Industry Linkages in Brazil: An Empirical Analysis. *Australasian Journal of Regional Studies*, v. 19, n. 3, p. 433-456, 2013.
- GERTLER, M. S. Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, v. 3, n. 1, p. 75-99, 2003.
- GERTLER, M. S. Tacit knowledge in production systems: how important is geography? the economic geography of innovation. In: POLENSKE, K. (Ed.). *The economic geography of innovation*. Cambridge, Cambridge University Press, 2007. p. 87-111.
- GLAESER, E. L.; KALLAL, H. D.; SCHEINKMAN, J. A.; SHLEIFER, A. Growth in cities. In: GLAESER, E.; KALLAL, H. D.; SCHEINKMAN, J. A.; SHLEIFER, A. (Orgs.) *The Journal of Political Economy*, v. 100, n. 6, p. 1126-1152, 1992.
- GONÇALVES, E.; ALMEIDA, E. Innovation and spatial knowledge spillovers: evidence from Brazilian patent data. *Regional Studies*, v. 43, n. 4, p. 513-528, 2009.
- GRILICHES, Z. Issues in assessing the contribution and development of research to productivity growth. *The Bell Journal of Economics*, v. 10, n. 1, p. 92-116, 1979.
- IBGE – Instituto Brasileiro de Geografia e Estatística. *Pesquisa de Inovação Tecnológica 2005*. Rio de Janeiro: IBGE, 2005.
- JACOBS, J. *The economy of cities*. London: Jonathan Cape, 1969.
- JAFFE, A. B. Real effects of academic research. *The American Economic Review*, v. 79, n. 5, p. 957-970, 1989.
- JAFFE, A. B. Technological opportunity and spillovers of R & D: evidence from firms' patents, profits, and market value. *The American Economic Review*, v. 76, n. 5, p. 984-1001, 1986.
- JAFFE, A. B.; TRAJTENBERG, M.; HENDERSON, R.; NARIN, F. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, v. 108, n. August, 1993. p. 577-598.

- LAURSEN, K.; REICHSTEIN, T.; SALTER, A. Exploring the Effect of Geographical Proximity and University Quality on University–Industry Collaboration in the United Kingdom. *Regional Studies*, v. 45, n. 4, p. 507-523, 2011.
- MARSHALL, A. Industry and Trade. *Journal of the Royal Statistical Society*, v. 83, n. 3, p. 292-301, 1920.
- MONTENEGRO, R. L.; BETARELLI JUNIOR, A. Análise e investigação dos fatores determinantes da inovação nos municípios de São Paulo. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 9, n. 1, p. 7-31, 2009.
- MONTENEGRO, R. L.; GONÇALVES, E.; ALMEIDA, E. Dinâmica espacial e temporal da inovação no estado de São Paulo: uma análise das externalidades de diversificação e especialização. *Estudos Econômicos*, v. 41, n. 4, p. 743-776, 2011.
- MORENO, R.; PACI, R.; USAI, S. Spatial spillovers and innovation activity in European regions. *Environment and Planning A*, v. 37, n. 10, p. 1793-1812, 2005.
- PONDS, R.; VAN OORT, F. G.; FRENKEN, K. Innovation, spillovers and university--industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography*, v. 10, n. 2, p. 231-255, 2010.
- RODRÍGUEZ-POSE, A.; VILLARREAL PERALTA, E. M. Innovation and regional growth in Mexico: 2000-2010. *Growth and Change*, v. 46, n. 2, p. 172-195, 2015.
- SINGH, J.; MARX, M. Geographic constraints on knowledge spillovers: political borders vs. spatial proximity. *Management Science*, v. 59, n. 9, p. 2056-2078, 2013.
- STERLACCHINI, A.; VENTURINI, F. R&D and Productivity in High-tech manufacturing: a comparison between Italy and Spain. *Industry and Innovation*, v. 21, n. 5, p. 359-379, 2014.
- STORPER, M.; VENABLES, A. J. Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography*, v. 4, n. 4, p. 351-370, 2004.
- TEECE, D.; PISANO, G. The Dynamic Capabilities of Firms: an Introduction. *Industrial and Corporate Change*, v. 3, n. 2, p. 537-556, 1994.
- THOMPSON, P. Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations. *Review of Economics and Statistics*, v. 88, n. 2, p. 383-388, 2006.
- VAN DER PANNE, G. Agglomeration externalities: Marshall versus Jacobs. *Journal of Evolutionary Economics*, v. 14, n. 5, p. 593-604, 2004.

Acknowledgments


Suelene Mascarini acknowledges financial support by master scholarship from FAPESP (process number 2010/13734-4) and CAPES (7 months).

ORCID

Suelene Mascarini  <https://orcid.org/0000-0002-9926-7877>

Renato Garcia  <https://orcid.org/0000-0001-9739-1658>

José Eduardo Roselino  <https://orcid.org/0000-0001-8066-8024>

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



Revista Brasileira de Estudos Regionais e Urbanos (RBERU)

v. 13, n. 2, p. 201-222, 2019

<http://www.revistaaber.org.br>

DETERMINANTES DA INOVAÇÃO AMBIENTAL: UMA ANÁLISE EMPÍRICA SOBRE A INDÚSTRIA DE TRANSFORMAÇÃO DE MINAS GERAIS*

Maria Alice Ferreira

Doutora em Economia Aplicada pela Universidade Federal de Viçosa (UFV)
Departamento de Economia da Universidade Estadual de Montes Claros (UNIMONTES)
E-mail: maria_alice16@hotmail.com

Maria Ivanilde Pereira Santos

Doutora em Ciências da Saúde pelo Programa de Pós-graduação em Ciências da Saúde (PPGCS/
UNIMONTES)
Departamento de Economia e Departamento de Saúde Mental e Saúde Coletiva da Universidade Estadual de
Montes Claros (UNIMONTES)
E-mail: ivanildps@yahoo.com.br

Emerson Costa dos Santos

Doutor em Economia Aplicada pela Universidade Federal de Viçosa (UFV)
Departamento de Economia da Universidade Estadual de Montes Claros (UNIMONTES)
E-mail: emersoco@gmail.com

RESUMO: O tema inovação ambiental ou “ecoinovação” ganha destaque no Brasil a partir da Conferência Rio-92, passando a ser amplamente debatido no cenário contemporâneo, tanto em nível acadêmico quanto governamental. Entretanto, essa discussão se intensifica a partir da inclusão do termo inovação no sistema de ciência e tecnologia brasileiro a partir de 2011. Nessa perspectiva, o objetivo do trabalho foi identificar possíveis heterogeneidades dos determinantes em relação à decisão das firmas da indústria de transformação de Minas Gerais em realizar inovações convencionais e/ou inovações ambientais. Para tanto, utilizou-se um modelo *probit* bivariado para captar o grau de relevância de cada determinante da adoção desses dois tipos de inovação na indústria mineira. A base de dados utilizada foi proveniente da Pesquisa de Inovação Tecnológica para o período de 1998 a 2011. Como principais resultados, o estudo aponta que a regulação ambiental, a cooperação, o apoio do Governo e as técnicas de gestão ambiental impactam de forma mais expressiva sobre a probabilidade de as firmas em adotarem inovação ambiental. Destaca-se que o resultado deste estudo, além de identificar fatores motivadores de investimentos em inovações convencionais e ambientais, pode subsidiar políticas públicas no sentido de ampliar ações inovativas e estratégias sustentáveis na indústria mineira.

Palavras-chave: Inovação ambiental; Indústria; Minas Gerais.

Classificação JEL: O31; Q55; C23.

*Recebido em: 14/06/2018. Aceito em: 01/05/2019.

DETERMINANTS OF ENVIRONMENTAL INNOVATION: AN EMPIRICAL ANALYSIS ON THE TRANSFORMATION INDUSTRY OF MINAS GERAIS

ABSTRACT: The environmental innovation or “eco-innovation” theme has been highlighted in Brazil since the Rio-92 Conference, being widely debated in the contemporary scenario, both at the academic and governmental levels. However, this discussion intensifies from the inclusion of the term innovation in the Brazilian science and technology system from 2011. From this perspective, the objective of the work was to identify possible determinant heterogeneities in relation to the decision of Minas Gerais manufacturing industry firms to carry out conventional and/or environmental innovations. Therefore, a bivariate probit model was used to capture the degree of relevance of each determinant of the adoption of these two types of innovation in the mineira industry. The database used came from the Technological Innovation Survey for the period from 1998 to 2011. As main results, the study points out that environmental regulation, cooperation, government support and environmental management techniques have a more significant impact on the likelihood of firms adopting environmental innovation. It is noteworthy that the result of this study, besides identifying motivating factors for investments in conventional and environmental innovations, can subsidize public policies in order to expand innovative actions and sustainable strategies in the mineira industry.

Keywords: Environmental innovation; Industry; Minas Gerais.

JEL Codes: O31; Q55; C23.

1. Introdução

A questão da mudança dos padrões insustentáveis de produção e de consumo que geram grandes impactos sobre a utilização dos recursos naturais disponíveis no mundo é foco de intenso debate atual. Para uma melhor atuação do ser humano para minimizar os efeitos ambientais desses padrões, novas perspectivas socioeconômicas, tecnológicas, políticas e ambientais devem estar em consonância com as discussões desenvolvidas para melhorar a qualidade de vida da sociedade. Assim, a busca por alternativas que reduzam as pressões ambientais e atendam às necessidades básicas da humanidade é objetivo a ser atingido por todos que estão comprometidos com a questão ambiental.

A partir do final da década de 1980 e na década de 1990, a questão ambiental ganhou espaço no debate acadêmico, governamental e também nas preocupações sociais das empresas brasileiras ao se perceber o crescente interesse e preocupação da sociedade em relação ao meio ambiente. Metas ambientais passaram a ser definidas em convenções globais como as de Montreal (1987), Rio de Janeiro (1992) e Kyoto (1997). A meta do desenvolvimento sustentável ganhou comprometimento global, vinte anos após a reunião em Estocolmo, na Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento (Rio-92), que foi convocada a partir do reconhecimento de que os padrões de produção e de consumo, principalmente nos países desenvolvidos, haviam alcançado níveis insustentáveis, colocando em risco a estabilidade de diversos bens e serviços ambientais que garantem tanto a continuidade das atividades produtivas como a própria qualidade de vida (PNUMA, 2005).

Na conferência Rio-92, estabeleceu-se a Agenda 21, uma ação global para o século XXI, com uma visão de longo prazo para equilibrar necessidades econômicas e sociais com os recursos naturais do planeta. Desde então, setores da sociedade iniciaram um processo de adoção de políticas impondo requisitos ambientais a inúmeras atividades econômicas e a demanda por produtos ambientalmente menos agressivos cresceu em paralelo. Os padrões internacionais de eficiência ambiental se elevaram gradativamente e algumas instituições passaram a atrelar financiamentos de projetos aos resultados de avaliações ambientais (UNITED NATIONS, 1992). Desse modo, crescentes exigências do mercado, pressões dos consumidores e competitividade entre as empresas

fazem com que a adoção de princípios ambientais torne-se condição necessária à sobrevivência das empresas.

Nesse período, surge o conceito de inovação ambiental ou *eco-innovation* que segundo Kemp e Pearson (2008) é resultante da inovação com base no Manual de Oslo com a adição do fator ambiental. Assim, a “ecoinovação” é a produção, assimilação e exploração de um produto e/ou processo, serviços ou gestão, método que é novo para a organização e que resulta, ao longo de seu ciclo de vida, na redução de poluição e de outros efeitos negativos decorrentes da utilização dos recursos, em comparação com alternativas relevantes. Portanto, a inovação ambiental é basicamente uma inovação que reduz ou elimina impactos ambientais causados pela firma, sendo esses em âmbito local ou global e a inovação pode ser desenvolvida de maneira intencional e premeditada pela empresa ou acidental.

Dentre as diferentes organizações envolvidas com a questão ambiental, as empresas desempenham um papel importante, uma vez que seu produto alcança diretamente a população influenciando no seu consumo. Entretanto, o setor industrial é um dos que mais provoca danos ao meio ambiente, seja por seus processos produtivos, seja pela fabricação de produtos poluentes e/ou que tenham problemas de disposição final após sua utilização (PORTER; LINDE, 1995). Por esse motivo, a adequação dos padrões de produção para a minimização de impactos ambientais se torna de fundamental importância – por pressões externas da população, dos investidores, das organizações não governamentais que passaram a exigir das empresas soluções para o controle da poluição, do desmatamento e da degradação, e responsabilidade pelos seus processos de produção; ou internas, quando ocorre a percepção por parte dos agentes de que essas tecnologias favorecem os lucros.

O estado de Minas Gerais, em particular, especializou-se, desde a década de 1970, em atrair indústrias sujas e mais intensivas no uso de recursos naturais. O crescimento de tais indústrias na região estava ligado ao tipo de inserção brasileira na nova divisão internacional do trabalho, em que os países periféricos tornaram-se atraentes para a instalação de indústrias que exercem fortes impactos no meio ambiente, com consequências para a qualidade do meio ambiente e de vida local. Nesse contexto, em 1977, foi criada a Comissão de Política Ambiental do estado de Minas Gerais (COPAM), com competência para definir a política de meio ambiente no estado e o consequente repasse de recursos, no orçamento estadual, para as atividades de qualidade e de controle ambiental e sua implementação com evoluções na organização administrativa e orçamentária. A criação da COPAM, renomeada em 1988, como Conselho Estadual de Política Ambiental, constituiu-se um marco para início da implantação de políticas de meio ambiente no estado (RIBEIRO, 2006).

Assim, Minas Gerais foi um dos estados pioneiros na implantação da legislação ambiental e a preocupação governamental com questões ambientais tornou-se mais evidente. Organizações públicas estaduais que se concentram na temática do meio ambiente se tornaram muito mais frequentes, podendo ser citada como exemplo a Fundação Estadual do Meio Ambiente (FEAM) que é um dos órgãos seccionais de apoio do COPAM e atua vinculado à Secretaria de Estado de Meio Ambiente e Desenvolvimento Sustentável (SEMAD) (SERRANO; BARBIERI, 2008).

Por apresentar muitas características peculiares, tanto em termos locais e infraestruturais bem como fisiográficos, com grande disponibilidade de recursos naturais e matérias-primas, o estado de Minas Gerais concentra muitas empresas importantes para o desenvolvimento econômico e regional. Poucos estudos na literatura têm analisado, de forma específica, os efeitos da adoção de inovações ambientais por parte dessas empresas. Nesse sentido, o presente trabalho buscou identificar possíveis heterogeneidades entre os determinantes em relação à decisão da firma de realizar inovações convencionais¹ e inovações ambientais das firmas da indústria de transformação de Minas Gerais no período de 1998 a 2011.

A motivação do trabalho é contribuir de alguma maneira para o debate sobre a inovação ambiental e convencional das empresas mineiras, no que se refere à necessidade de identificar o que

¹ Para distinguir os conceitos, o presente estudo define a inovação tecnológica como inovação “convencional” e essa terminologia será utilizada ao longo do trabalho.

motiva as empresas de Minas Gerais a realizarem inovação que permita reduzir o impacto ambiental, de modo a possibilitar o desenvolvimento de políticas e estratégias mais sustentáveis. Para tanto, utilizou-se o método de estimativa *probit* bivariado, em que são incluídas no modelo duas variáveis dependentes binárias e considera-se a possibilidade de existência de um fator endógeno que relaciona essas variáveis, de maneira que a escolha por uma determinada decisão pode afetar, de forma direta, a probabilidade de a firma decidir pela outra opção. A hipótese sustentada pelo trabalho é de que a regulação é um importante determinante para a adoção tanto de inovação ambiental quanto de inovação convencional.

Este estudo está dividido em três seções, além desta introdução e das conclusões. A seção 2 compreende a abordagem teórica e da literatura empírica sobre a conceituação de inovação ambiental. A seção 3 apresenta a metodologia e a fonte das variáveis utilizadas e a seção 4 analisa os resultados encontrados por meio de estatísticas descritivas e estimação econométrica.

2. Teoria da inovação e inovação ambiental

O enfoque neoclássico sobre a inovação se inicia pela função de produção, em que insumos como trabalho e capital ou outros recursos são transformados em produtos. Assim, várias combinações de fatores são tecnicamente viáveis e os agentes econômicos são movidos pela racionalidade substantiva, para os quais o comportamento das empresas é pré-determinado pelo princípio da maximização e supõe-se um perfeito conhecimento do mercado pelos agentes econômicos. Metodologicamente, o argumento centra-se em dois aspectos fundamentais: equilíbrio e análise estática. A análise de equilíbrio significa que o sistema econômico apenas ajusta-se às mudanças geradas externamente, sem alterar os seus parâmetros estruturais. Alcançar o equilíbrio significa que as forças econômicas que anteriormente causaram sua perturbação cessaram de produzir seus efeitos, não havendo tendência a mudanças; se estas ocorrem, não se devem a movimentos endógenos, mas a eventos estranhos a ela (WALRAS, 1983).

Em contraponto, o enfoque evolucionário da firma vem sendo desenvolvido com grande vigor ao longo das décadas de 1970 e 1980, a partir de trabalhos pioneiros de Nelson e Winter (1977) e Dosi (1988) e surge de uma preocupação em se estabelecer um marco teórico alternativo à economia neoclássica, capaz de tratar de forma mais ampla o problema da mudança tecnológica. Trata-se, portanto, do que se pode chamar de “economia da mudança tecnológica” (DOSI, 1988) e que coloca a firma como elemento central na dinâmica capitalista.

O ponto inicial da teoria evolucionária está no abandono das premissas da teoria neoclássica de “racionalidade maximizadora”, “tendência de equilíbrio dos mercados” e “mecanismo de preço” como instrumento principal da concorrência entre firmas. Os evolucionários destacam o papel da inovação como determinante fundamental dos saltos de produtividade do sistema econômico e atribuem à ação da empresa privada, em busca do lucro, a função de agente propulsora da inovação. A teoria evolucionária insiste na necessidade de analisar a empresa como organização constituída por indivíduos distintos e com características cognitivas próprias. A diversidade leva à ideia de racionalidade limitada dos agentes econômicos que possuem características diferenciadas, resultado de um processo de aprendizado acumulado ao longo do tempo (TIGRE, 2005). A conceituação de inovação convencional e inovação ambiental pelas duas vertentes está apresentada no Quadro 1.

O que se pode observar por meio do Quadro 1 é que a teoria neoclássica se mostra limitada quando se busca estudar fenômenos complexos e, dadas as premissas da teoria evolucionária apresentadas, essa teoria apresenta-se como um referencial teórico consistente para abordar as questões que envolvem tecnologia e meio ambiente, pois trata o sistema econômico como inerentemente dinâmico e evolutivo baseado em um processo constante de inovação tecnológica, organizacional e institucional. Entretanto, cabe ressaltar que os pressupostos dos modelos neoclássicos são mais complacentes com os dados geralmente disponíveis, de modo que eles se mostram adequados para uma introdução ao tema, embora apresentem limitações para tratar de estudos de fenômenos complexos.

Quadro 1 – Conceituação Inovação Convencional e Inovação Ambiental

Visão Neoclássica	Visão Evolucionária
A firma é vista como uma “caixa preta” que combina fatores de produção disponíveis no mercado para produzir produtos comercializáveis. A análise interna da firma não constitui uma questão relevante, pois, em situação de concorrência perfeita e na ausência de progresso técnico, a firma teria pouca escolha a fazer	A firma realiza esforços inovativos para a introdução de novos produtos e/ou processos com a finalidade de buscar vantagens competitivas. Em caso de sucesso da inovação, a firma pode dispor de um lucro de monopólio, mesmo que temporário. O processo de difusão tecnológica tende a reduzir os lucros extraordinários advindos da inovação. As firmas inovadoras reagem à perda de lucros ou à ameaça buscando outras inovações.
A inovação é vista sob a perspectiva das falhas de mercado, particularmente em relação à produção de conhecimento, que se caracteriza por ser incerta, indivisível, assumir aspectos de bem público, apresentar externalidades e ausência de garantias de apropriabilidade, não havendo incentivos para que as firmas desenvolvam inovação.	O processo inovativo é como busca, descoberta, experimentação, desenvolvimento, imitação e adoção de novos produtos, processos produtivos e formas organizacionais; e ele é caracterizado por um elevado grau de incerteza, associada à impossibilidade de prever <i>ex ante</i> o resultado do esforço inovativo e os desdobramentos da inovação.
As inovações ambientais fornecem, além do efeito de <i>spillover</i> inerente às inovações, mais uma externalidade, ou seja, reduzem a carga ambiental beneficiando a sociedade a custos elevados.	As inovações ambientais podem gerar uma mudança tecnológica que seja voltada à sustentabilidade. Caso sejam incorporadas de forma significativa no processo de desenvolvimento econômico, o meio ambiente pode se tornar um elemento importante na reconfiguração de um novo paradigma tecnoeconômico “verde”.
Não há preocupação em relação às especificidades setoriais ou das firmas, ao tratar a firma como agente individual, sem reconhecê-la como entidade coletiva, dotada de objetivos e regras diferenciadas.	A existência de um grande número de tecnologias e suas respectivas lógicas internas e trajetórias evolutivas, em determinados setores, indústrias ou firmas, dá-se em meio a incentivos e oportunidades diferentes, devido aos distintos graus de complexidade e especificidade de cada conhecimento tecnológico base.

Fonte: Elaboração própria com base em Tigre (2005), Dosi (1988) e Freeman (1996).

2.1. Evidências empíricas: determinantes da inovação ambiental

A literatura empírica sobre os determinantes de inovações ambientais teve sua sustentação no fato de que essas inovações diferem de outras inovações como externalidades e impulsores da sua introdução, destacando, principalmente, a importância da regulação para acioná-los. Alguns autores viam apenas nas regulações a solução para o desenvolvimento de tecnologias mais limpas e métodos de produção que gerassem menos resíduos ambientais. No entanto, há autores que acreditam que a inovação não deve ser associada de forma sistemática às regulações ambientais, mas deve ser vista como resultado de um processo complexo e interativo (OLTRA, 2008). Assim, desde os anos de 1990, vários estudos empíricos tentam identificar outros determinantes desse tipo de inovação. Entretanto, essas pesquisas são muito heterogêneas em termos de metodologias e resultados, uma vez que uma das principais dificuldades é encontrar dados e indicadores adequados sobre inovações ambientais.

Nos estudos empíricos, verificaram-se algumas formas de mensurar a inovação ambiental. A discussão clássica aponta que a inovação ambiental pode ser medida pelo investimento direto em Pesquisa & Desenvolvimento (P&D) ou pelas patentes. Desse modo, enquanto alguns trabalhos consideram modelos de contagem em que a inovação ambiental é medida por variáveis contínuas, como, por exemplo, o investimento em P&D ou o número de patentes concedidas pela firma; outros estudos consideram modelos binomiais em que a inovação ambiental é representada por uma variável binária, que assume valor igual a 1 (um) se as firmas adotam inovação ambiental e 0 (zero) caso contrário.

Horbach e Rennings (2012) testaram se a inovação ambiental é movida por diferentes determinantes. Com base na Pesquisa de Inovação da Comunidade Alemã em 2009, com 7.061

empresas dos setores de mineração e pedreiras, manufatura, fornecimento de energia e água e de serviços, os autores estimaram um modelo *probit* e concluíram que, exceto para as inovações ambientais que têm impacto sobre a redução de material e consumo de energia, a regulação ambiental apresenta correlação positiva e significativa com a adoção dos outros tipos diferentes de inovação ambiental.

Del Rio et al. (2011) utilizaram um modelo de dados em painel das Indústrias de transformação espanholas no período 2000-2006 e confirmaram a relevância da regulação ambiental e da intensidade de capital físico para explicar o investimento em tecnologias ambientais. Os autores verificaram que os determinantes de inovação ambiental podem ser diferentes entre a adoção de tecnologias limpas e *end-of-pipe*. Nesse estudo, inovação ambiental é medida como um investimento em tecnologias ambientais e a regulação ambiental é representada pela intensidade de gastos com proteção ambiental. Ademais, a intensidade de exportação apresentou uma correlação negativa com inovação ambiental, indicando, de acordo com os autores, que a proteção ambiental é sempre feita nos setores que são protegidos e altamente regulados, como, por exemplo, o de energia.

Brunnermeier e Cohen (2003) realizaram um estudo dos determinantes da inovação ambiental nas indústrias de transformação dos Estados Unidos entre 1983 e 1992. Os autores utilizaram o número de patentes ambientais de sucesso concedidas à indústria como uma *proxy* para a inovação ambiental. Os autores estavam interessados em entender como a adoção de inovação ambiental responderia em caso de mudanças nos gastos de redução de poluição e em caso de mudanças nas atividades de monitoramento do Governo, *proxies* para a política ambiental. Eles estimaram quatro modelos diferentes (efeitos fixos, Poisson, binomial negativo de efeitos fixos e efeitos aleatórios) e obtiveram resultados robustos concluindo que há uma relação positiva e significativa entre gastos de redução de poluição e número de patentes “verdes” concedidas pela firma. Não houve efeito de monitoramento das atividades sobre a inovação ambiental. Além disso, os autores encontraram uma correlação positiva entre as atividades inovativas e a intensidade de exportação. Eles interpretaram a intensidade da exportação como indicador da demanda externa por produtos mais ecológicos.

Wagner (2009) utilizou dados combinados e de patentes da pesquisa das empresas da indústria alemã para explorar a ligação entre a inovação ambiental, as patentes e as medidas de gestão ambiental. Ele encontrou que o nível de implementação de sistemas de gestão ambiental tem um efeito positivo exclusivamente sobre a inovação de processos ambientais, considerando que é negativamente associado com o nível geral de atividade de patentes de uma empresa. Em sua análise, uma combinação de pesquisas e dados de patentes ajudam a evitar o problema da autopercepção das atividades ambientais.

De Marchi (2012) explora a relação entre as estratégias de cooperação em P&D das empresas e sua propensão a introduzir inovações ambientais. Com base em dados do *Community Innovation Survey* das firmas espanholas (PITEC), a autora investigou especificidades que afetam mais como as inovações são desenvolvidas e, em particular, a maior importância da cooperação em P&D com parceiros externos. As estimativas econométricas, controlando o viés de seleção, sugerem que as empresas inovadoras ambientais cooperam em inovação com parceiros externos em maior escala do que com outras empresas inovadoras. Ademais, a cooperação com fornecedores e universidades é mais relevante do que para outros inovadores e conclui que há um efeito de substituição entre as atividades de cooperação e o esforço interno de P&D.

Especificamente relatando estudos sobre as firmas da indústria brasileira, Lucchesi et al. (2014), utilizando dados em painel entre 1998 e 2008 para o Brasil, apontam alguns determinantes da adoção de inovações ambientais. As variáveis utilizadas foram: regulação ambiental, tamanho, exportação, capital nacional e estrangeiro, *dummies* para os setores e para os estados. A hipótese de viés ambiental é verificada, o que indica que a regulação ambiental apresenta um papel importante para influenciar as firmas brasileiras a adotar tanto inovações ambientais tecnológicas quanto inovações ambientais organizacionais. Em relação especificamente ao processo inovativo em países em desenvolvimento, os resultados confirmam que empresas de capital estrangeiro têm maior probabilidade de adotar inovações “verdes”, geralmente através de transferência tecnológica e

acordos de licenciamento. O tamanho da firma, o fator da empresa ser exportadora e a intensidade de capital físico também são importantes determinantes da inovação ambiental nas indústrias de transformação brasileiras.

Oliveira et al. (2012) analisaram uma pesquisa inédita realizada com 2.400 empresas atuando no mercado brasileiro em diversos setores da indústria de serviços no período de 2009-2011. Alguns dos resultados obtidos foram: a maioria das empresas utiliza recursos próprios como a principal fonte financiadora da atividade inovativa e a pouca utilização de financiamentos e subsídios governamentais para a realização da inovação ambiental contrasta com a evidência de que os elevados custos para inovar são a maior barreira enfrentada pelas empresas. Os autores concluem que um maior conhecimento das fontes de financiamento e subvenção governamentais pode estimular a ocorrência de inovação ambiental no contexto das empresas brasileiras.

Gonçalves e Simões (2005) aplicaram técnicas de análise estatística multivariada a variáveis da PINTEC, referentes aos esforços de aprendizagem e incorporação tecnológica, com o objetivo de compreender as características dos esforços setoriais de inovação. Entre outras particularidades nacionais, os autores destacaram que o nível de gastos em atividades inovadoras é maior em setores que foram objetos da política industrial nas décadas de 1960 e 1970 – aeronáutica e siderurgia. Além disso, verificaram que setores com maiores oportunidades tecnológicas apresentam esforços reduzidos em P&D, o que revela a característica do sistema nacional de passividade e dependência de canais internacionais de transferência tecnológica.

De modo geral, os estudos apontaram ainda a importância do tamanho da firma, a intensidade de exportações, as forças internas (imagem e pressão corporativa) e a heterogeneidade setorial como características que influenciam a realização de inovação ambiental. Ademais, verificou-se uma controvérsia sobre a influência das forças de mercado na adoção de inovação ambiental, de modo que, além da oferta, a demanda também pode influenciar a adoção de inovações ambientais.

2.2. Conceituação da inovação ambiental

A inovação tecnológica tem sido definida como “a introdução de novos produtos, processos ou serviços para o mercado” (UNCTAD, 2006, p. 10). O Manual de Oslo apresentou uma definição mais ampla, segundo a qual a inovação é a implementação de um produto e/ou processo novo ou significativamente melhorado (bem ou serviço), um novo método de *marketing* ou organizacional nas práticas de negócios (OCDE, 2005). Essa definição incorpora a difusão de tecnologia ao realçar que uma inovação não precisa ser nova para o mercado para ser qualificada como tal, basta que seja nova para a empresa que a implementa (OZUSAGLAM, 2012).

A inovação convencional pode ser classificada de formas distintas com base em diferentes dimensões, sendo uma delas a distinção entre inovação incremental e radical. A inovação radical está relacionada ao desenvolvimento de um novo produto, processo ou forma de organização da produção inteiramente nova, ou seja, quando acontece um salto na tecnologia do produto e/ou processo. Já a inovação incremental associa-se à introdução de qualquer espécie de melhoria em produto, processo ou organização da produção sem qualquer mudança significativa na estrutura industrial, podendo resultar em aumento da produtividade e da qualidade, redução de custos e maior abrangência das aplicações de um produto ou processo (CASSIOLATO; LASTRES, 2005).

Outra dimensão é a inovação ambiental ou a “ecoinovação”, que muitas vezes é apresentada como uma forma específica de inovação. Esse conceito é recente na literatura, com conceito inicial apresentado por Fussler e James (1996), que definiram as “ecoinovações” como novos produtos e processos que proporcionam valor ao consumidor e aos negócios, mas que significativamente reduzem os impactos ambientais. Assim, uma vez que esse termo foi posto em debate, várias definições têm sido propostas na literatura. No entanto, a literatura apresenta uma divisão fundamental nos critérios que definem a inovação ambiental, de modo que, enquanto alguns estudos consideram uma definição ampla de inovação ambiental focada no desempenho e, portanto, um

possível subproduto da inovação convencional, outros estudos consideram a definição mais restrita com base em um objetivo *a priori* de redução de impactos ambientais.

A inovação ambiental pode ser definida como a introdução de novos procedimentos técnicos e organizacionais, no âmbito da produção industrial, que levam à maior proteção do meio ambiente. Por sua vez, diversos autores (por exemplo, Arundel et al. (2007), Kemp e Pearson (2008) e Oltra (2008)) reiteram que grande parte das inovações têm impactos positivos no meio ambiente, independentemente das suas motivações *a priori*. Assim, os autores definem inovação ambiental de forma ampla: qualquer inovação que reduz os danos ambientais, que pode ser desenvolvida de maneira intencional e premeditada pela empresa ou mesmo acidental. Para a OCDE (2010), o âmbito da inovação ambiental pode ir além dos limites convencionais das empresas em inovar e envolver um regime social mais amplo que provoca alterações das normas socioculturais e estruturas institucionais.

O conceito de inovação ambiental apresentado inclui não apenas a inovação que visa reduzir os impactos ambientais, mas também casos em que a inovação convencional conduz à redução de impactos sem que isso seja um objetivo explícito. Nesse sentido, as inovações convencionais que têm efeito ambiental positivo também são consideradas como “ecoinovações”. Como mencionado pela OCDE (2010), a “ecoinovação” pode ser motivada ambientalmente, mas também pode ocorrer como resultado de outros objetivos, tais como o cumprimento dos regulamentos e normas, o aumento da produtividade e a redução dos custos de entrada. De acordo com essa definição, uma inovação convencional deve também ser considerada como uma potencial “ecoinovação” (OZUSAGLAM, 2012).

A definição do “*Measuring Eco-Innovation*” Project (Projeto MEI, 2002-2006) adota a ideia de desempenho ambiental, ao invés de apontá-lo como o critério de definição fundamental. Assim, as “ecoinovações” não precisam ter um objetivo ambiental em qualquer fase de desenvolvimento ou utilização do produto/processo (SPEIRS et al., 2008). Conforme o projeto MEI, o conceito de “ecoinovação” não deve ser limitado a novas ou melhores tecnologias ambientais, o que resulta na contabilização de cada produto ou serviço ambientalmente melhorado como uma “ecoinovação”. A justificativa para isso é que levar em conta apenas o objetivo de inovação limitaria muito o termo.

Os contribuintes para o projeto MEI reconhecem que a “ecoinovação” está suscetível a ocorrer de diversas maneiras em toda a economia (KEMP; PEARSON, 2008). Diante disso, é necessário fazer a distinção entre as inovações que terão um grande impacto positivo sobre o meio ambiente e aquelas que terão apenas um impacto limitado. Assim, embora o projeto MEI afirme que “uma empresa que adote bem, serviço, gestão de processos de produção ou método de negócio com benefício ambiental é uma ‘ecoinovação’” (KEMP; PEARSON, 2008, p. 17), também reconhece a necessidade de distinguir quatro tipos deecoinovadores: i) Ecoinovadores estratégicos são ativos em setores de ecoequipamentos e serviços, e/ou desenvolvimento de “ecoinovações” para venda para outras empresas; ii) Ecoinovadores estratégicos implementam intencionalmente “ecoinovações”, sejam elas desenvolvidas dentro da empresa, adquiridas de outras empresas, ou ambos; iii) Ecoinovadores passivos não têm nenhuma estratégia específica para a “ecoinovação”, embora possam implementar acidentalmente inovações que resultam em benefícios ambientais; iv) Empresas que não são ecoinovadores e não desenvolvem nem tipo de inovação com benefícios ambientais.

Devido à natureza abrangente da definição da “ecoinovação” pelo projeto MEI, a distinção entre quatro tipos de eco-inovadores não é suficiente para captar a diversidade do conceito. É importante classificar as “ecoinovações” de acordo com a natureza das inovações envolvidas. Outra classificação é proposta por Kemp e Foxon (2007) que distinguem as “ecoinovações” entre: as tecnologias ambientais, inovações organizacionais para o meio ambiente, as inovações em produtos e processos que oferecem benefícios ambientais, o sistema de inovações verde e as tecnologias de propósito geral.

As tecnologias ambientais ou limpas se referem às mudanças que ocorrem nos processos de produção, a fim de reduzir a emissão de poluentes ou diminuir o uso de recursos ou o consumo de energia (OOSTERHUIS, 2006). As inovações organizacionais para o meio ambiente se referem à

introdução de métodos de organização e sistemas de gestão que tratam de questões ambientais encontradas no processo de produção. Teoricamente, a gestão ambiental deve garantir qualidade e preservação do meio ambiente nos limites intra e interorganizacionais, revelando técnicas e procedimentos com potencial suficiente para a mitigação dos problemas no setor industrial (OZUSAGLAM, 2012). As inovações em produtos e processos que oferecem benefícios ambientais incluem todos os produtos novos ou ambientalmente melhorados, bem como os serviços ambientalmente benéficos.

Kemp e Foxon (2007) definem o sistema de inovações verde como sistemas alternativos de produção e consumo que respeitam mais o ambiente do que os existentes. Esse sistema envolve um conjunto de mudanças em tecnologias de produção, conhecimento, organização, instituições e infraestruturas e, possivelmente, mudanças no comportamento do consumidor, como, por exemplo, agricultura biológica e energias renováveis. Já as tecnologias de propósito geral são parte importante do cenário tecnológico. Elas não são rotuladas como tecnologias completamente verdes, mas com certas configurações e tipos de usos ambientais, tais como a biotecnologia (KEMP; FOXON, 2007).

Em relação às inovações convencionais, a geração de inovação ambiental depende em grande medida dos benefícios recebidos pelo inovador. As inovações bem-sucedidas devem fornecer maior valor ou reduzir custos e, em última análise, aumentar as receitas dos clientes existentes ou atrair novos clientes (CARRILLO-HERMOSILLA et al., 2010). O inovador ambiental pode se beneficiar da sua atividade inovadora de forma direta ou indireta. Os benefícios diretos para uma “ecoinovação” consistem em vantagens operacionais, tais como redução de custos de uma maior produtividade dos recursos, melhor logística e vendas de comercialização. Benefícios indiretos incluem uma imagem melhor, melhores relações com fornecedores, clientes e autoridades, uma capacidade geral de inovação reforçada pelos contatos com os detentores de conhecimento, benefícios de saúde e segurança e maior satisfação do trabalhador. Assim, os benefícios indiretos que em sua maioria têm valor no longo prazo podem ser os condutores mais importantes para o comportamento verde proativo.

Nesse sentido, uma vez que os benefícios indiretos para a geração de inovação ambiental ocorrem no longo prazo e a mudança do padrão tecnológico é extremamente complexa, ela depende de muitas variáveis e, mesmo induzida por qualquer tipo de política, não se sabe *a priori* todas as consequências que irão decorrer delas. Desse modo, o estudo da mudança tecnológica envolve diversos aspectos, sendo um processo evolucionário. Ademais, a mudança tecnológica na direção da sustentabilidade ambiental depende de outros fatores não econômicos, como desenvolvimento de capacidades específicas da firma, infraestrutura e mudanças institucionais.

3. Metodologia

3.1. Base de dados e descrição de variáveis

Para a implementação do modelo proposto, utilizaram-se os dados da Pesquisa de Inovação Tecnológica (PINTEC), que coleta informações relevantes sobre inovação das empresas. As informações da PINTEC são divulgadas a partir da Classificação Nacional de Atividades Econômicas – CNAE versão 2.0.² A PINTEC é conduzida a cada 3 (três) anos, seguindo o Manual de Oslo e a recomendação da *Community Innovation Survey* (CIS). A PINTEC é composta por empresas formais com dez ou mais pessoas ocupadas, pertencentes aos segmentos da indústria extrativa, manufatureira e de alguns serviços selecionados. No caso da indústria de transformação, a PINTEC é censitária para o grupo de empresas industriais com quinhentas ou mais pessoas ocupadas e aleatória para as demais. Para a composição do estrato aleatório, além da adoção de

² O acesso aos microdados ocorreu mediante aprovação de projeto encaminhado ao IBGE, sob o número de processo 03605.001848/2015-20 para as bases da PINTEC de 1998 a 2011 que estavam disponíveis para o processo. A manipulação dos dados e as estimações foram feitas diretamente na sala de sigilo do IBGE.

técnicas de amostragem estratificada proporcional ao tamanho, assume a premissa de que a inovação constitui um fenômeno raro, o que justifica atribuir maior probabilidade de seleção àquelas empresas que têm maior potencial inovador, sendo que isso é expresso por meio de algumas características observadas da firma, como ter lançado patente, recebido financiamento/subvenção para inovação, sido inovadora, conduzido P&D, entre outras. Essa maior probabilidade não causa viés nos resultados, uma vez que o peso de cada empresa na amostra é inversamente proporcional à sua probabilidade de seleção (CARVALHO; AVELAR, 2015).

Na construção dos dados em painel, utilizaram-se cinco edições³ da PINTEC: PINTEC 2000 (referente ao período entre 1998 e 2000), PINTEC 2003 (referente ao período entre 2001 e 2003), PINTEC 2005 (referente ao período entre 2003 e 2005), PINTEC 2008 (referente ao período entre 2006 e 2008) e PINTEC 2011 (referente ao período entre 2009 e 2011). Desse modo, foi possível criar um painel desbalanceado de 12 anos referente ao período de 1998 a 2011, com 6.637 observações. Os modelos de regressão com dados em painel são também chamados de dados combinados, por agregarem uma combinação de séries temporais e de observações em corte transversal multiplicadas por T períodos de tempo. Nesse caso, há muito mais informação para se estudar fenômeno e graus de liberdade adicionais. Podem-se destacar algumas vantagens dos dados em painel em relação ao uso específico do corte transversal ou das séries temporais (BALTAGI, 2001).

As variáveis extraídas dessas fontes podem ser vistas de forma sistemática no Quadro 2 e estão organizadas em três grupos. No primeiro grupo, estão os dados referentes à inovação ambiental e convencional, que são as variáveis dependentes e os gastos com atividades inovativas:

- A inovação ambiental está associada ao conceito amplo de inovação ambiental, ao considerar a inovação ambiental como um subproduto da inovação convencional. Essa variável é definida com base na questão v87 do questionário da PINTEC 2000 e v105 dos questionários das PINTECs 2003, 2005, 2008 e 2011. Essa questão diz respeito à importância dos resultados das inovações de produto e/ou processo, implementadas pelas empresas que permitiram reduzir o impacto sobre o meio ambiente, classificada por níveis: alta, média, baixa e não relevante. Para o modelo, a variável assumiu 1 (um) para importância alta e média e 0 (zero) para baixa e não relevante. Ademais, a variável inovação representa a inovação convencional.
- Os gastos com atividades inovativas é representado pelo cálculo dos gastos em Pesquisa & Desenvolvimento (P&D), do gasto em aquisição externa de P&D, aquisição de outros conhecimentos externos voltados para a inovação, aquisição de *software* voltado para a inovação, aquisição de máquinas e equipamentos voltados para a inovação e treinamento orientado ao desenvolvimento de produtos e processos novos diretamente ligados ao lançamento de um produto novo ou aperfeiçoado. A abrangência do investimento em atividades inovativas possui vantagem sobre o investimento em P&D usual, pois absorve um número maior de tipos de esforços inovativos por parte da empresa, uma vez que considera não somente as atividades inovadoras provenientes dos orçamentos formais das empresas, mas também uma variedade maior desses investimentos realizados pelas firmas. A definição proposta pelo estudo está fundamentada em Cohen e Levinthal (1989) que apresentam um conceito mais amplo de atividades inovativas que pode ser gerada por meio do investimento em P&D e também da capacidade de absorção tecnológica externa à firma.

No segundo grupo, estão as variáveis relacionadas ao ambiente concorrencial da firma, caracterizado por variáveis como:

³ Os microdados para o ano de 2014 não estavam disponíveis para o processo no momento da pesquisa, sendo possível utilizar apenas os dados disponibilizados no site da PINTEC para a análise descritiva das variáveis. Para maiores detalhes, ver: <https://ww2.ibge.gov.br/home/estatistica/economia/industria/pintec/2014/default.shtm>.

- Regulação ambiental, representada por uma variável binária, com valores unitários para as empresas que responderam que a inovação possibilitou o enquadramento em regulações e normas padrão relativas ao mercado interno ou externo com importância alta ou média; condição de gerenciamento da inovação por parte da firma é representada pela variável controle e gerenciamento, que abrange o gerenciamento formal (marcas) e informal (segredo industrial, complexidade do desenho do produto) visando atender normas de certificação; e técnicas de gestão que correspondem à informação de implementação de alguma técnica de gestão ambiental pela firma.

Quadro 2 – Variáveis utilizadas

Códigos	Variáveis	Descrição
Atividades Inovativas e Inovação		
G	Gastos com Atividades Inovativas	Valor do investimento total em atividades inovativas (em Reais).
IA	Inovação Ambiental	Variável binária que assume 1 (um) caso a firma tenha realizado inovação com redução de impacto sobre o meio ambiente e 0 (zero) caso contrário.
IC	Inovação Convencional	Variável binária que assume 1 (um) caso a firma tenha realizado inovação e 0 (zero) caso contrário.
Ambiente Concorrencial		
REG	Regulação Ambiental	Variável binária que assume 1 (um) caso a firma se enquadre em regulações e normas padrão relativas ao mercado interno ou externo e 0 (zero) caso contrário.
Controle	Controle e Gerenciamento	Variável binária que assume 1 (um) caso a firma realize gerenciamento formal e informal e 0 (zero) caso contrário.
TecGA	Técnicas de Gestão Ambiental	Variável binária que assume 1 (um) se a empresa implementou alguma técnica de gestão ambiental e 0 (zero) caso contrário.
Intec	Intensidade Tecnológica	4 (quatro) variáveis binárias de intensidade tecnológica, que assumem 1 (um) caso a firma pertença à atividades econômicas com intensidade tecnológica baixa, média-baixa, média-alta e alta.
Cooperação, Capacitação e Apoio do Governo		
Coop	Arranjo Cooperativo	Variável binária que assume 1 (um) caso a firma esteve envolvida em arranjos cooperativos com outra (s) organização (ões) e 0 (zero) caso contrário.
Quali	Qualidade da Mão de Obra	Variável binária que assume 1 (um) caso a firma possua mão de obra com nível de qualificação superior e 0 (zero) caso contrário.
Gov	Apoio do Governo	Variável binária que assume 1 (um) caso a firma utilize algum programa de apoio do Governo para as atividades inovativas e 0 (zero) caso contrário.

Fonte: Elaboração própria.

No terceiro grupo, estão as variáveis que representam características mais voltadas para o processo inovativo em si e de capacitação:

- Arranjo cooperativo refere-se à firma que cooperou com outras empresas para o desenvolvimento de alguma inovação. A variável qualidade da mão de obra está relacionada ao fato da empresa possuir trabalhadores de nível superior nas atividades de P&D. A variável apoio do governo está relacionada à participação das empresas em algum programa governamental para desenvolver suas atividades inovativas. Esses programas podem ser de incentivos fiscais, de financiamentos para projetos de P&D ou para aquisição de máquinas e equipamentos para inovar, e de subvenção econômica a atividades de P&D e à inserção de pesquisadores.

Cabe destacar ainda o recorte dado por nível de intensidade tecnológica e de atividade econômica a que a empresa pertence. Em geral, para definir os níveis de intensidade tecnológica, utiliza-se a classificação da OCDE que, essencialmente apoiada na relação entre os gastos em P&D e o valor agregado ou nos gastos em P&D e a Receita Líquida de Vendas, reúne os setores da indústria de transformação em quatro grupos principais de intensidade tecnológica: alta, média-alta, média-baixa e baixa (OCDE, 2010). Para os setores industriais brasileiros, a classificação da OCDE é pertinente, dado que a PINTEC apresenta um padrão semelhante ao mesmo (FURTADO, 1994). A definição em quatro níveis está representada pela variável “Intec” especificada no Quadro 2.

Já o recorte territorial (Minas Gerais) se faz importante neste trabalho, uma vez que cada estado brasileiro possui estruturas econômicas e de apoio às atividades inovativas distintas e, por isso, analisar, especificamente, o estado de Minas Gerais representa um avanço no sentido de explorar a dimensão temporal e regional, além de investigar a dinâmica de especialização e diversificação direcionada pelo conhecimento tecnológico estadual, sob o período em análise.

3.2. Modelo analítico

Para analisar os determinantes da inovação ambiental, buscou-se analisar possíveis heterogeneidades entre os determinantes em relação à decisão da firma em realizar inovações convencionais e inovações ambientais, uma vez que alguns fatores que levam à adoção de inovação ambiental podem ser diferentes dos fatores relacionados à inovação convencional. Para tanto, foi utilizado o método do *probit* bivariado. Esse método de estimativa remove o viés de seleção da amostra e também apresenta parâmetros mais precisos através da inclusão de empresas não inovadoras (CHUN; MUN, 2012). Por exemplo, Heckman (1979) sugeriu o procedimento *Heckit* para variáveis dependentes contínuas. No entanto, nesse caso, as variáveis dependentes são discretas e a utilização da razão inversa Mills não é uma escolha apropriada. Assim, a melhor estratégia empírica seria utilizar o *probit* bivariado que apresenta a característica de não impor qualquer formato específico para a tomada de decisão, mas de supor que as duas opções se relacionam entre si de alguma maneira (MACHADO; OLIVEIRA, 2014).

Desse modo, entende-se que as decisões entre inovar e inovar ambientalmente fazem parte de um mesmo processo decisório da firma, mas podem apresentar determinantes distintos. Diante disso, verificaram-se os fatores que se associam à probabilidade de escolha pela inovação ambiental, considerando a possível simultaneidade das decisões entre inovar e inovar com redução de impacto ambiental. Portanto, têm-se duas variáveis dependentes binárias e considera-se a possibilidade de existência de um fator endógeno que relaciona essas variáveis, de maneira que a escolha por uma determinada decisão afeta, diretamente, a probabilidade da firma decidir-se pela outra opção (CAMERON; TRIVEDI, 2005).

O *probit* bivariado pode ser expresso da seguinte forma:

$$Y_{1it}^* = X_{it}\beta_i + \alpha_i + \mu_{1it} \quad (1)$$

$$Y_{1it} = 1 \text{ se } Y_{1it}^* > 0; Y_{1i} = 0 \text{ caso contrário}$$

$$Y_{2it}^* = X_{it}\beta_i + \alpha_i + \mu_{2it} \quad (2)$$

$$Y_{2it} = 1 \text{ se } Y_{2it}^* > 0; Y_{2it} = 0 \text{ caso contrário}$$

em que a primeira variável binária (Y_{1i}) indica se a firma adotou inovação ambiental ou não, enquanto a segunda variável binária (Y_{2i}) informa se a firma adotou inovação convencional ou não; Y_{1it}^* e Y_{2it}^* são variáveis latentes que determinam os dois resultados, de modo que as variáveis binárias resultantes dessas decisões (Y_{1it}^* e Y_{2it}^*) assumem valor igual a 1 caso a firma adote inovação ambiental ($Y = 1$ se $Y_{1it}^* > 0$) e valor 0 (zero) caso contrário; e assumem valor igual a 1 caso a firma adote inovação convencional ($Y = 1$ se $Y_{2it}^* > 0$) e valor 0 (zero) caso contrário; X_i

representa o vetor contendo as variáveis explicativas referentes aos determinantes das firmas que afetam ambas as decisões; μ_{1it} e μ_{2it} são os termos de erro, α_i é a heterogeneidade setorial não observada, $i = 1, 2, \dots, N$ – são unidades de firmas e $t = 1, 2, \dots, T$ – período de tempo (anos). A escolha simultânea dessas duas opções implica em quatro resultados diferentes e excludentes entre si:

(i) Não inova com redução de impacto ambiental e não inova de forma convencional:

$$[Y_{1i} = 0 \text{ e } Y_{2i} = 0] \quad (3)$$

(ii) Apenas inova com redução de impacto ambiental:

$$[Y_{1i} = 1 \text{ e } Y_{2i} = 0] \quad (4)$$

(iii) Inova com redução de impacto ambiental e inova de forma convencional:

$$[Y_{1i} = 1 \text{ e } Y_{2i} = 1] \quad (5)$$

(iv) Apenas inova de forma convencional:

$$[Y_{1i} = 0 \text{ e } Y_{2i} = 1] \quad (6)$$

As suposições sobre os erros aleatórios são:

$$E(\mu_1) = E(\mu_2) = 0 \quad (7)$$

Cameron e Trivedi (2005) apontam que, se as duas decisões são correlacionadas, os erros dos dois modelos não são independentes entre si ($\rho = Cov(\mu_1, \mu_2) \neq 0$) e a probabilidade de uma opção depende da probabilidade da outra, sendo determinadas conjuntamente, ou seja, a tomada de decisão em inovar ambientalmente é afetada pela decisão de inovar convencionalmente e vice-versa. As condições de exogeneidade estão condicionadas aos termos do coeficiente de correlação ρ , interpretado como a correlação entre as variáveis explicativas não observáveis das duas equações propostas. Caso $\rho = 0$, Y_{1it}^* e μ_{2it} não estão correlacionados e Y_{1it}^* é exógeno para a segunda equação do modelo (1). Quando $\rho \neq 0$, há evidência de que Y_{1it}^* está correlacionado com μ_{2it} , sendo endógeno. Portanto, a hipótese nula é de que as equações são exógenas e a alternativa é de que há evidências para a endogeneidade. Caso a hipótese nula seja rejeitada, a aplicação do *probit* bivariado busca corrigir a endogeneidade existente entre inovação convencional e inovação ambiental. Ademais, os erros seguem distribuição normal bivariada.

4. Apresentação e discussão dos resultados

4.1. Descrição analítica dos dados

Com base na PINTEC, é possível calcular a taxa de inovação⁴ da indústria de transformação de Minas Gerais. A Tabela 1 registra as taxas de inovação convencional e as taxas de inovação para as empresas que implementaram inovação ambiental.

⁴ A taxa de inovação corresponde ao quociente entre o número de empresas que declararam ter introduzido pelo menos uma inovação no período considerado e o número total de empresas nos setores pesquisados pela PINTEC. Assim, a taxa de inovação é uma medida de resultado dos esforços inovativos das empresas (CAVALCANTI; DE NEGRI, 2011).

Tabela 1 – Taxa de inovação na indústria de transformação de Minas Gerais – 1998-2014

Período de referência	Total de empresas industriais	Total de empresas inovadoras	Taxa de inovação (%)	Total de empresas inovadoras ambientais	Taxa de inovação (%)
1998-2000	8272	2303	27,84	1008	43,78
2001-2003	7944	3388	42,66	1185	34,98
2003-2005	10446	3111	29,78	814	26,18
2006-2008	12113	5111	42,19	1727	33,79
2009-2011	13923	5647	40,56	2372	42,01
2012-2014	13559	4624	34,10	1753	37,91

Fonte: Elaboração própria a partir das bases do IBGE (2018).

Conforme a Tabela 1, houve um crescimento sistemático da taxa de inovação em quatro pesquisas (de 27,84% para 42,66% e de 29,78% para 42,19%) e um decréscimo nas últimas pesquisas para 34,10%. No período 2006 e 2014, as taxas de inovação convencional diminuíram sistematicamente passando de 42,19% para 34,10%. As taxas de inovação ambiental foram maiores apenas na primeira edição da PINTEC, que abrangia o período 1998-2000. Segundo Cavalcanti e De Negri (2011), isso pode ser justificado pela novidade dos conceitos e, portanto, pela pouca familiaridade dos respondentes.

A Tabela 2 apresenta as estatísticas descritivas das variáveis binárias e da variável contínua para as empresas da indústria de transformação de Minas Gerais no período de 1998 a 2014. Analisando a amostra de empresas inovadoras pelos anos das PINTECs, verifica-se que 54% das empresas no período de 2012-2014 se enquadram em regulações e normas padrão relativas ao mercado interno ou externo. Esse resultado pode estar relacionado ao fato de que muitas firmas são reguladas ambientalmente em decorrência do setor de atuação e do porte, além de abrangerem o mercado externo que já é significativamente regulado na esfera ambiental, o que pode induzi-las a melhorar a qualidade ambiental de seus produtos e processos.

Tabela 2 – Descrição das variáveis binárias e da variável contínua para as empresas inovadoras de Minas Gerais no período de 1998 a 2014

Variáveis Binárias	1998-2000 (%)	2001-2003 (%)	2003-2005 (%)	2006-2008 (%)	2009-2011 (%)	2012-2014 (%)
Inovação Ambiental	43,8	42,7	29,8	42,2	40,6	37,9
Inovação Convencional	27,8	34,9	26,2	33,8	42,0	34,1
Regulação Ambiental	33,7	26,4	23,0	38,2	43,3	54,0
Controle	19,8	9,59	7,7	5,7	0,8	19,1
Técnicas de Gestão Ambiental	33,9	1,4	17,6	36,4	35,1	45,1
Cooperação	4,26	3,90	3,5	11,5	9,1	18,4
Qualidade da Mão de Obra	47,28	29,6	40,4	37,6	138,4	102,5
Apoio do Governo	25,8	24,3	18,6	27,4	33,4	41,0
Variável Contínua (Média)	1998-2000	2001-2003	2003-2005	2006-2008	2009-2011	2012-2014
Gasto com Atividades Inovativas - Valor (1000 R\$)	374.960	323.131	149.689	434.177	1.568.181	1.397.333

Fonte: Elaboração própria a partir das bases do IBGE (2018).

Quanto à variável controle, nota-se que, do total de empresas inovadoras, 19,8% realizaram algum tipo de controle e gerenciamento no período 1998-2000 e 19,1% no período de 2012-2014, seja formal (marcas) ou informal (segredo industrial, complexidade do desenho do produto) visando atender normas de certificação.

A variável técnicas de gestão ambiental também apresentou maior percentual para o período de 2012-2014 (45,1%), seguido do período de 2009-2011 (35,1%). A justificativa para esse resultado é de que o aumento da conscientização dos consumidores por produtos verdes tem forçado as empresas a se adequarem a esse novo mercado. Ademais, a gestão visa respeitar as leis ambientais, desenvolver formas de preservação do local em que a indústria está inserida, reestruturar a empresa, de modo que diminua o desperdício e promova a reutilização e a preservação dos recursos naturais ligados a essa conscientização, ao desenvolvimento de processos e ao aprimoramento da produção e da qualidade do produto ofertado (CAMPOS, 2006).

Quanto à variável cooperação, que representa o arranjo cooperativo formado pelas firmas da indústria mineira, verifica-se que, dentre as empresas que realizaram inovação, aproximadamente 4% em 1998-2000, 3,9% em 2001-2003, 3,5% em 2003-2005, 11,5% em 2006-2008, 9,1% em 2009-2011 e 18,4% em 2012-2014, declararam ter realizado atividades de cooperação. Diante disso, percebe-se que a estratégia de cooperação entre agentes, por mais importante que seja para o processo inovativo, está sendo inserida aos poucos na prática da maioria das empresas. Esse resultado levanta a hipótese de que as firmas podem não buscar estabelecer cooperação pelo fato de essa ação ser vista como algo bastante complexo, pois envolve conflitos de interesse, capacidade de governança, risco, definição de modalidades de apropriação de conhecimento, etc. (MACULAN, 2010).

Um dado interessante diz respeito à variável qualidade de mão de obra. Observa-se que o percentual de empresas inovadoras com mão de obra com um nível de qualificação superior exercendo as atividades internas de P&D cresceu significativamente durante o período analisado. Esse fato pode ser favorável ao desenvolvimento ou aperfeiçoamento de inovações, sejam convencionais ou ambientais, ao se considerar uma melhor qualificação dos empregados das firmas envolvidos nessas atividades.

Considerando o apoio do Governo, verifica-se que o percentual de firmas que utilizou algum programa de apoio do Governo para as atividades inovativas aumentou durante o período analisado. Esse resultado evidencia que as empresas que realizam tanto inovação convencional quanto ambiental utilizam, em média, um maior percentual de fontes de financiamento público, ainda que baixo, que pode ser justificado pelo fato de que realizar inovação pode ser caro, pois, muitas vezes, depende de tecnologias de ponta e, por isso, as firmas recorrem às fontes governamentais.

4.2. Resultados das estimações

Com o objetivo de analisar os determinantes dos dois tipos de inovação, estimou-se o modelo *biprobit*. A Tabela 3 apresenta a estimação dos efeitos marginais do modelo com as variáveis inovação ambiental e inovação convencional como dependentes. Verifica-se que a correlação $\rho = 0,299$ entre as duas equações é estatisticamente significativa, indicando que o modelo *probit* bivariado pode ser aplicado para a estimação.

Pela Tabela 3, verifica-se que os efeitos marginais das variáveis incluídas apresentam, em geral, os sinais esperados e foram estatisticamente significativos para as firmas que adotaram inovação ambiental ou adotaram inovação convencional separadamente. As firmas que investiram em gastos com atividades inovativas aumentam em 5,5 p.p. a probabilidade de adotar inovação ambiental, enquanto esse coeficiente não foi significativo para as firmas que investiram em inovações convencionais. Observa-se ainda que um aumento em gastos com atividades inovativas aumenta a probabilidade de adotar inovação ambiental e inovação convencional em 1,9 p.p.

De forma bastante expressiva, a regulação se mostra como importante indutora de inovação ambiental, de modo que as firmas que se enquadram em regulamentos ambientais aumentam em 75,3 p.p. a probabilidade de realizar inovação ambiental, enquanto reduzem em 55,1 p.p. a probabilidade de realizar inovação convencional. Analisando os efeitos marginais individuais, o impacto da regulação ambiental sobre a probabilidade de adotar apenas inovação ambiental (0,6 p.p) foi expressivamente maior do que o impacto (negativo) sobre adoção de apenas inovação

convencional (-27,6 p.p). Portanto, a variável regulação ambiental se configura como o determinante mais importante em relação às demais variáveis devido à magnitude apresentada por esses coeficientes. Esse resultado confirma a importância da regulação ambiental em influenciar a probabilidade das empresas em realizar inovação ambiental e corresponde aos resultados apresentados por Horbach e Rennings (2012), Del Rio et al. (2011) e Brunnermeier e Cohen (2003).

Tabela 3 – Estimações: Inovação Ambiental (IA) e Inovação Convencional (IC) – Probit Bivariado – Efeitos Marginais

Variáveis	y = Pr[IA =1]	y = Pr[IC=1]	y = Pr[IA=1 IC=1]	y=Pr[Apenas IA =1 e IC=0]	y=Pr[Apenas IC=1 e IA=0]	y=Pr[Não IA=0 e IC=0]
<i>lnG</i>	0,055*** (0,021)	-0,018 (0,055)	0,019*** (0,007)	-0,0003 (0,0003)	0,019** (0,007)	0,0004 (0,0009)
REG	0,753** (0,034)	-0,551*** (0,101)	0,263*** (0,011)	0,006*** (0,0015)	-0,276*** (0,011)	0,007*** (0,002)
Coop	0,218*** (0,046)	0,026 (0,150)	0,077*** (0,016)	0,0004 (0,001)	-0,076*** (0,016)	-0,001 (0,002)
Gov	0,321*** (0,069)	0,221* (0,129)	0,142*** (0,025)	0,027*** (0,005)	-0,039 (0,026)	-0,074*** (0,010)
Controle	0,171*** (0,054)	0,198*** (0,089)	0,088*** (0,019)	-0,028*** (0,005)	0,013 (0,021)	-0,073*** (0,009)
TecGA	0,383*** (0,033)	0,201** (0,095)	0,137*** (0,012)	-0,0002 (0,0006)	-0,132*** (0,011)	-0,004*** (0,002)
Quali	0,248** (0,103)	0,698*** (0,265)	0,092** (0,036)	-0,004** (0,002)	-0,076** (0,036)	-0,013** (0,005)
Média-baixa	0,0752* (0,040)	0,243* (0,127)	-	-	-	-
Média-alta	0,115** (0,047)	-0,0339 (0,138)	-	-	-	-
Alta	0,031*** (0,014)	0,410 (0,271)	-	-	-	-
<i>Athp</i>	0,309***					
<i>ρ</i>	0,299***					
Nº de Obs.	6.637					

Nota: *** significativo a 1%, ** significativo a 5% e * significativo a 10%. Os erros-padrão são robustos, encontram-se entre parênteses e foram estimados por *bootstrap* com mil replicações. As constantes foram incluídas nas equações e não estão reportadas.

Fonte: Elaboração própria a partir das bases citadas no Quadro 2.

Pela variável arranjo cooperativo, verifica-se que as firmas que participam de alguma rede de cooperação aumentam em 21,8 p.p. a probabilidade em adotar inovação ambiental. O efeito é bem menor sobre a probabilidade de adotar inovação convencional (2,6 p.p.). Esse resultado confirma os argumentos de autores, como De Marchi (2012), os quais afirmam que o pertencimento a algum tipo de rede aparece como um indutor positivo da inovação ambiental, ressaltando a importância de “economias horizontais de escala” e de estratégias cooperativas para esse tipo de inovação. A formação de arranjos cooperativos também lida com o fator incerteza, que pode ser ainda maior para as inovações ambientais e, portanto, envolve maiores riscos, principalmente quando essas são radicais ou com elevado grau de novidade (BARBIERI et al., 2010) e, portanto, podem ser atenuados através de parcerias para realizar inovação.

As estimativas demonstram que a variável apoio do Governo foi positiva e significativa, de modo geral, confirmando a relevância do papel exercido pelo contexto institucional representado

pela espera pública na promoção de iniciativas de investimento em inovações ambientais nas indústrias. Verifica-se ainda que o apoio do Governo aumenta a probabilidade de adotar inovação ambiental em 32,1 p.p. para as firmas que adotam esse tipo de inovação e em 22,1 p.p. a probabilidade de a firma adotar inovação convencional. Os resultados também sugerem que as empresas que utilizam apoio do Governo têm obrigação maior de mostrar resultados, a ele e à sociedade em geral, o que aumentaria sua preocupação com resultados ambientais.

Em termos da variável controle, observa-se que a probabilidade de uma firma adotar inovação ambiental aumenta em 17,1 p.p. pelo fato de a firma ter realizado algum tipo de gerenciamento formal (marcas) e/ou informal (segredo industrial, complexidade do desenho do produto) visando atender normas de certificação e em 19,8 p.p. a probabilidade de a firma adotar inovação convencional. Em relação à variável de técnicas de gestão ambiental, observa-se que a probabilidade em adotar inovação ambiental é aumentada em 38,3 p.p. e a probabilidade de adotar inovação convencional é aumentada em 20,1 p.p. com o uso dessas técnicas. Para a variável qualidade de mão de obra, verifica-se que a probabilidade em adotar inovação ambiental é aumentada em 24,8 p.p. e a probabilidade em adotar inovação convencional é aumentada em 69,8 p.p. Esse resultado aponta a importância das firmas empregarem mão de obra altamente qualificada em atividades inovativas.

Os efeitos marginais sobre a probabilidade de inovação conjunta também estão apresentados na Tabela 3 oferecendo *insights* interessantes sobre como uma escolha de adoção dos dois tipos de inovação conjuntamente é feita pelas empresas. Verifica-se que a probabilidade de uma firma adotar inovação ambiental dado que ela também adotou inovação convencional é positiva e significativa em relação aos gastos com as atividades inovativas (1,9 p.p.). Esse resultado mostra que, em geral, as empresas que combinam inovações ambientais e convencionais são tipicamente mais dotadas de capacidades de inovação do que as empresas que introduzem exclusivamente inovações ambientais.

Em relação à regulação ambiental, a probabilidade de adoção de inovação ambiental conjuntamente com a inovação convencional tem sido expressivamente impulsionada por forças reguladoras (26,3 p.p.). Assim, a regulação se mostra um importante determinante quando a firma adota conjuntamente inovação ambiental e inovação convencional. Essa variável gera um sinal para as firmas de possíveis ineficiências de recursos criando uma possibilidade de melhora tecnológica. Através da geração de pressão nas firmas e setores, a regulação pode auxiliar a superar a inércia e romper o *lock-in* de determinadas trajetórias tecnológicas. Ademais, a regulação pode gerar informação que ajude a reduzir as incertezas inerentes aos processos inovativos que são ainda maiores quando se trata das inovações ambientais. E a regulação também pode gerar vantagens competitivas para as firmas que tenham se adaptado mais rápido às regras, pois a geração de uma resposta inovativa anterior às demais concorrentes acarreta vantagens de ser um *first-mover*, tais como o incremento da capacidade de *learning-by-doing* pela firma.

Por meio da variável arranjo cooperativo, verifica-se que as firmas que participam de alguma rede de cooperação aumentam em 7,7 p.p. a probabilidade em adotar inovação ambiental e inovação convencional de forma conjunta, de modo que, à medida que as firmas trocam experiências pela participação em arranjos cooperativos, elas podem melhorar suas capacidades de absorção ou implementação tanto de inovação ambiental quanto de inovação convencional. Por meio da variável apoio do Governo, percebe-se que as firmas que recebem algum tipo de incentivo do Governo aumentam em 14,2 p.p. a probabilidade de realizarem conjuntamente inovação ambiental e inovação convencional.

Ao observar a variável controle, verifica-se que a probabilidade em adotar inovação ambiental e inovação convencional aumenta em 8,8 p.p. dado que a firma realizou algum tipo de controle e gerenciamento internamente à firma. Em relação à implementação de técnicas de gestão ambiental, observa-se que a probabilidade em adotar inovação ambiental e inovação convencional, de forma conjunta, aumenta em 13,7 p.p. com o uso de técnicas de gestão ambiental, correspondendo aos resultados apresentados por Horbach (2008) e Wagner (2009). Para a variável qualidade de mão de obra, verifica-se que a probabilidade em adotar inovação ambiental e inovação convencional é

aumentada em 9,2 p.p. pelo fato de a mão de obra empregada em atividades internas de P&D possuir um nível de qualificação superior, fato que se mostra favorável ao desenvolvimento ou aperfeiçoamento desse tipo de inovação ao se considerar uma melhor qualificação dos empregados das firmas envolvidos nessas atividades.

Em relação à variável de intensidade tecnológica, verifica-se que a participação em setores de alta tecnologia aumenta a probabilidade de a firma adotar inovação ambiental de forma mais expressiva do que em relação à inovação convencional. Esse resultado sugere que empresas que pertencem a setores mais intensivos em tecnologia têm maior propensão a adotar inovação ambiental do que empresas de setores menos intensivos em tecnologia demonstrando a força das atividades econômicas mais intensivas em tecnologia para esse tipo de inovação.

De modo geral, os efeitos marginais conjuntos da adoção de inovação ambiental e inovação convencional se mostraram significativos e maiores para todos os determinantes apresentados. Em particular, a regulação impacta de forma expressiva quando a empresa adota apenas inovação ambiental. Assim, esse modelo permitiu verificar as heterogeneidades existentes entre os dois tipos de inovação, de modo que foi possível perceber algumas especificidades dos determinantes da inovação ambiental em relação à inovação convencional. Verificou-se que a regulação ambiental, a cooperação, o apoio do Governo e as técnicas de gestão ambiental impactam de maneira mais expressiva sobre a probabilidade de as firmas adotarem inovação ambiental. Já os gastos com atividades inovativas, o controle e gerenciamento e a qualidade de mão de obra impactam mais sobre a probabilidade de adotar inovação convencional. Cabe ressaltar que, quando se analisou a probabilidade conjunta em realizar inovação ambiental e inovação convencional, todas as variáveis apresentaram resultados bastante significativos, evidenciando que esses fatores podem beneficiar positivamente a adoção combinada desses dois tipos de inovação.

5. Conclusões

A combinação de diversos fatores tem se tornado um desafio para as empresas realizarem mudanças em suas práticas no sentido de equacionar os impasses econômicos, ambientais e sociais na direção do desenvolvimento sustentável. De um lado, as empresas enfrentam o aumento da concorrência e a escassez de recursos, e de outro as empresas se encontram diante da maior reação social e política frente aos problemas das externalidades ambientais causadas pelos impactos ambientais das suas atividades. Por essas questões, evidenciou-se o debate sobre a necessidade das empresas em mudar o seu comportamento para uma produção mais sustentável.

Nesse sentido, o estudo buscou identificar possíveis heterogeneidades entre os determinantes em relação à decisão da firma em realizar inovações convencionais e inovações ambientais. O principal objetivo dessa análise foi verificar a existência de especificidades em relação aos determinantes dos dois tipos de inovação, uma vez que alguns fatores que levam à adoção de inovação ambiental podem ser diferentes dos fatores relacionados à inovação convencional. Para tanto, utilizou-se o método do *probit* bivariado, levando-se em conta que as decisões entre inovar e inovar com redução de impacto ambiental fazem parte de um mesmo processo decisório da firma, mas podem apresentar determinantes distintos.

Os resultados apontaram que a regulação, a cooperação e o apoio do Governo influenciaram de forma mais expressiva a adoção de inovação ambiental em comparação com a inovação convencional, indicando, portanto, especificidades da inovação ambiental em relação à inovação convencional. Os resultados mostraram ainda que os determinantes são mais expressivos quando há a combinação de inovação ambiental e inovação convencional, evidenciando que esses determinantes podem ser importantes quando há a adoção combinada desses dois tipos de inovação.

De modo geral, concluiu-se que a regulação ambiental, o apoio do Governo, os arranjos cooperativos, os controles de gerenciamento e as técnicas de gestão ambiental são importantes determinantes para a adoção de inovação ambiental das firmas da indústria de transformação de Minas Gerais. Dentre todos os determinantes, a regulação ambiental se destaca como um fator expressivo e importante para impulsionar a inovação ambiental, de modo a influenciar em uma

produção mais limpa e apresentar instrumentos mais específicos de iniciativas empresariais que possam induzir as firmas a introduzirem produtos e/ou processos menos danosos ao meio ambiente, confirmando parcialmente a hipótese levantada por este estudo de que a regulação é um importante determinante para a adoção tanto de inovação ambiental quanto de inovação convencional.

Ao analisar o impacto da regulação ambiental apenas sobre a inovação convencional, verificou-se que esse impacto foi negativo para a realização desse tipo de inovação. Uma justificativa para esse resultado pode ser analisada pelo fato de que a inovação é vista sob a perspectiva das falhas de mercado, particularmente em relação à produção de conhecimento que se caracteriza por ser incerto, indivisível, com aspectos de bem público e ausência de garantias de apropriabilidade, não havendo, assim, incentivos para que as firmas desenvolvam qualquer tipo de inovação. E, embora a necessidade por parte da firma em se enquadrar em normas e regulamentos que levem em conta a questão ambiental seja um critério cada vez mais importante para a sua permanência no mercado, os benefícios dessa inclusão são, em sua maioria, percebidos no longo prazo e, portanto, podem ser negligenciados pelas empresas à procura de lucros de curto prazo.

Ademais, constatou-se, ainda, em relação a hipótese do estudo, que, quando as empresas adotam, conjuntamente, a inovação ambiental e a inovação convencional, a regulação ambiental se mostra novamente como importante determinante, gerando um sinal para as firmas de possíveis ineficiências de recursos com a possibilidade de melhora tecnológica. Assim, concluiu-se que a adoção de inovação ambiental e inovação convencional depende tanto da indução dessas inovações pela regulação quanto da capacidade de resposta de cada firma da indústria de transformação mineira, evidenciando o caráter amplo e sistêmico do processo inovativo.

O tema abordado neste estudo tem relevância para o contexto atual no nível das firmas como unidade de análise da pesquisa, visto que proporciona um debate pautado em evidências empíricas e um modelo alternativo capaz de diferenciar detalhes importantes para as conclusões da análise proposta. Assim, o estudo contribui para o debate no sentido de apresentar um panorama capaz de dotar gestores privados de dados e informações sistematizadas sobre indústrias inovadoras pautadas na redução de danos ambientais, permitindo a minimização das externalidades negativas dos negócios e ainda possíveis ganhos de competitividade ao adotar esse tipo de inovação.

Por fim, é necessária uma constante mobilização das empresas e do Estado em relação às questões ambientais por meio da implementação e do aprimoramento de discussões sobre o assunto, a fim de que exista uma regulação e um incentivo efetivos por parte do Estado para as empresas na adoção de inovação ambiental. Nesse sentido, o fortalecimento das agências de proteção ambiental é essencial, mas os instrumentos regulatórios devem ser capazes de fazer com que as firmas se adaptem às normas de maneira mais eficiente. Assim, seria possível concretizar a participação de empresas mineiras ambientalmente sustentáveis e ratificar o compromisso do Estado em promover políticas públicas que garantam a adoção de inovação ambiental.

Referências

- ARUNDEL, A.; KEMP, R.; PARTO, S. Indicators for environmental innovation: what and how to measure. In: MARINOVA, D.; ANNANDALE, D. E.; PHILIMORE, J.; ELGAR, E. (Eds.) *International Handbook on Environment and Technology Management*. Cheltenham, 2007.
- BALTAGI, B. H. *Econometrics analysis of panel data*. 2 ed. Chichester, UK: Wiley & Sons, 2001.
- BARBIERI, J. C.; VASCONCELOS, I. F.; ANDREASSI, T. Inovação e sustentabilidade: novos modelos e proposições. *Revista de Administração de Empresas*, v. 50, n. 2, p. 146-154, 2010.
- BRUNNERMEIER, S. B.; COHEN, M. A. Determinants of environmental innovation in US manufacturing industries, *Journal of Environmental Economics and Management*, v. 45, n. 2, p. 278-293, 2003.

- CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics: methods and applications*. Cambridge: Cambridge University Press, 2005.
- CAMPOS, L. M. S. Sistemas de Gestão Ambiental para Pequenas Empresas: uma comparação entre as visões das grandes empresas certificadas, dos implementadores e das pequenas empresas. In: Encontro de Estudos Organizacionais, 4, 2006. *Anais...* Porto Alegre: EnEO, 2006.
- CARRILO-HERMOSILLA, J.; GONZÁLEZ, P. del R.; KÖNNÖLÄ, T. *Eco-innovation: when sustainability and competitiveness shake hands*. New York: Palgrave Macmillan, 2010. p.28-50.
- CARVALHO, L.; AVELLAR, A. P. M. Inovação e Capacidade Exportadora: Evidências para Empresas Brasileiras, 2015. In: Encontro Nacional de Economia, 40, 2015. *Anais...* Florianópolis: ANPEC, 2015.
- CASSIOLATO, J. E.; LASTRES, H. M. M. Sistema de inovação e desenvolvimento: as implicações políticas. *São Paulo em Perspectiva*, v. 19, n. 1, p. 34-45, 2005.
- CAVALCANTE, L. R.; DE NEGRI, F. *Trajatória recente dos indicadores de inovação no Brasil*. Rio de Janeiro: IPEA, 2011 (Texto para Discussão, n. 1659).
- CHUN, H.; MUN, S. Determinants of R&D cooperation in small and medium-sized enterprises, *Small Business Economics*, v. 39, n. 2, p. 419-436, 2012.
- COHEN, W. M.; LEVINTHAL, D. A. Innovation and learning: two faces of R&D. *Economic Journal*, v. 99, n. 397, p. 569-596, 1989.
- DE MARCHI, V. Environmental innovation and R&D cooperation: empirical evidence from Spanish manufacturing firms. *Research Policy*, v.41, n. 3, p. 614-623, 2012.
- DEL RIO, P.; MORÁN, M.A. T.; ALBIÑANA, F. C. Analysing the determinants of environmental technology investments. a panel-data study of spanish industrial sectors. *Journal of Cleaner Production*, v. 19, p. 1170-1179, 2011.
- DOSI, G. Sources, procedures, and microeconomic effects of innovation. *Journal of Economic Literature*, v. 26, n. 3, p. 1120-1171, 1988.
- FREEMAN, C. The greening of technology and models of innovation. *Technological Forecasting and Social Change*, v. 53, n. 1, p. 27-39, 1996.
- FURTADO, A. *Capacitação tecnológica, competitividade e política industrial: uma abordagem setorial e por empresas líderes*. Brasília: IPEA, 1994. (Texto para Discussão, n. 348).
- FUSSLER, C.; JAMES, P. *Driving eco-innovation: a breakthrough discipline for innovation and sustainability*. London: Pitman Publishing, 1996.
- GONÇALVES, E.; SIMÕES, R. Padrões de esforço tecnológico da indústria brasileira: uma análise setorial a partir de técnicas multivariadas. *Revista Economia*, v. 6, n. 2, p. 391-433, 2005.
- HECKMAN, J. J. Sample selection bias as a specification error. *Econometrica*, v. 47, n. 1, p. 153-161, 1979.
- HORBACH, J., Determinants of environmental innovation - new evidence from German panel data sources, *Research Policy*, v. 37, n. 1, p. 163-173. 2008.
- HORBACH, J.; RENNINGS, K. Environmental innovation and employment dynamics in different technology fields – an analysis based on the German Community Innovation Survey 2009. *Centre for European Economic Research*, Mannheim, Germany, 2012. (ZEW Discussion Paper, n. 12-006).

- IBGE – Instituto Brasileiro de Geografia e Estatística, 2018. Disponível em: <<http://www.ibge.gov.br/home/estatistica/economia/industria/pia/atividades/conceitos>> Acessado em: março de 2018.
- KEMP, R.; FOXON, T. *Typology of eco-innovation*. Maastricht, MEI (Measuring Eco-Innovation), 2007.
- KEMP, R.; PEARSON, P. Final report MEI project about measuring eco-innovation: Deliverable 15 of MEI project (D15). *Project Report*. 2008.
- LUCCHESI, A.; COLE, M. A.; ELLIOT, J. R. R.; MENEZES-FILHO, N. A. Determinants of environmental innovation in Brazilian manufacturing industries. In: Encontro Nacional de Economia, 39, 2014. *Anais...* Natal: ANPEC, 2014.
- MACHADO, G. C.; OLIVEIRA, C. A. Programas de transferência de renda e trabalho infantil no Rio Grande do Sul. *Economic Analysis of Law Review*, v. 5, n. 1, p. 99-117, 2014.
- MACULAN, A. M. A importância das interações para a inovação e a busca por indicadores. In: CGEE (Orgs.). *Bases conceituais em pesquisa, desenvolvimento e inovação: implicações políticas no Brasil*. 1ª ed. Brasília: Centro de Gestões e Estudos Estratégicos, 2010. p. 165-186.
- NELSON, R.; WINTER, S. In search of a useful theory of innovations. *Research Policy*, v. 6, n. 1, p. 36-76, 1977.
- OLIVEIRA, C. A.; CARVALHO, F. P.; DUTRA, H. D. Eco-inovações em empresas brasileiras: determinantes e resultados. In: Simpósio de Gestão da Inovação tecnológica, 27, 2012. *Anais...* Salvador, 2012.
- OLTRA, V. *Environmental innovation and industrial dynamics: the contributions of evolutionary economics*. Groupe de Recherche en Economie Théorique et Appliquée (GREThA). 2008. (Cahiers du GREThA, n. 2008-28).
- OOSTERHUIS, F. Innovation dynamics induced by environmental policy: Final report. *Institute of Environmental Studies Vrije Universiteit*. Amsterdã, 2006.
- OCDE – Organização para a Cooperação e o Desenvolvimento Econômico. *OECD Science, Technology and Industry Outlook*. Paris: OECD Publishing, 2010.
- OCDE – Organização para a Cooperação e o Desenvolvimento Econômico. *OSLO Manual: guidelines for collecting and interpreting innovation data*. 3. ed. Paris: Statistical Office of the European Communities – Eurostat, p. 163, 2005.
- OZUSAGLAM, S. Environmental innovation: a concise review of the literature. *Vie & Sciences de l'entreprise*, n. 191-192, p. 15-38, 2012.
- PINTEC – *Pesquisa de Inovação Tecnológica 2000, 2003, 2005, 2008 e 2011*. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, 2017.
- PNUMA – Programa das Nações Unidas para o Meio Ambiente. *A Produção mais Limpa e o Consumo Sustentável na América Latina e Caribe*. São Paulo: Governo do Estado de São, 2005.
- PORTER, M. E.; LINDE, C. V. D. Toward a new conception of the environment – competitiveness relationship. *Journal of Economic Perspectives*, v. 9, n. 4, p. 97-118, 1995.
- RIBEIRO, J. C. J. *Desafios do licenciamento ambiental*. In: Seminário Estadual sobre Licenciamento Ambiental. *Anais...* Belo Horizonte: Amda/Semad, 2006.


- SERRANO, L. M.; BARBIERI, A. F. Meio ambiente e desenvolvimento sustentável no Brasil: uma descrição de indicadores de sustentabilidade ambiental aplicáveis à realidade brasileira. In: Encontro Nacional de Estudos Populacionais, 16, 2008. *Anais...* Caxambu: ABEP, 2008.
- SPEIRS, J.; PEARSON, P.; FOXON, T. *Adapting innovation systems indicators to assess eco-innovation*. DIME International Conference “Innovation, sustainability and policy”, GREThA, University Montesquieu Bordeaux IV, France, p. 11-13, 2008.
- TIGRE, P. B. Paradigmas tecnológicos e teorias econômicas da firma. *Revista Brasileira de Inovação*, v. 4, n. 1, p. 35-60, 2005.
- UNCTAD – United Nations Conference on Trade and Development. *World investment report 2006*. FDI from developing and transition economies: implications for development. New York; Geneva: United Nations, 2006.
- UNITED NATIONS. *Earth Summit Agenda 21*. United Conference on Environment and Development – UNCED. Rio de Janeiro, 1992.
- WAGNER, M. Empirical influence of environmental management on innovation: evidence from Europe. *Ecological Economics*, v. 66, n. 2, p. 392-402, 2009.
- WALRAS, L. *Compêndio dos elementos de economia política pura*. Tradução: João Guilherme Vargas Netto. Série: Os Economistas. São Paulo: Abril Cultural, 1983.

ORCID

Maria Alice Ferreira  <https://orcid.org/0000-0001-5296-642X>

Maria Ivanilde Pereira Santos  <https://orcid.org/0000-0002-4828-7931>

Emerson Costa dos Santos  <https://orcid.org/0000-0002-6493-7236>

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

DIFERENCIAIS DE CUSTOS DE VIDA ENTRE OS ESTADOS BRASILEIROS: UMA APLICAÇÃO DE PREÇOS HEDÔNICOS A PARTIR DOS DADOS DA PNAD*

Marcello Luiz de Souza Junior

Departamento de Economia, Administração e Sociologia. Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo (ESALQ/USP)
E-mail: marcello.luiz.souza@usp.br

Alexandre Nunes de Almeida

Departamento de Economia, Administração e Sociologia - ESALQ /USP
E-mail: alex.almeida@usp.br

João Gomes Martines Filho

Departamento de Economia, Administração e Sociologia - ESALQ /USP
E-mail: martines@usp.br

Maria Cristina Galvão

Departamento de Economia, Administração e Sociologia - ESALQ /USP
E-mail: galvao.mariacristina@usp.br

RESUMO: Este estudo tem como objetivo estimar o diferencial de custo de vida entre as unidades federativas do país a partir de um modelo de preços hedônicos. Para isso, utilizam-se as informações da Pesquisa Nacional por Amostra de Domicílios de 2015 disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística. O modelo proposto apresenta uma abordagem de regressão quantílica e tem como variável dependente os gastos com aluguel, considerados uma *proxy* para o custo de vida, e variáveis independentes relacionadas às características físicas e locacionais dos imóveis. Os resultados obtidos estão de acordo com a literatura existente e revelam divergências significativas nos custos de se viver nos mais diferentes locais, considerando todos os quantis analisados. Notou-se que as unidades federativas que compõem a região nordeste apresentam os menores custos de vida do país para todos os quantis. Já Distrito Federal, São Paulo, Rio de Janeiro e Santa Catarina despontam como os estados mais caros. As estimativas aqui apresentadas podem ser utilizadas por agentes econômicos como deflatores regionais, de modo a contribuir para análises futuras acerca da paridade de poder de compra nas diferentes localidades.

Palavras-chave: Custos de vida; Aluguel; Preços hedônicos; Regressão quantílica.

Classificação JEL: O18; R10; R32.

COST OF LIVING DIFFERENTIALS BETWEEN BRAZILIAN STATES: AN APPLICATION OF THE HEDONIC THEORY FROM A NATIONAL HOUSEHOLD SAMPLE SURVEY

ABSTRACT: This paper aims to evaluate the cost of living differentials between Brazilian states using a hedonic price model and data from National Household Sample Survey for 2015. The identification strategy proposed is a quantile regression. Rent expenses are considered as a dependent variable, proxy for cost of living, while physical and locational characteristics are considered as independent variables. For all the quantiles analyzed, the results are in line with the literature and reveal significant divergences in the costs of living in different places. States located in the Northeast region have the lowest life costs of the country for any quantile, while Distrito Federal, São Paulo, Rio de Janeiro and Santa Catarina are considered as the most expensive states. These findings can be used by economic agents as regional deflators, contributing to analyses regarding to purchasing power parity in different locations.

Keywords: Cost of living; Rent; Hedonic prices; Quantile regression.

JEL Codes: O18; R10; R32.

1. Introdução

Ao se deslocar entre as diversas localidades do país, é possível identificar diferenças significativas no que diz respeito ao clima, à disponibilidade de recursos naturais, à dotação de fatores de produção, aos aspectos culturais, aos hábitos de consumo e à renda da população. Em conjunto, esses fatores contribuem para a formação de sistemas de produção e estruturas de demanda peculiares a cada região, acarretando assim dinâmicas econômicas distintas entre as unidades federativas (FERREIRA; DINIZ, 1995; GODIM et al., 2007).

De fato, quando se compara o PIB *per capita* dos estados brasileiros, a disparidade de renda entre as várias regiões do país é expressiva. Por exemplo, se a renda *per capita* no Distrito Federal é de aproximadamente R\$ 58 mil por ano; em Minas Gerais, esse valor é de cerca de R\$ 18,3 mil; enquanto no Piauí não ultrapassa R\$ 8 mil¹ (IPEA, 2011). Contudo, como aponta Menezes (1999), para ser possível afirmar que regiões mais ricas são, de fato, as mais caras para se viver, maiores investigações são necessárias.

Nesse sentido, o diferencial de custo de vida entre localidades não só representa uma área dinâmica de estudo para discussões acadêmicas, como também apresenta desdobramentos importantes no processo de decisão de alocação de recursos por parte dos agentes econômicos e no oferecimento de subsídios para políticas públicas (ALMEIDA; AZZONI, 2016). Assim, analisar os diferentes custos de vida pode ser útil para: auxiliar indivíduos na decisão de migrar para outro local; fornecer informações para que o Estado defina um nível de salário mínimo ideal; amparar a decisão de empresas em investir em determinadas regiões; servir de referência para o repasse de recursos federais para as demais esferas públicas; e tornar justo o mecanismo de tributação progressiva e políticas de transferência de renda (WILLIAMS, 1940).

Diante dos benefícios advindos da estimação do diferencial de custo de vida entre regiões, alguns países têm construído índices regionais que captam essas informações. Os Estados Unidos têm sido referência neste aspecto, uma vez que o *Bureau of Economic Analysis* (BEA), a partir de dados do índice de preços ao consumidor e do *Bureau of Labor Statistics* (BLS), estima, anualmente, a paridade regional de preços para 38 áreas no país².

No que diz respeito à experiência brasileira, Cavalcanti (2014) cita o Índice de Custo de Vida entre Regiões (ICVR) calculado pela Fundação Instituto de Pesquisas Econômicas (FIPE). Além do autor mencionado, alguns outros trabalhos têm sido dedicados ao estudo do custo de vida no Brasil, como, por exemplo: Azzoni, Carmo e Menezes (2003), Menezes, Moreira e Azzoni (2007), Seabra e Azzoni (2015), Almeida e Azzoni (2016) e Azzoni e Seabra (2018).

Embora exista um número significativo de trabalhos que abordaram o tema, a ausência de dados atualizados dificulta a realização de novas pesquisas. Isso ocorre porque o cômputo de um índice de custo de vida requer informações de dispêndio domiciliar disponibilizadas pelas onerosas e pouco frequentes Pesquisas de Orçamentos Familiares (POF) e informações dos escassos vetores de preços disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para regiões metropolitanas selecionadas (ALMEIDA; AZZONI, 2016).

A fim de superar a indisponibilidade de dados supracitada, Menezes, Moreira e Azzoni (2007), Cavalcanti (2014) e Seabra e Azzoni (2015), adotando embasamento teórico que aponta a alta correlação entre gastos com habitação e custo de se viver em determinada região, utilizaram aluguéis como uma *proxy* para o custo de vida, com base nas informações da POF 2002/2003 e do Censo 2010.

Nesse sentido, o objetivo deste estudo consiste em estimar o diferencial de custo de vida entre as 27 unidades federativas do país, considerando os níveis distintos de renda da população, a partir dos valores de aluguéis obtidos na Pesquisa Nacional por Amostra de Domicílios (PNAD-Domicílios) em 2015³ (IBGE, 2015). Para tanto, a estratégia empírica incide em adotar um método de estimação

¹ Valores a preços de 2010.

² Para maiores informações, ver Aten, Figueroa e Martin (2012).

³ A PNAD-Domicílios de 2015 corresponde à pesquisa mais recente em termos de informações acerca de aluguéis e características físicas de imóveis em âmbito nacional.

quantílico, de modo a melhor representar a distribuição condicional do aluguel em relação às variáveis independentes utilizadas. O uso desse método de estimação não só reduz problemas relacionados à má especificação da forma funcional, como também permite identificar como a localização de um imóvel influencia no valor do aluguel e, conseqüentemente, no custo de vida, nas diferentes faixas de aluguel presentes na amostra.

Ressalta-se que este trabalho contribui à literatura de diferenciais de custos de vida, uma vez que explora uma base de dados recente e leva em conta diferentes níveis de renda. Não obstante, a abrangência geográfica aqui empregada gera, adicionalmente, informações relevantes para populações de áreas onde não há levantamento de preços sistemáticos realizados pelo IBGE, sendo essa uma das principais motivações do presente estudo.

Este trabalho se divide em mais 5 seções além desta introdução. A Seção 2 apresenta uma breve revisão sobre o uso de aluguel como *proxy* do custo de vida e os aspectos microeconômicos relacionados à teoria de preços hedônicos. A terceira seção detalha a estratégia empírica adotada para calcular o índice de custo de vida proposto. A quarta seção apresenta a base de dados utilizada. A Seção 5 traz e discute os principais resultados obtidos nesta pesquisa e, por fim, a Seção 6 reúne as considerações finais do trabalho.

2. Revisão bibliográfica

Esta seção contém duas subseções. A primeira trata da relação existente entre gasto com aluguel e custo de se viver em determinada localidade, demonstrando, a partir de estudos anteriores e evidências empíricas, como tal gasto tem sido empregado como *proxy* do custo de vida. Já a segunda apresenta aspectos da teoria microeconômica acerca do modelo de preços hedônicos e sua aplicação no mercado imobiliário.

2.1. O aluguel como *proxy* do custo de vida

O uso de aluguéis como uma *proxy* do real custo de vida tem sido recorrente na literatura nacional e internacional, como pode ser observado nos trabalhos de Menezes, Moreira e Azzoni (2007), Martin, Aten e Figueroa (2011), Cavalcanti (2014) e Seabra e Azzoni (2015).

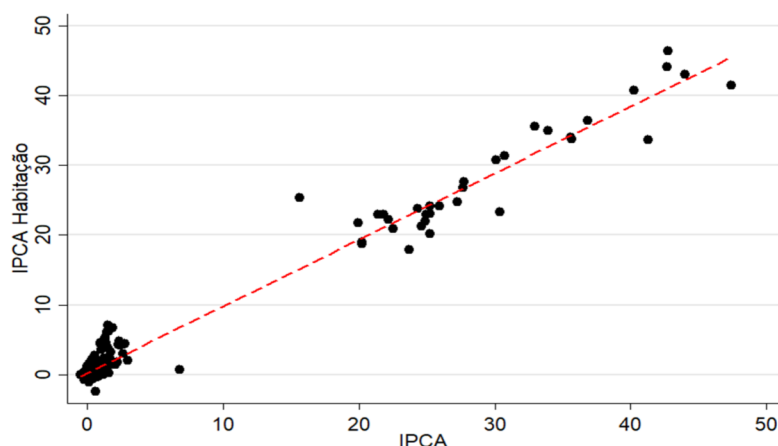
Na década de 1980, Fava (1984) apontou para a correlação existente entre os gastos com habitação e o custo de vida de uma população. Segundo a autora, a evolução do montante dispendido com moradia impacta direta e indiretamente famílias de todas as classes de renda. Diretamente, pois o aluguel é um item de grande representatividade na cesta básica de consumo e indiretamente em virtude dos efeitos marginais ocasionados nas demais rubricas dos orçamentos familiares, como transporte, alimentação e outros serviços.

No que se refere às conseqüências verificadas nos custos com transporte, Galvão et al. (2016) argumentam que à medida que os aluguéis aumentam, as populações de mais baixa renda são forçadas a ocuparem as regiões periféricas dos grandes centros, gerando assim maiores distâncias para deslocamento e conseqüente aumento nos gastos com mobilidade urbana.

Quanto à alimentação e demais serviços, Biderman (2001) esclarece que os aluguéis, por representarem parte dos custos de empresas e estabelecimentos prestadores de serviços, apresentam movimentos diretamente proporcionais aos preços dos bens finais. Essa dinâmica caracteriza um efeito de repasse dos aluguéis nos preços finais da economia.

Diante desse embasamento teórico, Azzoni, Menezes e Moreira (2007) afirmam que o aluguel pode ser considerado uma aproximação eficiente do custo de vida de uma população, sobretudo em regiões urbanas. A título de ilustração, a Figura 1 apresenta a dispersão ente o IPCA e o IPCA – Habitação entre setembro de 1991 e julho de 2017. Para o intervalo analisado, a correlação registrada foi de 0,9751.

Figura 1 – Relação entre IPCA Habitação e IPCA entre Set. de 1991 e Jul. de 2017



Fonte: Elaborado pelos autores a partir de dados do IBGE (2017).

O estudo desenvolvido por Menezes, Moreira e Azzoni (2007), ao utilizar dados de gastos com aluguéis da POF 2002/2003, estimou um índice de custo de vida entre os estados brasileiros a partir de um modelo de preços hedônicos. Além dos atributos físicos empregados como controles, os pesquisadores incluíram também *dummies* para áreas geográficas (rural, urbana, regiões metropolitanas e capitais), estados e faixas de renda (baixa, média, alta). Os autores concluíram que o custo de vida para as capitais são os mais elevados, seguidos das regiões metropolitanas, das outras áreas urbanas e da zona rural, independentemente do nível de renda familiar considerado. Outra constatação relevante foi o fato de que indivíduos com diferentes rendas apresentam percepções distintas quanto ao diferencial de custo de vida entre os estados.

Já Cavalcanti (2014) calculou o diferencial de custo de vida entre 283 municípios brasileiros a partir da aplicação da metodologia *Weighted-Country-Product-Dummy* (WCPD). Os dados utilizados foram os valores de aluguéis publicados pelo Censo 2010. O autor apontou São Caetano do Sul-SP, Rio das Ostras-RJ e Balneário Camboriú-SC como as três cidades mais caras do país. Em contrapartida, os três municípios com menores custos de vida foram Garanhuns-PE, Codó-MA e Paulo Afonso-BA.

Por fim, o estudo realizado por Seabra e Azzoni (2015) mensurou o diferencial de custo de vida intracitadino nos 100 municípios mais populosos do país. Os autores empregaram uma variação da metodologia CPD a partir de dados de aluguéis disponibilizados pelo Censo de 2010. Dentre os principais resultados, destacam-se aqueles observados para a cidade de São Paulo. Segundo os autores, o centro expandido da capital paulista apresenta um custo de vida até 468% mais caro do que a média nacional, ao passo que em regiões periféricas das zonas norte e sul esse indicador é 5,4% menor do que no restante do país.

2.2. O modelo de preços hedônicos

Em virtude da estratégia empírica adotada no presente estudo, apresenta-se nesta seção a definição de preços hedônicos, sua origem e os aspectos microeconômicos relacionados a sua aplicação na literatura econômica (LANCASTER, 1966, 1971; ROSEN, 1974; GOODMAN, 1978; WITTE et al., 1979; CAN; MEGBOLUGBE, 1997).

O modelo de preços hedônicos surge a partir de um esforço realizado para incorporar a importância dos atributos presentes em um bem para a decisão de consumo e produção na análise microeconômica. Nesse contexto, a concepção de que o ferramental neoclássico tradicional se mostra ineficiente para representar as reais interações entre bens e consumidores levou Lancaster (1966) a propor uma interpretação alternativa da problemática. Essa nova interpretação baseia-se na hipótese de que a utilidade advinda do hábito de consumo é derivada das propriedades intrínsecas do bem e não do simples fato de adquirir o produto, tal como prega a teoria neoclássica.

Apesar de Lancaster (1966) ser um dos principais autores que atuaram na estruturação da teoria hedônica, a literatura atribui a Waught (1928) o pioneirismo na valoração de atributos de um bem. Em seu estudo, Waught buscou mensurar o impacto que certas características de tomates, aspargos e pepinos apresentavam em seus preços finais. Para tanto, o autor inspecionou 200 locais de vendas em Boston, Estados Unidos, entre os meses de maio e junho de 1927. Cabe ressaltar que, já naquela época, o pesquisador havia antecipado o que seria sugerido anos depois por Rosen (1974), ao afirmar que a preocupação em monetizar as propriedades intrínsecas de um bem não é benéfica apenas para compreender a demanda de uma população, mas também para planejar a oferta das firmas, uma vez que o ajuste entre oferta e demanda se dá pela quantidade e qualidade dos produtos transacionados.

Rosen (1974) também contribuiu para estruturar a teoria dos preços hedônicos, contextualizando-a ao ambiente de mercado. Esse autor define preços hedônicos como sendo preços implícitos de características que são reveladas aos agentes econômicos a partir de cotações de bens relativamente heterogêneos e quantidades de atributos relacionadas a cada produto.

Assim como a teoria neoclássica, o modelo de preços hedônicos assume que indivíduos são maximizadores de utilidade e que, nesse processo de otimização, enfrentam restrições orçamentárias. Contudo, para esse particular caso, a função utilidade definida apresenta como variáveis independentes n propriedades, tal que $U(\mathbf{x}, z_1, z_2, \dots, z_n)$, em que \mathbf{x} representa todos os demais bens consumidos e o vetor \mathbf{z} consiste em uma série de atributos.

De maneira simétrica, Rosen (1974) admite que as firmas maximizam o lucro a partir da produção de cestas de características. Desse modo, o autor assume $\mathbf{M}(\mathbf{z})$ como sendo a quantidade de atributos \mathbf{z} produzida por uma firma. Tem-se ainda, segundo os mesmos princípios da teoria tradicional, que uma função custo é estimada a partir do processo de minimização de custos sujeito a um determinado nível de produção. A equação de custos encontrada para esse modelo é representada por $C(\mathbf{M}, \mathbf{z}; \boldsymbol{\omega})$, em que \mathbf{M} e \mathbf{z} são a quantidade e os atributos produzidos e $\boldsymbol{\omega}$ é um parâmetro que faz referência aos fatores de produção empregados no processo produtivo e seus respectivos preços. Por fim, define-se que o lucro será maximizado quando a receita marginal se igualar ao custo marginal, tal que $\partial Mp(\mathbf{z}) = \partial C(\mathbf{M}, \mathbf{z}; \boldsymbol{\omega})$.

Para prosseguir com a análise de equilíbrio, torna-se necessário igualar a equação de demanda $Q^d(\mathbf{z})$, derivada do procedimento de maximização da utilidade, à equação de oferta $Q^s(\mathbf{z})$, encontrada a partir do comportamento de otimização do lucro das firmas. Nessa condição de equilíbrio, os preços de mercado de um bem em questão podem ser relacionados aos diversos atributos que o compõem, de modo a estabelecer uma regressão de preços hedônicos, tal como explicitado na Expressão 1:

$$p(\mathbf{z}) = p(z_1, z_2, z_3, \dots, z_n). \quad (1)$$

No contexto do mercado imobiliário, Can e Megbolugbe (1997) apresentam uma forma geral para a regressão hedônica, explicitando os componentes físicos e locacionais que impactam no preço de mercado dos imóveis, de modo que:

$$\mathbf{P} = f(S\boldsymbol{\beta}, N\boldsymbol{\gamma}) + \boldsymbol{\varepsilon} \quad (2)$$

em que, especificamente, \mathbf{P} refere-se a um vetor contendo o valor de mercado das moradias, S é uma matriz que representa as características estruturais do domicílio, N é uma matriz que contempla atributos geográficos e amenidades locais relacionadas à localização do imóvel, $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são os respectivos vetores de coeficientes a serem estimados e $\boldsymbol{\varepsilon}$ é o erro.

Assim como em Menezes, Moreira e Azzoni (2007), o presente estudo adota uma variação do modelo de Rosen (1974), com o intuito de definir a influência que a localidade do imóvel em determinado estado brasileiro exerce no valor do aluguel, uma vez controlados os atributos físicos.

3. Metodologia

A estratégia empírica empregada neste trabalho consiste na aplicação de um modelo de preços hedônicos a partir de dados de aluguéis disponibilizados pela PNAD-Domicílios. Essa estratégia possibilita a construção de um índice de custo de vida que identifica o diferencial de aluguéis existente entre as 27 unidades federativas do país, considerando ainda as diversas faixas de aluguéis observadas na amostra.

Vale argumentar também que, assim como em Menezes, Moreira e Azzoni (2007) e Seabra e Azzoni (2015), o escopo da análise foi delimitado a imóveis alugados em virtude da disponibilidade de informações. Desse modo, apesar de existir a possibilidade de ocorrer um problema de seleção amostral decorrente do fato de domicílios alugados diferirem em característica daquelas que não o são, tal viés impactaria somente o coeficiente das características estruturais do imóvel. Assim, uma vez que essas variáveis são incorporadas ao modelo apenas para controle, pode-se concluir que a estratégia empírica não apresenta nenhum prejuízo à confiabilidade dos resultados que são apresentados posteriormente (MENEZES; MOREIRA; AZZONI, 2007). Logo, o modelo econométrico a ser estimado é descrito na Expressão 3:

$$\ln(p) = \alpha + \beta Z + \delta \ln(y) + \theta G + \varepsilon \quad (3)$$

em que o termo dependente é o logaritmo neperiano do preço do aluguel dos 19.984 imóveis analisados, Z é uma matriz que contém as 9 características físicas empregadas como mecanismos de controle, β é o vetor de parâmetros referentes à matriz Z , $\ln(y)$ é o logaritmo neperiano da renda domiciliar e δ é seu respectivo parâmetro, G é uma matriz que retrata 27 variáveis *dummies* de propriedades locais dos imóveis analisados - sendo 26 referentes ao estado em que a moradia está localizada e 1 representando a situação censitária do imóvel (urbano ou rural), θ é o vetor de parâmetros relacionado à matriz G e ε é o termo de erro aleatório.

Com o intuito de melhor caracterizar a relação entre as *dummies* estaduais e o logaritmo neperiano dos aluguéis nos diferentes segmentos da distribuição condicional da variável independente, a Expressão 3 será estimada a partir de uma regressão quantílica, tal como explicitado por Koenker e Bassett (1978)⁴. Segundo Koenker (2005), além de mitigar questões relacionadas à má especificação da forma funcional, esse método reduz o efeito de *outliers* sob os coeficientes obtidos e atenua os problemas relacionados à heterocedasticidade.

Especificamente, a estimação do modelo se dá a partir da minimização da soma ponderada dos desvios absolutos, em que o critério de ponderação consiste em aplicar penalizações em situações de super e sub predições, tal como explicitado na Expressão 4.

$$\min_{\beta_\tau \in R} \left[\sum_{y \geq X'\beta} \tau |y - X'\beta_\tau| + \sum_{y < X'\beta} (1 - \tau) |y - X'\beta_\tau| \right] \quad (4)$$

em que τ é o quantil analisado e $\tau \in [0,1]$, y se refere à variável dependente ($\ln(p)$) e $X'\beta_\tau$ é a especificação ($\alpha + \beta Z + \delta \ln(y) + \theta G$) descrita na Expressão 3, analisada no quantil τ . O cômputo da Expressão 4 é realizado por meio de um algoritmo de otimização proposto por Hunter e Lange (2000).

A Expressão 3 é estimada considerando $\tau = \{0,10; 0,25; 0,50; 0,75; 0,90\}$, de modo a permitir a identificação do impacto das variáveis independentes, em especial as *dummies* locais, em toda a distribuição do logaritmo neperiano dos aluguéis e não apenas na média condicional. Não obstante, testes de Wald são realizados para investigar a possibilidade de igualdade entre os coeficientes obtidos nos diferentes quantis reportados (WOOLDRIDGE, 2010).

⁴ Agradecemos a um dos pareceristas anônimos por sugerir métodos alternativos que melhor se adequam à distribuição dos dados.

Por fim, tem-se que o uso da regressão quantílica como estratégia de estimação permite identificar a variabilidade do custo de vida apurado para um estado entre os diferentes níveis de aluguéis observados e, conseqüentemente, nas distintas faixas de renda da população. Não obstante, os parâmetros calculados a partir do método de Mínimos Quadrados Ordinários são também reportados a título de comparação. As estimações foram obtidas usando o *software* estatístico Stata 14.

4. Base de dados

No ano de 2015, a PNAD entrevistou 151.189 unidades domiciliares distribuídas por todo o território nacional. Desse total, 19.984 (13,2%) imóveis são considerados na presente análise. Ao adotar o fator de expansão disponibilizado pela pesquisa, a amostra utilizada neste estudo representa aproximadamente 92.123.169 domicílios em todo o país.

Dentre as características dos imóveis empregadas no modelo a ser estimado, destacam-se as propriedades físicas, tais como número de cômodos, dormitórios e banheiros, e locais, como a situação censitária do domicílio. A Tabela 1 descreve os atributos selecionados para análise⁵.

Tabela 1 – Características utilizadas como controles no modelo

Variável	Descrição	Média	Desvio Padrão
<i>Dependente</i>			
Valor do aluguel ⁺	Aluguel mensal pago no mês de referência.	586,74	501,41
<i>Independentes</i>			
Nº cômodos	Número de cômodos no imóvel.	5,77	1,99
Dormitórios	Número de cômodos que servem como dormitórios.	1,84	0,79
Banheiros	Número de banheiros no imóvel.	1,35	0,71
Renda domiciliar ⁺	Rendimento mensal familiar (exclusive rendimentos mensais de pensionistas e empregados domésticos).	3.144,16	4.365,94

Nota: ⁺ Valores referentes a setembro de 2015.

Fonte: Elaborado pelos autores com dados da PNAD-Domicílios 2015.

Além dos atributos representados na Tabela 1, foram utilizados como variáveis independentes *dummies* relacionadas ao material de construção do imóvel, acesso à água, rede de esgoto, internet e tipo de domicílio (casa, apartamento ou cômodo). Ademais, foram incluídas *dummies* locais referentes à situação censitária e às unidades federativas em que o domicílio está situado, a fim de captar a diferença de custos entre os estados brasileiros.

Vale ressaltar que o plano amostral realizado pela PNAD utiliza como unidade de análise o domicílio, de modo que o processo de seleção da amostra realizado em 3 estágios (municípios, setores censitários e, por fim, domicílios) busca contemplar a heterogeneidade intraestadual, minimizando eventuais erros de medida que os dados possam apresentar.

5. Resultados econométricos

Os resultados obtidos a partir da estimação da Expressão 4 representam o diferencial do custo de vida observado entre os estados brasileiros nos diferentes quantis analisados (q_{10} , q_{25} , q_{50} , q_{75} , q_{90}) e na média condicional, Mínimos Quadrados Ordinários (MQO), para o ano de 2015. Cabe mencionar ainda que a análise considera o Distrito Federal (DF) como base de

⁵ A definição de cada variável e a forma como estão descritas e apresentadas na Tabela 1 são baseadas nas notas explicativas da PNAD-Domicílios 2015.

comparação⁶ e que os coeficientes retratados na Tabela 2 encontram-se em sua forma exponencial⁷. Os erros padrão, robustos para heterocedasticia, foram estimados pelo método de *bootstrap* usando 100 replicações, seguindo sugestão de Davidson e MacKinnon (2000).

Apesar de o presente estudo ter como variáveis de interesse as *dummies* estaduais, é interessante discutir os coeficientes das características físicas dos imóveis. Primeiramente, cabe destacar que todas as relações encontradas para os diferentes quantis da distribuição condicional estão em concordância com a literatura (MENEZES; MOREIRA; AZZONI, 2007; CAVALCANTI, 2014). Em resumo, percebe-se que maiores números de cômodos, dormitórios e banheiros atuam no sentido de elevar o valor do aluguel. Nota-se também que propriedades construídas em alvenaria ou madeira são mais caras do que aquelas edificadas a partir de materiais não duráveis, sendo que essa diferença se reduz à medida que são considerados aluguéis mais elevados. As amenidades relacionadas à infraestrutura da vizinhança também contribuem para a elevação do aluguel. Tem-se ainda que apartamentos apresentam aluguéis entre 33 e 46% mais elevados do que os de casas e que o incremento de um ponto percentual na renda domiciliar aumenta o preço do aluguel entre 0,20% a 0,23%.

Considerando a situação censitária dos domicílios, os resultados obtidos revelam que a zona rural é mais barata do que a urbana, independente do quantil analisado. Esse resultado está de acordo com a literatura, dada a falta de amenidades existentes, entre outros atributos exclusivos das áreas urbanas (WINTERS, 2009). Para efeitos de comparação, pode-se afirmar que, dada uma determinada tipologia de imóvel, seu valor é 59,43% mais caro caso se situe em uma cidade e não no meio rural, no caso de q_{10} e 27,73% em q_{90} .

Admitindo o gasto de aluguel como uma *proxy* do custo de vida, tem-se que os coeficientes encontrados para as *dummies* estaduais sugerem divergências significativas nos custos de se viver nos mais diferentes estados, considerando os variados níveis de aluguéis e, por consequência, de renda.

Com o intuito de verificar a igualdade entre os valores estimados para cada estado nos diferentes quantis, foram realizados testes F de significância conjunta para as variáveis de interesse, cujos resultados são reportados na Tabela A.1 (apêndice). Os testes apontam que há diferenças no impacto que a localização de um domicílio exerce sobre o valor do aluguel para as seguintes unidades federativas: Amazonas, Pará, Tocantins, Maranhão, Piauí, Ceará, Paraíba, Pernambuco, Bahia, Minas Gerais, Rio de Janeiro, São Paulo, Santa Catarina, Rio Grande do Sul e Mato Grosso do Sul. Para os demais estados, não houve rejeição da hipótese nula de igualdade dos coeficientes de cada quantil, indicando, portanto, que o efeito da localização do imóvel em relação ao valor do aluguel (i.e., custo de vida) se mantém ao longo de toda a distribuição condicional da variável dependente. A título de representação gráfica, a Figura A.1 (apêndice) apresenta os coeficientes e intervalos de confiança obtidos a partir do modelo quantílico para os casos em que houve rejeição de H_0 no teste F de significância conjunta.

A Figura 2 exibe o índice de custo de vida computado a partir dos exponenciais dos coeficientes das *dummies* estaduais para q_{10} , q_{50} , q_{90} e a média condicional, com o intuito de representar o comportamento desse indicador para as diferentes faixas de aluguel presentes na amostra.

Nota-se que as curvas apresentam formato similar, sendo que, para grande parte dos estados, a linha referente ao quantil 90 se sobrepõe às demais, indicando que, de fato, indivíduos de maior renda possuem custos de vida mais elevados, conforme o esperado. Contudo, há situações em que a linha relacionada ao quantil 10 se encontra acima das restantes, como é o caso do Mato Grosso do Sul⁸. Tal fato traz evidências de que, para esse estado, a localização do imóvel impacta em maior intensidade o custo de vida da população de mais baixa renda em comparação a indivíduos que arcam com um maior valor de aluguel (q_{90}).

⁶ A escolha do DF como base de comparação considera conclusões de estudos anteriores, que apontam essa como a região mais cara do país e tem sido utilizada como *default* desde então.

⁷ Com exceção aos coeficientes da renda domiciliar (já se encontra em elasticidade) e do intercepto.

⁸ Tal situação também ocorre em AP, AL e ES, porém os resultados do teste F indicam não existir diferença estatística nos coeficientes obtidos para cada quantil.

Tabela 2 – Resultados das regressões múltipla e quantílica

Continua

Variável	MQO	Q10	Q25	Q50	Q75	Q90
	Ln aluguel	Ln aluguel	Ln aluguel	Ln aluguel	Ln aluguel	Ln aluguel
Constante	3,1262*** (0,0013)	2,4006*** (0,1377)	2,7080*** (0,1514)	3,0041*** (0,1035)	3,4644*** (0,1161)	3,8926*** (0,1714)
Nº Cômodos	1,0583*** (0,0000)	1,0514*** (0,0046)	1,0560*** (0,0039)	1,0631*** (0,0037)	1,0661*** (0,0036)	1,0698*** (0,0048)
Nº Dormitórios	1,0137*** (0,0000)	1,0075 (0,0091)	1,0042 (0,0066)	1,0178*** (0,0055)	1,0237*** (0,0061)	1,0192** (0,0082)
Nº Banheiros	1,2375*** (0,0001)	1,2133*** (0,0106)	1,2048*** (0,0105)	1,2192*** (0,0086)	1,2429*** (0,0108)	1,2811*** (0,0136)
Material Constr.	1,3787*** (0,0011)	1,6445*** (0,1097)	1,4958*** (0,1302)	1,5520*** (0,0837)	1,3869*** (0,0987)	1,2346 (0,1617)
Acesso à água	1,0260*** (0,0002)	1,0469** (0,0189)	1,0602*** (0,0176)	1,0524*** (0,0143)	1,0420** (0,0199)	0,9951 (0,0226)
Acesso à internet	1,1347*** (0,0000)	1,1733*** (0,0129)	1,1382*** (0,0082)	1,1198*** (0,0073)	1,1118*** (0,0072)	1,1041*** (0,0094)
Rede de esgoto	1,1811*** (0,0001)	1,2333*** (0,0139)	1,1969*** (0,0090)	1,1708*** (0,0089)	1,1818*** (0,0093)	1,1730*** (0,0114)
Apartamento	1,3929*** (0,0001)	1,3446*** (0,0130)	1,3368*** (0,0081)	1,3544*** (0,0083)	1,4027*** (0,0107)	1,4681*** (0,0140)
Cômodo	1,0217*** (0,0004)	1,0190 (0,0560)	0,9805 (0,0485)	1,0167 (0,0370)	0,9850 (0,0357)	0,9635 (0,0470)
Renda Domiciliar	0,2232*** (0,0000)	0,2304*** (0,0097)	0,2384*** (0,0070)	0,2293*** (0,0063)	0,2187*** (0,0067)	0,2013*** (0,0075)
Zona urbana	1,4715*** (0,0005)	1,5943*** (0,0485)	1,5068*** (0,0322)	1,4065*** (0,0326)	1,2815*** (0,0356)	1,2773*** (0,0423)
RO	0,7352*** (0,0003)	0,6907*** (0,0457)	0,7510*** (0,0265)	0,7487*** (0,0265)	0,7373*** (0,0302)	0,7289*** (0,0301)
AC	0,7258*** (0,0004)	0,6728*** (0,0644)	0,7304*** (0,0512)	0,7602*** (0,0380)	0,7269*** (0,0396)	0,7029*** (0,0351)
AM	0,8002*** (0,0003)	0,7329*** (0,0337)	0,7653*** (0,0283)	0,7875*** (0,0306)	0,8374*** (0,0320)	0,8758** (0,0648)
RR	0,7771*** (0,0004)	0,7638*** (0,0492)	0,7933*** (0,0354)	0,7771*** (0,0383)	0,7894*** (0,0428)	0,7982*** (0,0560)
PA	0,7365*** (0,0004)	0,7112*** (0,0381)	0,7430*** (0,0235)	0,7525*** (0,0216)	0,7812*** (0,0243)	0,8213*** (0,0350)
AP	0,7963*** (0,0005)	0,8178*** (0,0581)	0,8366*** (0,0640)	0,7844*** (0,0560)	0,8128*** (0,0391)	0,7414*** (0,0721)
TO	0,7071*** (0,0003)	0,6423*** (0,0486)	0,6932*** (0,0486)	0,7004*** (0,0271)	0,7477*** (0,0296)	0,7458*** (0,0233)
MA	0,6537*** (0,0004)	0,6149*** (0,0565)	0,6299*** (0,0336)	0,6394*** (0,0329)	0,6983*** (0,0496)	0,7211*** (0,0405)
PI	0,5262*** (0,0005)	0,4453*** (0,0869)	0,5014*** (0,0770)	0,5566*** (0,0477)	0,5478*** (0,0475)	0,6272*** (0,0629)

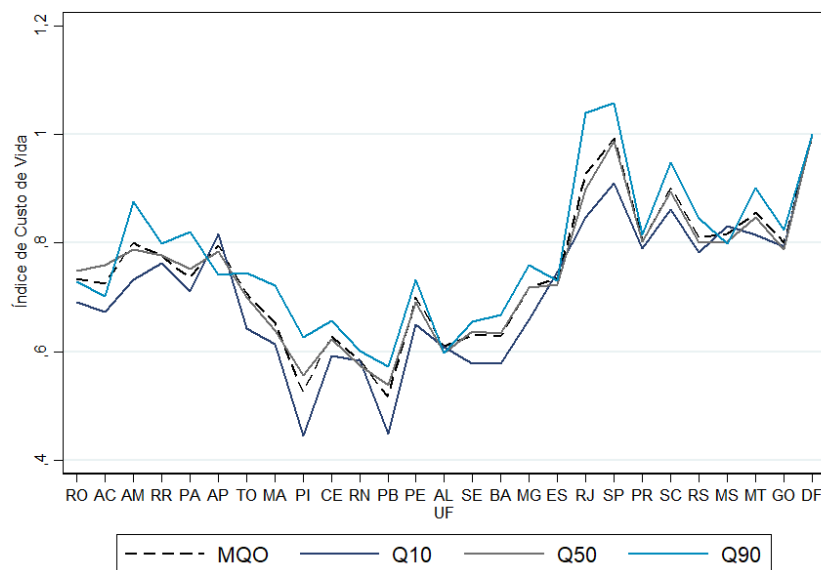
Tabela 2 – Resultados das regressões múltipla e quantílica

Variável	<i>Conclusão</i>					
	OLS Ln aluguel	Q10 Ln aluguel	Q25 Ln aluguel	Q50 Ln aluguel	Q75 Ln aluguel	Q90 Ln aluguel
CE	0,6290*** (0,0002)	0,5923*** (0,0309)	0,6263*** (0,0204)	0,6236*** (0,0206)	0,6441*** (0,0213)	0,6569*** (0,0226)
RN	0,5849*** (0,0003)	0,5829*** (0,0518)	0,5893*** (0,0418)	0,5749*** (0,0298)	0,5889*** (0,0304)	0,6008*** (0,0432)
PB	0,5176*** (0,0003)	0,4493*** (0,0430)	0,5082*** (0,0445)	0,5385*** (0,0263)	0,5535*** (0,0271)	0,5723*** (0,0357)
PE	0,6995*** (0,0002)	0,6492*** (0,0286)	0,6846*** (0,0218)	0,6917*** (0,0211)	0,7075*** (0,0232)	0,7329*** (0,0194)
AL	0,6097*** (0,0003)	0,6079*** (0,0544)	0,6192*** (0,0344)	0,5983*** (0,0302)	0,6166*** (0,0303)	0,5980*** (0,0352)
SE	0,6312*** (0,0003)	0,5787*** (0,0471)	0,6062*** (0,0391)	0,6372*** (0,0239)	0,6440*** (0,0251)	0,6546*** (0,0398)
BA	0,6287*** (0,0002)	0,5789*** (0,0326)	0,6288*** (0,0202)	0,6332*** (0,0197)	0,6267*** (0,0216)	0,6686*** (0,0205)
MG	0,7190*** (0,0002)	0,6595*** (0,0306)	0,7141*** (0,0154)	0,7191*** (0,0186)	0,7358*** (0,0169)	0,7603*** (0,0199)
ES	0,7368*** (0,0003)	0,7476*** (0,0415)	0,7278*** (0,0265)	0,7239*** (0,0256)	0,7247*** (0,0338)	0,7314*** (0,0270)
RJ	0,9277*** (0,0002)	0,8483*** (0,0260)	0,8866*** (0,0178)	0,8990*** (0,0172)	0,9439*** (0,0198)	1,0400* (0,0211)
SP	0,9929*** (0,0001)	0,9111*** (0,0230)	0,9615*** (0,0154)	0,9868 (0,0172)	1,0004 (0,0157)	1,0569*** (0,0172)
PR	0,8039*** (0,0002)	0,7909*** (0,0316)	0,8037*** (0,0160)	0,8033*** (0,0187)	0,7888*** (0,0211)	0,8152*** (0,0222)
SC	0,9017*** (0,0002)	0,8625*** (0,0289)	0,8532*** (0,0272)	0,8946*** (0,0203)	0,9101*** (0,0230)	0,9472* (0,0310)
RS	0,8100*** (0,0002)	0,7821*** (0,0301)	0,8192*** (0,0190)	0,8004*** (0,0163)	0,7987*** (0,0193)	0,8455*** (0,0197)
MS	0,8169*** (0,0002)	0,8315*** (0,0480)	0,8383*** (0,0192)	0,8025*** (0,0261)	0,7889*** (0,0222)	0,7984*** (0,0275)
MT	0,8566*** (0,0003)	0,8152*** (0,0555)	0,8548*** (0,0193)	0,8467*** (0,0239)	0,8770*** (0,0248)	0,9008*** (0,0270)
GO	0,8006*** (0,0002)	0,7933*** (-0,0290)	0,7985*** (0,0171)	0,7879*** (0,0172)	0,7931*** (0,0196)	0,8237*** (0,0232)
R ²	0,6064	-	-	-	-	-
Pseudo R ²	-	0,3435	0,3557	0,3814	0,4111	0,4375
<i>Bootstrapping</i>	-	100	100	100	100	100
Nº Observações	19.984	19.984	19.984	19.984	19.984	19.984

Nota: Erro-padrão entre parênteses. Os símbolos ***, **, * representam valores significantes a 1%, 5% e 10%, respectivamente, no teste *t*.

Fonte: Elaborado pelos autores.

Figura 2 – Diferencial de custo de vida entre os estados, considerando a média condicional e os quantis q_{10} , mediana (q_{50}) e q_{90}



Fonte: Elaborado pelos autores.

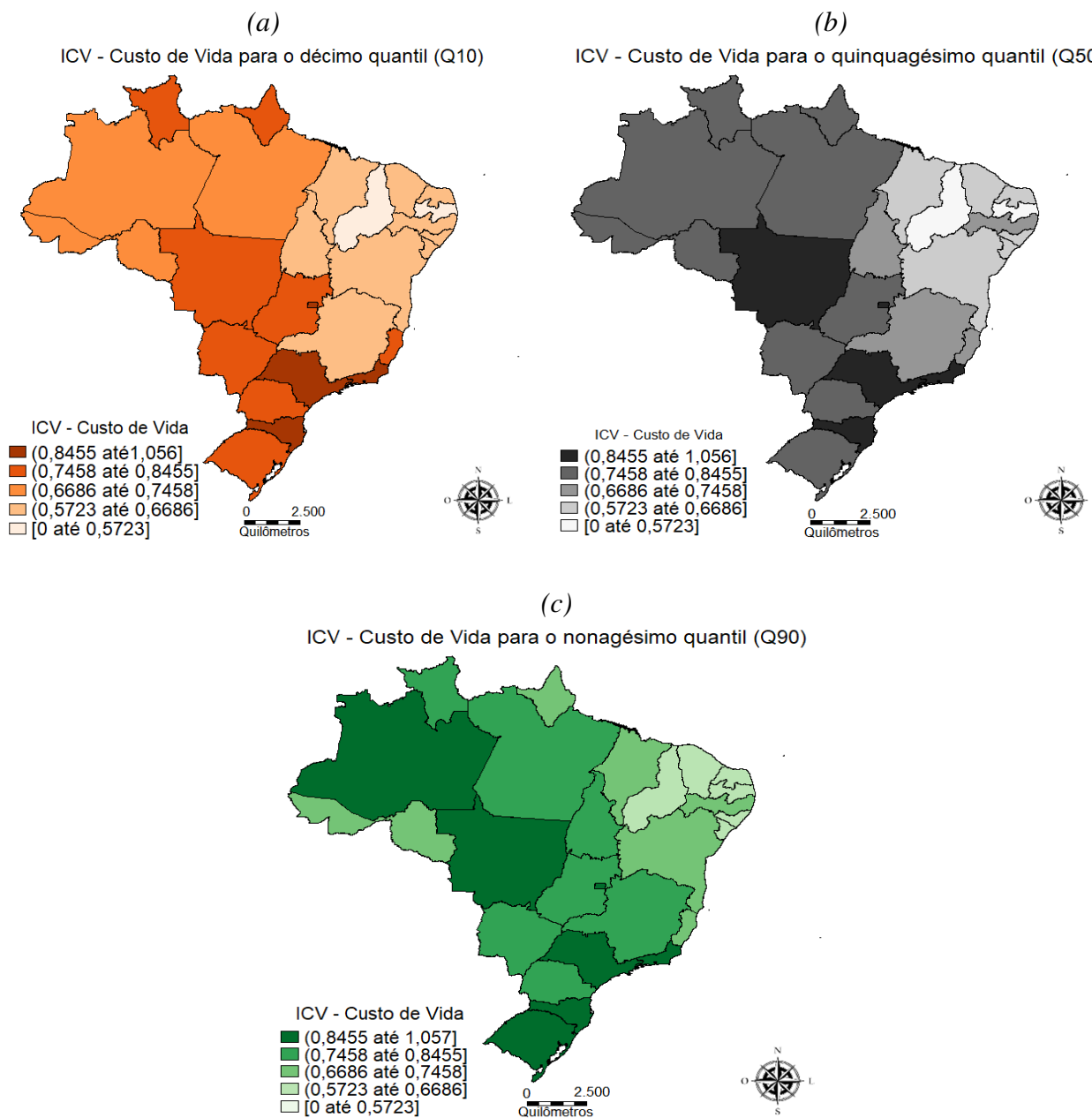
Para as famílias que pagam baixos aluguéis (q_{10}), os estados com maiores custos de vida são Distrito Federal⁹ (1,00), São Paulo (0,911), Santa Catarina (0,862), Rio de Janeiro (0,848) e Mato Grosso do Sul (0,831), ao passo que os mais baratos são Sergipe (0,579), Paraíba (0,449) e Piauí (0,445). Ao considerar o *ranking* das unidades federativas na mediana (q_{50}), tem-se a seguinte classificação: Distrito Federal (1,00), São Paulo (0,987), Rio de Janeiro (0,899), Santa Catarina (0,895) e Mato Grosso (0,847) nas primeiras posições e Rio Grande do Norte (0,575), Piauí (0,557) e Paraíba (0,539) nas últimas. Por fim, ao analisar os indivíduos com maiores gastos com aluguel (q_{90}), nota-se que São Paulo (1,057) assume a posição de estado mais caro, seguido de Rio de Janeiro (1,040), Distrito Federal (1,00), Santa Catarina (0,947) e Mato Grosso (0,901). Já para os estados mais baratos, tem-se Rio Grande do Norte (0,601), Alagoas (0,598) e Paraíba (0,572).

Cabe destacar que, dentre os estados em que H_0 do teste F de significância conjunta foi rejeitada, os maiores desvios-padrão nos coeficientes estimados foram observados em São Paulo, Piauí, Amazonas e Rio de Janeiro. Já os menores desvios foram apurados em Tocantins, Rio Grande do Sul e Mato Grosso do Sul.

A análise dos mapas (Figura 3) revela que os custos de vida mais elevados se concentram na região Centro-Sul do país. Para o quantil q_{10} , destacam-se os custos de vida observados no Distrito Federal, na região Sudeste (representada por São Paulo e Rio de Janeiro), no estado de Santa Catarina, no Sul, e na região Centro-Oeste como um todo. Cabe ressaltar o elevado custo de vida apurado para Amapá e Roraima, ambos na região Norte. Esse fator pode ser explicado em virtude da menor infraestrutura (acesso à internet, água e rede de esgoto) existente nestes dois estados, de modo que, ao controlar o valor dos aluguéis por essas variáveis, obteve-se um coeficiente mais elevado do que a média regional.

⁹ O valor do índice para o Distrito Federal é a unidade, visto que é a base de comparação.

Figura 3 – Custo de vida entre os estados brasileiros considerando os quantis q_{10} (a), q_{50} (b) e q_{90} (c)



Nota: A legenda do mapa se refere ao exponencial dos coeficientes das *dummies* estaduais apresentados na Tabela 2.
 Fonte: Elaborado pelos autores.

Ao considerar o quantil q_{50} , nota-se que os estados de Mato Grosso do Sul, Goiás, Roraima e Amapá apresentam um menor custo de vida quando comparados ao caso anterior, entre 15,45% e 25,42% mais baratos do que a capital. Verifica-se ainda que os estados de Minas Gerais, Tocantins e Pernambuco tornam-se relativamente mais caros, passando a participar da faixa de estados que são entre 25,42% e 33,14% mais baratos do que o Distrito Federal.

Para o último quantil analisado (q_{90}), destaca-se a posição assumida pelo estado do Amazonas, cujo valor encontrado indica um custo de vida 12,42% mais barato do que o DF. Vale ressaltar que, como mencionado anteriormente, esse estado apresentou elevado desvio-padrão entre os coeficientes estimados nos diferentes quantis, de modo a indicar uma elevada amplitude na percepção do custo de vida para a população residente. Por fim, a região mais barata do Brasil para todos os casos anteriores é a região Nordeste.

6. Considerações finais

A alocação eficiente de recursos em um sistema econômico depende da existência de um fluxo de informações capaz de auxiliar os agentes econômicos a tomarem suas decisões da melhor forma possível. Uma dessas informações diz respeito às diferenças observadas no custo de se viver em localidades distintas. Nesse sentido, ao se embasar no desenvolvimento teórico que atesta a alta correlação existente entre os custos de vida e os gastos com habitação, esta pesquisa construiu um índice de custo de vida, a partir da aplicação de um modelo de preços hedônicos aos dados de gastos com aluguel disponibilizados pela PNAD-Domicílios de 2015.

O modelo quantílico adotado permitiu tratamento mais adequado da distribuição condicional da variável dependente, de modo a aumentar a robustez dos resultados e possibilitar a avaliação dos diferenciais de custos de vida entre os estados brasileiros para diferentes faixas de aluguel e, conseqüentemente, níveis de renda da população. Os resultados revelaram divergências significativas nos custos de se viver nos mais diferentes estados, considerando todos os quantis analisados. Verificou-se ainda que os coeficientes das variáveis estruturais utilizadas como controle apresentaram os sinais esperados.

No que se refere ao coeficiente das *dummies* estaduais, notou-se que o custo de vida apurado para o quantil q_{90} é mais elevado do que nos demais, indicando que famílias com maiores gastos com aluguel enfrentam um maior custo de vida. Nesse cenário, São Paulo e Rio de Janeiro são considerados mais caros do que o Distrito Federal. Santa Catarina e os estados da região Centro-Oeste também apresentam elevados coeficientes. Por fim, notou-se que as unidades federativas que compõem a região nordeste são as que apresentam os menores custos de vida do país para qualquer quantil analisado. Tais efeitos encontram-se em linha com aqueles obtidos por Menezes, Moreira e Azzoni (2007) e Cavalcanti (2014).

A presente pesquisa contribuiu para a literatura do tema ao empregar uma base de dados recente e pouco utilizada para estimar custos de vida. Assim, foi possível gerar informações relevantes para populações residentes em regiões que não são acompanhadas pelo levantamento de preços do IBGE e diagnosticar empiricamente a heterogeneidade existente no custo de vida das diversas unidades federativas do país, nos diferentes níveis de renda.

Não obstante, os resultados aqui apresentados oferecem suporte para a decisão de agentes econômicos em questões que contemplam desigualdades regionais, uma vez que oferece informações valiosas para avaliar a paridade de poder de compra entre diferentes estados. No que diz respeito à esfera pública, o índice estimado nesta pesquisa traz possibilidades de melhorias para as políticas de transferência de renda existentes e de crédito habitacional, visto que torna possível direcionar recursos para as mais diversas unidades federativas de acordo com o custo de vida apurado em cada região.

Por outro lado, cabe mencionar a limitação apresentada pelo estudo em utilizar aluguel como uma *proxy* do custo de vida em alternativa ao conjunto de gastos com alimentação, transporte, saúde, lazer, educação e habitação realizados por uma família. Contudo, em virtude da inexistência de tais informações para a abrangência regional, optou-se por utilizar o embasamento existente na literatura econômica que sustenta a alta correlação entre aluguel e custo de vida. Por fim, registra-se a importância de estudos futuros que contemplem uma abordagem geográfica municipal, de modo a melhor retratar a heterogeneidade intraestadual existente em cada unidade federativa do país.

Referências

- ALMEIDA, A. N.; AZZONI, C. R. Custo de vida comparativo das regiões metropolitanas brasileiras: 1996–2014. *Estudos Econômicos*, v. 46, n. 1, p. 253-276, 2016.
- ATEN, B.; FIGUEROA, E.; MARTIN, T. Regional price parities for states and metropolitan areas, 2006-2010. *Survey of Current Business*, v. 92, n. 8, p. 229-242, ago. 2012.
- AZZONI, C. R.; CARMO, H. E.; MENEZES, T. Comparação da paridade do poder de compra entre

- idades: aspectos metodológicos e aplicação ao caso brasileiro. *Pesquisa e Planejamento Econômico*, v. 33, n. 1, p. 91-126, 2003.
- AZZONI, C. R.; SEABRA, D. M. S. Custo comparativo de aluguéis para cidades e áreas de ponderação. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 12, n. 2, p. 242-260, 2018.
- BIDERMAN, C. Forças de atração e expulsão na Grande São Paulo. Tese (Doutorado em Economia). Fundação Getúlio Vargas. São Paulo, 2001.
- CAN, A.; MEGBOLUGBE, I. Spatial dependence and house price index construction. *Journal of Real Estate Finance and Economics*, v. 14, n. 1-2, p. 203-222, 1997.
- CAVALCANTI, E. M. *Diferencial de custo de vida entre as regiões: um índice baseado em aluguel*. Dissertação (Mestrado em Economia). Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco. Recife, 2014.
- DAVIDSON, R.; MACKINNON, J.G. Bootstrap tests: How many bootstraps? *Econometric Review*, v. 19, n. 1, p. 55-68, 2000.
- FAVA, V. L. *Urbanização, custo de vida e pobreza no Brasil*. São Paulo: IPE-USP, p. 1-203, 1984.
- FERREIRA, A.; DINIZ, C. Convergência entre as rendas per capita estaduais no Brasil. *Revista de Economia Política*, v. 15, n. 4, p. 38-56, 1995.
- GALVÃO, M. C.; ALMEIDA, A. N.; SPOLADOR, H. F. S.; AZZONI, C. R. Custo de vida, amenidades e salários nas regiões metropolitanas brasileiras. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 10, n. 2, p. 199-216, 2016.
- GODIM, J. L. B.; BARRETO, F. A.; CARVALHO, J. R. Condicionantes de clubes de convergência no Brasil. *Estudos Econômicos*, v. 37, n. 1, p. 71-100, 2007.
- GOODMAN, A. C.; Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, v. 5, n. 4, p. 471-484, 1978.
- HUNTER, D. R.; LANGE, K. Quantile regression via an MM algorithm. *Journal of Computational and Graphical Statistics*, v. 9, n. 2, p. 60-77, 2000.
- IBGE – Instituto Brasileiro de Geografia e Estatística. *Microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD)*. Rio de Janeiro, IBGE, 2015.
- IBGE – Instituto Brasileiro de Geografia e Estatística. *Índice Nacional de Preço ao Consumidor Amplo (IPCA)*. Disponível em: <https://www.ibge.gov.br/estatisticas-novoportal/economicas/precos-e-custos>. Acesso em: 02 dez. 2017.
- IPEA – Instituto de Pesquisa Econômica Aplicada. *PIB estadual per capita*. Disponível em: <http://ipeadata.gov.br>. Acesso em: 02 dez. 2017.
- KOENKER, R. *Quantile Regression*. Cambridge: University Press, 2005. p. 349
- KOENKER, R.; BASSET, G. Regression Quantiles. *Econometrica*, v. 46, n. 1, p. 33-50, 1978.
- LANCASTER, K. J. A New Approach to Consumer Theory. *Journal of Political Economy*, v. 74, n. 2, p. 132–157, 1966.
- LANCASTER, K. J. *Consumer demand: a new approach*. New York: Columbia U. Press, 1971.
- MARTIN, T.; ATEN, B.; FIGUEROA, E. *Estimating the price of rents in regional price parities*. Bureau of Economic Analysis, 2011. (BEA Working Papers, n. 1278).
- MENEZES, T. *Custo de vida regional e convergência de renda per capita entre as regiões* *Revista Brasileira de Estudos Regionais e Urbanos*, v. 13, n. 2, p. 223-242, 2019

metropolitanas brasileiras. Tese (Doutorado em Economia). Programa Pós-Graduação do Departamento de Economia da Universidade de São Paulo. São Paulo, 1999.

MENEZES, T. A.; MOREIRA, G. R. C.; AZZONI C. R. Diferenças em gastos com aluguel entre estados, tipos de área e níveis de renda familiar no Brasil. In: SILVEIRA, F. G.; SERVO L. M.; MENEZES, T.; PIOLA, S. F. (Orgs.) *Gasto e consumo das famílias brasileiras contemporâneas*. Brasília: IPEA, v. 2, 2007. p. 329-346.

ROSEN, S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, v. 82, n. 1, p. 34-55, 1974.

SEABRA, D. M. S.; AZZONI, C. R. Custo de vida comparativo para os distritos das 100 maiores cidades brasileiras. *Informações FIPE, Temas em Economia Aplicada*, n. 417, p. 12-24, 2015.

WAUGH, F. V. Quality factors influencing vegetable prices. *Journal of Farm Economics*, v. 10, n. 2, p. 185-196, 1928.

WILLIAMS, F. M. Factors to be considered in measuring intercity and interregional differences in living costs. *Journal of the American Statistical Association*, v. 35, n. 211, p. 471-482, 1940.

WINTERS, J. V. Wages and prices: are workers fully compensated for cost of living differences? *Regional Science and Urban Economics*, v. 39, n. 5, p. 632-643, 2009.

WITTE, A. D.; SUMKA, H. J.; EREKSON, H. An estimate of a structural hedonic price model of the housing market: an application of Rosen's Theory of implicit markets. *Econometrica*, v. 47, n. 5, p. 1151-73, 1979.

WOOLDRIDGE, J. M. *Econometric analysis of cross section and panel data*. 2 Ed. Cambridge, MIT Press, 2010.


ORCID

Marcello Luiz de Souza Junior  <https://orcid.org/0000-0001-6039-0029>

Alexandre Nunes de Almeida  <https://orcid.org/0000-0002-0680-5446>

João Gomes Martines Filho  <https://orcid.org/0000-0002-1320-8586>

Maria Cristina Galvão  <https://orcid.org/0000-0001-6241-9171>

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

Apêndice

Tabela A.1 – Teste F de significância conjunta da regressão quantílica

Continua

Teste	RO	AC	AM	RR	PA	AP	TO
$H_0[q_{10}=q_{25}=q_{50}=q_{75}=q_{90}]$	F(4, 19946) = 0,91 Prob> F = 0,4540	F(4, 19946) = 1,76 Prob> F = 0,1343	F(4, 19946) = 2,74 Prob> F = 0,0269	F(4, 19946) = 0,33 Prob > F = 0,8566	F(4, 19946) = 2,38 Prob > F = 0,0494	F(4, 19946) = 0,89 Prob > F = 0,4669	F(4, 19946) = 2,46 Prob > F = 0,0430
$H_0 [q_{10}=q_{25}]$	F(1, 19946) = 3,45 Prob> F = 0,0632	F(1, 19946) = 2,46 Prob> F = 0,1167	F(1, 19946) = 2,01 Prob> F = 0,1560	F(1, 19946) = 0,79 Prob > F = 0,3755	F(1, 19946) = 1,66 Prob > F = 0,1975	F(1, 19946) = 0,17 Prob > F = 0,6826	F(1, 19946) = 2,48 Prob > F = 0,1152
$H_0 [q_{10}=q_{50}]$	F(1, 19946) = 2,66 Prob> F = 0,1028	F(1, 19946) = 4,18 Prob> F = 0,0410	F(1, 19946) = 3,54 Prob> F = 0,0601	F(1, 19946) = 0,11 Prob > F = 0,7394	F(1, 19946) = 2,22 Prob > F = 0,1366	F(1, 19946) = 0,40 Prob > F = 0,5266	F(1, 19946) = 3,54 Prob > F = 0,0598
$H_0 [q_{10}=q_{75}]$	F(1, 19946) = 1,73 Prob> F = 0,1888	F(1, 19946) = 1,48 Prob> F = 0,2245	F(1, 19946) = 9,91 Prob> F = 0,0016	F(1, 19946) = 0,32 Prob > F = 0,5688	F(1, 19946) = 4,74 Prob > F = 0,0294	F(1, 19946) = 0,01 Prob > F = 0,9201	F(1, 19946) = 8,35 Prob > F = 0,0039
$H_0 [q_{10}=q_{90}]$	F(1, 19946) = 1,05 Prob> F = 0,3047	F(1, 19946) = 0,43 Prob> F = 0,5116	F(1, 19946) = 6,06 Prob> F = 0,0138	F(1, 19946) = 0,44 Prob > F = 0,5076	F(1, 19946) = 8,17 Prob > F = 0,0043	F(1, 19946) = 1,25 Prob > F = 0,2638	F(1, 19946) = 7,50 Prob > F = 0,0062
$H_0 [q_{25}=q_{50}]$	F(1, 19946) = 0,01 Prob> F = 0,9091	F(1, 19946) = 1,22 Prob> F = 0,2685	F(1, 19946) = 1,26 Prob> F = 0,2625	F(1, 19946) = 0,35 Prob > F = 0,5558	F(1, 19946) = 0,31 Prob > F = 0,5790	F(1, 19946) = 1,24 Prob > F = 0,2663	F(1, 19946) = 0,06 Prob > F = 0,8034
$H_0 [q_{25}=q_{75}]$	F(1, 19946) = 0,26 Prob> F = 0,6103	F(1, 19946) = 0,01 Prob> F = 0,9234	F(1, 19946) = 7,45 Prob> F = 0,0063	F(1, 19946) = 0,01 Prob > F = 0,9136	F(1, 19946) = 3,11 Prob > F = 0,0778	F(1, 19946) = 0,23 Prob > F = 0,6318	F(1, 19946) = 2,24 Prob > F = 0,1344
$H_0 [q_{25}=q_{90}]$	F(1, 19946) = 0,56 Prob> F = 0,4555	F(1, 19946) = 0,53 Prob> F = 0,4660	F(1, 19946) = 4,02 Prob> F = 0,0451	F(1, 19946) = 0,01 Prob > F = 0,9149	F(1, 19946) = 6,50 Prob > F = 0,0108	F(1, 19946) = 2,01 Prob > F = 0,1565	F(1, 19946) = 1,82 Prob > F = 0,1778
$H_0 [q_{50}=q_{75}]$	F(1, 19946) = 0,38 Prob> F = 0,5386	F(1, 19946) = 1,64 Prob> F = 0,2001	F(1, 19946) = 4,91 Prob> F = 0,0267	F(1, 19946) = 0,19 Prob > F = 0,6613	F(1, 19946) = 2,91 Prob > F = 0,0882	F(1, 19946) = 0,62 Prob > F = 0,4315	F(1, 19946) = 4,58 Prob > F = 0,0324
$H_0 [q_{50}=q_{90}]$	F(1, 19946) = 0,59 Prob> F = 0,4429	F(1, 19946) = 3,86 Prob> F = 0,0494	F(1, 19946) = 2,99 Prob> F = 0,0838	F(1, 19946) = 0,27 Prob > F = 0,6046	F(1, 19946) = 6,44 Prob > F = 0,0112	F(1, 19946) = 0,70 Prob > F = 0,4034	F(1, 19946) = 3,25 Prob > F = 0,0714
$H_0 [q_{75}=q_{90}]$	F(1, 19946) = 0,13 Prob> F = 0,7220	F(1, 19946) = 1,01 Prob> F = 0,3161	F(1, 19946) = 0,68 Prob> F = 0,4098	F(1, 19946) = 0,06 Prob > F = 0,8114	F(1, 19946) = 2,46 Prob > F = 0,1167	F(1, 19946) = 2,66 Prob > F = 0,1032	F(1, 19946) = 0,01 Prob > F = 0,9305

Tabela A.1 – Teste F de significância conjunta da regressão quantílica

Teste	<i>Continua</i>						
	MA	PI	CE	RN	PB	PE	AL
$H_0 [q_{10}=q_{25}=q_{50}=q_{75}=q_{90}]$	F(4, 19946) = 2,54 Prob > F = 0,0377	F(4, 19946) = 4,22 Prob > F = 0,0020	F(4, 19946) = 2,77 Prob > F = 0,0259	F(4, 19946) = 0,32 Prob > F = 0,8616	F(4, 19946) = 6,40 Prob > F = 0,0000	F(4, 19946) = 3,76 Prob > F = 0,0047	F(4, 19946) = 0,77 Prob > F = 0,5469
$H_0 [q_{10}=q_{25}]$	F(1, 19946) = 0,28 Prob > F = 0,5979	F(1, 19946) = 2,62 Prob > F = 0,1055	F(1, 19946) = 5,32 Prob > F = 0,0211	F(1, 19946) = 0,07 Prob > F = 0,7869	F(1, 19946) = 7,45 Prob > F = 0,0064	F(1, 19946) = 4,67 Prob > F = 0,0308	F(1, 19946) = 0,17 Prob > F = 0,6787
$H_0 [q_{10}=q_{50}]$	F(1, 19946) = 0,52 Prob > F = 0,4691	F(1, 19946) = 6,77 Prob > F = 0,0093	F(1, 19946) = 2,94 Prob > F = 0,0862	F(1, 19946) = 0,07 Prob > F = 0,7865	F(1, 19946) = 18,39 Prob > F = 0,0000	F(1, 19946) = 4,49 Prob > F = 0,0340	F(1, 19946) = 0,10 Prob > F = 0,7531
$H_0 [q_{10}=q_{75}]$	F(1, 19946) = 4,10 Prob > F = 0,0428	F(1, 19946) = 4,84 Prob > F = 0,0278	F(1, 19946) = 6,58 Prob > F = 0,0103	F(1, 19946) = 0,04 Prob > F = 0,8456	F(1, 19946) = 21,52 Prob > F = 0,0000	F(1, 19946) = 7,83 Prob > F = 0,0051	F(1, 19946) = 0,06 Prob > F = 0,8106
$H_0 [q_{10}=q_{90}]$	F(1, 19946) = 6,06 Prob > F = 0,0138	F(1, 19946) = 13,34 Prob > F = 0,0003	F(1, 19946) = 8,09 Prob > F = 0,0044	F(1, 19946) = 0,22 Prob > F = 0,6395	F(1, 19946) = 22,47 Prob > F = 0,0000	F(1, 19946) = 12,91 Prob > F = 0,0003	F(1, 19946) = 0,08 Prob > F = 0,7838
$H_0 [q_{25}=q_{50}]$	F(1, 19946) = 0,23 Prob > F = 0,6313	F(1, 19946) = 2,67 Prob > F = 0,1022	F(1, 19946) = 0,05 Prob > F = 0,8195	F(1, 19946) = 0,52 Prob > F = 0,4729	F(1, 19946) = 2,20 Prob > F = 0,1380	F(1, 19946) = 0,30 Prob > F = 0,5840	F(1, 19946) = 1,91 Prob > F = 0,1670
$H_0 [q_{25}=q_{75}]$	F(1, 19946) = 4,69 Prob > F = 0,0303	F(1, 19946) = 1,42 Prob > F = 0,2334	F(1, 19946) = 1,36 Prob > F = 0,2438	F(1, 19946) = 0,00 Prob > F = 0,9865	F(1, 19946) = 3,37 Prob > F = 0,0665	F(1, 19946) = 1,94 Prob > F = 0,1640	F(1, 19946) = 0,01 Prob > F = 0,9080
$H_0 [q_{25}=q_{90}]$	F(1, 19946) = 7,89 Prob > F = 0,0050	F(1, 19946) = 7,69 Prob > F = 0,0056	F(1, 19946) = 3,29 Prob > F = 0,0695	F(1, 19946) = 0,11 Prob > F = 0,7456	F(1, 19946) = 4,51 Prob > F = 0,0336	F(1, 19946) = 7,61 Prob > F = 0,0058	F(1, 19946) = 0,70 Prob > F = 0,4041
$H_0 [q_{50}=q_{75}]$	F(1, 19946) = 4,97 Prob > F = 0,0258	F(1, 19946) = 0,13 Prob > F = 0,7234	F(1, 19946) = 3,24 Prob > F = 0,0717	F(1, 19946) = 0,55 Prob > F = 0,4588	F(1, 19946) = 1,10 Prob > F = 0,2940	F(1, 19946) = 1,31 Prob > F = 0,2518	F(1, 19946) = 1,05 Prob > F = 0,3046
$H_0 [q_{50}=q_{90}]$	F(1, 19946) = 7,63 Prob > F = 0,0057	F(1, 19946) = 3,73 Prob > F = 0,0534	F(1, 19946) = 4,82 Prob > F = 0,0282	F(1, 19946) = 0,71 Prob > F = 0,4002	F(1, 19946) = 2,25 Prob > F = 0,1338	F(1, 19946) = 8,37 Prob > F = 0,0038	F(1, 19946) = 0,00 Prob > F = 0,9914
$H_0 [q_{75}=q_{90}]$	F(1, 19946) = 0,62 Prob > F = 0,4325	F(1, 19946) = 6,32 Prob > F = 0,0120	F(1, 19946) = 0,72 Prob > F = 0,3955	F(1, 19946) = 0,25 Prob > F = 0,6166	F(1, 19946) = 1,10 Prob > F = 0,2944	F(1, 19946) = 2,94 Prob > F = 0,0863	F(1, 19946) = 0,97 Prob > F = 0,3253

Tabela A.1 – Teste F de significância conjunta da regressão quantílica

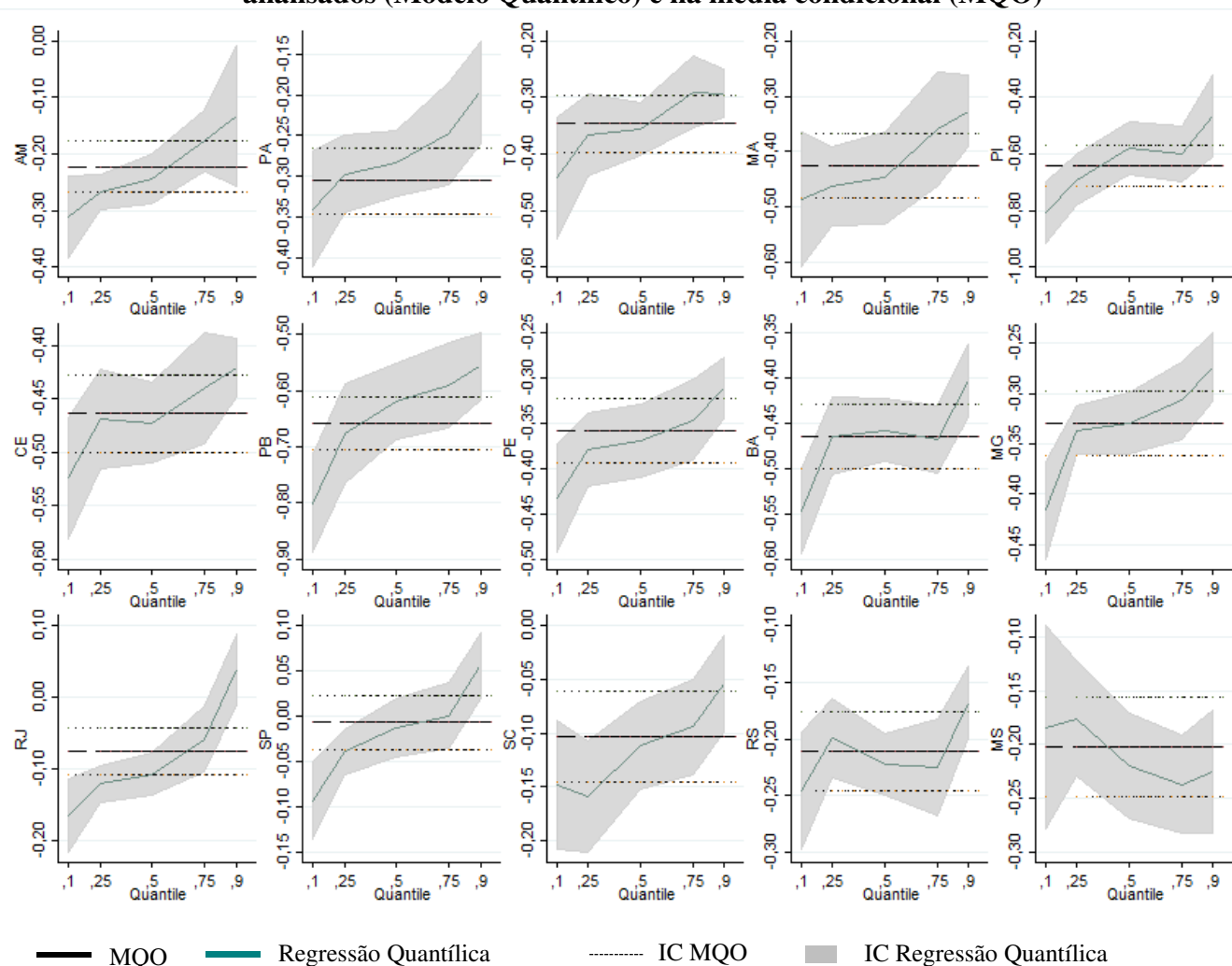
Teste	<i>Continua</i>						
	SE	BA	MG	ES	RJ	SP	PR
$H_0 [q_{10}=q_{25}=q_{50}=q_{75}=q_{90}]$	F(4, 19946) = 1,58 Prob > F = 0,1753	F(4, 19946) = 5,99 Prob > F = 0,0001	F(4, 19946) = 4,14 Prob > F = 0,0024	F(4, 19946) = 0,23 Prob > F = 0,9226	F(4, 19946) = 13,57 Prob > F = 0,0000	F(4, 19946) = 7,67 Prob > F = 0,0000	F(4, 19946) = 0,96 Prob > F = 0,4305
$H_0 [q_{10}=q_{25}]$	F(1, 19946) = 1,35 Prob > F = 0,2458	F(1, 19946) = 9,27 Prob > F = 0,0023	F(1, 19946) = 10,35 Prob > F = 0,0013	F(1, 19946) = 0,65 Prob > F = 0,4219	F(1, 19946) = 4,66 Prob > F = 0,0309	F(1, 19946) = 8,19 Prob > F = 0,0042	F(1, 19946) = 0,38 Prob > F = 0,5373
$H_0 [q_{10}=q_{50}]$	F(1, 19946) = 5,20 Prob > F = 0,0226	F(1, 19946) = 8,19 Prob > F = 0,0042	F(1, 19946) = 8,63 Prob > F = 0,0033	F(1, 19946) = 0,68 Prob > F = 0,4092	F(1, 19946) = 6,11 Prob > F = 0,0135	F(1, 19946) = 12,72 Prob > F = 0,0004	F(1, 19946) = 0,27 Prob > F = 0,6032
$H_0 [q_{10}=q_{75}]$	F(1, 19946) = 5,31 Prob > F = 0,0212	F(1, 19946) = 6,31 Prob > F = 0,0120	F(1, 19946) = 12,76 Prob > F = 0,0004	F(1, 19946) = 0,41 Prob > F = 0,5228	F(1, 19946) = 15,64 Prob > F = 0,0001	F(1, 19946) = 13,73 Prob > F = 0,0002	F(1, 19946) = 0,01 Prob > F = 0,9399
$H_0 [q_{10}=q_{90}]$	F(1, 19946) = 4,63 Prob > F = 0,0315	F(1, 19946) = 20,21 Prob > F = 0,0000	F(1, 19946) = 16,21 Prob > F = 0,0001	F(1, 19946) = 0,21 Prob > F = 0,6448	F(1, 19946) = 45,45 Prob > F = 0,0000	F(1, 19946) = 28,79 Prob > F = 0,0000	F(1, 19946) = 0,73 Prob > F = 0,3927
$H_0 [q_{25}=q_{50}]$	F(1, 19946) = 2,69 Prob > F = 0,1011	F(1, 19946) = 0,18 Prob > F = 0,6697	F(1, 19946) = 0,20 Prob > F = 0,6577	F(1, 19946) = 0,06 Prob > F = 0,8104	F(1, 19946) = 0,73 Prob > F = 0,3914	F(1, 19946) = 3,75 Prob > F = 0,0528	F(1, 19946) = 0,00 Prob > F = 0,9780
$H_0 [q_{25}=q_{75}]$	F(1, 19946) = 2,59 Prob > F = 0,1078	F(1, 19946) = 0,03 Prob > F = 0,8719	F(1, 19946) = 2,95 Prob > F = 0,0859	F(1, 19946) = 0,02 Prob > F = 0,8927	F(1, 19946) = 8,02 Prob > F = 0,0046	F(1, 19946) = 4,94 Prob > F = 0,0263	F(1, 19946) = 0,71 Prob > F = 0,3980
$H_0 [q_{25}=q_{90}]$	F(1, 19946) = 2,38 Prob > F = 0,1230	F(1, 19946) = 8,36 Prob > F = 0,0038	F(1, 19946) = 8,79 Prob > F = 0,0030	F(1, 19946) = 0,03 Prob > F = 0,8733	F(1, 19946) = 38,47 Prob > F = 0,0000	F(1, 19946) = 21,14 Prob > F = 0,0000	F(1, 19946) = 0,36 Prob > F = 0,5505
$H_0 [q_{50}=q_{75}]$	F(1, 19946) = 0,19 Prob > F = 0,6622	F(1, 19946) = 0,35 Prob > F = 0,5520	F(1, 19946) = 2,11 Prob > F = 0,1461	F(1, 19946) = 0,00 Prob > F = 0,9671	F(1, 19946) = 9,29 Prob > F = 0,0023	F(1, 19946) = 0,83 Prob > F = 0,3633	F(1, 19946) = 0,86 Prob > F = 0,3547
$H_0 [q_{50}=q_{90}]$	F(1, 19946) = 0,51 Prob > F = 0,4733	F(1, 19946) = 6,74 Prob > F = 0,0094	F(1, 19946) = 5,80 Prob > F = 0,0160	F(1, 19946) = 0,14 Prob > F = 0,7046	F(1, 19946) = 42,40 Prob > F = 0,0000	F(1, 19946) = 10,92 Prob > F = 0,0010	F(1, 19946) = 0,48 Prob > F = 0,4900
$H_0 [q_{75}=q_{90}]$	F(1, 19946) = 0,24 Prob > F = 0,6270	F(1, 19946) = 10,91 Prob > F = 0,0010	F(1, 19946) = 3,02 Prob > F = 0,0824	F(1, 19946) = 0,10 Prob > F = 0,7542	F(1, 19946) = 19,64 Prob > F = 0,0000	F(1, 19946) = 8,89 Prob > F = 0,0029	F(1, 19946) = 3,11 Prob > F = 0,0778

Tabela A.1 – Teste F de significância conjunta da regressão quantílica

Teste						Conclusão
	SC	RS	MS	MT	GO	
$H_0 [q_{10}=q_{25}=q_{50}=q_{75}=q_{90}]$	F(4, 19946) = 2,06 Prob > F = 0,0829	F(4, 19946) = 2,68 Prob > F = 0,0297	F(4, 19946) = 2,19 Prob > F = 0,0674	F(4, 19946) = 1,30 Prob > F = 0,2693	F(4, 19946) = 0,90 Prob > F = 0,4599	
$H_0 [q_{10}=q_{25}]$	F(1, 19946) = 0,16 Prob > F = 0,6865	F(1, 19946) = 3,40 Prob > F = 0,0650	F(1, 19946) = 0,04 Prob > F = 0,8359	F(1, 19946) = 0,93 Prob > F = 0,3351	F(1, 19946) = 0,07 Prob > F = 0,7852	
$H_0 [q_{10}=q_{50}]$	F(1, 19946) = 1,51 Prob > F = 0,2192	F(1, 19946) = 0,65 Prob > F = 0,4191	F(1, 19946) = 0,62 Prob > F = 0,4293	F(1, 19946) = 0,54 Prob > F = 0,4641	F(1, 19946) = 0,06 Prob > F = 0,8036	
$H_0 [q_{10}=q_{75}]$	F(1, 19946) = 2,99 Prob > F = 0,0837	F(1, 19946) = 0,45 Prob > F = 0,5012	F(1, 19946) = 1,29 Prob > F = 0,2563	F(1, 19946) = 1,68 Prob > F = 0,1945	F(1, 19946) = 0,00 Prob > F = 0,9929	
$H_0 [q_{10}=q_{90}]$	F(1, 19946) = 5,06 Prob > F = 0,0245	F(1, 19946) = 4,70 Prob > F = 0,0303	F(1, 19946) = 0,63 Prob > F = 0,4260	F(1, 19946) = 2,90 Prob > F = 0,0884	F(1, 19946) = 1,02 Prob > F = 0,3133	
$H_0 [q_{25}=q_{50}]$	F(1, 19946) = 4,85 Prob > F = 0,0277	F(1, 19946) = 1,80 Prob > F = 0,1793	F(1, 19946) = 4,90 Prob > F = 0,0269	F(1, 19946) = 0,23 Prob > F = 0,6315	F(1, 19946) = 0,69 Prob > F = 0,4049	
$H_0 [q_{25}=q_{75}]$	F(1, 19946) = 5,66 Prob > F = 0,0174	F(1, 19946) = 1,18 Prob > F = 0,2780	F(1, 19946) = 6,74 Prob > F = 0,0094	F(1, 19946) = 0,93 Prob > F = 0,3338	F(1, 19946) = 0,10 Prob > F = 0,7484	
$H_0 [q_{25}=q_{90}]$	F(1, 19946) = 6,79 Prob > F = 0,0092	F(1, 19946) = 1,56 Prob > F = 0,2111	F(1, 19946) = 2,88 Prob > F = 0,0899	F(1, 19946) = 2,88 Prob > F = 0,0896	F(1, 19946) = 1,51 Prob > F = 0,2185	
$H_0 [q_{50}=q_{75}]$	F(1, 19946) = 0,68 Prob > F = 0,4088	F(1, 19946) = 0,02 Prob > F = 0,9017	F(1, 19946) = 0,52 Prob > F = 0,4711	F(1, 19946) = 2,40 Prob > F = 0,1210	F(1, 19946) = 0,13 Prob > F = 0,7159	
$H_0 [q_{50}=q_{90}]$	F(1, 19946) = 2,71 Prob > F = 0,0995	F(1, 19946) = 6,18 Prob > F = 0,0129	F(1, 19946) = 0,02 Prob > F = 0,8763	F(1, 19946) = 4,36 Prob > F = 0,0368	F(1, 19946) = 3,18 Prob > F = 0,0748	
$H_0 [q_{75}=q_{90}]$	F(1, 19946) = 1,92 Prob > F = 0,1663	F(1, 19946) = 8,07 Prob > F = 0,0045	F(1, 19946) = 0,21 Prob > F = 0,6479	F(1, 19946) = 1,21 Prob > F = 0,2719	F(1, 19946) = 2,62 Prob > F = 0,1055	

Fonte: Elaborado pelos autores.

Figura A.1 – Coeficiente e intervalo de confiança das *dummies* estaduais nos diferentes quantis analisados (Modelo Quantílico) e na média condicional (MQO)



Nota: Resultados reportados apenas para os casos em que houve a rejeição de H_0 no teste F de significância conjunta.
 Fonte: Elaborado pelos autores.

**SERVIÇOS DE SANEAMENTO BÁSICO EM MINAS GERAIS E SEUS
DETERMINANTES LOCACIONAIS, DEMOGRÁFICOS E SOCIOECONÔMICOS***

Juliana Brito de Oliveira

Mestranda em Economia na Universidade Federal de Viçosa (UFV)

E-mail: juliana.brito93@outlook.com

Gabriel Teixeira Ervilha

Doutor em Economia Aplicada pela Universidade Federal de Viçosa (UFV)

E-mail: gabriel.ervilha@ufv.br

RESUMO: O saneamento básico é fator fundamental para o desenvolvimento de uma região e, neste contexto, o objetivo do presente estudo é mensurar indicadores de serviços de saneamento básico em 794 municípios de Minas Gerais e identificar seus determinantes locais, demográficos e socioeconômicos. Especificamente, pretende-se, com dados de 2015 e através da Análise Multivariada de Dados, mensurar um Indicador de Acesso e Qualidade do Saneamento Básico (IAQSB) e apontar quais os fatores que discriminam os municípios com maiores e menores IAQSB. Os resultados obtidos permitem identificar dois municípios de destaque no IAQSB: Contagem e Juiz de Fora. Tal análise também evidencia a existência de expressivas desigualdades entre os municípios das regiões Norte, vales do Jequitinhonha e Mucuri. A análise dos fatores sociodemográficos que mais discriminam o IAQSB aponta para a existência de quatro variáveis com capacidade de discriminação, relacionadas ao desenvolvimento tributário e econômico; à pobreza; à taxa de urbanização; e ao desenvolvimento municipal. Assim, todos os resultados obtidos e analisados no presente estudo fortalecem o argumento de que a efetividade das políticas públicas para a melhoria na qualidade nos serviços de saneamento básico, assim como a de redução nas desigualdades de acesso dos mesmos à população, requer ações coordenadas nas diversas frentes de ação governamental.

Palavras-chave: Análise multivariada de dados; Saneamento básico; Minas Gerais.

Classificação JEL: C43; I18.

**BASIC SANITATION SERVICES IN MINAS GERAIS STATE AND THEIR LOCAL,
DEMOGRAPHIC AND SOCIO-ECONOMIC DETERMINANTS**

ABSTRACT: The basic sanitation is a fundamental factor for the development of a region and, in this context, the present study aims to measure indicators of basic sanitation services in 794 municipalities of Minas Gerais State, as well as to identify their local, demographic and socio-economic determinants. Specifically, it is intended with data from 2015 and through the multivariate data analysis, the study aimed to measure the Indicator of Access and Quality of the Basic Sanitation (IAQSB) and to point out which factors discriminate the municipalities with larger and smaller IAQSB. The results obtained identify two prominent municipal districts within IAQSB: Contagem and Juiz de Fora. Such analysis also demonstrates the existence of expressive inequalities among the municipalities of the northern areas, Jequitinhonha and Mucuri valleys. The analysis of socio-demographic factors that discriminated IAQSB more, present as four variables with discrimination capacity, related to the tax and economic development; to the poverty; to the urbanization tax and the economic development. Thus, all the results obtained and analyzed in the present study strengthen the argument that the effectiveness of public policies for the improvement of the quality of sanitation services, as well as, the reduction of access inequalities for the population, request actions coordinated in the different fronts of the government action.

Keywords: Multivariate data analysis; Basic sanitation; Minas Gerais.

JEL Codes: C43; I18.

1. Introdução

O saneamento básico é fator fundamental para o desenvolvimento de uma região, visto que impacta social, ambiental e economicamente os municípios atendidos. Ao implementar as condições sanitárias necessárias ao bem-estar coletivo, os serviços de saneamento garantem qualidade de vida e saúde para a população, disponibilizando água tratada, coleta e tratamento de esgotos e acondicionamento adequado dos resíduos sólidos urbanos.

Segundo a Organização Mundial da Saúde (WHO, 2004), saneamento é o conjunto de medidas adotadas em uma região para melhorar a qualidade de vida e a saúde dos indivíduos, impedindo que fatores físicos possam exercer efeitos nocivos sobre o bem-estar físico, mental ou social. De outra forma, pode-se dizer que saneamento caracteriza o conjunto de ações socioeconômicas que tem como objetivo a salubridade ambiental¹ (GUIMARÃES; CARVALHO; SILVA, 2007).

No Brasil, ao findar a primeira década do século XXI, uma parcela significativa da população ainda não possuía acesso ao abastecimento de água e esgotamento sanitário e, se os possuía, era de forma precária e insuficiente. Dados da Pesquisa Nacional de Amostra de Domicílios (PNAD), em 2008, apontaram que cerca de 8% da população brasileira não tinha acesso ao abastecimento de água, sendo que a Região Nordeste registrou o maior déficit absoluto, atingindo 7,7 milhões de pessoas (IBGE, 2010). Os dados do IBGE (2012) revelam, ainda, que 1.514.992 domicílios não tinham banheiros nem sanitários (2,64% do total de domicílios) e 7.218.079 (12,59% do total) lançavam seus resíduos sólidos diretamente no meio ambiente de forma inadequada no ano de 2010.

Entretanto, o país atravessava condições potencialmente favoráveis ao desenvolvimento das ações de saneamento², sustentadas pelo marco legal criado nesse período e pela crescente institucionalização do setor. Através da Lei dos Consórcios Públicos n.º 11.107/2005 e da Lei das Diretrizes Nacionais para o Saneamento Básico n.º 11.445/2007³, a nova Política Nacional de Saneamento pode ser implementada com grandes possibilidades de avanço, conduzindo a prestação dos serviços de saneamento a um novo patamar e ampliando as chances de universalização de tais serviços com integridade e equidade (LIMA; MARQUES, 2012).

Com o reconhecimento dos serviços de saneamento básico como dever público, é de fundamental importância para os gestores proverem, de forma eficaz, por meios próprios ou através de concessões, as necessidades da população no que tange ao saneamento básico e, também, a outros serviços sociais que são relevantes para a melhoria da qualidade de vida. Destarte, surge a necessidade de mensurar as ações governamentais ao provimento da qualidade de vida para a população; sendo assim, despontam os indicadores sociais.

Segundo Rouquayrol (1993), sob o ponto de vista sanitário, esses indicadores de desempenho são parâmetros utilizados com o objetivo de avaliar a higidez de agregados humanos, bem como fornecer subsídios aos planejamentos, permitindo o estabelecimento de padrões e o acompanhamento de sua evolução ao longo dos anos.

O uso de indicadores no Brasil e no mundo vem se tornando uma prática cada vez mais frequente. Como uma importante ferramenta de suporte à gestão pública, a utilização de indicadores de desempenho, no setor de saneamento, pode ser definida como uma medida quantitativa da

¹ Salubridade ambiental é entendida como o estado de higidez (estado de saúde normal) em que vive a população urbana e rural, tanto no que se refere à sua capacidade de inibir, prevenir ou impedir a ocorrência de endemias ou epidemias veiculadas ao meio ambiente, como no tocante ao seu potencial de promover o seu aperfeiçoamento de condições mesológicas (que diz respeito ao clima e/ou ambiente) favoráveis ao pleno gozo de saúde e bem-estar (GUIMARÃES; CARVALHO; SILVA, 2007).

² De acordo com Lima e Marques (2012), é preciso enfatizar a dinâmica populacional brasileira nesse período, a qual atravessou um momento de impulso à coletivização do acesso aos bens públicos, graças a um bônus demográfico favorável às melhorias em saneamento. O crescimento da economia nos anos anteriores também influenciou as previsões otimistas de investimentos para o saneamento e suas interfaces.

³ Para efeitos da Lei, considera-se Saneamento Básico como conjunto de serviços, infraestruturas e instalações operacionais de abastecimento de água potável, esgotamento sanitário, limpeza urbana e manejo de resíduos sólidos e drenagem e manejo das águas pluviais, limpeza e fiscalização preventiva das respectivas redes urbanas.

eficiência de uma entidade gestora, sendo uma forma importante de avaliação dos serviços prestados à população.

O Brasil possui um desenvolvimento regional extremamente desigual, tanto nos aspectos econômicos quanto nos sociais. Neste contexto, o estudo do saneamento básico, também através de indicadores, adquire importância e complexidade relevantes, caracterizada pela extensão territorial: regiões diferenciadas pelo clima, pelo relevo e pela natureza do solo e também pela população com diferentes estágios sociais e culturais.

Em Minas Gerais, por exemplo, encontram-se regiões com indicadores socioeconômicos de alto nível, como é o caso das regiões Centro-Sul e Triângulo Mineiro; áreas altamente dinamizadas economicamente e com a maior parcela de produto interno. Por outro lado, há regiões menos desenvolvidas, como os vales do Jequitinhonha e Mucuri e as zonas Norte e Leste de Minas, caracterizadas por menores dinamismo e expressividade econômica e com indicadores sociais mais pobres. Em relação a outros estados, Minas Gerais é o retrato-síntese do Brasil, pois representa a heterogeneidade que se observa no país (QUEIROZ; GOLGHER; AMARAL, 2010).

Nesse contexto, as diferenças entre as regiões mais desenvolvidas e as menos desenvolvidas estão muito acentuadas, tanto no que diz respeito à oportunidade de renda e emprego quanto no acesso à educação, ao saneamento básico, dentre outros serviços.

No âmbito do saneamento básico, de acordo com o último levantamento do Sistema Nacional de Informações sobre o Saneamento (SNIS 2015), 82,47% da população do estado de Minas Gerais é atendida com o abastecimento de água tratada; 69,11% da população tem acesso à coleta de esgoto (referido aos municípios atendidos com água); 39,91% da população de Minas Gerais tem o esgoto tratado em sua residência (esgoto tratado referido à água consumida); e 57,71% da população urbana é atendida por sistemas de tratamento e/ou disposição final de resíduos sólidos urbanos (ITB, 2015; FEAM, 2015).

Essa ausência considerável de coleta e tratamento de esgoto pode ocasionar impactos ao meio ambiente, à saúde, e, portanto, à qualidade de vida da população. De acordo com estudo do Instituto Trata Brasil (2015), estima-se que 65% das internações em hospitais mineiros, de crianças com menos de 10 anos, sejam provocadas por males ou inexistência de tratamento de esgoto e de água limpa e potável. Tal deficiência também afeta o rendimento escolar, pois crianças que vivem em áreas sem saneamento básico apresentaram desempenho escolar 18% inferior ao daquelas que possuem acesso aos serviços (ITB, 2015).

Sob a ótica da estrutura de provisão do setor, os principais modelos de gestão do saneamento básico em Minas Gerais podem ser caracterizados através do modelo centralizado, ou administração direta, o qual trata dos serviços de saneamento prestados diretamente pela Prefeitura Municipal; pelo modelo descentralizado ou administração indireta, o qual corresponde aos serviços organizados sob a forma de autarquias municipais; e, por fim, pelas empresas classificadas pelo SNIS de prestadores regionais⁴ – as quais se organizam no estado de forma capilar e descentralizada e a atuação nos municípios é regulada por contratos de concessão (COPASA, 2007).

A literatura apresenta vários estudos relacionados à aplicação de indicadores de desempenho no âmbito do saneamento básico. Os que merecem destaque em Minas Gerais são apontados por Heller, Von Sperling e Heller (2009), que avaliaram, por meio da análise multicritério, o desempenho tecnológico dos serviços de água e esgotamento sanitário em quatro municípios de Minas Gerais; e Costa et al. (2013), que, por meio de um estudo exploratório e qualitativo, analisaram a evolução dos prestadores de serviços de saneamento do estado de Minas Gerais com

⁴ São empresas que atendem a mais de um município, distribuídas pelo estado e que geralmente atendem as regiões metropolitanas. Em Minas Gerais, a Companhia de Saneamento de Minas Gerais (COPASA) é um exemplo desse grupo. Tais empresas são compostas por diretorias que são subdivididas em superintendências, essas em distritos, e por último, em escritórios locais.

base em indicadores do Sistema Nacional de Informações em Saneamento Básico (SNIS), trazendo à tona o monitoramento do progresso das ações do setor no estado.⁵

Tendo em vista a contribuição para a melhoria do sistema de gestão no setor de saneamento básico e o auxílio na elaboração de políticas públicas no estado de Minas Gerais, a presente pesquisa tem por objetivo analisar as condições sanitárias do estado, utilizando-se um indicador criado, que sintetiza as variáveis operacionais e de investimento para o acesso desses serviços. Posteriormente, permite-se analisar seus determinantes locacionais, demográficos e socioeconômicos para identificar as necessidades da provisão do saneamento básico entre os municípios mineiros.

Através de uma análise quantitativa para o estado de Minas Gerais, o artigo pretende contribuir para a literatura sobre o tema ao fomentar a discussão sobre o acesso e a qualidade dos serviços de saneamento básico e identificar os principais determinantes que diferem os grupos de municípios quanto às condições no estado. Do ponto de vista metodológico, a contribuição se dá pela mensuração de um indicador capaz de sintetizar as variáveis relevantes de acesso e qualidade desses serviços, utilizando-se dados de 2015.

Uma das razões para a escolha do estado de Minas Gerais, conforme já discutido, é sua significativa heterogeneidade geográfica, que reproduz, em grande medida, as características percebidas no Brasil como um todo; além da expressiva dimensão geográfica e da sua quantidade de municípios (853). Ademais, este trabalho poderá servir de referência para a realização de estudos semelhantes que possam testar a mesma metodologia adotada para outros estados da federação.

2. Indicadores sociais e sua atuação nas políticas públicas de saneamento básico

Ao avaliar as condições do desenvolvimento socioeconômico brasileiro, percebe-se que há disparidades regionais que são verificadas através das condições econômicas e sociais. O serviço de saneamento básico constitui um indicador social, pois, ao ser mensurado, consegue retratar realidades de uma região, mostrando essas disparidades regionais que incluem, também, diferenças significativas relacionadas a fatores culturais e ambientais (OLIVEIRA; LAGES; DANTAS, 2010).

Segundo Jannuzzi (2002), o aparecimento de indicadores sociais está ligado à consolidação de atividades de planejamento do setor público ao longo do século XX. O desenvolvimento desses indicadores vem acompanhando as transformações sociais e a aferição do impacto das políticas sociais nas sociedades desenvolvidas e subdesenvolvidas. A definição de indicadores sociais, na perspectiva de Jannuzzi (2003), é uma medida usada para substituir, quantificar ou operacionalizar um conceito abstrato de interesse teórico ou pragmático. Seria o elo entre os modelos explicativos da teoria social e a evidência empírica dos fenômenos sociais observados.

Indicadores sociais informam sobre um aspecto da realidade social, subsidiam as atividades de planejamento público e de formulação de políticas sociais nas diferentes esferas de Governo, permitindo uma definição mais adequada de prioridades sociais para a alocação de recursos do orçamento público. Assim, os indicadores sociais podem ser classificados segundo as diversas aplicações a que se destinam (JANNUZZI, 2002).

Para a análise e formulação de políticas sociais, a diferenciação dos indicadores quanto à natureza é bastante relevante. Na avaliação de programas sociais, há um sistema de classificação que os diferencia em três aspectos: indicadores para a avaliação da eficiência dos meios e recursos empregados; indicadores para a avaliação da eficácia no cumprimento das metas; e, por fim, indicadores para a avaliação da efetividade social do programa, ou seja, em termos da contribuição para o aumento do engajamento político, dos efeitos do programa na geração do bem-estar da sociedade (JANNUZZI, 2002).

⁵ Outras importantes constatações sobre o saneamento em Minas Gerais podem ser obtidas em Heller, Coutinho e Mingoti (2006), Rossoni et al. (2014), Carmo (2015) e Miranda (2015). Para análises a nível nacional, consultar Toneto Júnior e Saiani (2006), Scriptore e Toneto Junior (2012), Silveira, Heller e Rezende (2013), Borja (2014), Santos (2016), Britto e Rezende (2017), Ponciano (2017) e Rocha, Mattos e Saiani (2017).

A literatura que busca avaliar a eficiência dos serviços de saneamento básico é bastante disseminada. Para o estado de Minas Gerais, ao sintetizar indicadores operacionais, de investimento e qualidade, Miranda (2015) contribuiu para uma melhor compreensão das disparidades regionais no estado. O autor identificou os principais determinantes dos escores de eficiência dos prestadores do estado, inclusive relacionados a subsídio cruzado na prestação regional de economias de escala. Os resultados apresentaram pura eficiência entre os municípios mineiros.

Já Ponciano (2017) utilizou dados do SNIS para medir a eficiência do setor nos municípios brasileiros analisando o efeito da Lei Nacional de Saneamento Básico (LNSB), discriminando pela natureza do capital (público e privado) e da cobertura territorial (regional e local) entre os anos de 2006 e 2013. O Índice de Malmquist foi utilizado para uma análise dinâmica e, como resultado, encontrou escores baixos, o que possibilita grande espaço para melhorias de produtividade. A análise de eficiência de escala apontou que as prestadoras regionais deveriam reduzir de tamanho para se aproximar da escala ótima. Já a análise dinâmica indicou que as prestadoras regionais obtiveram maiores avanços do que as locais e que não houve diferença entre as públicas e as privadas. Por fim, o autor concluiu que não ocorreu relação entre a presença de agência reguladora e o aumento de eficiência técnica - talvez pelo curto período de implementação da LNSB.

Costa et al. (2013) realizaram uma análise dos indicadores sobre a prestação dos serviços de saneamento básico em Minas Gerais, com dados do SNIS de 2005 a 2010 por meio da Regulação *Sunshine*⁶. Os resultados foram considerados satisfatórios em termos de qualidade, atendimento e perdas dos serviços de água para todo o estado. Os resultados dos serviços de esgotamento sanitário apresentaram níveis inferiores aos desejados, com a exceção do Triângulo Mineiro, Oeste e Zona da Mata. Porém, ao analisar a abrangência das prestadoras no estado, os autores identificaram que essas parecem influenciar a qualidade dos serviços, uma vez que o indicador de qualidade da água apresentou resultados médios abaixo dos permitidos pela regulamentação do Ministério da Saúde para prestadores locais, apontando para a ausência de um controle mais rígido sobre a qualidade da água e até mesmo para riscos de contaminação da população.

A disponibilidade de um sistema amplo de indicadores sociais válidos, relevantes e confiáveis potencializa o sucesso na implementação de políticas públicas e permite diagnósticos sociais e monitoramento de ações e avaliações de resultados mais abrangentes. Como toda atividade sócio-política, é essencial garantir a participação social no processo, a fim de legitimá-lo perante a sociedade, garantir o compromisso dos agentes implementados e potencializar a efetividade almejada pelas políticas públicas (JANNUZZI, 2002).

Portanto, as políticas de saneamento são imprescindíveis quando se trata de “saúde ambiental”⁷, sobretudo quando se considera o modelo socioeconômico, com sua má distribuição de renda: enriquecimento das minorias e abandono da maioria às condições de pobreza. É essa exclusão, perpetuação e permanência da pobreza que confirma a pertinência das preocupações com a relação ao saneamento e à saúde, haja vista que as consequências da insalubridade ambiental afetam com maior intensidade estratos de população mais vulneráveis, com menor nível socioeconômico (HELLER, 1998).

3. A análise multivariada de dados

A estatística multivariada de dados consiste em um conjunto de métodos estatísticos utilizados quando as variáveis são medidas concomitantemente em cada elemento amostral. Tem como

⁶ Esse esquema de regulação firma-se na publicação dos resultados do desempenho dos prestadores de serviços e na sua comparação com as demais entidades do mesmo setor. A conscientização da sua atuação é obtida por meio da pressão dos consumidores por intermédio de seus grupos de defesa e de representação, da mídia e da classe política.

⁷ A expressão “saúde ambiental”, segundo Heller (1998), é considerada a chave para orientar a organização institucional e sensibilizar comunidade, técnicos e governos sobre a necessidade de uma abordagem que articule ambas as esferas. Há, segundo o autor, a percepção de que saúde e ambiente se aproximem no conceito e na prática. Para a área da saúde, seria valorizar o meio ambiente como fator determinante de agravos à saúde, enquanto, para a área ambiental, seria visualizar os efeitos das alterações ambientais sobre a saúde humana.

principal objetivo investigar as relações de dependência entre a variável resposta ao fenômeno que está sendo estudado e as variáveis explicativas (MINGOTI, 2005).

Na construção de indicadores, utiliza-se a Análise Multivariada de Dados para sintetizar diversas variáveis que descrevem um fenômeno em uma única variável, com o intuito de construir algo específico para quantificar o objeto que está sendo estudado. Os indicadores desenvolvidos são calculados pela razão entre variáveis da mesma natureza ou de natureza distinta, sendo adimensionais ou não (VON SPERLING; VON SPERLING, 2013).

A análise de dados através de indicadores para avaliar os serviços de saneamento básico é considerada uma medida quantitativa da eficiência e da eficácia pelas quais esses serviços são providos por uma entidade gestora para ponderar seu desempenho (VON SPERLING; VON SPERLING, 2013).

Portanto, atendendo aos objetivos deste trabalho, através da Análise Fatorial, pretende-se determinar as variáveis relevantes na criação de um Indicador de Acesso e Qualidade do Saneamento Básico (IAQSB). Após mensurá-lo, através da Análise Discriminante, pretende-se identificar os fatores demográficos, locais e socioeconômicos que discriminam os municípios do estado de Minas Gerais com melhores e piores indicadores de acesso e qualidade do saneamento básico.

3.1. Análise fatorial

A Análise Fatorial é uma técnica de interdependência que permite explicar e definir uma estrutura de análise entre as variáveis, ou seja, através de métodos estatísticos, destina-se condensar as informações contidas em um conjunto de variáveis originais em um conjunto menor de fatores, ocorrendo uma perda mínima de informação (REZENDE; FERNANDES; RODRIGUES, 2007).

Para avaliar a adequabilidade desta técnica, utiliza-se o teste de Kaiser-Meyer-Olkin (KMO). Segundo Hair et al. (2009), o teste KMO é uma estatística que indica a proporção da variância dos dados que pode ser comum a todas as variáveis (pode ser atribuído um fator comum). Os valores do índice de KMO considerados adequados são entre 0,5 e 1,0. Quanto mais próximo de 1, mais adequada é a amostra à aplicação da Análise Fatorial.

Já o teste de Esfericidade de Barlett pode ser definido como uma estatística de teste para verificar se as variáveis não são correlacionadas na amostra. Sendo assim, o teste verifica a hipótese de a matriz de correlação ser igual à matriz identidade, ou seja, admitir a ausência de associação linear entre as variáveis selecionadas (HAIR et al., 2009).

Considerando os objetivos propostos no presente estudo, após a construção da matriz de correlação, a construção do indicador de qualidade e acesso do saneamento básico para os municípios mineiros é realizada utilizando o método Varimax. Aplica-se a rotação Varimax com o objetivo de reduzir as variáveis iniciais em subconjuntos com o maior grau de independência possível (componentes principais) (HAIR et al., 2009).

Ao definir os fatores que justificam as correlações observadas entre as variáveis, com o intuito de reduzi-las em dimensões latentes menores, pode-se identificá-las separadamente da estrutura das variáveis originais e determinar o grau que cada variável é explicada por cada dimensão. A redução dos dados pode ser calculada através de escores para cada dimensão latente e, assim, poder-se-á substituir as variáveis originais pelos mesmos (HAIR et al., 2009).

A mensuração do IAQSB em cada município de Minas Gerais é feita através da ponderação dos escores dos fatores com a comunalidade das variáveis presentes em cada fator. O IAQSB para avaliar o acesso e a qualidade nos serviços de saneamento nos municípios mineiros analisados é calculado de acordo com a seguinte equação:

$$IAQSB = \sum_{i=1}^n F_i P_i \quad (1)$$

em que n é igual ao número de fatores; F_i é igual ao escore fatorial do Fator i ; e P_i é igual ao peso do Fator i , calculado pela porcentagem da variância explicada pelo Fator i .

Posteriormente, com o cálculo do IAQSB, procede-se com a interpretação e discussão dos resultados obtidos.

3.2. Análise discriminante

A Análise Discriminante consiste em um método de estatística multivariada que possibilita a classificação de elementos de uma dada amostra de acordo com grupos previamente conhecidos, sendo também possível elaborar uma regra de classificação a ser utilizada para enquadrar eventuais novas observações nos grupos existentes (MINGOTI, 2005).

Dependendo do número de categorias da variável dependente, a técnica de Análise Discriminante pode ser de dois grupos, em que é deduzida somente uma função discriminante, ou múltipla, em que pode ser estimada mais de uma função. Neste estudo, a técnica utilizada será a de dois grupos, ou seja, um grupo com os municípios de melhores IAQSB e outro grupo com os municípios de piores IAQSB.

A Análise Discriminante reduz o número de variáveis para um número menor de parâmetros, que são funções discriminantes linearmente dependentes das variáveis originais. Os coeficientes da função discriminante indicam a contribuição da variável original para a função. O modelo é dado de acordo com a equação:

$$Y_{jk} = \beta_0 + \beta_1 \cdot X_{1k} + \beta_2 X_{2k} + \dots + \beta_n X_{nk} \quad (2)$$

em que Y_{jk} é o escore discriminante da função discriminante j para o município k ; β_n é o peso discriminante atribuído a cada variável independente i ; e X_{ik} são as variáveis que mais foram discriminadas na função para os municípios. Os coeficientes β_n são estimados de modo a diferenciar os grupos ao máximo.

Utiliza-se o método *stepwise*, para selecionar as variáveis para entrar na análise, baseando-se nas suas capacidades de discriminação. O processo inicia-se selecionando a variável que apresenta maior valor de discriminação. Essa variável é pareada com as demais variáveis, uma de cada vez, e o critério de seleção é novamente comparado. A variável que, em conjunto com a primeira selecionada, produzir o melhor valor para o critério é a segunda variável escolhida para entrar na análise. Sequencialmente, de forma semelhante, todas as demais variáveis são testadas através do critério de seleção, até que todas sejam ordenadas pelas suas capacidades de discriminação (SHARMA, 1996).

O método de decisão estatística para o caso de dois grupos classifica uma observação no Grupo 1 se:

$$Y_k \geq \frac{\bar{Y}_1 + \bar{Y}_2}{2} + \ln \left[\frac{p_2 C(1/2)}{p_1 C(2/1)} \right] \quad (3)$$

e outra observação será classificada no Grupo 2 se:

$$Y_k < \frac{\bar{Y}_1 + \bar{Y}_2}{2} + \ln \left[\frac{p_2 C(1/2)}{p_1 C(2/1)} \right] \quad (4)$$

em que Y_k é o valor discriminante para uma dada observação k ; \bar{Y}_j é valor discriminante médio para o grupo j ; p_j é a probabilidade prévia do grupo j ; e $C(1/2)$ é o custo de classificação incorreta dentro do grupo i de uma observação que pertence ao grupo j .

Para Hair et al. (2009), um dos métodos mais utilizados para identificar o poder discriminatório de uma variável é o teste Lambda de Wilks (L^*). Esse teste considera como critério de seleção de variáveis o valor da Estatística F Multivariada, para o teste da diferença entre os centroides dos grupos. A variável que maximiza o valor da estatística F também minimiza o L^* , que é uma medida de discriminação entre os grupos.

Posteriormente, é realizada a interpretação dos resultados e, por fim, procede-se com a avaliação da validade da Análise Discriminante, em que os coeficientes β_n estimados são multiplicados pelos valores das variáveis independentes na amostra retida, a fim de gerar valores discriminantes para os casos nessa amostra (MALHOTRA, 2011).⁸

3.3. Base de dados e variáveis utilizadas

Para atender aos objetivos específicos presentes neste trabalho, através da Análise Fatorial, pretende-se determinar as variáveis relevantes na criação de um Indicador de Acesso e Qualidade do Saneamento Básico (IAQSB). Após mensurá-lo, através da Análise Discriminante, pretende-se identificar os fatores demográficos, locacionais e socioeconômicos que discriminam os municípios do estado de Minas Gerais com melhores e piores indicadores de acesso e qualidade do saneamento básico.

A base de dados para a construção do Indicador de Qualidade e Acesso do Saneamento Básico (IAQSB) é o Sistema Nacional de Informações sobre Saneamento (SNIS) que é, atualmente, a base de dados mais completa sobre o setor no Brasil. O SNIS tem sido desenvolvido, desde a sua criação, pelo Programa de Modernização do Setor Saneamento (PMSS), vinculado à Secretaria Nacional de Saneamento Ambiental do Ministério das Cidades.

A base de dados reúne informações de prestadores estaduais, regionais e municipais de serviços de acesso à água, coleta e tratamento de esgoto, além de resíduos sólidos. De acordo com a base de dados mais recente disponível (2015), a amostra deste estudo foi realizada com 794 municípios de Minas Gerais que responderam, de forma voluntária, os questionários por parte das operadoras de saneamento. Após a análise dos dados, foram selecionadas 43 variáveis direcionadas às questões de acesso e qualidade dos serviços de saneamento básico municipal⁹.

A partir do IAQSB, utiliza-se a técnica multivariada de análise discriminante a fim de se verificar quais são os principais determinantes que diferenciam os grupos de municípios de Minas Gerais em relação à condição de saneamento básico municipal. Para a realização do presente estudo, foram selecionadas 23 variáveis, das quais 18 encontram-se no Índice Mineiro de Responsabilidade Social (IMRS, 2015), quatro no Índice Firjan de Desenvolvimento Municipal (IFDM, 2015) e uma variável no Instituto Brasileiro de Geografia e Estatística (IBGE, 2012)¹⁰.

Por fim, para melhor compreensão dos procedimentos metodológicos propostos, a Figura 1 apresenta as etapas, os métodos, os dados e os resultados para se alcançarem os dois objetivos inicialmente definidos.

4. Resultados e discussão

4.1. O IAQSB

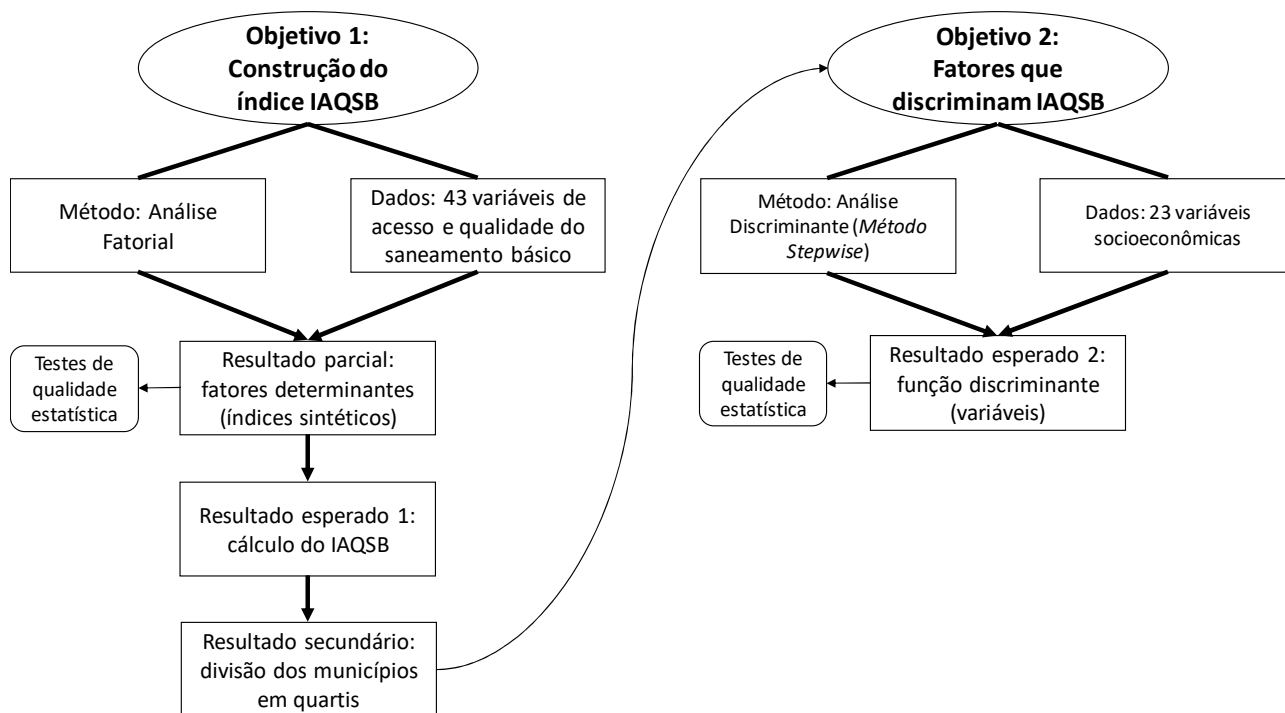
Como o objetivo deste estudo foi determinar as variáveis relevantes na criação de um Indicador de Acesso e Qualidade do Saneamento Básico (IAQSB) em Minas Gerais e, em seguida, mensurá-lo, precedeu-se com a técnica multivariada de Análise Fatorial.

⁸ Para consultas mais detalhadas sobre a análise multivariada de dados, recomendam-se referências como: Hair et al. (2009), Malhotra (2011) e Mingoti (2005).

⁹ A tabela com as variáveis selecionadas para a construção do IAQSB encontra-se no Anexo A.

¹⁰ A tabela com as 23 variáveis encontra-se no Anexo B.

Figura 1 – Síntese dos procedimentos metodológicos propostos



Fonte: Elaboração própria.

No que tange à qualidade estatística dos resultados do modelo, o teste KMO obteve um valor de 0,894, indicando que os dados se inter-relacionam e, portanto, a análise dos componentes principais pode ser realizada. Por sua vez, o teste de Esfericidade de Barlett é significativo a um nível de significância de 1%, rejeitando a hipótese nula de que a matriz de correlações entre as 43 variáveis originais do modelo é uma matriz identidade, o que leva a aceitar a existência de correlações significativas entre as variáveis.

Através da estimação, esse processo delineou as 43 variáveis em 12 fatores determinantes, que explicam 78,265% da variância total e mostram que, para cada componente principal, as cargas fatoriais indicam a relação das variáveis com os fatores extraídos¹¹. A Tabela 1 apresenta os fatores e a participação dos fatores para explicar a variância das variáveis relativas ao saneamento básico em Minas Gerais.

Constata-se que tanto o Fator 1 quanto o Fator 2, subjacente à base de dados associada aos serviços de saneamento básico em Minas Gerais, são significativamente representados por seis variáveis originais. O primeiro está relacionado à concentração e manejo de resíduos sólidos urbanos e o segundo à gestão desses resíduos, população atendida com a coleta de resíduos sólidos e frequência desse atendimento.

Já os Fatores 3 e 7 estão relacionados com a qualidade nos serviços de saneamento executados. O Fator 3 é representado por cinco variáveis e refere-se ao atendimento e às reclamações nos serviços de abastecimento de água e coleta de esgoto executados. Já o Fator 7 é representado por três variáveis originais e refere-se às paralisações e interrupções nos serviços de saneamento.

Os Fatores 5, 6 e 11 são identificados por variáveis referentes aos serviços de abastecimento de água. O Fator 4 extraído é composto por cinco variáveis originais referentes aos serviços de esgotamento sanitário prestados.

As variáveis que mostram o índice de produtividade nos serviços de saneamento estão correlacionadas com o Fator 9. Já os investimentos realizados nos serviços de saneamento básico são atributos do Fator 10, composto por três variáveis originais.

¹¹ As tabelas com as relações das variáveis originais com os fatores extraídos encontram-se no Anexo A.

Por fim, devido à inspeção da carga dos fatores na matriz dos componentes principais rotacionada, os Fatores 8 e 12 não manifestaram características de um serviço de saneamento em específico, mas são compostos por duas variáveis principais (cada) e explicam, respectivamente, 3,888% e 3,007% da variância total.

Tabela 1 – Fatores extraídos e suas comunalidades

Fator	Eigenvalues	% da variância explicada	Cumulatividade (%)
Fator 1	6,197	16,808	16,808
Fator 2	5,109	13,858	30,665
Fator 3	3,437	9,322	39,988
Fator 4	2,729	7,402	47,390
Fator 5	1,945	5,277	52,667
Fator 6	1,625	4,407	57,074
Fator 7	1,569	4,257	61,330
Fator 8	1,434	3,888	65,219
Fator 9	1,316	3,568	68,787
Fator 10	1,249	3,388	72,175
Fator 11	1,137	3,083	75,258
Fator 12	1,109	3,007	78,265

Nota: Variância explicada (*Eigenvalue*): 78,265%

Fonte: Resultados da pesquisa.

De posse desses 12 Fatores inerentes ao saneamento básico, a mensuração do IAQSB em cada município de Minas Gerais é realizada através da ponderação dos escores dos fatores para cada município com a comunalidade das variáveis presentes em cada fator. Os escores fatoriais passam, a partir de então, a ter qualidade estatística suficiente para que sejam utilizados como índices sintéticos das 43 variáveis originais. Assim, atendendo a um dos objetivos da pesquisa e reiterando a fundamentação teórica adotada neste estudo, o IAQSB é calculado da seguinte forma:

$$\begin{aligned}
 IAQSB = & 0,16808P_{01} + 0,13858P_{02} + 0,09322P_{03} + 0,07402P_{04} + 0,05277P_{05} \\
 & + 0,04407P_{06} + 0,04257P_{07} + 0,03888P_{08} + 0,03568P_{09} + 0,03388P_{10} \\
 & + 0,03083P_{11} + 0,03007P_{12}
 \end{aligned} \quad (5)$$

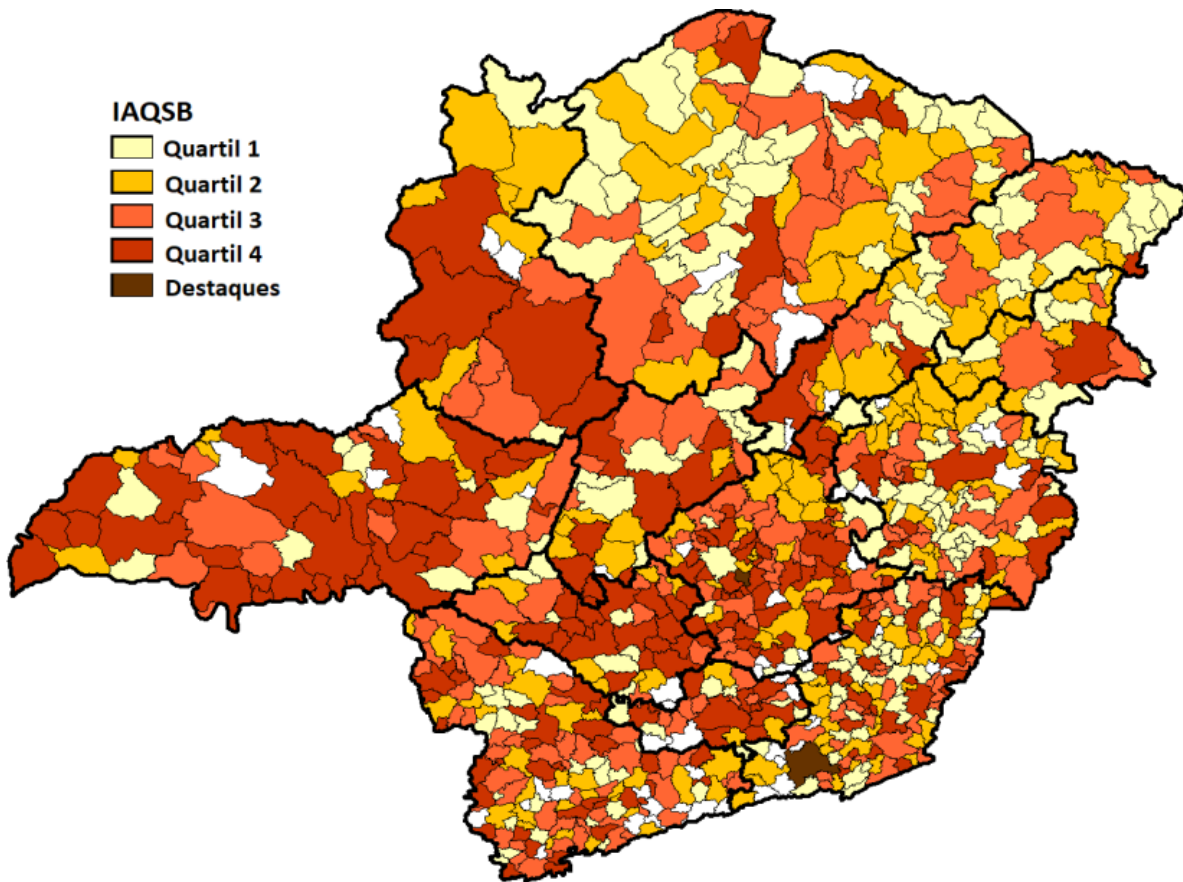
em que P_i são os escores fatoriais e os valores numéricos são os coeficientes de cada escore ponderado que dizem respeito ao percentual da variância explicada por cada fator.

Com o cálculo do IAQSB, é possível mapear o estado de acordo com os resultados obtidos e analisar quais municípios apresentam os melhores e os piores indicadores nos serviços de saneamento básico municipal. Após a mensuração, padronizou-se o valor 10 para o melhor indicador e os demais IAQSB foram calculados e ranqueados com base na padronização realizada.

A Figura 2 apresenta a distribuição dos municípios mineiros em quartis, abarcando 198 municípios cada quartil. O Quartil 1 representa os que obtiveram o indicador IAQSB até 1,65; no Quartil 2, entre o total, estão os municípios que apresentaram o indicador entre 1,66 e 2,01; no Quartil 3 estão os municípios mineiros obtiveram o indicador entre 2,02 e 2,26; e o Quartil 4 representa os que apresentaram o IAQSB entre 2,27 e 5,00. Os municípios de Contagem e Juiz de Fora são tratados como destaques na figura.

É relevante destacar a ocorrência de dois *outliers* que contribuem para o resultado desses respectivos cálculos: os municípios de Contagem e Juiz de Fora. Esses apresentaram os melhores indicadores de acesso e qualidade nos serviços de saneamento básico, com os valores 10 e 5,58, respectivamente, e foram considerados destaques no presente estudo.

Figura 2 – Distribuição espacial dos municípios mineiros para o IAQSB (em quartis)



Fonte: Elaboração própria.

Nos anos que antecederam 2015, Contagem recebeu recursos que somaram mais de R\$280 milhões através de um convênio entre o município, a Companhia de Saneamento de Minas Gerais (COPASA) e os governos Estadual e Federal. Com o montante, obras de tratamento e canalização de esgoto, construção de avenidas sanitárias (canalização de córregos, construção de redes de drenagem pluvial e coletoras de esgoto), urbanização e pavimentação foram feitas na cidade (JORNAL O TEMPO CONTAGEM, 2013).

Já o município de Juiz de Fora, de acordo com a Companhia de Saneamento Municipal (CESAMA, 2014), foi um dos 12 municípios, dentre as 100 maiores cidades brasileiras quanto à elaboração dos Planos Municipais de Saneamento Básico, que atenderam a todos os requisitos da pesquisa do Instituto Trata Brasil acerca da qualidade e integridade do Plano Municipal de Saneamento¹².

Ao analisar os dois melhores IAQSB em Minas Gerais, Contagem e Juiz de Fora, esses dados indicam os resultados de um projeto de infraestrutura de magnitude abrangente nos serviços de saneamento básico nos dois municípios¹³. Segundo os estudos de Scriptorre e Toneto Junior (2012), os quais compararam o desempenho dos provedores públicos e privados dos serviços de saneamento básico no Brasil, não há evidências fortes para que um grupo seja superior ao outro na maior parte dos indicadores; porém, os grupos revelaram superioridade em indicadores específicos. Em relação ao grupo de indicadores de investimento e qualidade, tanto as empresas privadas quanto

¹² A pesquisa avaliou os planos de saneamento básico levando alguns quesitos em consideração, dentre os quais: os componentes, a viabilidade, o conteúdo mínimo e a participação social. O Plano de Saneamento Básico de Juiz de Fora, elaborado entre 2012 e 2014, atendeu a todas as exigências.

¹³ Em relação à estrutura de provisão à qual cada município pertence, sabe-se que Contagem é abastecida pela COPASA, de origem regional, e Juiz de Fora pela CESAMA, de origem local.

as empresas locais públicas apresentaram investimentos inferiores, em comparação ao nível de investimento realizado pelas empresas regionais, como a COPASA.

Contudo, para avaliar o comportamento do investimento, deve-se ponderar uma série de dificuldades. No caso das empresas privadas, esse pode ser fortemente condicionado pela conjuntura econômica, por exemplo. No caso dos provedores públicos, o investimento pode ser fortemente condicionado por variáveis de natureza política: ano eleitoral, coincidência partidária entre governos regionais, locais e federal, entre outros aspectos. Logo, os resultados obtidos para a variável investimento devem ser avaliados com bastante cautela (SCRIPTORE; TONETO JUNIOR, 2012). Dessa forma, esses fatores podem justificar os destaques dados a Juiz de Fora e, principalmente, a Contagem, em detrimento a outros municípios, mesmo localidades com o mesmo operador de saneamento, a COPASA.

Os resultados obtidos para todo o estado apontaram evidências de assimetria quando o assunto é saneamento básico. A Figura 2 permite avaliar os municípios com os melhores e piores números do IAQSB. Nessa perspectiva, pode-se observar que grande parte dos municípios das mesorregiões Noroeste, Centro/Metropolitana, Oeste, Triângulo Mineiro e Campo das Vertentes apresentaram elevados valores do IAQSB; na situação intermediária, estão grande parte dos municípios das regiões Sul/Sudeste, Zona da Mata e Vale do Rio Doce. Na situação oposta, vale destacar os municípios do Norte de Minas, vales do Jequitinhonha e Mucuri, os quais apresentaram os piores valores do indicador social calculado¹⁴. A autocorrelação espacial sugerida pelo mapa da Figura 2 é corroborada por uma análise de dependência espacial *I* de Moran, que apresentou coeficiente significativo de 0,2834.¹⁵ Tal relação espacial positiva indica a existência de um padrão de concentração, em que municípios com determinado indicador de saneamento possuem vizinhos ligeiramente semelhantes (IAQSBs altos “vizinhos” de IAQSBs altos e IAQSBs baixos “vizinhos” de IAQSBs baixos).

Esse resultado encontrado é, ainda, corroborado pelas informações da Pesquisa por Amostra de Domicílios (PAD-MG, 2013), a qual aponta para o Norte de Minas, mostrando que essa é a mesorregião com o menor percentual de domicílios com rede de esgoto ou pluvial (44,4%, enquanto a média do estado é de 73,9%) e 34,3% dos municípios dependem da fossa rudimentar como forma de escoadouro do banheiro. Os vales do Jequitinhonha e Mucuri apresentam 69,5% dos seus domicílios ligados à rede coletora de esgotos; porém, 5,0% dos domicílios jogam seus dejetos diretamente em rio ou lago. Esse número é considerado contraproducente, uma vez que a média de Minas Gerais é de 3,2% e essa é uma das formas de escoadouro mais agressivas tanto para o homem quanto para o meio ambiente. Problemas de saúde causados pelo inadequado escoamento dos dejetos denotam impasse para o desenvolvimento socioeconômico da região.

O cálculo do IAQSB para os municípios mineiros mostra as disparidades das condições de vida nas mesorregiões de Minas Gerais. Ao compará-las com os resultados obtidos, essas disparidades também podem ser percebidas sob a ótica da distribuição da população pelas cidades no estado. De acordo com dados da publicação Perfil de Minas Gerais, em 2010, a média populacional dos municípios mineiros é de aproximadamente 20 mil habitantes. Por outro lado, apenas 20% deles possuem uma população acima desse valor, o que demonstra uma grande concentração populacional em alguns centros urbanos.

A concentração populacional em áreas urbanas faz com que haja investimentos para a expansão e a melhoria da infraestrutura urbana, à medida que aumenta a taxa de urbanização dos municípios. Assim, é importante ressaltar que a concentração populacional está correlacionada com o acesso aos serviços de saneamento básico municipal, pois quanto maior é o aumento da população nos centros urbanos, maior será a demanda por esses serviços, além da maior facilidade de prover esses serviços em aglomerações / concentrações do que em populações (domicílios) dispersos – economias de densidade. Na divisão em quartis ora proposta, percebe-se a predominância de municípios

¹⁴ O mapa de Minas Gerais com a identificação das mesorregiões encontra-se no Anexo C.

¹⁵ O indicador *I* de Moran avalia a significância de agrupamentos espaciais locais ao redor de um dado município e indica focos de autocorrelação espacial. A matriz de contiguidade espacial utilizada valeu-se da convenção binária da rainha de primeira ordem. Para detalhes sobre o método, ver Anselin (1988) e LeSage e Pace (2009).

enquadrados nos quartis 3 e 4, nas mesorregiões Central/Metropolitana e Triângulo Mineiro, apresentando essas relações.

Vale mencionar, ainda, a importância de se analisarem os indicadores sociais nas mesorregiões mineiras. Outra ocorrência é o aumento da parcela da população com níveis mais elevados de educação e, em função disso, com maior acesso a oportunidades e influência política. Rezende, Fernandes e Rodrigues (2007) encontraram evidências de uma relação positiva entre o acesso aos serviços de saneamento e o nível educacional. Os autores argumentaram que um maior nível de educação pode tornar os indivíduos mais conscientes em relação a questões ambientais e de saúde, demandando serviços adequados e tentando influenciar as decisões dos governantes. Nesse sentido, através dos quartis apresentados na Figura 2, pode-se perceber que as consequências positivas entre o acesso aos serviços de saneamento e o nível educacional estão presentes também nos quartis 3 e 4.

De acordo com o Instituto Trata Brasil (2015), crianças sem acesso ou com condições precárias aos serviços de saneamento básico não têm um bom aproveitamento escolar. O estudo realizado pelo instituto também evidenciou que metade das internações por doenças causadas pela precariedade nos serviços de saneamento como diarreia, por exemplo, é de crianças com menos de cinco anos, fase em que a atividade cerebral está sendo desenvolvida, comprometendo a capacidade de aprendizado. Assim, as relações entre educação e condições de saneamento são multidirecionais.

Nesse sentido, pode-se perceber que o IAQSB reflete os precários índices educacionais. As diferenças regionais são marcantes, sendo reconhecidas pela predominância de municípios enquadrados nos dois primeiros quartis para as mesorregiões Norte de Minas, Jequitinhonha e Mucuri.

As informações complementares acima citadas mostraram o panorama das desigualdades regionais em Minas Gerais, reforçado pelos problemas inerentes ao saneamento básico. Essas devem ser consideradas aos propósitos deste estudo que visa tentar identificar os fatores locais que influenciam nos resultados do indicador de acesso e qualidade dos serviços de saneamento. Ademais, na próxima seção, são identificadas e discutidas as variáveis que mais discriminam o IAQSB dentro do contexto do saneamento básico municipal, através da Análise Discriminante.

4.2. Os fatores discriminantes do IAQSB

Os 794 municípios mineiros analisados anteriormente foram ranqueados em termos do valor obtido para o IAQSB. Com exceção dos *outliers*, os municípios de Contagem e Juiz de Fora, os quais obtiveram os valores do IAQSB considerados destaques no presente estudo, os demais 792 municípios foram divididos em quartis, cada um composto por 198 municípios. O primeiro quartil foi formado pelos 198 municípios com os menores valores do IAQSB e assim sucessivamente. Logo, o quarto quartil foi composto pelos 198 municípios com os maiores valores.

A Análise Discriminante foi aplicada no intuito de se verificar quais fatores sociodemográficos mais discriminam o acesso e a qualidade do saneamento básico nos municípios mineiros. De modo específico, essa técnica foi aplicada para a discriminação dos municípios alocados nos quartis 1 e 4, ou seja, os municípios com os piores e melhores níveis de acesso e qualidade dos serviços de saneamento básico, respectivamente.

Deve-se destacar que a função discriminante apresenta boa qualidade de ajustamento, corroborando estatisticamente a capacidade de discriminação desta função. O autovalor, que descreve a capacidade de discriminação da função, foi de 0,602, ao passo que a correlação canônica foi de 0,613. Ademais, o Lambda de Wilks, que é um teste multivariado para a significância global da análise, apresentou um valor de 0,624. O teste qui-quadrado apontou que a função é significativa a 1%, validando a utilização da Análise Discriminante.

A partir do método *stepwise*, foram selecionadas as variáveis com maior capacidade de discriminação. Essas foram, em ordem relevante de capacidade discriminatória, as variáveis selecionadas: i) IDTE: Índice de Desenvolvimento Tributário e Econômico; ii) POBR: percentual de famílias com renda *per capita* de até R\$ 154,00 (linha de pobreza); iii) URBA: taxa de urbanização;

e iv) IFDM: Índice FIRJAN de Desenvolvimento Municipal. A Tabela 2 apresenta estatísticas descritivas das variáveis sociodemográficas com maior capacidade de discriminação, apontando os resultados em termos da subamostra analisada e de cada um dos quartis avaliados, 1 e 4.

Tabela 2 – Estatísticas descritivas das variáveis sociodemográficas mais relevantes dos municípios dos quartis 1 e 4

Variável	Quartis 1 e 4		Quartil 1		Quartil 4	
	Média	D.P.*	Média	D.P.*	Média	D.P.*
IDTE	29,72	18,56	19,89	10,18	39,55	19,82
POBR	33,98	15,81	41,76	15,03	26,19	12,40
URBA	72,03	18,61	62,96	16,88	81,11	15,63
IFDM	0,68	0,09	0,64	0,08	0,72	0,07

Nota: *Desvio-padrão.

Fonte: Resultados da pesquisa.

A análise da Tabela 2 indica que, ao menos em termos médios das variáveis analisadas, os dois quartis apresentam expressivas diferenças. Essa é uma indicação de que essas variáveis podem possivelmente ser boas discriminantes do acesso e da qualidade do saneamento básico nos municípios mineiros.

Em termos médios, os municípios que apresentam um IAQSB elevado também apresentam um nível de desenvolvimento tributário e econômico aproximadamente duas vezes superior ao apresentado pelos municípios que apresentam o IAQSB baixo. Quanto à taxa de urbanização, percebe-se também uma discrepância entre os valores médios dos maiores e dos menores valores do IAQSB analisados. Em média, os municípios com menores índices do IAQSB possuem uma taxa de urbanização de, aproximadamente, 63%, ao passo que essa taxa é de pouco mais de 81% para os municípios com maiores indicadores.

Dando continuidade à análise da estatística descritiva (Tabela 2), enquanto os municípios com menores índices do IAQSB possuem mais de 41% da população na linha de pobreza, essa proporção é pouco maior do que 26% para os municípios que possuem maiores IAQSB. Em consonância com esses resultados, o nível médio de desenvolvimento municipal também foi relativamente distinto entre os grupos, pois as médias do IFDM para os municípios com menores e maiores valores do IAQSB foram, respectivamente, de 0,64 e 0,72.

Visto isso, foi possível especificar a função discriminante para os municípios mineiros, conforme segue:

$$Y = -0,313POBR + 0,344URBA + 0,477IDTE + 0,202IFDM \quad (6)$$

Por se estar analisando dois grupos, apenas uma função discriminante foi gerada. Os coeficientes da função discriminante padronizada indicam a contribuição parcial de cada variável para a função discriminante. Observa-se claramente que, em termos absolutos, a ordem de grandeza dos coeficientes das variáveis respeita a ordem de capacidade discriminatória apontada pelo método *stepwise*.

Entre as quatro variáveis que mais discriminam os municípios com menores e maiores valores do IAQSB, o Índice FIRJAN de Desenvolvimento Municipal (IFDM) e o Índice de Desenvolvimento Econômico e Tributário (IDTE) merecem destaque.

Verifica-se que existe uma relação positiva entre os municípios com altos desempenhos tributários (IDTE) e o IFDM – e dos dois indicadores com a função discriminante. Essa relação pode ser verificada, pois, uma vez apresentados elevados desempenhos tributários, os municípios exibem índices de desenvolvimento socioeconômico maiores do que as outras unidades municipais de desempenho tributários inferiores. No presente estudo, essa diferença é evidenciada nos municípios mineiros que apresentam os menores valores e, portanto, esses apresentam IAQSB

inferior em relação aos municípios que apresentam os valores mais elevados do IDTE e IFDM. Nesse caso, pode-se perceber que as unidades municipais com melhores gestões fiscais conduzem melhorias na qualidade dos serviços públicos considerados essenciais à população, dentre eles o saneamento básico¹⁶.

Quanto à variável discriminada percentual de famílias com renda *per capita* de até R\$ 154,00, percebe-se que há uma relação negativa entre o indicador com a função discriminante. O baixo montante que as famílias recebem inviabiliza o acesso dos mais pobres aos serviços de saneamento, pois as famílias não conseguem arcar com a cobrança de tarifas impostas pelas prestadoras dos serviços, por exemplo. Logo, a capacidade de pagamento mais reduzida faz com que o retorno esperado dos investimentos, pela ótica econômica, não seja o bastante. Portanto, uma renda *per capita* baixa contribui para a desigualdade entre as mesorregiões mineiras e tal diferença também é vista entre os municípios com menores e maiores IAQSB.

Já a taxa de urbanização, não menos importante, tem uma relação positiva com a função discriminante. As relações da população com a taxa de urbanização podem sinalizar a existência de economias de escala e de densidade do setor na provisão dos serviços de saneamento básico, pois quanto maior o número de beneficiários e mais concentrados estiverem, menor será o custo desses serviços. No caso da população, essa relação pode decorrer também do número de contribuintes, que viabilizaria a provisão em função do pagamento de tarifas e de tributos.

Tais resultados vão ao encontro dos estudos e das estatísticas apresentados por Toneto Júnior e Saiani (2006), os quais analisaram o déficit de acesso a serviços de saneamento básico no Brasil e discutiram as principais restrições existentes para a expansão dos investimentos no setor. Os indicadores encontrados pelos autores mostraram que o país possui um significativo déficit de acesso aos serviços de saneamento básico, concentrados principalmente em regiões mais pobres, na zona rural, nos municípios de menor porte, com menor taxa de urbanização, menor renda *per capita* e nos municípios de baixa renda.

Portanto, seguindo as diferenças entre os grupos, conforme anteriormente apontado, a análise da função discriminante aponta que quanto menor a proporção da população na linha da pobreza e quanto maiores o nível de desenvolvimento tributário e econômico, a taxa de urbanização e o nível de desenvolvimento municipal, maior será o indicador de acesso e qualidade dos serviços de saneamento básico.

Indo além, pode-se verificar a capacidade da função discriminante em alocar corretamente os municípios em seus respectivos grupos. Para tanto, utiliza-se a tabela de classificação, na qual é possível mensurar a taxa de sucesso na classificação dos municípios. Esses resultados estão apresentados na Tabela 3.

Tabela 3 – Tabela de classificação da Análise Discriminante

Classificação original	Classificação predita		Total
	Quartil 1	Quartil 4	
Quartil 1	160 (80,81)	38 (19,19)	198 (100,00)
Quartil 4	51 (25,76)	147 (74,24)	198 (100,00)
Total	211 (53,28)	185 (46,72)	396 (100,00)

Nota: Valores percentuais em parênteses.

Fonte: Resultados da pesquisa.

¹⁶ Para consultas mais detalhadas sobre a arrecadação e a gestão fiscal nos municípios mineiros, recomendam-se referências como Fernandes, Brunozi Júnior e Leroy (2018).

A tabela de classificação aponta que o modelo apresenta melhor capacidade preditiva para os municípios que possuem menores valores do IAQSB do que para aqueles com maiores. Enquanto quase 81% dos municípios do Quartil 1 foram corretamente classificados com menores IAQSB, esse percentual foi de pouco mais de 74% para os municípios do IAQSB elevados. Em geral, 77,53% dos 396 municípios foram corretamente alocados para seus grupos, indicando uma boa capacidade de previsão da função discriminante.

Apesar de apenas dois dos quatro quartis terem sido utilizados na Análise Discriminante, o estudo não é inviabilizado, pois essa escolha foi feita de modo a reforçar as diferenças do indicador, IAQSB, criado. Todavia, é preciso destacar que a proximidade de certos indicadores pode ter impactado negativamente o poder preditivo da análise, embora não tenha havido perda de compreensão dos fatores discriminantes.

Os resultados obtidos e analisados através deste estudo evidenciam uma maior responsabilidade na elaboração de políticas públicas de saneamento básico, como forma de assegurar a promoção da saúde e qualidade de vida da população dos municípios mineiros. Contudo, por mais que exista uma nova Política Nacional de Saneamento, o acesso a esses serviços, assim como sua implementação e fiscalização, não é tão simples.

A fiscalização se configura como uma das atividades de suma importância para a agência reguladora. Em Minas Gerais, a Agência Reguladora de Serviços de Abastecimento de Água e Esgotamento Sanitário (ARSAE-MG) foi criada com a finalidade de regular e fiscalizar as atividades da COPASA e da COPANOR¹⁷; além de outras prestadoras de serviços conveniadas. Porém, no relatório da ARSAE-MG, elaborado por Carmo (2015), a agência apresentou um resultado tímido na fiscalização: 254 sistemas em 164 municípios mineiros (22,67% de abrangência)¹⁸. Partindo dessas informações, apenas oito municípios tiveram fiscalizações mais de uma vez, seja para ações de acompanhamento, seja para complementação de fiscalização ou demandadas pelo Ministério Público.¹⁹ Além do mais, essas fiscalizações não contemplavam todas as localidades previstas no contrato do programa.

No aspecto regulatório, há duas razões para se regular uma empresa: a primeira remete à correção das falhas de mercado (natureza monopólica); e a segunda se dá pela garantia do interesse público na prestação de serviços públicos de qualidade (GALVÃO JUNIOR; SILVA; QUEIROZ, 2006). No caso do saneamento básico, os envolvidos são o município, a prestadora de serviço e a população consumidora. A ARSAE-MG disponibiliza 56 Resoluções Normativas em seu sítio eletrônico. De acordo com o relatório elaborado por Carmo (2015), ao todo, 35 resoluções são referentes a “reajuste/revisão tarifária e prazos”, o que mostrou que a ARSAE atuou com foco muito mais no viés econômico do que no viés normalizador. A longo prazo, essa relação dificulta o processo de aprimoramento da gestão dos serviços de abastecimento de água e esgotamento sanitário.

O IAQSB obtido no presente estudo está suscetível às diferenças regulatórias e o resultado corroborou para a existência de expressivas desigualdades entre os municípios mineiros. Apesar do intuito de contribuir para a regulação e fiscalização dos serviços, a abrangência da ARSAE-MG ainda é limitada. Como consequência, por exemplo, os municípios com maiores índices de doenças de veiculação hídrica não são fiscalizados e continuam em péssimas condições sanitárias. Para tal, é de suma importância a intervenção estatal, não só interferindo na competência municipal, mas também tratando o saneamento básico como interesse comum e, portanto, de competência estadual e regional.

As variáveis socioeconômicas no estudo não são independentes e interagem de forma imbricada na determinação do quadro geral de necessidades em saneamento dos municípios. Essa constatação requer ações coordenadas nas diversas frentes de ação governamental.

¹⁷ COPANOR – Copasa Serviços de Saneamento Integrado do Norte e Nordeste de Minas Gerais.

¹⁸ Considera-se que o universo de fiscalização da agência compreende 635 municípios com concessão da COPASA/COPANOR, mais os municípios conveniados, Itabira, Passos, Juiz de Fora e Ituiutaba, totalizando 639.

¹⁹ Segundo Carmo (2015), há três formas de acionar a fiscalização desses serviços: via denúncia na ouvidoria especializada, através da solicitação do Ministério Público ou outros órgãos de controle e da fiscalização programada.

Nesse sentido, os resultados sugerem que a simples ampliação da oferta com qualidade nos serviços de saneamento básico é uma ação importante, porém coadjuvante de uma política eficaz de promoção do saneamento à população municipal. Articula-se que, diante da gestão dos serviços de saneamento básico não só em Minas Gerais, como também nos demais estados do Brasil, a universalização dos serviços constitui o maior dos desafios que se impõem às políticas públicas aplicáveis ao setor.

5. Considerações finais

O saneamento básico é um serviço que tem por objetivo principal a manutenção da vida com qualidade, através do abastecimento de água potável, do desenvolvimento de soluções ao esgotamento sanitário com sua coleta e tratamento, e o manejo e a destinação correta de resíduos sólidos.

A partir do novo Plano Nacional de Saneamento (BRASIL, 2007), com a universalização do acesso ao saneamento no Brasil, e do novo Marco Regulatório, a institucionalização do uso de indicadores de desempenho passou a integrar o processo de planejamento, regulação e fiscalização desses serviços. No intuito de auxiliar a elaboração de políticas públicas pró-equitativas no estado de Minas Gerais, a presente pesquisa objetivou a elaboração de um indicador de acesso aos serviços de saneamento básico, além de identificar os fatores locais, demográficos e socioeconômicos das condições de acesso e qualidade desses serviços no estado.

Para identificar quais municípios apresentam os melhores e os piores indicadores, o presente estudo elaborou, através da técnica multivariada da Análise Fatorial, um indicador de acesso e qualidade do saneamento básico (IAQSB). O IAQSB foi capaz de sintetizar o conjunto de variáveis direcionadas às questões de acesso e qualidade dos serviços de saneamento básico municipal com os dados de 2015 obtidos através do Sistema Nacional de Informações sobre Saneamento (SNIS).

A amostra foi realizada com 794 municípios de Minas Gerais e o IAQSB permitiu também avaliar os possíveis fatores locais que culminaram nas desigualdades e necessidades do acesso ao saneamento básico entre os municípios mineiros. Em seguida, utilizou-se a técnica multivariada de Análise Discriminante a fim de se verificar quais foram os principais determinantes que diferenciaram os grupos de municípios de Minas Gerais em relação à condição de saneamento básico.

Os resultados obtidos pela mensuração do IAQSB permitiram mapear o estado (em que os municípios analisados foram divididos em quartis) e identificar quais municípios apresentaram os melhores e os piores indicadores nos serviços de saneamento básico municipal. No estudo, foi relevante destacar a ocorrência de dois *outliers*: os municípios de Contagem e Juiz de Fora, que apresentaram os melhores indicadores de acesso e qualidade nos serviços de saneamento básico, com os valores 10 e 5,58, respectivamente.

A análise dos IAQSBs obtidos evidencia a existência de expressivas desigualdades entre os municípios mineiros, sugerindo cenário de iniquidades. No entanto, as desigualdades são mais expressivas em desfavor das regiões Norte, vales do Jequitinhonha e Mucuri, podendo observar que grande parte dos municípios dessas mesorregiões apresentou os piores valores do indicador social calculado. Na situação oposta, grande parte dos municípios das regiões Noroeste, Centro/Metropolitana, Oeste, Triângulo Mineiro e Campos Vertentes apresentou elevados valores do IAQSB, enquanto que na situação intermediária estão grande parte dos municípios das regiões Sul/Sudeste, Zona da Mata e Vale do Rio Doce.

Em seguida, a análise dos fatores sociodemográficos que mais discriminam o IAQSB foi aplicada para os municípios com os piores e os melhores níveis de acesso e qualidade dos serviços de saneamento básico. O resultado apontou para a existência de quatro variáveis com capacidade de discriminação, a saber: Índice de Desenvolvimento Tributário e Econômico (IDTE); o percentual de famílias com renda *per capita* de até R\$ 154,00; a taxa de urbanização; e, por fim, o Índice FIRJAN de Desenvolvimento Municipal (IFDM).

Os resultados evidenciam que quanto menor a proporção da população na linha da pobreza e quanto maiores o nível de desenvolvimento tributário e econômico, a taxa de urbanização e o nível de desenvolvimento municipal, maior será o indicador de acesso e qualidade dos serviços de saneamento básico. Esses resultados convergem para estudos e estatísticas citados e com as análises locais realizadas no presente estudo.

Assim, todos os resultados obtidos e analisados no presente estudo fortalecem o argumento de que a efetividade das políticas públicas para a melhoria na qualidade nos serviços de saneamento básico, assim como a redução nas desigualdades de acesso dos mesmos à população em requer ações coordenadas nas diversas frentes de ação governamental.

A efetividade de tais políticas públicas que podem impactar no desenvolvimento ambiental, urbano e social para a melhoria nas condições dos municípios mineiros depende do aumento e da melhoria dos serviços de saneamento, saúde, do ensino e de ações que exigem a coparticipação das demais esferas de Governo, de maneira a ampliar, por exemplo, a habitação e a renda média domiciliar das populações dos municípios mineiros.

Também é necessário o fortalecimento da ARSAE-MG no sentido de ampliar a sua capacidade de fiscalização e regulação para assegurar sua autonomia e independência, buscando a excelência na regulação dos serviços e contribuindo para a prestação dos serviços públicos com qualidade. Portanto, a retomada de investimentos no setor em Minas Gerais requer a superação das indefinições regulatórias, a criação de mecanismos de financiamento adequados e o aumento da eficiência do setor.

Apesar de os resultados encontrados possibilitarem a avaliação e condições de acesso aos serviços de saneamento básico por parte dos municípios mineiros, é importante ressaltar algumas limitações deste estudo. Pode-se citar, por exemplo, a ausência de dados referentes a outros fatores que poderiam exercer influência no indicador criado, uma vez que alguns indicadores na plataforma de dados estavam incompletos. Ademais, dos 853 municípios de Minas Gerais, 7% deixou de responder aos questionários do Sistema Nacional de Informações por parte das operadoras de saneamento, deixando de ser analisados no IAQSB. A respeito do IAQSB mensurado, é preciso destacar que a proximidade de certos indicadores pode ter impactado negativamente o poder preditivo da análise, embora não tenha havido perda de compreensão dos fatores discriminantes.

Os resultados obtidos na aplicação deste estudo contribuíram para fomentar a literatura em torno deste tema, especialmente, para corroborar a literatura concernente aos determinantes sociais de saneamento básico nos municípios mineiros. A pesquisa poderá ser utilizada também como instrumento de auxílio nos processos de gestão nas esferas estadual e federal em estudos futuros. Do ponto de vista metodológico, o método pode ser aplicado ao contexto nacional, evidenciando as disparidades regionais e, alternativamente, pode-se comparar o acesso ao saneamento básico e os respectivos determinantes no âmbito regional.

Referências

- ANSELIN, L. *Spatial econometrics: methods and models*. Dordrecht: Kluwer Academic Publishers, 1988.
- BRASIL. *Lei nº 11.445 de 05/01/2007*. Estabelece diretrizes nacionais para o saneamento básico e para a política federal de saneamento básico. Brasil, 2007.
- BORJA, P. C. Política pública de saneamento básico: uma análise da recente experiência brasileira. *Saúde e Sociedade*, v. 23, n. 2, p. 432-447, 2014.
- BRITTO, A. L.; REZENDE, S. C. A política pública para os serviços urbanos de abastecimento de água e esgotamento sanitário no Brasil: financeirização, mercantilização e perspectivas de resistência. *Cadernos Metrópole*, v. 19, n. 39, p. 557-581, 2017.
- CARMO, D. D. Os avanços da regulação e fiscalização dos serviços de saneamento abastecimento de água e esgotamento sanitário de Minas Gerais: estudo de caso da ARSAE-MG. In: Congresso

- do Conselho Nacional de Secretários de Estado da Administração, 8, 2015. *Anais...* Brasília: CONSAD, 2015.
- COPASA – Companhia de Saneamento de Minas Gerais. *A História da Copasa*. 2007. Disponível em: <http://www.copasa.com.br>. Acesso em 15 de abr. de 2017.
- CESAMA – Companhia de Saneamento Municipal. *Notícias Cesama*, 2014. Disponível em: <https://goo.gl/ML19XF>. Acesso em 10 nov. 2017.
- COSTA, S. A. B.; CORTÊS, L. S.; COELHO, T.; FREITAS JR, M. M. Indicadores em Saneamento: avaliação da prestação dos serviços de água e de esgoto em Minas Gerais. *Revista da Universidade Federal de Minas Gerais*, v. 20, n. 2, p. 334-357, 2013.
- FERNANDES, C. M.; BRUNOZI JÚNIOR, A. C.; LEROY, R. S. D. Desempenho Tributário e Desenvolvimento Econômico e Social em Municípios Mineiros. *Perspectivas Contemporâneas*, v. 13, n. 1, p. 74-93, 2018.
- FEAM - Fundação Estadual do Meio Ambiente. *Panorama da destinação dos resíduos sólidos urbanos no Estado de Minas Gerais em 2015*. Belo Horizonte: FEAM, 2016. 63p.
- FJP – Fundação João Pinheiro. *Perfil de Minas Gerais*. Relatório elaborado pelo Centro de Estatística e Informações (CEI) com a colaboração do Centro de Políticas Públicas (CEPP). Belo Horizonte: FJP, 2010.
- GALVÃO JÚNIOR, A. C.; SILVA, A. D.; QUEIROZ, E. A. D. *Regulação: Procedimentos de Fiscalização em Sistemas de Abastecimento de Água*. Fortaleza: Expressão Gráfica Ltda./Arce, 2006.
- GUIMARÃES A. J. A.; CARVALHO D. F.; SILVA L. D. B. *Saneamento e meio ambiente*. 2007. Disponível em: <https://goo.gl/b5zcKD>. Acesso em 06 abr. 2017.
- HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C.; BALBIN, B. J. *Análise Multivariada de Dados*. 6 ed. Porto Alegre: Bookman, 2009. 688p.
- HELLER, L. Relação entre saúde e saneamento na perspectiva do desenvolvimento. *Ciência & Saúde Coletiva*, v. 3, n. 2, p. 73-84, 1998.
- HELLER, L.; COUTINHO, M. L.; MINGOTI, S. A. Diferentes modelos de gestão de serviços de saneamento produzem os mesmos resultados? Um estudo comparativo em Minas Gerais com base em indicadores. *Engenharia Sanitária e Ambiental*, v. 11, n. 4, p. 325-336, 2006.
- HELLER, P. G. B.; VON SPERLING, M.; HELLER, L. Desempenho tecnológico dos serviços de abastecimento de água e esgotamento sanitário em quatro municípios de Minas Gerais: uma análise comparativa. *Engenharia Sanitária e Ambiental*, v. 14, n. 1, p. 109-118, 2009.
- IMRS – Índice Mineiro de Responsabilidade Social. *Saneamento e Habitação*, 2015. Disponível em: <http://imrs.fjp.mg.gov.br>. Acesso em: 14 mai. 2017.
- IFDM – Índice Firjan de Desenvolvimento Municipal. *Dados*, 2015. Disponível em: <https://www.firjan.com.br/ifdm/>. Acesso em: 14 mai. 2017.
- IBGE – Instituto Brasileiro de Geografia e Estatística. *Um panorama da saúde no Brasil: acesso de utilização dos serviços e condições de saúde e fatores de risco e proteção à saúde, 2008*. Pesquisa Nacional por Amostra de Domicílios. Rio de Janeiro: IBGE, 2010.
- IBGE – Instituto Brasileiro de Geografia e Estatística. *Resultados do universo do Censo Demográfico 2010*. 2012. Disponível em: <https://goo.gl/zYN2YQ>. Acesso em: 28 ago. 2018.

- IGTEC – Instituto de Geoinformação e Tecnologia. *Mapa Geopolítico de Minas Gerais - Mesorregiões*. Disponível em: <https://www.iga.br/mapas/>. Acesso em: 20 ago. 2017.
- ITB – Instituto Trata Brasil. *Saneamento no Brasil*. 2015. Disponível em: <http://www.tratabrasil.org.br/saneamento-no-brasil>. Acesso em: 06 abr. 2017.
- JANNUZZI, P. M. *Indicadores sociais no Brasil*. Campinas: Alinea, 2003.
- JANNUZZI, P. M. Considerações sobre o uso, mau uso e abuso dos indicadores sociais na formulação e avaliação de políticas públicas municipais. *Revista de Administração Pública*, v. 36, n. 1, p. 51-72, 2002.
- JORNAL O TEMPO CONTAGEM. *Saneamento básico recebe investimentos*. 2013. Disponível em: <https://goo.gl/RxFwuv>. Acesso em: 10 nov. 2017.
- LESAGE, J.; PACE, R. K. *Introduction to Spatial Econometrics*. 1 ed. New York: Chapman and Hall/CRC, 2009.
- LIMA, S. C. R. B.; MARQUES, D. H. F. *Evolução e perspectivas do abastecimento de água e do esgotamento sanitário no Brasil*. Brasília: CEPAL/IPEA, 2012. (Texto para Discussão, n. 47).
- MALHOTRA, N. K. *Pesquisa de marketing: uma orientação aplicada*. 6 ed. Porto Alegre: Bookman, 2011.
- MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG, 2005.
- MIRANDA, R. J. S. *A Eficiência dos serviços de saneamento e o impacto na geração de emprego e renda em Minas Gerais*. Dissertação (Mestrado em Economia). Programa de Mestrado em Economia da Universidade Federal de Viçosa. Viçosa, 2015.
- OLIVEIRA, J. M.; LAGES, A. M. G.; DANTAS, N. F. Indicadores de desenvolvimento: uma resenha em construção. *Revista de Economia Mackenzie*, v. 8, n. 1, p. 76-101, 2010.
- PONCIANO, F. Análise dinâmica da eficiência técnica dos serviços de água e esgotos no Brasil após a Lei de Saneamento. In: Congresso Nacional de Saneamento e Meio Ambiente, 29, 2017. *Anais...* Rio de Janeiro: ABES, 2017
- QUEIROZ, B. L.; GOLGHER, A. B.; AMARAL, E. Mudanças demográficas e condições econômicas e sociais em Minas Gerais. In: OLIVEIRA, F. A.; SIQUEIRA, W. B. (Orgs.). *As muitas Minas: ensaios sobre a economia mineira*. 1 ed. Belo Horizonte: Conselho Regional de Economia, 2010.
- REZENDE, M. L.; FERNANDES L. P. S.; RODRIGUES, A. M. Utilização da análise fatorial para determinar o potencial de crescimento econômico em uma região do sudeste do Brasil. *Economia e Desenvolvimento*, v. 19, n. 1, p. 92-109, 2007.
- ROCHA, M. S. B.; MATTOS, E. H. C.; SAIANI, C. C. S. Descentralização e provisão de serviços públicos: evidências a partir da criação dos municípios brasileiros no setor de saneamento básico. *Pesquisa e Planejamento Econômico*, v. 47, n. 1, p. 105-150, 2017.
- ROSSONI, H. A. V.; BRIZON, E.; CHAIB, A.; SANTOS, L. A.; HELLER, L.; LIMA, S. C. R. B. Avaliação da política estadual de saneamento para a Região Norte e Nordeste do Estado de Minas Gerais: uma análise do Projeto Vida no Vale. *Revista de Políticas Públicas*, v. 18, n. 1, p. 241-254, 2014.
- ROUQUAYROL, M. Z. *Epidemiologia e Saúde*. 4 ed. Rio de Janeiro: Editora MEDSI, 1993.

SANTOS, F. O. *Saneamento básico no Brasil: avaliação dos fatores determinantes do investimento com base nos dados da pesquisa do Sistema Nacional de Informações sobre Saneamento (SNIS)*. Dissertação (Mestrado em Administração Pública). Fundação Getúlio Vargas. Rio de Janeiro: FGV, 2016.

SCRIPTORE, J. S.; TONETO JUNIOR, R. A estrutura de provisão dos serviços de saneamento básico no Brasil: uma análise comparativa do desempenho dos provedores públicos e privados. *Revista de Administração Pública*, v. 46, n. 6, p. 1479-1504, 2012.

SHARMA, S. *Applied multivariate techniques*. New York: John Wiley & Sons, 1996.

SILVEIRA, R. B.; HELLER, L.; REZENDE, S. C. Identificando correntes teóricas de planejamento: uma avaliação do Plano Nacional de Saneamento Básico (Plansab). *Revista de Administração Pública*, v. 47, n. 3, p. 601-622, 2013.

SNIS – Sistema Nacional de Informações sobre Saneamento. *Dados – Série histórica*. Disponível em: <http://www.snis.gov.br/>. Ministério do Desenvolvimento Regional, Secretaria Nacional de Saneamento, 2015.

TONETO JÚNIOR, R.; SAIANI, C. C. S. Restrições à expansão dos investimentos no saneamento básico brasileiro. *Revista Econômica do Nordeste*, v. 37, n. 4, p. 572-591, 2006.


VON SPERLING, T. L.; VON SPERLING M. Proposição de um sistema de indicadores de desempenho para avaliação da qualidade dos serviços de esgotamento sanitário. *Revista de Engenharia Sanitária e Ambiental*, v. 18, n. 4, p. 313-322, 2013.

WHO – World Health Organization. *Water, sanitation and hygiene links to health: facts and figures*. Geneva: WHO, 2004.

ORCID

Juliana Brito de Oliveira  <https://orcid.org/0000-0002-2615-5540>

Gabriel Teixeira Ervilha  <https://orcid.org/0000-0002-0952-3663>

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

Anexo**Anexo A****Quadro A.1 – Especificação dos fatores, identificados pela Análise Fatorial***Continua*

FATOR 1:	
RS003	Veículos para coleta de RDO e RPU (veículo/mil hab. atend. RS)
RS002	Quantidade <i>per capita</i> anual de RCC (ton/ano/ind atend. RS)
RS004	Despesa <i>per capita</i> com serviços de manejo de RSU (R\$/ano/ind atend RS)
RS005	Receita <i>per capita</i> arrecadada referentes à gestão e ao manejo de RSU (R\$/ano/ind atend RS)
RS006	Quantidade de RDO e RPU coletada por todos os agentes (ton/ano/mil ind atend. RS)
RS007	Quantidade de materiais recicláveis recuperados (ton/ano/mil ind atend RS)
FATOR 2:	
RS001	Trabalhadores alocados no serviço de coleta e manejo de RDO e RSU, varrição, roçada e capina (trab./mil hab.)
DC001	Despesa Corrente <i>per capita</i> da Prefeitura durante o ano com TODOS os serviços do município (R\$/ano/hab.)
RS008	População urbana atendida com o serviço regular de coleta seletiva (%)
RS009	Percentual da população atendida com frequência diária, 2 a 3 vezes por semanas e 1 vez por semana (% pond.)
RS010	População urbana atendida pelo serviço de coleta domiciliar direta, ou seja, porta a porta (%)
RS011	População total atendida por coleta de lixo no município (%)
FATOR 3:	
AG001	População total atendida com abastecimento de água (%)
ES001	População total atendida com esgotamento sanitário (%)
ES006	População urbana atendida com esgotamento sanitário (%)
QD001	Quantidade de reclamações ou solicitações de serviços (Reclam/ano/lig A E)
IN_AE009	Duração média dos serviços executados (H/serv.)
FATOR 4:	
ES002	Extensão da rede de esgoto por ligação (m/lig.)
ES003	Receita operacional direta de esgoto por ligação
ES005	Volume de esgotos tratado (%)
ES007	Tarifa média de esgoto (R\$/m ³)
ES008	Índice de consumo de energia elétrica em sistemas de esgotamento sanitário (kWh/m ³)
FATOR 5:	
AG002	Volume de água consumido que é tratado em ETAs (%)
AG004	Receita operacional direta de água por ligação
AG007	Tarifa média de água (R\$/m ³)
IN_AE003	Indicador de desempenho financeiro (%)
FATOR 6:	
AG003	Extensão da rede de água por ligação (m/lig.)
IN_AE002	Despesa total com os serviços por m ³ faturado (R\$/m ³)
AG008	Consumo médio <i>per capita</i> de água (L/hab./dia)
AG009	Índice de consumo de água (%)

Quadro A.1 – Especificação dos fatores, identificados pela Análise Fatorial

Conclusão

FATOR 7	
IN_AE005	Economias atingidas por paralisações (eco/par)
IN_AE006	Duração média das paralisações (H/par)
IN_AE008	Economias atingidas por intermitências (eco/interrup.)
FATOR 8:	
IN_AE001	Densidade de economias de água por ligação (eco/lig)
RS012	Autossuficiência financeira da prefeitura com o manejo de RSU (%)
FATOR 9:	
IN_AE007	Índice de produtividade: economias ativas por pessoal próprio
IN_AE004	Índice de produtividade: empregados próprios por 1000 ligações de água e esgoto (empreg./mil lig.)
FATOR 10	
AG006	Investimento <i>per capita</i> realizado em abastecimento de água
ES004	Investimento <i>per capita</i> realizado em esgotamento sanitário
SB001	Investimento <i>per capita</i> realizado com saneamento básico geral
FATOR 11:	
AG010	Índice de consumo de energia elétrica em sistemas de abastecimento de água (kWh/m ³)
FATOR 12:	
AG005	População urbana atendida com abastecimento de água (%)
RS013	Taxa de recuperação de materiais recicláveis em relação à quantidade de resíduos coletados(%)

RSU: Resíduos Sólidos Urbanos; RCC: Resíduos Sólidos da Construção Civil; RDO: Resíduos Domiciliares; RPU: Resíduos Públicos; ETA: Estação de Tratamento de Água.

Fonte: Resultados da pesquisa.

Anexo B**Quadro B.1 – Variáveis pré-selecionadas para a discriminação dos municípios com maiores e menores IAQSB**

Área (A1)	Área do município em km ²
População Total (P1)	População total (estimadas ajustadas) do ano de 2014
Densidade populacional (P2)	Densidade populacional (hab/km ²) do ano de 2014
Taxa de urbanização (P3)	Taxa de urbanização dos municípios do ano de 2014
Pobreza (R1)	Percentual da população residente com renda domiciliar mensal <i>per capita</i> até meio salário mínimo
Extrema Pobreza (R2)	Percentual da população residente com renda domiciliar mensal <i>per capita</i> até R\$ 154,00 reais mensais
Receita corrente (R3)	Receita corrente líquida <i>per capita</i> dos municípios do ano de 2013
IDTE (T1)	Índice de Desenvolvimento Tributário e Econômico dos municípios do ano de 2013
Endividamento (ED1)	Participação da dívida consolidada líquida na receita corrente líquida do ano de 2013
Gasto com infraestrutura (G1)	Gasto com infraestrutura urbana em geral, serviços urbanos, transportes em geral, comunicações postais e telecomunicações por município no ano de 2013
Esforço Orçamentário em infraestrutura (G2)	Esforço Orçamentário em infraestrutura dos municípios no ano de 2013
Taxa de emprego no setor formal (TE1)	Número de empregados no setor formal dos municípios no ano de 2013
Taxa bruta de mortalidade padronizada (M1)	Razão entre os óbitos esperados por faixa etária e a população no ano de 2013
Taxa de crianças que frequentam a escola (E1)	Taxa de crianças que frequentam o ensino fundamental na série adequada no ano de 2014
Taxa de adolescentes que frequentam a escola (E2)	Taxa de adolescentes que frequentam o ensino médio na série adequada no ano de 2014
Escolarização líquida do ensino fundamental (E3)	Taxa de Escolarização Líquida do ensino fundamental no ano de 2014
Escolarização líquida do ensino médio (E4)	Taxa de Escolarização Líquida do ensino médio no ano de 2014
IFDM Saúde (IFDM_1)	Compreende o levantamento dos indicadores de saúde, como número de consultas, pré-natal, óbitos em geral e internações em geral, no ano de 2014
IFDM Emprego & Renda (IFDM_2)	Compreende o levantamento dos indicadores de emprego e renda, como o número de empregados formais, absorção da mão de obra local, geração de renda, salários médios, do emprego, etc., para o ano de 2014
IFDM Educação (IFDM_3)	Compreende o levantamento dos indicadores de educação, como matrículas na educação infantil e os seguintes dados sobre o ensino fundamental: abandono, distorção idade/série, docentes com ensino superior, médias hora/aula, entre outros, no ano de 2014
IFDM (IFDM_4)	Compreende o levantamento dos três indicadores juntos no ano de 2014
Índice de Gini (IG)	Índice de Gini dos municípios no ano de 2010
Rendimento <i>per capita</i> no setor formal (TE2)	Valor total do rendimento dos empregados do setor formal no ano de 2013

Fonte: Elaboração própria, baseado na base de dados do IMRS, IFDM e IBGE.

Anexo C

Figura A.1 – Mesorregiões de Minas Gerais



Fonte: Instituto de Geoinformação e Tecnologia – IGTec (2017).

**EFEITOS DE LONGO PRAZO DO AJUSTE FISCAL SOBRE A ECONOMIA
SERGIPANA***

Luiz Carlos de Santana Ribeiro

Doutor em Economia – CEDEPLAR/UFMG. Professor Adjunto do Departamento de Economia da Universidade Federal de Sergipe. Bolsista de Produtividade em Pesquisa do CNPq e Coordenador do Laboratório de Economia Aplicada e Desenvolvimento Regional da UFS - LEADER
E-mail: ribeiro.luiz84@gmail.com

Kênia Barreiro de Souza

Doutora em Economia – CEDEPLAR/UFMG. Professora Adjunta no Departamento de Economia da Universidade Federal do Paraná. Bolsista de Produtividade em Pesquisa do CNPq e Pesquisadora do Núcleo de Estudos em Desenvolvimento Urbano e Regional da UFPR - NEDUR
E-mail: keniadesouza@gmail.com

RESUMO: O objetivo deste artigo é estimar os efeitos de longo prazo do ajuste fiscal, conduzido pelo Governo Federal, sobre a economia sergipana. Para tanto, são utilizadas simulações com um modelo dinâmico de Equilíbrio Geral Computável, ano base 2013, constituído de 27 estados e 67 setores. Nessa abordagem, considera-se que os investimentos privados reagem apenas a alterações nas taxas de retorno do capital e não a mudanças nas expectativas. Os principais resultados indicam que o ajuste fiscal atenuaria, no ano de 2035, o crescimento acumulado do PIB real de Sergipe em 12% em relação ao cenário de referência. Além disso, a nível municipal, os resultados são bastante heterogêneos em decorrência dos diferentes níveis de dependência dos gastos públicos.

Palavras-chave: Ajuste fiscal; Simulações em EGC; Economia sergipana.

Classificação JEL: C68; E62; R15.

LONG-TERM EFFECTS OF TAX ADJUSTMENT ON THE SERGIPE'S ECONOMY

ABSTRACT: This paper aims to estimate the long-term effects of the fiscal cut, conducted by the federal government, on the Sergipe economy. To do so, we use simulations with a dynamic Computable General Equilibrium model, base year 2013, built for the 27 Brazilian states and 67 sectors. In this approach, it is considered that private investments react only to changes in rates of return on capital rather than changes in expectations. The results indicate that the fiscal adjustment attenuated, in 2035, the accumulated growth of Sergipe's real GDP by 12% in relation to the baseline scenario. Furthermore, at municipality level, the results are quite heterogeneous due to the different levels of dependence on public expenditures.

Keywords: Fiscal adjustment; Simulations in EGC; Sergipe economy.

JEL Codes: C68; E62; R15.

1. Introdução

O Brasil vivenciou entre 2015-2016 a pior e mais longa crise econômica da sua história, com retração do PIB de 3,77% e 3,62%, respectivamente. Não obstante, a crise política agravou ainda mais a situação econômica do país. Como uma das medidas adotadas para o combate à crise, o Governo Federal anunciou um novo regime fiscal, conhecido como a Proposta de Emenda Constitucional (PEC) dos Gastos ou o Teto dos Gastos Públicos. Grosso modo, de acordo com esse regime, os gastos e investimentos públicos ficam limitados aos mesmos valores do ano anterior, corrigidos pela inflação, por até 20 anos (BRASIL, 2016). Com efeito, esse ajuste fiscal provocará impactos heterogêneos sobre as economias estaduais no curto, médio e longo prazo.

A forte queda dos gastos em formação bruta de capital fixo em 2014, para Oreiro (2017), foi o estopim da crise, embora esse autor destaque dois outros fatores como amplificadores da recessão: i) ajuste abrupto de preços administrados; e ii) o comportamento pró-ciclo da política fiscal entre 2003-2015.

A crise provocou piores resultados para a economia sergipana quando comparada à brasileira. A variação real do PIB brasileiro, em 2016, foi de -3,6%, enquanto a variação de Sergipe foi de -4,6%. Além disso, a taxa de desemprego do Brasil aumentou de 9% para 12%, comparando os últimos trimestres de 2015 e 2016, e a taxa sergipana aumentou de 9,9% para 15% no mesmo período. O principal fator de crescimento da economia sergipana esteve associado ao longo da história à produção de *commodities* minerais, com destaque para o petróleo. No entanto, os investimentos da Petrobrás anunciados no Plano de Negócios da empresa (2013-2018) para o estado, os quais somavam R\$ 42 bilhões, não se concretizaram. Em relação à indústria de transformação, Sergipe tradicionalmente concentra sua produção em segmentos de baixo conteúdo tecnológico como, por exemplo, alimentos e bebidas, têxtil e calçados (CAMARGOS et al., 2017).

Nesse contexto, o objetivo deste artigo é estimar os efeitos de longo prazo do ajuste fiscal, conduzido pelo Governo Federal, sobre as economias regionais e setores econômicos, com foco especial na economia sergipana. Em outras palavras, busca-se isolar o efeito do corte de gastos federal sobre Sergipe, assumindo que nenhuma outra medida de recuperação econômica seja realizada simultaneamente. Para tanto, são utilizadas simulações com um modelo dinâmico e inter-regional de Equilíbrio Geral Computável (EGC), ano base 2013, constituído de 27 estados e 67 setores econômicos. A principal contribuição do artigo é identificar os efeitos indiretos que mapeiam as conexões dos setores públicos com o restante da economia. Mais do que isso, alguns municípios sergipanos seriam mais influenciados do que outros, porque dependem em maior proporção dos gastos públicos.

O artigo que segue, além desta introdução, é composto de mais quatro seções. A próxima seção traz o debate econômico sobre política fiscal e gasto público que permeia a ementa de ajuste fiscal do Governo Federal. A terceira seção descreve o modelo utilizado de EGC e sua base de dados. A quarta seção apresenta as estratégias de simulações e discute os resultados, e é seguida pelas considerações finais.

2. O debate econômico da política fiscal e do gasto público e os mecanismos causais em um modelo de EGC

A visão dominante sobre a política fiscal, que prevaleceu na condução da política econômica no Brasil no período pós plano real, fundamenta-se, segundo Montes e Alves (2012), na condução da política econômica de forma a gerar expressivos superávits primários e fazer frente à dívida pública. Nesse sentido, haveria a necessidade de um permanente equilíbrio orçamentário intertemporal do Governo, o que promoveria uma “economia saudável”. Tal argumento, segundo os autores, apoia-se na noção de que o déficit público é um mal em si, ao desencadear processos inflacionários.

Essa visão está associada a concepção neoclássica que identifica na criação de moeda para o financiamento de gastos públicos a principal causa do processo inflacionário. Nessa perspectiva, a

monetização do déficit público não teria qualquer efeito sobre a economia real e o equilíbrio de longo prazo, mas apenas sobre os preços e a inflação. Ou seja, o viés inflacionário seria resultado do uso de emissões monetárias para financiar o déficit público. Assim, o equilíbrio orçamentário seria prerrogativa para a estabilidade de preços (MOLLO, 2004; GOBETTI, 2008; MONTES e ALVES, 2012).

Essa abordagem surgiu em resposta a um período de gastos elevados e de desequilíbrio orçamentário que se instaurou principalmente a partir da década de 1950, com o “estado de bem-estar social”. Nesse contexto, Friedman (1953) sugere que as políticas econômicas não devem ter caráter discricionário (i.e. formuladas para situações específicas) e propõe que a autoridade monetária adote regras para a gestão da moeda. Essas regras deveriam ser pré-anunciadas, evitando, segundo o autor, perturbações amplas e irregulares na economia. Assim, a melhor forma de lidar tanto com a política fiscal quanto com a política monetária seria criar processos automáticos e pré-estabelecidos, com a manutenção de um orçamento sempre equilibrado independente do ciclo (GOBETTI, 2008).

Neste sentido, Sargent e Wallace (1981) abordam a questão da interação entre as políticas fiscal e monetária na determinação do nível de preços. Quando o Governo decide elevar a dívida para manter os gastos públicos, os agentes antecipariam o aumento na arrecadação futura, ajustando seu consumo no presente, de tal forma que o efeito multiplicador dos gastos do Governo seria igual a zero e o gasto sofreria efeito *crowding-out*, ou seja, deslocar-se-iam os gastos das famílias para o Governo.

A melhor resposta do Governo, nesta situação, seria manter o orçamento equilibrado a cada período, permitindo que as variações nos gastos fossem idênticas às variações na arrecadação. Para alguns autores, como Sargent e Wallace (1981) e Woodford (1994, 1995), não existe espaço para qualquer relaxamento da política fiscal, a qual deve, período a período, manter suas contas equilibradas, com gastos inferiores à arrecadação.

Outro argumento contrário ao uso da política fiscal para afetar o lado real, e favorável ao equilíbrio fiscal desenvolvido nos anos 1970 e 1980, está na relação causal entre o déficit público e o déficit externo, conhecidos como déficits gêmeos. Essa discussão foi formalizada por Barro (1989) com sua Teoria Padrão do Déficit Orçamentário. Para o autor, o surgimento de um déficit público seria acompanhado de redução na poupança pública e elevação na renda disponível, conseqüentemente, parte dessa renda seria destinada ao consumo e outra parte à poupança. Assim, a poupança pública não seria totalmente compensada por poupança privada e ocorreria redução na poupança nacional.

Em um equilíbrio compatível com a situação de pleno emprego, em que vigora a igualdade entre poupança e investimento, o aumento no déficit público provocaria uma escassez de poupança frente aos investimentos e ao mesmo tempo excesso de demanda agregada frente à oferta doméstica, resultando em um inevitável déficit em transações correntes. Logo, o controle do déficit externo passa necessariamente pelo controle do déficit público.

Com a abertura financeira e o crescimento do capital especulativo, a abordagem convencional passou a incorporar o comportamento futuro da inflação e o problema dos déficits gêmeos (LOPREATO, 2006; MONTES; ALVES, 2012). Conforme ressalta Lopreato (2006), as aplicações financeiras ao redor do mundo e as constantes reavaliações de portfólio redefiniram os espaços de valorização do capital e acentuaram seu caráter especulativo, ou seja, os capitais ganharam autonomia e a volatilidade e a sensibilidade ao risco tornaram-se mais relevantes na economia mundial.

Nessa concepção, os investidores, cujas expectativas são racionais, *forward-looking*, provocam movimentos especulativos que desencadeiam flutuações nas taxas de câmbio e juros atuais e futuros. O resultado é o condicionamento do comportamento dos Estados aos interesses e expectativas de risco de agentes econômicos do mundo todo. Assim, quando o Governo aparenta dificuldades em garantir a sustentabilidade intertemporal da política fiscal e do próprio orçamento público de forma mais geral, as expectativas de que o Governo fará emissões para cobrir o déficit geram uma crise de confiança que tem como consequência a instabilidade da taxa de câmbio.

Os efeitos na taxa de câmbio, por sua vez, provocam flutuações macroeconômicas que afetam a inflação e, por conseguinte, o próprio déficit público. Tal argumento reforça a ideia de que são necessárias tanto medidas conjunturais quanto estruturais que garantam a sustentabilidade fiscal, ou seja, não apenas a situação fiscal atual importa, mas também as expectativas quanto à situação fiscal futura (LOPREATO, 2006; MONTES; ALVES, 2012).

Adicionalmente, os fluxos de capitais e prêmios de risco associados aos títulos da dívida pública relacionam-se diretamente com a situação fiscal futura e a relação entre a dívida líquida do setor público e o PIB. Assim, o esforço fiscal e a garantia de estabilidade intertemporal das contas públicas seriam a única forma de garantir taxas de juros baixas e estáveis, além de controlar a inflação e o déficit externo. Os benefícios do controle persistente das contas públicas garantiriam a estabilidade no curto prazo e permitiriam que no longo prazo fosse observada a redução no risco país, a queda na taxa de juros e o acesso mais barato ao capital internacional (MONTES; ALVES, 2012).

Uma das principais críticas a essa concepção de política fiscal vem da corrente pós-keynesiana. Nessa corrente, a intervenção fiscal do Estado também passa pelo objetivo de reduzir as incertezas que prejudicam as decisões dos agentes, em especial as decisões de investimentos, levando à insuficiência da demanda agregada. Porém, a política fiscal deve ser ativa com o papel de incrementar a demanda de forma anticíclica, seja diretamente via gastos públicos, ou via efeito multiplicador. Dentro dessa proposta, a política fiscal deve ao mesmo tempo gerar equilíbrio intertemporal do orçamento público e realizar investimentos como estratégia preventiva de estabilização da demanda efetiva (MONTES; ALVES, 2012). Em termos práticos, conforme ressalta Carvalho (1992), para regular a demanda agregada, o Governo deve possuir planos de investimento de longo prazo que sinalizem sua disposição em compensar reduções de investimento do setor privado com seus próprios gastos.

De forma geral, pode-se afirmar que as duas concepções de política fiscal têm como ponto comum a necessidade de estabilidade dos gastos públicos para que se crie um ambiente propício ao investimento, porém os caminhos e fundamentações teóricas são bastante distintos. No que tange à “PEC dos gastos”, suas preposições alinham-se com a primeira abordagem com ampla redução nos gastos públicos e imposição de limites ao orçamento com regras práticas e bastante restritivas, conforme já sugeria Friedman (1953). Nesse sentido, a aprovação da ementa tem como pano de fundo a garantia de estabilidade fiscal das contas públicas e da confiabilidade da economia brasileira, visando principalmente à retomada de investimentos e à maior estabilidade financeira.

Ao delimitar todos os gastos, independentemente de sua natureza, a política afeta diretamente os setores de saúde, educação, administração pública e seguridade social, os quais, por sua vez, possuem elevada participação em diversas regiões do país, principalmente no interior. Desta forma, muito embora a política tenha sido desenhada ao nível nacional e macroeconômico, seus efeitos regionais e setoriais são eminentes. Nesse contexto, a simulação proposta tem como objetivo analisar os efeitos regionais e setoriais da redução nos gastos públicos para os próximos anos, com foco na economia sergipana. Vale ressaltar que a efetividade da política no cumprimento de seus objetivos e de suas regras de aplicação não serão discutidas, uma vez que a política foi aprovada em 2016 e deverá ser colocada em prática, pela força da lei, a partir de 2017.

Neste sentido, o exercício proposto neste artigo se limita a verificar os efeitos da fixação dos gastos, sem qualquer outro tipo de política ou mesmo qualquer alteração no comportamento dos agentes relacionada a expectativas ou confiança. Seguindo teoricamente a concepção *walrasiana* incorporada nos modelos de equilíbrio geral computável da tradição australiana, apenas as alterações nos preços relativos afetam as decisões de maximização de utilidade e lucro, sem qualquer efeito puramente monetário sobre o lado real. Em outras palavras, o modelo não captura o efeito de recuperação da produção e do investimento (também conhecido como *crowding in*), via redução na taxa de juros ou via recuperação na confiança dos agentes nacionais e estrangeiros.

Não obstante, a principal contribuição da simulação proposta está na possibilidade de mensurar o quão dependente as economias e os setores regionais estão da realização dos gastos públicos e, conseqüentemente, o quanto e de que maneira poderão ser afetadas pela política. Esse diagnóstico

torna-se possível por meio da integração de uma base de dados inter-regional e intersetorial detalhada¹, articulada por meio de relações econômicas comportamentais. Tais informações determinam a importância desses gastos nas economias subnacionais, assim como os efeitos diretos da política de variação real zero nos gastos públicos. Os efeitos indiretos, por sua vez, são capturados por meio das relações de insumo-produto que revelam as conexões dos setores públicos com o restante da economia.

Algumas regiões serão mais influenciadas do que outras, porque dependem em maior proporção dos gastos públicos. Por exemplo, municípios menores, com setores agropecuário e industrial pouco desenvolvidos, possuem grande participação no PIB dos setores relacionados ao Poder Público (saúde, educação, administração pública e seguridade social) e serão necessariamente mais afetados.

3. Metodologia

As simulações apresentadas neste trabalho foram realizadas a partir de um modelo Dinâmico e Inter-Regional de Equilíbrio Geral Computável, no qual constam informações detalhadas regional e setorialmente para 27 unidades da federação e 67 atividades econômicas². A fim de traçar um cenário macroeconômico e seus efeitos, foram utilizados dados históricos e conjunturais da economia brasileira, permitindo a projeção de um cenário alternativo em que as tendências macrorregionais são mantidas, enquanto os gastos públicos são controlados continuamente.

Conforme representado na Figura 1, esse modelo foi construído a partir de um modelo nacional de EGC calibrado para o ano de 2013³, a partir das Tabelas de Usos e Recursos (TRU) do Instituto Brasileiro de Geografia e Estatística (IBGE) e dos dados secundários regionais (por Unidade da Federação) da produção setorial e dos componentes da demanda final referentes ao ano base (2013)⁴. De posse desse conjunto de dados, a base de dados do modelo nacional de EGC foi regionalizada conforme procedimento criado por Horridge (2006)⁵.

Por sua vez, as simulações levaram em consideração variáveis macroeconômicas históricas e tendenciais, além do cenário de redução dos gastos públicos, que formam um conjunto de choques exógenos ao modelo de equilíbrio geral. A partir desses choques, são traçadas trajetórias para a economia sergipana, com detalhamentos que chegam ao nível municipal e setorial.

4. O modelo de EGC

Os modelos de EGC têm sido utilizados por diversos países em diferentes áreas econômicas, a saber: crescimento de longo prazo e mudanças estruturais, política tributária, distribuição de renda, avaliação de estratégias de desenvolvimento e políticas comerciais. Vale ressaltar que tais modelos podem ser desenvolvidos para diferentes escalas espaciais (DOMINGUES, 2002; PEROBELLI, 2004).

¹ Durante a calibragem do modelo inter-regional, basicamente o vetor de gastos públicos é decomposto primeiro ao nível estadual e, em seguida, ao nível municipal por meio de informações das contas nacionais, regionais e municipais.

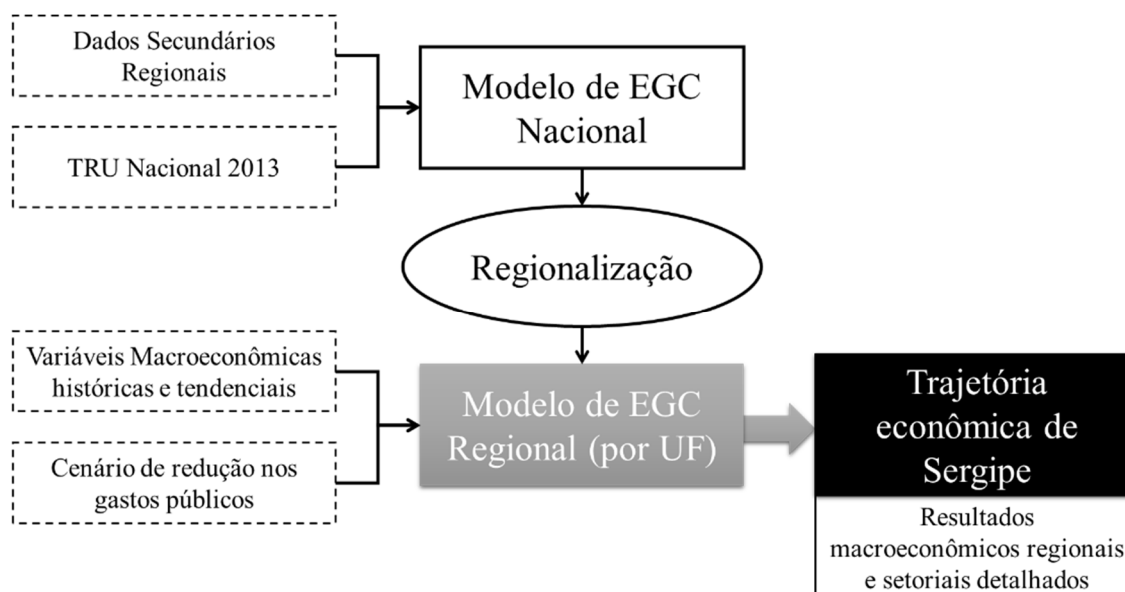
² Ver Anexo 1.

³ Matriz de Insumo-Produto (MIP) estimada para o Brasil a partir das TRU do IBGE referente ao ano de 2013, com 67 setores e 67 produtos com base no procedimento desenvolvido por Guilhoto e Sessa Filho (2005; 2010).

⁴ Foram utilizadas diversas fontes estatísticas, como: Contas Regionais do IBGE, massa salarial por setor de atividade e estados obtida junto à Relação Anual de Informações Sociais (RAIS) do Ministério do Trabalho e Emprego e informações sobre exportações por produto do sistema *Comex Stat* (antigo *AliceWeb*) do Ministério da Economia, Indústria, Comércio Exterior e Serviços (MDIC).

⁵ Esse procedimento já foi aplicado para dados da economia brasileira e pode ser visto com mais detalhe em Carvalho et al. (2017) e Ribeiro et al. (2018).

Figura 1 – Esquema metodológico



Fonte: Elaboração própria.

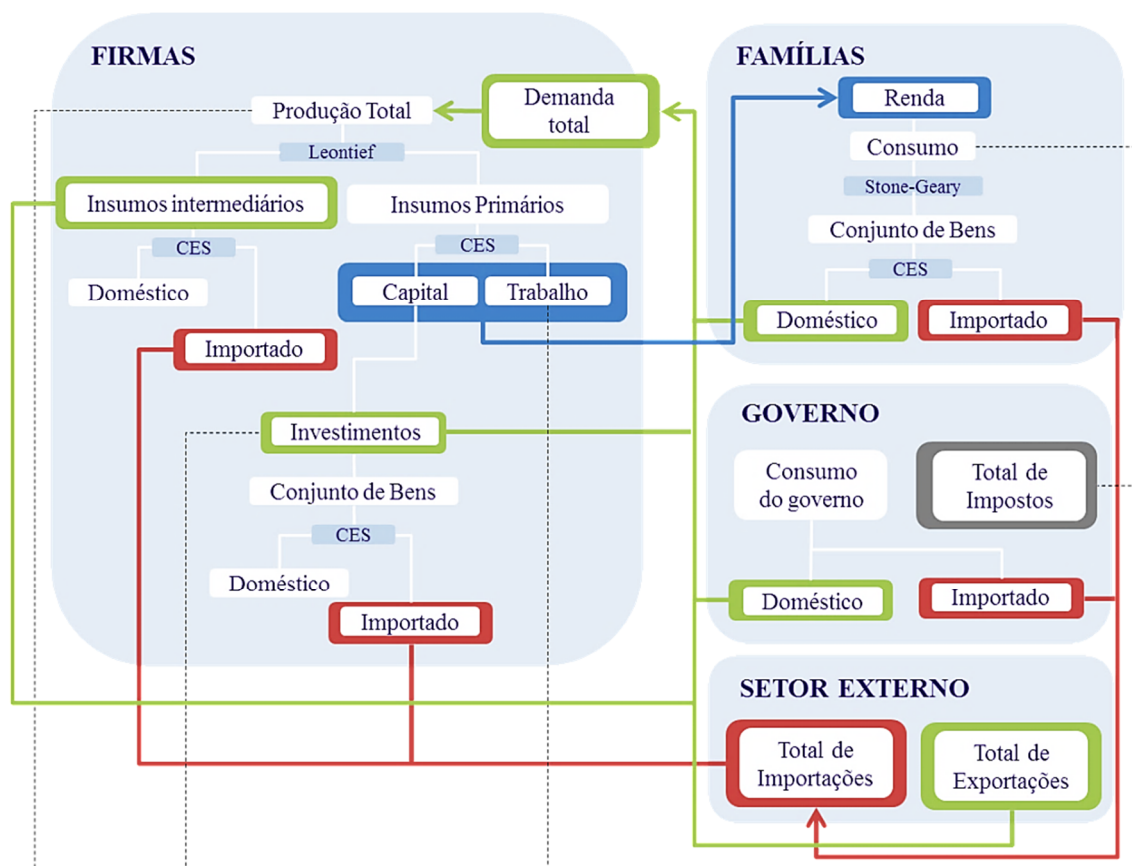
A estrutura teórica do modelo é baseada na estrutura do *The Enormous Regional Model - TERM* (HORRIDGE et al., 2005), com a incorporação de elementos de dinâmica recursiva. Essa especificação baseia-se na tradição australiana de modelos de EGC, cuja estrutura central está representada esquematicamente na Figura 2. Para cada um dos estados brasileiros, os setores produtivos minimizam custos de produção sujeitos a uma tecnologia de retornos constantes de escala, cuja combinação de insumos intermediários e fator primário (total de capital e trabalho) é determinada por coeficientes fixos (em uma função de produção do tipo Leontief). Na composição dos insumos, há substituição via preços entre produto doméstico e importado, por meio de funções de elasticidade de substituição constante (CES⁶). Na composição do fator primário também há substituição via preço entre capital e trabalho com elasticidades de substituição constante.

A demanda das famílias é especificada a partir de funções de utilidade não homotéticas Stone-Geary (PETER et al., 1996), dividindo o consumo dos bens e serviços em parcelas de “luxo” e “subsistência”, de tal forma que uma parcela fixa do gasto é reservada ao consumo de subsistência e à parcela residual em “gastos de luxo”. Essa especificação permite que modificações na renda causem modificações diferenciadas no consumo dos produtos, a depender do tipo de produto e do comportamento dos preços relativos. Ao mesmo tempo, novamente, a composição entre domésticos e importados é estabelecida por meio de funções de elasticidade de substituição constante.

As exportações setoriais respondem a curvas de demanda negativamente associadas aos custos domésticos de produção e positivamente afetadas pela expansão exógena da renda internacional, adotando-se a hipótese de país pequeno no comércio internacional, ou seja, o destino na produção, ao mercado doméstico ou externo, depende unicamente dos termos de comércio. O consumo do Governo é tipicamente exógeno, podendo estar associado ou não ao consumo das famílias. Os estoques se acumulam de acordo com a variação da produção.

A dinâmica recursiva do modelo implica soluções sequenciais, ano a ano, requerendo que o modelo possua dois tipos de equações: o primeiro grupo contém as equações referentes aos comportamentos anteriormente descritos, que são solucionadas como em um modelo de estática comparativa; já o segundo grupo determina o ajuste intertemporal do estoque de capital, permitindo a conexão e a atualização dos dados de forma dinâmica, a partir das soluções de cada ano.

⁶ Do inglês *constant elasticity of substitution*.

Figura 2 – Representação esquemática da estrutura do modelo de EGC

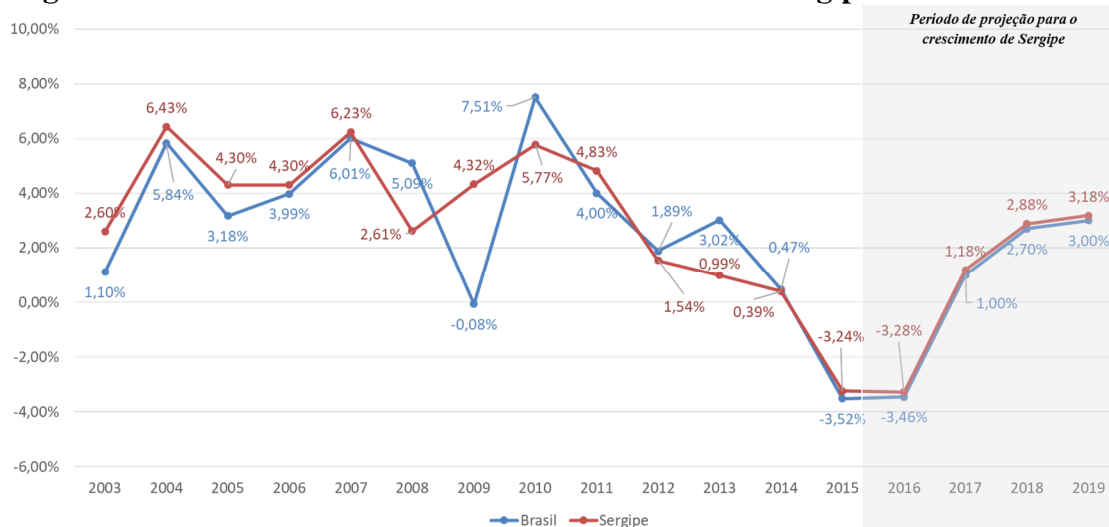
Fonte: Souza e Domingues (2014).

O investimento e o estoque de capital seguem mecanismos de acumulação e de deslocamento intersetorial a partir de regras pré-estabelecidas, associadas à taxa de depreciação e de retorno. Seguindo Dixon e Rimmer (2002), em cada ano de simulação, assume-se que as taxas de crescimento do capital da indústria i (e dessa forma, os níveis de investimento) são determinadas pela disposição dos investidores em fornecer fundos a essa indústria frente aos aumentos limitados em sua taxa de retorno esperada. Dessa forma, a taxa de crescimento do capital na indústria i no ano t só será maior do que sua taxa normal (estado estacionário do crescimento de capital) se a taxa de retorno esperada pelos investidores for superior à taxa de retorno normal (DIXON; RIMMER, 2002).

Todas as relações entre agentes e setores econômicos descritas nesta seção formam um sistema de m equações e n variáveis cuja solução inicial pode ser obtida por meio dos dados referentes ao ano base, neste caso, 2013. Por sua vez, nas simulações, são calculadas as variações em m variáveis do modelo (endógenas) provocadas por mudanças nas demais $(n - m)$ variáveis (exógenas) em comparação ao ano base (DIXON; RIMMER, 2002). O completo detalhamento das equações pode ser encontrado em Horridge (2012) e os parâmetros utilizados encontram-se detalhados no Anexo 1.

5. Base de dados

Para a construção do cenário tendencial, inicialmente, apresenta-se na Figura 3 a taxa de crescimento observada, extraída das Contas Regionais do IBGE, referente ao PIB do Brasil e de Sergipe entre 2003-2015 e a projeção do crescimento entre 2016-2019.

Figura 3 – Taxas de Crescimento do PIB do Brasil e de Sergipe entre 2003 e 2019

Nota: *Para o período 2018-2019, as estimativas do Brasil correspondem aos valores projetados pelo Banco Central do Brasil (2018), enquanto as projeções para Sergipe (2016-2019) consideram que o estado mantém o crescimento ligeiramente acima do Brasil, conforme a média do período 2000-2015.

Fonte: Elaboração própria com base nos dados das Contas Regionais do IBGE (2018).

No período 2003-2015, a economia brasileira cresceu a uma taxa média anual de 3,2%, ao passo que Sergipe cresceu ligeiramente mais, 3,39%. Desse modo, assume-se para o cenário de referência, que no período de projeção (2016-2035), a economia sergipana crescerá 0,19% acima da brasileira⁷. Entre o período 2014-2017, são usadas variáveis macroeconômicas observadas para a economia brasileira e a taxa de crescimento do PIB de Sergipe observada e projetada. Esses dados foram obtidos nas Contas Regionais do IBGE, FUNCEX e IPEADATA.

A Tabela 1 apresenta os indicadores anuais utilizados no cenário de referência entre 2014 e 2017. Em 2018 e 2019, utilizaram-se as projeções de crescimento da economia brasileira divulgadas pelo Banco Central do Brasil (BCB), isto é, 2,8% e 3%, respectivamente⁸. Para o período 2020-2035, adotou-se uma taxa anual de crescimento do PIB real das economias brasileira e sergipana de 3% e 3,19%, respectivamente. Os demais componentes macroeconômicos, consumo das famílias, gastos do Governo, exportações e investimentos, por exemplo, são endógenos, ou seja, determinados pelo próprio modelo.

Tabela 1 – Variáveis do cenário de referência para o período 2014-2017 - em variação %

Período	2014	2015	2016	2017
PIB real	0,50	-3,55	-3,46	0,99
Consumo das Famílias	2,25	-3,22	-4,34	0,96
Consumo do Governo	0,81	-1,44	-0,06	-0,56
Formação Bruta de Capital Fixo	-4,22	-13,95	-10,27	-1,84
Exportação	-1,13	6,82	1,92	5,18
Índice de Preços do Consumidor (IPCA)	6,41	10,67	6,29	2,95
Índice de Preços das Importações	-1,97	-11,88	-8,94	4,06
PIB Sergipe	0,39	-3,24	-3,26*	1,18*

* Valores estimados com base na diferença histórica do PIB de Sergipe em relação ao Brasil.

Fonte: Elaboração própria com base nos dados do IBGE (2018), FUNCEX (2018) e IPEA (2018).

⁷ Vale mencionar que não há dados oficiais de projeção de crescimento do PIB de Sergipe.

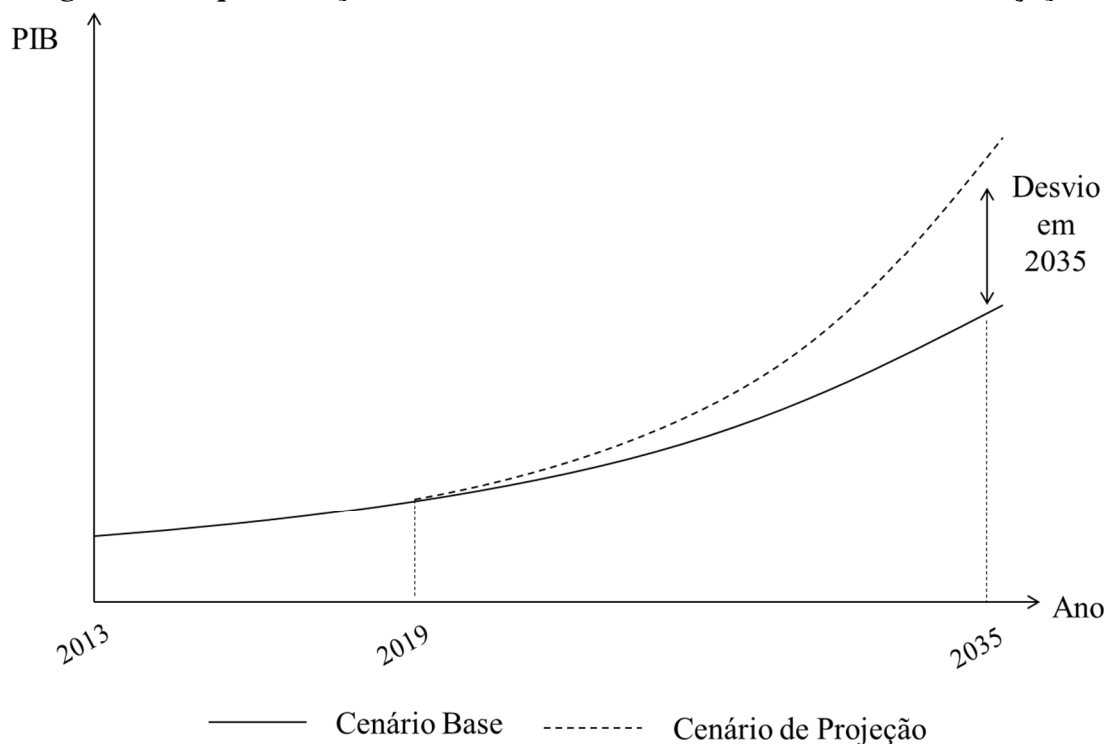
⁸ Boletim Focus de 16 de março de 2018 (Banco Central do Brasil, 2018).

No sentido de deixar o cenário tendencial mais consistente, também foram utilizadas taxas de crescimento da população de todos os estados brasileiros, conforme informações disponibilizadas pelo IBGE.

6. Resultados e discussões

Nesta seção, são apresentados os resultados das duas simulações realizadas: i) cenário tendencial da economia sergipana no período 2018-2035; e ii) impacto do ajuste fiscal sobre a economia sergipana. A primeira simulação representa o cenário de referência (*baseline*) que indica a trajetória de crescimento da economia sergipana, sem considerar nenhum efeito exógeno adicional ao longo do período de 2019 a 2035. Este cenário servirá de comparação com o cenário de política (ajuste fiscal) como ilustra a Figura 4.

Figura 4 – Representação do Cenário de Referência e do Cenário de Projeção



Fonte: Adaptado de Souza (2015).

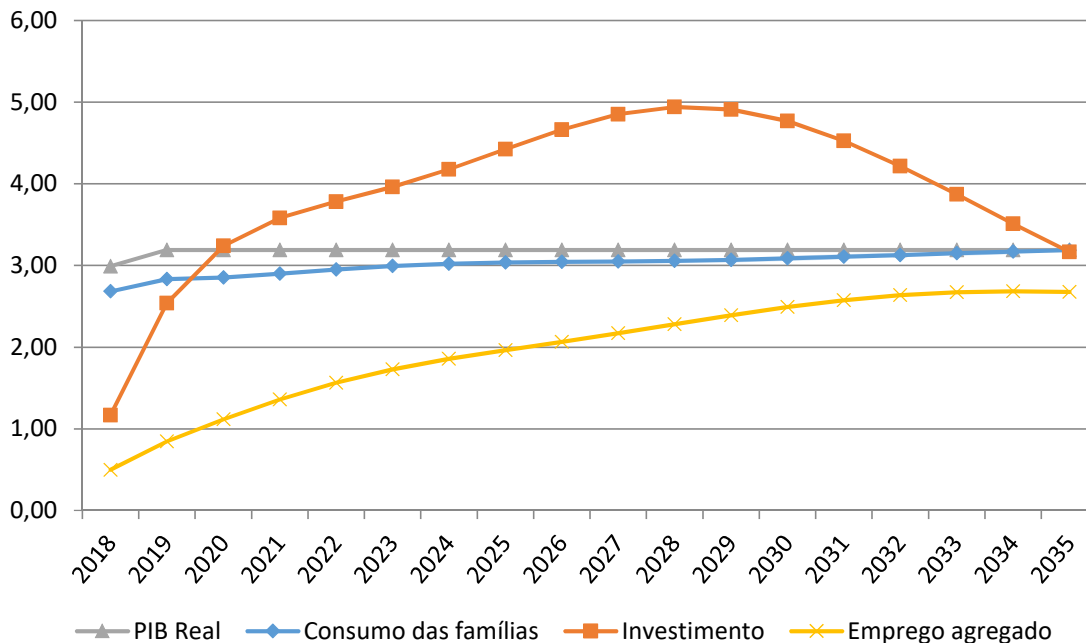
Ainda que a escolha do período da simulação de política (2019-2035) seja arbitrária, dentro do horizonte temporal escolhido, é possível verificar os efeitos de longo prazo da política, a qual terá vigência de até 20 anos, e como a economia sergipana reage e se estabiliza com a realocação de recursos. Além disso, dentro desse período, a Política de Teto dos Gastos ainda estaria vigente.

A Figura 5 apresenta os resultados do cenário tendencial para o PIB real, consumo das famílias, investimento real e emprego agregado da economia sergipana no período 2018-2035, sem qualquer alteração nos gastos públicos. A partir de 2020, o comportamento homogêneo do PIB é explicado pelo choque exógeno de 3,19% a.a.

O componente mais dinâmico é o investimento real que reage positivamente à recuperação econômica após o período de crise. A variação real do investimento é crescente até 2028 (quando chega a 4,94%) e desacelera até o final do período, com um crescimento de 3,17% em 2035. O período de crescimento acelerado entre 2018 a 2028 ocorre em decorrência de um reajuste dos investimentos após o período de crise (2015-2016), capturado pelos mecanismos de acumulação de capital. Até 2016, após um período de baixos investimentos, os níveis de estoque de capital estavam baixos (crescendo menos do que a economia), pressionando para cima as taxas de retorno na

economia e a razão investimento/capital. Com taxas de retorno mais altas, o investimento sobe e promove nova pressão sobre a taxa de retorno do capital, que volta a cair no final no período de projeção. A taxa de retorno mais baixa, por sua vez, faz com que o investimento volte aos níveis tendenciais até o final do período de simulação.

Figura 5 – Trajetória tendencial de variáveis macroeconômicas de Sergipe (%)



Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

Com a retomada do investimento, o emprego também se recupera e segue uma tendência de crescimento, seguindo a mesma trajetória de crescimento do PIB, já que a expansão da economia implica maior utilização de fatores primários na produção. Ao final do período de simulação, o investimento retoma níveis tendenciais próximos ao crescimento do PIB, indicando a manutenção do estoque de capital. Vale ressaltar que a estabilidade do PIB e do consumo das famílias no período é consequência direta dos choques exógenos da simulação⁹ e do fechamento do modelo, que implica estabilidade entre consumo e produção.

Conseqüentemente, além dos investimentos, o setor externo também absorve os componentes do ciclo econômico, como mostra a Figura 6. Entre 2019-2026, as exportações crescem mais do que as importações, o que ocorre em decorrência do período de crise, no qual todo o país se defrontou com uma desvalorização do real concomitante à redução no poder de compra das famílias brasileiras. A recuperação provocaria um período de ajuste, entre 2027-2031, em que o crescimento das importações superaria o crescimento das exportações, porém, no período seguinte, as exportações recobram força e indicam a retomada do crescimento, decorrente do fortalecimento da economia interna e dos ganhos de competitividade promovidos pelo período de fortes investimentos.

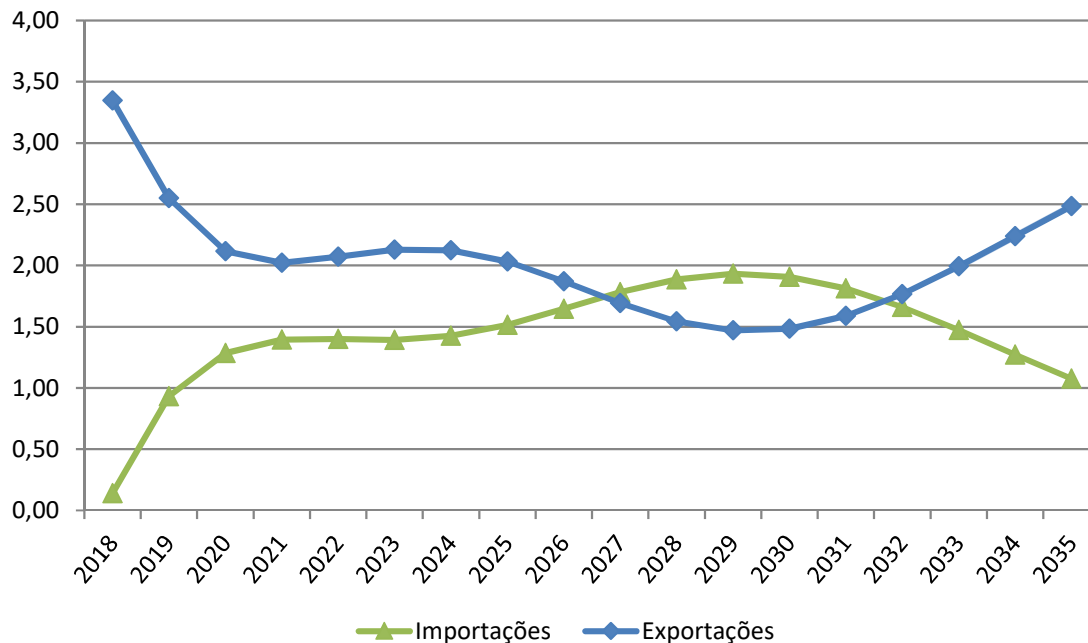
A trajetória dessas variáveis está diretamente alinhada à relação entre preços domésticos e importados. Os índices de preços apresentariam queda (elevação) em alguns períodos, o que tornaria os produtos nacionais mais (menos) atrativos e, conseqüentemente, provocaria aumento (redução) das exportações e queda (aumento) das importações.

Do ponto de vista setorial, as cinco maiores taxas de crescimento médio da produção e do emprego setorial entre 2018-2035, no cenário tendencial, seriam das seguintes atividades: i) Fabricação de máquinas e equipamentos mecânicos; ii) Serviços de arquitetura, engenharia, testes/análises técnicas e P&D; iii) Alojamento; iv) Atividades de vigilância, segurança e

⁹ A partir de 2019, o crescimento do PIB é homogêneo.

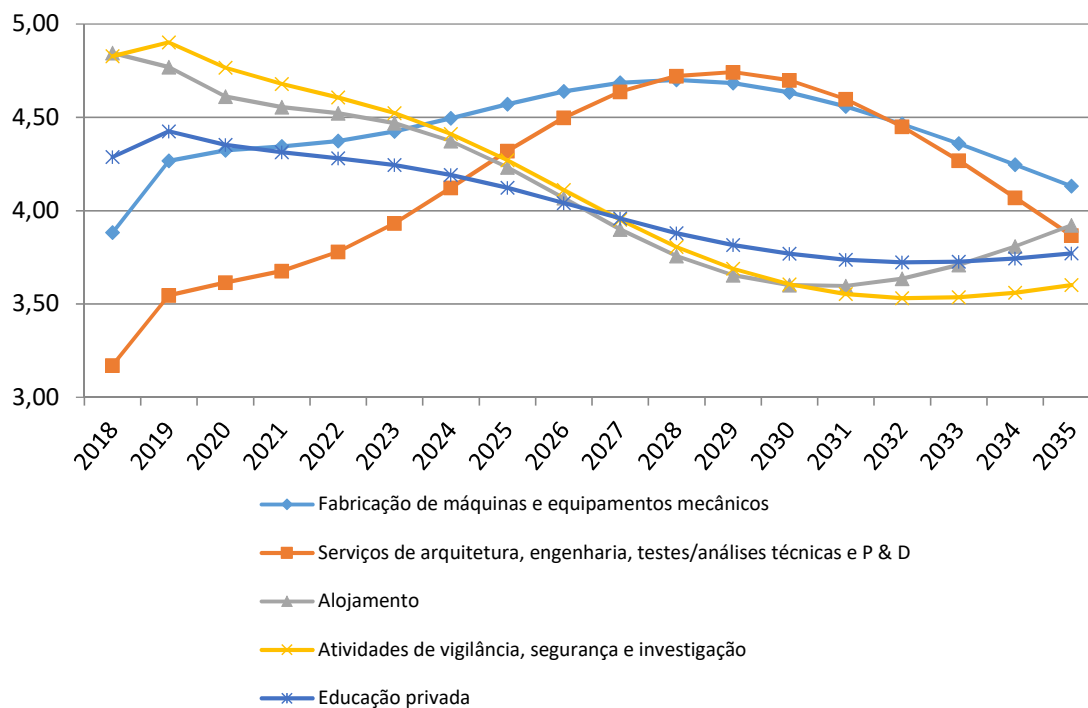
investigação; e v) Educação privada. Tais atividades estão relacionadas tanto ao elevado investimento no período quanto ao crescimento no consumo das famílias, que incentiva a produção regional. A Figura 7 revela a dinâmica tendencial da produção anual dessas atividades entre 2018-2035.

Figura 6 – Trajetória tendencial das exportações e importações agregadas em Sergipe (%)



Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

Figura 7 – Trajetória tendencial da produção setorial, atividades selecionadas de Sergipe (%)



Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

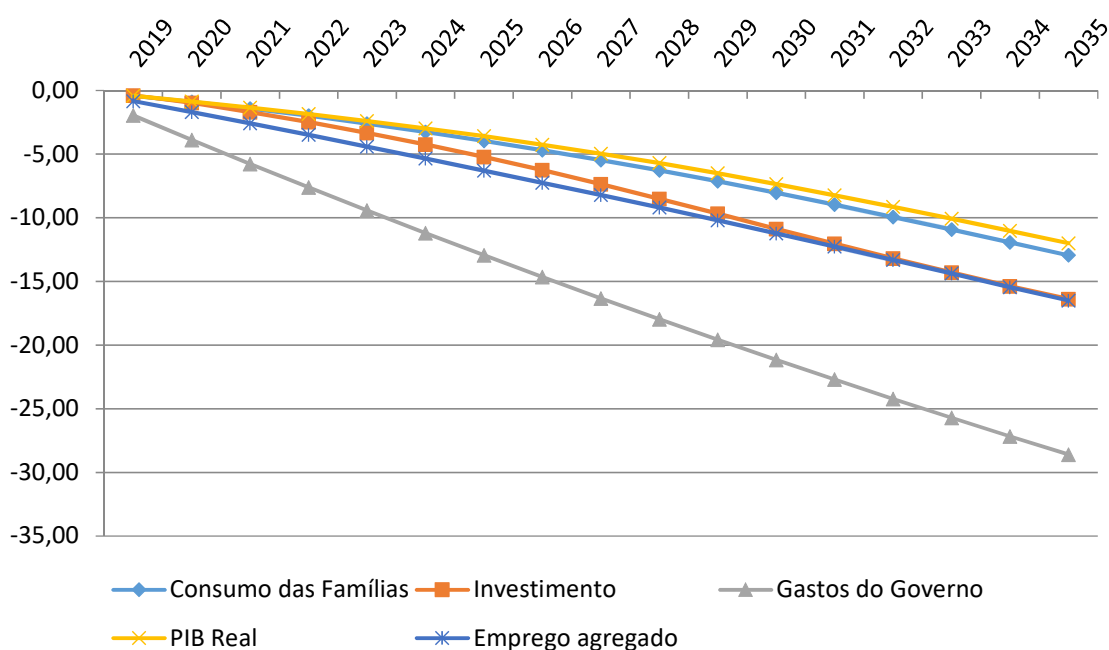
Na sequência, são apresentados os resultados da simulação do impacto do ajuste fiscal ou dos cortes dos gastos públicos do Brasil sobre o estado de Sergipe. Para isso, além das informações para

o Brasil utilizadas anteriormente, simulou-se a redução de 0,54%¹⁰ do gasto público brasileiro em 2017 e o crescimento real igual a zero durante o período 2019-2035, conforme define a PEC dos gastos.

Mais uma vez, é importante ressaltar que esta simulação busca isolar o efeito do ajuste fiscal sobre a economia sergipana, ou seja, assume-se implicitamente que nenhuma outra medida, como, por exemplo, a recuperação do investimento privado, é levada em consideração. Ou seja, diante da imprevisibilidade quanto aos efeitos de confiabilidade sobre os investidores, o resultado sobre os investimentos reflete apenas os incentivos econômicos, regionais e setoriais, quanto à atratividade de cada investimento em termos do retorno no capital, conforme define o comportamento endógeno do investimento no modelo, descrito na Seção 3. Não obstante, assume-se implicitamente que as expectativas dos agentes não seriam ajustadas a partir da PEC dos gastos.

Todos os resultados a seguir são reportados como o desvio acumulado entre 2019 e 2035 em relação ao cenário de referência do modelo (simulação 1). A Figura 8 revela os resultados sobre PIB real, consumo das famílias, investimento real, gastos do Governo e emprego agregado de Sergipe.

Figura 8 – Resultados do ajuste fiscal sobre principais variáveis macroeconômicas de Sergipe - desvio acumulado 2019-2035 em relação ao cenário de referência (%)



Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

O ajuste fiscal impactaria negativamente todas as principais variáveis macroeconômicas de Sergipe. O resultado acumulado em 2035 do PIB real, por exemplo, seria de -12%. Isto quer dizer que, em 2035, o PIB do estado estaria 12% menor do que o previsto no cenário tendencial, na ausência do ajuste fiscal. Em outras palavras, o ajuste fiscal realizado pelo Governo Federal brasileiro atenuaria o crescimento da economia sergipana. Não obstante, a arrecadação de impostos indiretos teria queda acumulada em 2035 de 20,78% em relação ao cenário tendencial de Sergipe, uma consequência não apenas da redução dos gastos do Governo, como também da queda na demanda agregada.

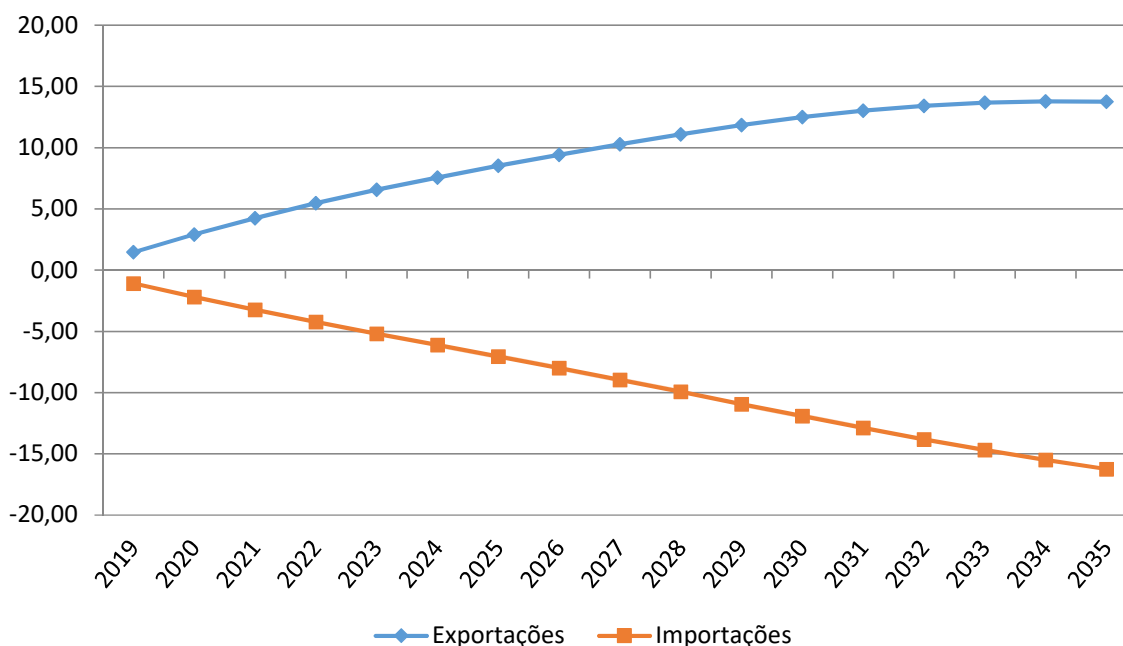
A Figura 9 mostra o impacto que o ajuste fiscal provocaria sobre as contas externas de Sergipe. Evidencia-se um impacto positivo sobre as exportações na maior parte do período e uma redução das importações, ou seja, favorável à balança comercial do estado. Tal cenário decorre da redução na renda interna e consequente queda no consumo tanto das famílias quanto entre empresas,

¹⁰ Dado observado já divulgado pelo IBGE.

reduzindo os preços nacionais em comparação aos internacionais e, conseqüentemente, incentivando as exportações e reduzindo as importações.

O desvio acumulado das exportações sergipanas, em 2035, seria de 13,76% em relação ao cenário de referência. As importações, por outro lado, apresentariam variações negativas, acumulando queda de 16,24% em relação ao cenário de referência. Em 2035, tais importações apresentariam queda de 33,01% em comparação ao cenário de referência, isto é, 0,80%. Vale salientar que, como mencionado anteriormente, a balança comercial no modelo não depende de fatores externos, mas apenas da alteração nos preços relativos que tornam os produtos brasileiros mais competitivos no cenário internacional.

Figura 9 – Resultados do ajuste fiscal sobre as contas externas de Sergipe - desvio acumulado 2020-2035 em relação ao cenário de referência (%)

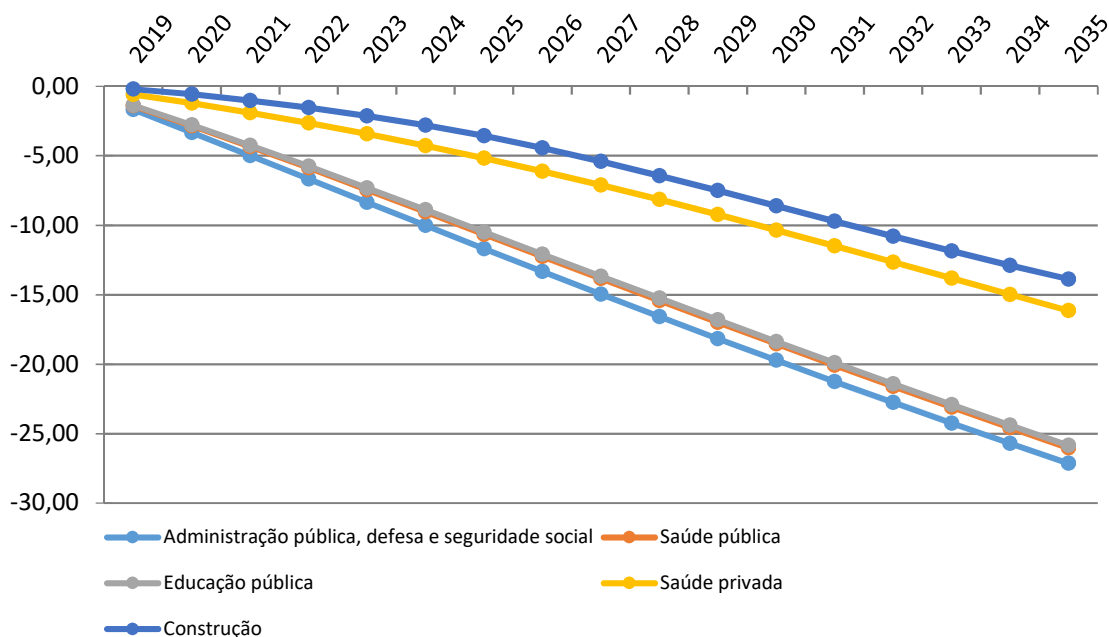


Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

Sob a ótica setorial, o ajuste fiscal do Governo Federal provocaria queda da produção e do emprego de quase todas as atividades de Sergipe em relação ao cenário de referência. No que remete à produção setorial, considerando o desvio acumulado em 2035, os cinco setores mais afetados seriam: i) Administração pública, defesa e seguridade social, queda de 27,13%; ii) Saúde pública, 26,03%; iii) Educação pública, 25,85%; iv) Saúde privada, 16,15; v) Construção, 13,87%. Vale ressaltar que principalmente os setores de Educação e Saúde públicas possuem despesas obrigatórias vinculadas em lei, as quais não foram discriminadas na presente análise. A Figura 10 apresenta o resultado anual do desvio acumulado da produção setorial das referidas atividades.

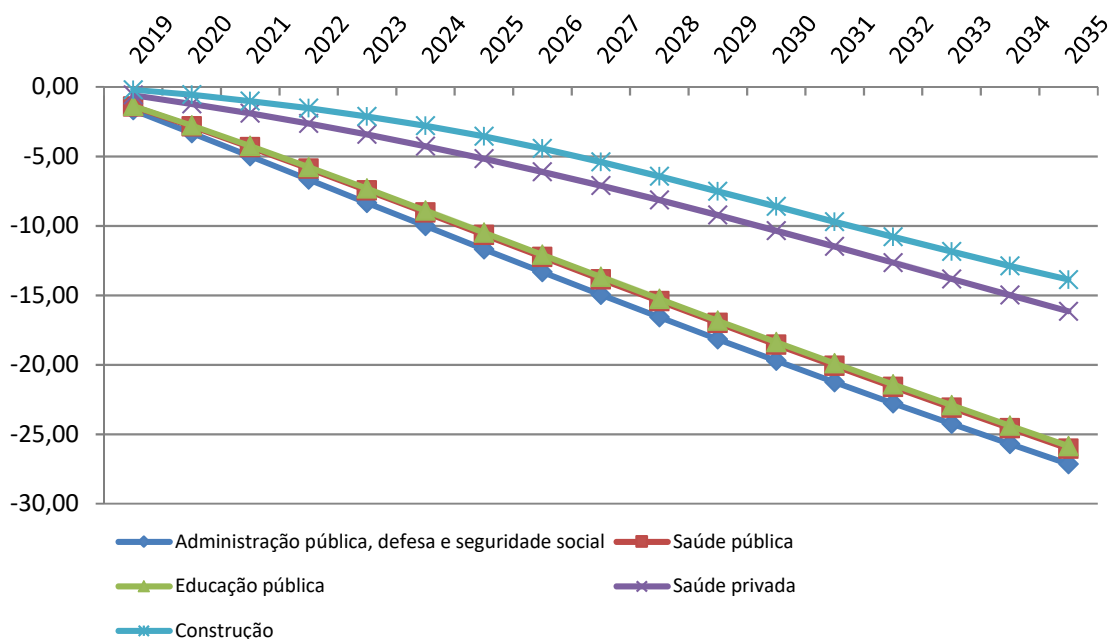
Em relação ao emprego setorial, considerando o desvio acumulado em 2035, os cinco setores econômicos mais afetados em Sergipe seriam: Administração pública, defesa e seguridade social (-43,08%), Fabricação de máquinas e equipamentos mecânicos (-42,96%), Educação pública (-42,96%), Fabricação de automóveis, caminhões e ônibus, exceto peças (-42,81%) e Saúde pública (-42,77%). Como esperado, tanto em relação à produção quanto ao emprego setorial, atividades que ofertam serviços públicos essenciais, saúde e educação estariam entre as mais fortemente impactadas pelo ajuste fiscal. A Figura 11 mostra o resultado anual do desvio acumulado do emprego setorial das referidas atividades.

Figura 10 – Resultados do ajuste fiscal sobre a produção setorial de Sergipe - desvio acumulado 2019-2035 em relação ao cenário de referência (%), atividades selecionadas.



Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

Figura 11 – Resultados do ajuste fiscal sobre o emprego setorial de Sergipe - desvio acumulado 2019-2035 em relação ao cenário de referência (%), atividades selecionadas.



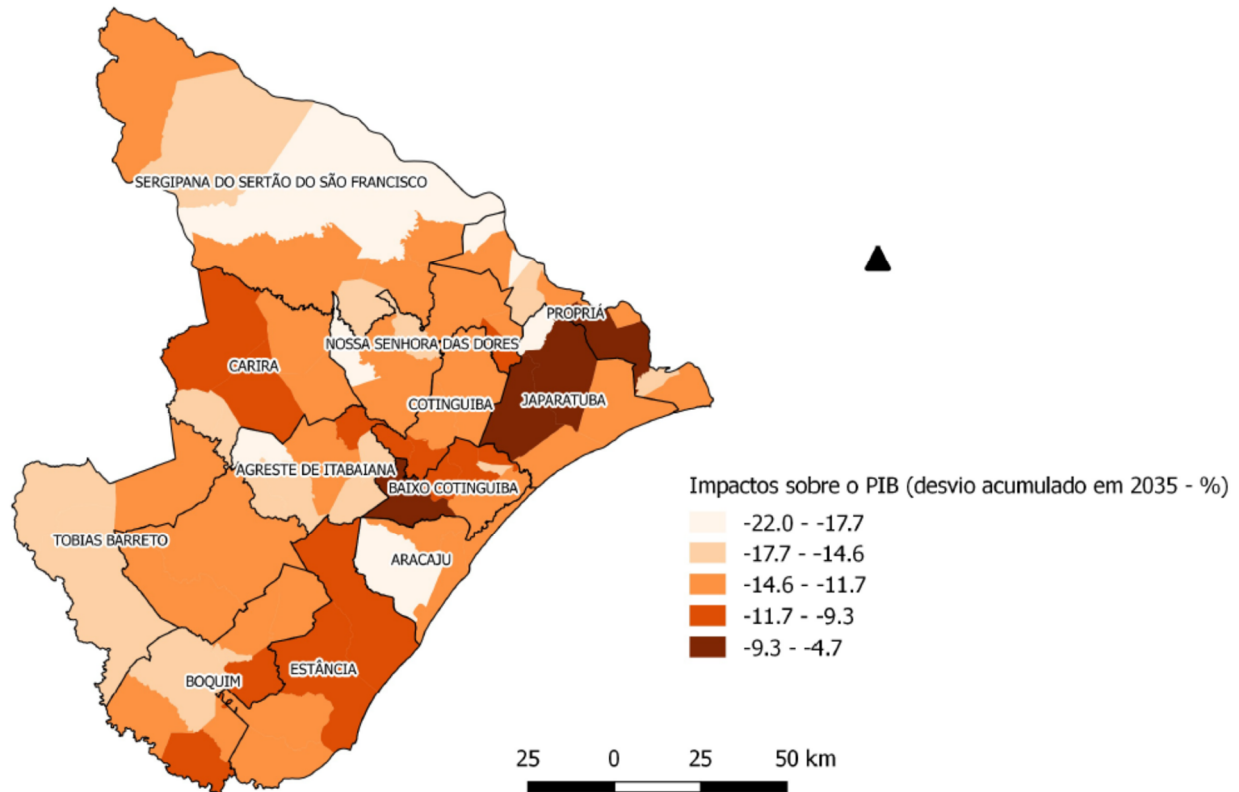
Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

Como se pode observar, as atividades mais prejudicadas seriam aquelas diretamente relacionadas ao consumo do Governo e que teriam seus gastos limitados pelo ajuste fiscal. No entanto, diversas outras atividades também sofreriam perdas, principalmente decorrentes da dependência econômica da renda proveniente do setor público e dos consequentes efeitos multiplicadores no restante da economia.

Esses resultados tornam-se heterogêneos ao nível municipal, em decorrência dos diferentes níveis de dependência dos gastos públicos, como ilustra a Figura 12. Percebe-se um efeito

heterogêneo entre o território sergipano com quedas do PIB dos seus municípios variando entre 4,7% e 22%. Os municípios mais impactados seriam aqueles localizados mais ao norte do estado como, por exemplo, Monte Alegre (-22,01%), Amparo de São Francisco (-20,5%), Gararu (-18,6%) e São Francisco (-18,25%); e na região metropolitana de Aracaju, o município de São Cristóvão (-19,18%).

Figura 12 – Resultados do ajuste fiscal sobre o PIB dos municípios de Sergipe - desvio acumulado 2019-2035 em relação ao cenário de referência (%)



Fonte: Elaboração própria com base nas simulações com o modelo de EGC.

O resultado para São Cristóvão pode ser explicado, em parte, em decorrência do município sediar o maior *campus* da Universidade Federal de Sergipe no estado. Amparo do São Francisco, São Francisco, Gararu e Monte Alegre de Sergipe são municípios pequenos e relativamente mais dependentes da Administração Pública. De acordo com a população estimada de 2018 do IBGE, os referidos municípios têm, respectivamente, 2.368, 3.987, 11.606 e 15.120 habitantes.

7. Considerações finais

O objetivo deste artigo foi mensurar o efeito de longo prazo do ajuste fiscal, conduzido pelo Governo Federal, sobre a economia sergipana. Para tanto, foram realizadas simulações com um modelo inter-regional e dinâmico de EGC, calibrado para 2013, constituído por 27 estados e 67 setores econômicos.

Os principais resultados revelam que o ajuste fiscal atenuaria o crescimento de todas as principais variáveis macroeconômicas do estado de Sergipe. O PIB real, por exemplo, cresceria 12% a menos do que no cenário de referência em 2035. No que remete aos impactos setoriais sobre a produção e o emprego, entre os principais setores mais afetados estariam Administração Pública, Educação e Saúde, os quais ofertam serviços públicos essenciais para a economia sergipana. Além disso, o impacto sobre o PIB dos municípios sergipanos seria bastante heterogêneo.

Como consequência dos pressupostos do modelo de EGC, os resultados apresentados reforçam o argumento de que o ajuste fiscal permite maior equilíbrio no balanço de pagamentos, conforme

prevê a teoria econômica. Porém, simultaneamente, os resultados permitem a construção de um cenário, que se reflete na forte redução do nível de atividade econômica e emprego, principalmente ao nível municipal para regiões altamente dependentes do consumo do Governo. Nesse sentido, a análise se torna relevante ao revelar um cenário de longo prazo referente ao corte dos gastos federais sobre um estado periférico e com resultados detalhados ao nível setorial e regional.

O exercício de simulação realizado possui como principal limitação o fato de não considerar eventuais ações que podem surgir a médio-prazo em decorrência do ajuste fiscal como, por exemplo, recuperação do investimento privado, que seria estimulado pelo aumento da confiança dos agentes e pelas alterações de suas expectativas, ou ainda alterações significativas no cenário internacional que permitam maiores expansões nas exportações.

No que tange ao modelo utilizado, as limitações estão relacionadas diretamente aos pressupostos teóricos nos quais as simulações se fundamentam. Por seguir a tradição *walrasiana* de equilíbrio geral, o modelo trata do lado real da economia, de tal forma que apenas os preços relativos importam e nenhuma alteração puramente monetária afeta os resultados. Tais considerações levam ao pressuposto implícito de que a demanda agregada é perfeitamente inelástica em relação às alterações na taxa juros e, conseqüentemente, não existe a possibilidade de um efeito *crowding in* após o ajuste fiscal. Por sua vez, pelo lado fiscal, não há uma conexão direta entre a arrecadação e os gastos públicos, que são determinados exogenamente.

Ademais, pela falta de informações detalhadas ao nível regional, não foi possível discriminar nas simulações os mínimos obrigatórios referentes aos setores de Educação e Saúde públicas, sendo assim, todas as simulações consideram uma redução no gasto total que recai de forma homogênea independente da natureza do gasto. Na medida em que novas informações se tornarem disponíveis, será possível atualizar o exercício e, conseqüentemente, oferecer novos resultados para a discussão. Mais especificamente, um passo adiante nesta pesquisa seria estimar a magnitude do aumento do investimento, após os primeiros anos de ajuste fiscal, no sentido de capturar com a simulação o efeito de recuperação da confiança e melhoria nas expectativas dos agentes.

Referências

- BARRO, R. J. The Ricardian approach to budget deficits. *Journal of Economic Perspectives*, v. 3, n. 2, p. 37-54, 1989.
- BCB – Banco Central do Brasil. Publicações. Focus - Relatório de Mercado. Disponível em: <http://www.bcb.gov.br/pec/GCI/PORT/readout/R20180316.pdf>. Acesso em: março de 2018.
- BRASIL. *Emenda Constitucional nº 95*, de 15 de dezembro de 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc95.htm. Acesso em 29 de março de 2018.
- CAMARGOS, L. R.; GOIS, R. M.; NÓBREGA, W. *Anuário Socioeconômico de Sergipe*. Ano 1, v. 1, 2017.
- CARVALHO, F. J. C. *Mr. Keynes and the post Keynesians*. Aldershot: Edward Elgar, 1992.
- CARVALHO, T. S.; DOMINGUES, E. P.; HORRIDGE, J. M. Controlling deforestation in the Brazilian Amazon: Regional economic impacts and land-use change. *Land Use Policy*, v. 64, p. 327-341, 2017.
- DIXON, P.; RIMMER, M. *Dynamic general equilibrium modelling for forecasting and policy*. A practical guide and documentation of MONASH. Cayton: Emerald, 2002.
- DOMINGUES, E. P. *Dimensão regional e setorial da integração brasileira na Área de Livre Comércio das Américas*. Tese (Doutorado em Economia). Programa de Pós-Graduação em Economia, Universidade de São Paulo. São Paulo, 2002.

- FRIEDMAN, M. *A monetary and fiscal framework for economic stability*. Essays in Positive Economics. Chiago: The University Press, 1953.
- FUNCEX – Fundação Centro de Estudos do Comércio Exterior. Disponível em: <http://www.funcedata.com.br>. Acesso em: março de 2018.
- GOBETTI, S. W. *Tópicos sobre a política fiscal e o ajuste fiscal no Brasil*. Tese (Doutorado em Economia). Programa de Pós-Graduação do Departamento de Economia da Universidade Federal de Brasília, Brasília, 2008.
- GUILHOTO, J. J. M.; 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.
- GUILHOTO, J. J. M.; SESSO FILHO, U. A. 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. 53-62, 2010.
- HORRIDGE, J. M. *Preparing a TERM bottom-up regional database*. Preliminary Draft, Centre of Policy Studies, Monash University, 2006.
- HORRIDGE, J. M. The TERM model and its data base. In: WITTEWER, G. (Ed.) *Economic modelling of water: the Australian CGE experience*. Springer, Dordrecht, Netherlands, Chapter 2, 2012.
- HORRIDGE, J. M.; MADDEN, J.; WITTEWER, G. The Impact of the 2002-2003 Drought on Australia. *Journal of Policy Modeling*, v. 27, n. 3, 2005/4, p. 285-308, 2005.
- IBGE – Instituto Brasileiro de Geografia e Estatística. Contas Nacionais. *Sistema de Contas Regionais*. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9054-contas-regionais-do-brasil.html?=&t=resultados>. Acesso em: março de 2018.
- IPEA – Instituto de Pesquisa Econômica Aplicada. *Ipeadata*. Macroeconômico. Disponível em: <http://www.ipeadata.gov.br/Default.aspx>. Acesso em: março de 2018.
- LOPREATO, F. L. C. *O papel da política fiscal: um exame da visão convencional*. Campinas: Unicamp, IE, 2006. (Texto para Discussão IE/UNICAMP, n. 119).
- MOLLO, M. L. R. Ortodoxia e heterodoxia monetárias: a questão da neutralidade da moeda. *Revista de Economia Política*, v. 24, n. 3, p. 323-343, 2004.
- MONTES, G. C.; ALVES, R. C. O debate acerca dos objetivos e condução da política fiscal: uma abordagem crítica à visão convencional. *Economia e Sociedade*, v. 21, n. 2, p. 363-386, 2012.
- OREIRO, J. L. A grande recessão brasileira: diagnóstico e uma agenda de política econômica. *Estudos Avançados*, v. 31, n. 89, p. 75-88, 2017.
- PEROBELLI, F. S. *Análise das interações econômicas entre os estados brasileiros*. Tese (Doutorado em Economia). Programa de Pós-Graduação em Economia, Universidade de São Paulo. São Paulo, 2004.
- PETER, M. W.; HORRIDGE, M.; MEAGHER, G. A.; NAQVI, F.; PARMENTER, B. R. *The theoretical structure of MONASH-MRF*. Centre of Policy Studies/Impact Centre: Monash University, 1996.
- RIBEIRO, L. C. S.; DOMINGUES, E. P.; PEROBELLI, F. S.; HEWINGS, G. J. D. Structuring investment and regional inequalities in the Brazilian Northeast. *Regional Studies*, v. 52, n. 5, p. 727-739, 2018.
- SARGENT, T.; WALLACE, N. Some unpleasant monetarist arithmetic. *Quarterly Review*, v. 5, n.

3, p. 1-17, 1981.

SOUZA, K. B.; DOMINGUES, E. P. Mudanças no mercado de serviços domésticos: uma análise da evolução dos salários no período 2006-2011. *Economia Aplicada*, v. 18, n. 2, p. 317-344, 2014.

SOUZA, K. B. *Impactos econômicos da participação da mulher no mercado de trabalho brasileiro: uma análise de equilíbrio geral*. Tese (Doutorado em Economia). Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR) da Universidade Federal de Minas Gerais (UFMG). Belo Horizonte, 2015.


WOODFORD, M. Monetary policy and price level determinacy in a cash-in-advance economy. *Economic Theory*, v. 4, n. 3, p. 345-80, 1994.

WOODFORD, M. Price-level determinacy with control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy*, p. 1-46, 1995.

ORCID

Luiz Carlos de Santana Ribeiro  <https://orcid.org/0000-0001-6374-3811>

Kênia Barreiro de Souza  <https://orcid.org/0000-0002-6306-2044>

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

Anexo

Anexo 1 – Parâmetros do modelo

Continua

Elasticidades setoriais		(a)	(b)	(c)	(d)
1	Agricultura, inclusive o apoio à agricultura e a pós-colheita	0,27	1,24	1,49	0,59
2	Pecuária, inclusive o apoio à pecuária	0,27	1,24	1,49	0,59
3	Produção florestal; pesca e aquicultura	0,27	1,24	1,49	0,59
4	Extração de carvão mineral e de minerais não metálicos	0,63	0,27	0,92	1,08
5	Extração de petróleo e gás, inclusive as atividades de apoio	1,12	0,27	0,92	1,08
6	Extração de minério de ferro, inclusive beneficiamentos e aglomeração	0,63	0,27	0,92	1,08
7	Extração de minerais metálicos não ferrosos, inclusive beneficiamentos	0,63	0,27	0,92	1,08
8	Abate e produtos de carne, inclusive os produtos do laticínio e da pesca	0,73	2,03	1,49	1,08
9	Fabricação e refino de açúcar	0,73	2,42	0,80	1,08
10	Outros produtos alimentares	0,73	3,59	0,80	0,59
11	Fabricação de bebidas	0,73	1,47	0,80	0,63
12	Fabricação de produtos do fumo	0,73	1,18	0,80	0,44
13	Fabricação de produtos têxteis	0,52	3,36	0,92	0,73
14	Confecção de artefatos do vestuário e acessórios	0,33	2,23	0,38	0,73
15	Fabricação de calçados e de artefatos de couro	0,63	2,23	0,85	0,73
16	Fabricação de produtos da madeira	1,24	1,86	0,38	0,83
17	Fabricação de celulose, papel e produtos de papel	1,24	1,01	0,38	1,08
18	Impressão e reprodução de gravações	1,24	1,01	1,13	1,08
19	Refino de petróleo e coquerias	0,66	1,18	1,13	1,08
20	Fabricação de biocombustíveis	0,63	0,27	1,08	1,08
21	Fabricação de químicos orgânicos e inorgânicos, resinas e elastômeros	0,63	0,27	1,08	1,08
22	Fabricação de defensivos, desinfetantes, tintas e químicos diversos	0,63	0,27	1,08	1,08
23	Fabricação de produtos de limpeza, cosméticos/perfumaria e higiene pessoal	0,63	0,40	1,08	1,08
24	Fabricação de produtos farmoquímicos e farmacêuticos	0,63	0,40	1,08	1,08
25	Fabricação de produtos de borracha e de material plástico	1,04	2,16	1,08	1,08
26	Fabricação de produtos de minerais não metálicos	0,63	0,75	0,99	1,08
27	Produção de ferro-gusa/ferroligas, siderurgia e tubos de aço sem costura	0,63	0,57	0,95	1,08
28	Metalurgia de metais não ferrosos e a fundição de metais	0,63	0,98	0,95	1,08
29	Fabricação de produtos de metal, exceto máquinas e equipamentos	0,63	1,50	0,95	1,08
30	Fabricação de equipamentos de informática, produtos eletrônicos e ópticos	0,63	0,16	1,32	1,08
31	Fabricação de máquinas e equipamentos elétricos	1,58	0,36	1,32	1,08
32	Fabricação de máquinas e equipamentos mecânicos	1,58	1,78	1,32	1,08
33	Fabricação de automóveis, caminhões e ônibus, exceto peças	0,63	1,43	0,96	1,08
34	Fabricação de peças e acessórios para veículos automotores	0,56	0,41	1,15	1,08
35	Fabricação de outros equipamentos de transporte, exceto veículos automotores	0,56	0,41	1,15	1,08
36	Fabricação de móveis e de produtos de indústrias diversas	0,56	1,86	1,11	1,08
37	Manutenção, reparação e instalação de máquinas e equipamentos	0,46	1,24	1,11	1,08
38	Energia elétrica, gás natural e outras utilidades	0,61	1,24	0,79	1,08
39	Água, esgoto e gestão de resíduos	0,61	1,24	0,79	1,08
40	Construção	0,63	1,24	1,04	1,08
41	Comércio e reparação de veículos automotores e motocicletas	0,45	1,24	1,04	1,08
42	Comércio por atacado e a varejo, exceto veículos automotores	0,45	1,24	1,04	1,08
43	Transporte terrestre	0,63	1,24	1,04	1,03
44	Transporte aquaviário	0,63	1,24	1,04	1,03
45	Transporte aéreo	0,63	1,24	1,04	1,03

Anexo 1 – Parâmetros do modelo

		<i>Conclusão</i>			
Elasticidades setoriais		(a)	(b)	(c)	(d)
46	Armazenamento, atividades auxiliares dos transportes e correio	0,63	1,24	1,04	1,03
47	Alojamento	0,63	1,24	1,04	1,08
48	Alimentação	0,63	1,24	1,04	0,88
49	Edição e edição integrada à impressão	0,91	1,24	1,04	1,08
50	Atividades de televisão, rádio, cinema e gravação/edição de som e imagem	0,91	1,24	1,04	1,08
51	Telecomunicações	0,91	1,24	1,04	1,08
52	Desenvolvimento de sistemas e outros serviços de informação	0,91	1,24	1,04	1,08
53	Intermediação financeira, seguros e previdência complementar	0,63	1,24	1,04	1,08
54	Atividades imobiliárias	0,63	1,24	1,04	1,08
55	Atividades jurídicas, contábeis, consultoria e sedes de empresas	0,63	1,24	1,04	1,08
56	Serviços de arquitetura, engenharia, testes/análises técnicas e P&D	0,46	1,24	1,04	1,08
57	Outras atividades profissionais, científicas e técnicas	0,46	1,24	1,04	1,08
58	Aluguéis não imobiliários e gestão de ativos de propriedade intelectual	0,63	1,24	1,04	1,08
59	Outras atividades administrativas e serviços complementares	0,63	1,24	1,04	1,08
60	Atividades de vigilância, segurança e investigação	0,63	1,24	1,04	1,08
61	Administração pública, defesa e seguridade social	0,58	1,24	1,04	1,08
62	Educação pública	0,58	1,24	1,04	1,11
63	Educação privada	0,63	1,24	1,04	1,11
64	Saúde pública	0,58	1,24	1,04	0,96
65	Saúde privada	0,63	1,24	1,04	0,96
66	Atividades artísticas, criativas e de espetáculos	0,63	1,24	1,04	0,99
67	Organizações associativas e outros serviços pessoais e domésticos	0,63	1,24	1,04	0,92
Taxa de depreciação do capital		8,00%			
Razão investimento/capital		10,00%			
Taxa de retorno normal bruta		20,00%			
Elasticidade do investimento		10,00			
Elasticidade de substituição entre produtos de diferentes origens		5,00			
Elasticidade de substituição entre regiões para margens		0,20			
(a)	Elasticidades de substituição entre fatores primários				
(b)	Elasticidade de substituição entre domésticos e importados				
(c)	Elasticidade de exportação				
(d)	Elasticidade de substituição dos gastos				

Fonte: Elaboração própria.

**SUICÍDIO NO RN E SUA RELAÇÃO COM DETERMINANTES ESPACIAIS,
URBANIZAÇÃO, DESENVOLVIMENTO E OUTROS FATORES SOCIOECONÔMICOS***

Stefany Silva Amaral

Doutorando do Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco (PIMES/UFPE)

E-mail: s.amaral@outlook.com

RESUMO: Este trabalho trata o suicídio como um fenômeno genérico, com efeitos e causas eminentemente econômicas e sociais. Pretende-se inicialmente fazer uma breve exposição teórica e empírica de material levantado na literatura acerca do suicídio e seus determinantes espaciais e socioeconômicos e testar parte dessas relações propostas para os municípios do estado do Rio Grande do Norte, utilizando a Análise Exploratória de Dados Espaciais (AEDE) e a econometria espacial. Nos resultados foram observados *clusters* da taxa de suicídio e foi verificado o efeito contágio. Em um modelo SAR, foram encontradas relações diretas do suicídio com pessoas com ensino superior e com a desigualdade de renda e relação inversa com o nível de riqueza e a razão de dependência.

Palavras-chave: Taxa de suicídio; AEDE; Econometria espacial.

Classificação JEL: C21; I10; Z13.

**SUICIDE IN RN AND ITS RELATIONSHIP WITH SPATIAL DETERMINANTS,
URBANIZATION, DEVELOPMENT AND OTHER SOCIOECONOMIC FACTORS**

ABSTRACT: This work deals with suicide as a generic phenomenon with eminently economic and social effects and causes. The objective of this study is to present a brief theoretical and empirical exposition of the literature about suicide and its spatial and socioeconomic determinants and to test part of these relationships proposed for the municipalities of the state of Rio Grande do Norte, using Exploratory Spatial Data Analysis (ESDA) and spatial econometrics. In the results, clusters of the suicide rate were verified and the contagious effect was verified and in a SAR model, direct suicide relations were found with people with higher education and income inequality and inversely with the level of wealth and the dependency ratio.

Keywords: Suicide rate; ESDA; Spatial econometrics.

JEL Codes: C21; I10; Z13.

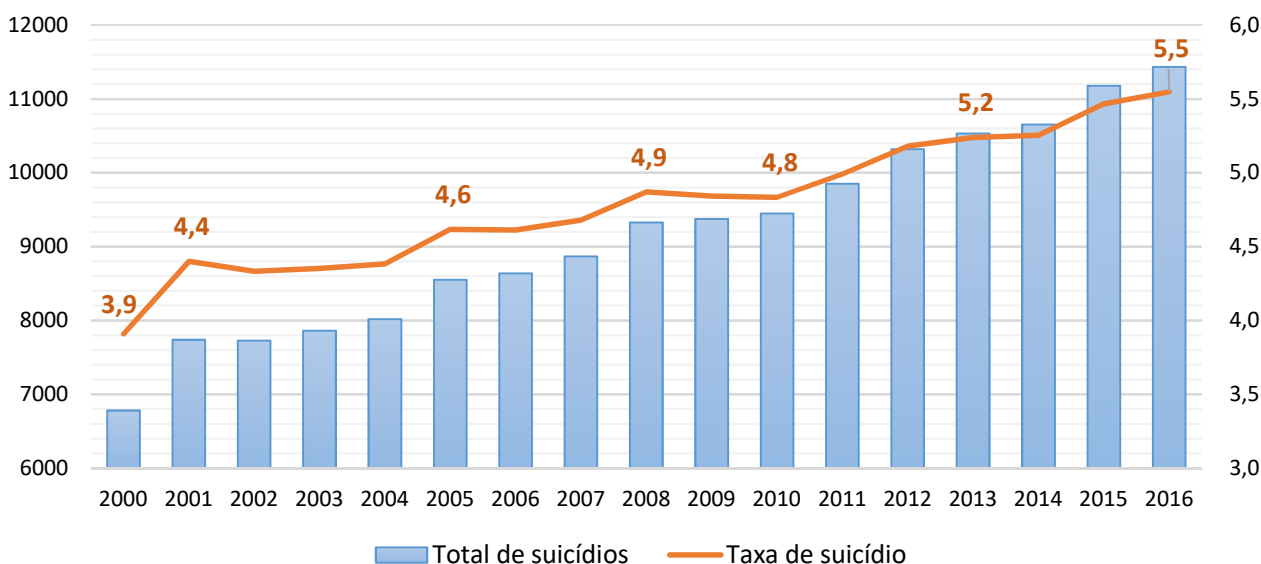
1. Introdução

O suicídio é um assunto que no meio acadêmico parece apresentar já uma larga literatura, observada não apenas na área da saúde, todavia constantemente explorado pelas ciências sociais. Em economia, a disposição de um largo ferramental empírico e profundo conhecimento das relações sociais e econômicas possibilita uma exploração produtiva dessa temática, visando à formulação de políticas e otimização da alocação de recursos da área da saúde, buscando eficiência mediante redução dos gastos e melhoria dos resultados.

No Brasil, os números do suicídio não são tão altos se comparados aos de outros países, sendo fácil constatar países desenvolvidos, ou mesmo emergentes, com taxas bem maiores (71ª posição no mundo, taxa bem mais baixa do que França, Alemanha, Reino Unido, Itália, Suíça, Bélgica, Áustria, Canadá, Estados Unidos e Japão) (SANTOS; BARBOSA, 2017). Isso não exclui o fato de as taxas serem crescentes e cada vez mais custosas aos sistemas públicos de saúde quando não passam de tentativas ou quando o óbito não é imediato.

Neste sentido, Sgobin (2013) realiza uma análise de custos, de um hospital público, referentes a internamentos decorrentes de tentativas de suicídio com alta intencionalidade e alta letalidade de método. Chega-se à conclusão de que os custos envolvidos nesses tipos de casos podem ser tão altos quanto os da doença com o maior nível de dispêndio econômico no mundo (síndromes coronarianas agudas). A Figura 1 expõe o crescimento anual do número de suicídios no Brasil entre 2000 e 2016, ocorreu uma elevação de 68,6%, vis-à-vis um crescimento populacional de apenas 18,8%. Isso equivale a uma taxa de crescimento média anual de 4%. A taxa de suicídio que era 3,9 casos para cada 100.000 habitantes, em 2000, passou para 5,5 em 2016, um aumento de aproximadamente 42% na proporção de suicídios.

Figura 1 – Evolução do total (eixo da esquerda) e da taxa de suicídio por 100.000 habitantes (eixo da direita) no Brasil de 2000 a 2016



Fonte: Datasus (mortes por suicídio) e IBGE (população).

Ademais, Durkheim (2000), que escreveu acerca do suicídio em 1897, e tantos outros pesquisadores anteriores e posteriores teorizaram e verificaram empiricamente as relações entre os suicídios de áreas específicas e fatores socioeconômicos, com Durkheim (2000) chegando a afirmar que esses fatores são preponderantes sobre todos os outros para explicar o suicídio. Assim, fica evidente a evitabilidade desse tipo de morte mediante políticas públicas.

O suicídio está entre as dez maiores causas de morte no mundo e entre as três maiores para a faixa de 15 a 35 anos, conforme expõem Santos e Barbosa (2017). No caso do Brasil, Gonçalves et

al. (2011) observam que as taxas de suicídio são crescentes e a quantidade ainda incipiente de trabalhos de pesquisa envolvendo o assunto não exaurem lacunas nesse entendimento, como as relacionadas ao efeito contágio do suicídio. Este último, segundo Durkheim (2000), é conceituado como a forma com a qual o ato de suicídio pode influenciar outros em situações psicológicas e socioeconômicas parecidas, geralmente próximas geograficamente, devendo ser explorada por técnicas de análise espacial. Gonçalves et al. (2011) afirmam que, não apenas na literatura nacional, mas também na internacional, o uso da econometria espacial para abordar esse tema é uma lacuna a ser preenchida.

Dessa forma, o objetivo desta pesquisa é explorar no Rio Grande do Norte os determinantes do suicídio no âmbito do que é enfatizado por Durkheim (2000) e Gonçalves et al. (2011). Apesar das maiores taxas de suicídio serem observadas na região Sul e Centro-Oeste, nos estados do Nordeste elas estão crescendo e, como mostram Santos e Barbosa (2017), dos dez municípios com maiores taxas de suicídio do Nordeste, quatro são do RN, sendo que três desses estão entre as cinco maiores taxas, e a maior é São José do Seridó-RN. O próprio RN é o terceiro estado em taxas de suicídio da região, cabendo, portanto, um estudo que explore mais profundamente suas causas e que forneça uma aplicação do ferramental econométrico-espacial.

Além desta introdução, a seguir é realizada uma revisão da literatura, inicialmente delineando aspectos teóricos e em seguida ressaltando outras análises empíricas na área. Na terceira seção, é explorada a metodologia da verificação empírica que se seguirá, revelando o ferramental utilizado. Na quarta seção, são expostas características referentes à base de dados trabalhada, bem como algumas análises descritivas iniciais. A quinta seção está destinada à exposição e interpretação dos resultados da pesquisa, que está dividida em, inicialmente, Análise Exploratória de Dados Espaciais (AEDE) e, por último, na análise do modelo econométrico-espacial. Ainda, são realizadas considerações finais acerca da metodologia, do modelo estimado e de sugestões derivadas dos resultados, em conclusão, na seção final.

2. Revisão da literatura

Nas subseções seguintes, será feita uma breve revisão da literatura sobre o suicídio. Na primeira, questões teóricas iniciais acerca do tipo de estudo serão delimitadas e, em seguida, as relações deste tipo serão enfatizadas com formulações teóricas compreendidas para o suicídio. Na segunda subseção, são exploradas investigações empíricas realizadas análogas a essa pesquisa, destacando aplicações nacionais e metodologicamente semelhantes.

2.1. Revisão teórica

Um primeiro ponto a ser levantado deve ser a definição do principal objeto deste estudo, o suicídio. Uma conceituação formal e rigorosa torna-se necessária à ciência para que não se incorra no risco de cair em dubiedades e má identificação do que é ou não enquadrado no objeto de estudo; tal como pode acontecer caso seja tomado apenas em seu sentido popular subjetivo, e não delineado e categorizado. Assim, Durkheim (2000) investiga as controvérsias que poderiam surgir dessa definição e chega a seguinte formulação: “Chama-se suicídio todo o caso de morte que resulta direta ou indiretamente de um ato, positivo ou negativo, realizado pela própria vítima, e que ela sabia que produziria esse resultado” (DURKHEIM, 2000, p. 14). Para além disso, no momento da análise empírica desta pesquisa, deve ser tomado um conceito ainda mais objetivo e pragmático, característico dos próprios dados, que caracteriza suicídio como a morte decorrente de lesão autoprovocada intencionalmente, cuja categorização é realizada a partir da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-10, X60-84).

Além disso, Durkheim (2000) identifica a possibilidade da ocorrência de suicídios em padrões definidos no espaço. O termo “espaço” tem um caráter problemático e complexo para se definir e para ser alcançada uma convenção ou um senso comum em torno do seu significado. Ao se observar as diversas vertentes de abordagem do espaço na ciência regional, como enfatizam Cruz et

al. (2011), é perceptível a adoção, por alguns, de uma ideia de espaço mais palpável ligado à matemática euclidiana e/ou às coordenadas geográficas, de forma mais aplicável às formulações de modelos teóricos e empíricos, enquanto outros tentam lançar mão de uma ideia de espaço mais abrangente e subjetiva que resulta de uma interação de diversos elementos.

Contudo, não será tema deste trabalho discutir questões elementares como a definição de espaço, já que, como afirmam Cruz et al. (2011), o conceito a ser usado depende do enfoque aplicado sobre o objeto de análise. Portanto, será adotada a definição de “espaços territorialmente contíguos inseridos em espaços nacionais sobre os quais há possibilidades concretas de intervenção e de levantamento de informações individualizadas” (CRUZ et al., 2011, p. 45). Tal definição será adotada, sobretudo, tendo em vista que carrega uma maior capacidade de aplicabilidade de políticas e de mensuração de seus efeitos com objetividade.

Praticamente todos os dados passíveis de estudo e referentes a algum fenômeno têm uma localização espacial, desde os dados com elementos sociais e econômicos mais comuns aos aqui estudados, em um espaço bidimensional no âmbito da geografia física, até os estudos em topografias mais abstratas e complexas, como dados de sinapses no espaço tridimensional do cérebro (ASSUNÇÃO, 2001). Seja como for, fica evidente a necessidade de constatação da localização espacial dos objetos tratados em estudo.

Entretanto, a mera característica dos dados provirem de uma localização específica não faz das informações estatísticas extraídas uma técnica de estatística espacial, assim como também dados associados ao tempo não necessariamente serão uma técnica de séries temporais, como enfatiza Assunção (2001). Para tal ser conceituado, as coordenadas espaciais devem estar no processo de coleta dos dados e também em sua descrição e análise. Assim, segundo Monteiro et al. (2004), a análise espacial deve realizar a verificação de propriedades e das características de relações do objeto de estudo, considerando de forma explícita a localização dos fenômenos estudados, incorporando o espaço no modelo utilizado.

A análise espacial pode ser definida como um “conjunto de procedimentos encadeados cuja finalidade é a escolha de um modelo inferencial que considere explicitamente os relacionamentos espaciais presentes no fenômeno” (MONTEIRO et al., 2004, p. N/A). Para que a espacialidade dos dados seja considerada em um modelo, de modo a torná-lo uma análise estatística espacial, deve ser verificado se suas características espaciais afetam as variáveis em observação, ou seja, se existe algum tipo de associação ou relação espacial significativa nos objetos de estudo.

À vista disso, no que concerne à variável suicídio, Émile Durkheim foi quem compôs uma das primeiras e mais importantes sistematizações teóricas realizadas a respeito, com o livro *Le Suicide*, publicado em 1897, no qual o autor busca explorar os fatores individuais e, principalmente, sociais que determinam o suicídio. Em Durkheim (2000), o fator espacial e o suicídio se relacionam por meio da imitação. Esse é um conceito que tem natureza psicológica e se revela quando um ato realizado sucede outro semelhante anteriormente representado, “sem que entre essa representação e a execução se intercale nenhuma operação intelectual, explícita ou implícita, sobre as características do ato reproduzido” (DURKHEIM, 2000, p.138). Assim, é enfatizado, portanto, o caráter irracional da motivação do suicídio pela imitação, em uma espécie de contágio que parece ficar mais evidente quando há proximidade espacial entre os atos.

Por mais que a imitação, como Durkheim (2000) afirma, por si só “raramente seja uma explicação suficiente” para o ato suicida, por isso mesmo chega-se à conclusão que ela não é um fator original, mas sim um fator de reforço a outros motivos. Logo, se essa relação existe, as ocorrências iniciais devem ter um centro de irradiação que influencia a propagação para regiões próximas, sendo sua distribuição perceptível geograficamente. Dessa forma, advoga-se que essa análise deve ser feita a partir do uso de mapas e métodos que o “interroguem” de maneira correta. Assim, ele o faz como exemplo, para a investigação do caso dos distritos franceses, e também para a Europa, apesar de ainda não gozar de ferramentas apropriadas, como as computacionais, e dos métodos mais avançados e eficazes como são os atuais.

Além das relações espaciais do suicídio, nesta pesquisa pretende-se verificar a hipótese de relação dessa variável com o nível de urbanização e desenvolvimento e, ademais, com outras variáveis socioeconômicas. A principal análise realizada por Durkheim (2000) é acerca dos fatores que vão adiante daqueles individuais que são considerados do campo da psicologia. Busca-se, portanto, explorar elementos que influenciam a realização do suicídio de uma forma genérica, conjunta, como característica de uma sociedade. A parcela de suicídios cometidos preponderantemente por causa destes fatores é o que passa a se chamar de taxa social do suicídio.

Assim, o autor numera variados tipos de causas sociais para o suicídio e constrói uma análise que evidencia os mecanismos pelos quais funcionam cada um desses fatores. A sistematização de Durkheim (2000) é realizada de tal forma que classifica os tipos de suicídio relacionando-os com suas causas sociais, nomeadamente: o suicídio egoísta, que está relacionado ao maior nível de independência, exclusão e isolamento social; o altruísta, que ressalta a integração e espírito de corpo social; e o anômico, relativo a questões patrimoniais, conjugais e a fatos específicos.

Mais especificamente, Durkheim (2000) observa que o aumento da independência pessoal de um indivíduo é um fator influenciador para cometer o suicídio e que um nível financeiro mais elevado de renda e um alto grau de educação são elementos que contribuem para essa autonomia. Por outro lado, uma situação mais crônica de carência financeira e educacional podem tornar os indivíduos mais resignados. Dessa forma, o nível de desenvolvimento de uma sociedade parece resumir de forma adequada fatores que levam a uma maior independência dos indivíduos no geral. Outro caso é a relação com a urbanização, nesse caso, afirma-se que, em zonas urbanas, aparentemente, ocorrem mais casos do que em zonas rurais, inevitavelmente ainda apresentando uma relação espacial, abreviada pela taxa de urbanização. Exploram-se ainda vários outros fatores: avanço da idade, sexo masculino e fatos específicos como guerras, fenômenos naturais e crises econômicas que têm uma relação positiva com o suicídio; elementos que levam à integração social e ao fortalecimento de laços familiares, e religiosidade, sobretudo quando essa é relacionada a um maior controle moral, apresentam relações inversas.

2.2. Revisão empírica

Podem se encontrados poucos trabalhos que buscam verificar a imitação como um determinante do suicídio, sobretudo observando o Brasil, apesar da existência de alguns que buscam verificar a relação com outras variáveis. Como já foi visto em Durkheim (2000), a verificação do contágio deve ser feita via observação de mapas e com métodos apropriados para sua análise. Atualmente, o método mais apropriado para isso é, certamente, a Análise Exploratória de Dados Espaciais (AEDE) e a econometria espacial.

Neste sentido, Gonçalves et al. (2011) realizam uma profusa análise espacial do suicídio para o Brasil até o ano 2000, evidenciando o efeito contágio nas microrregiões brasileiras a partir da metodologia de AEDE e econometria espacial. Verifica-se um índice de autocorrelação espacial, *I* de Moran, para o suicídio sempre positivo e significativo que está em torno de 0,52 e 0,57. Além de perceber, a partir de indicadores locais (LISA), a presença de *clusters* com padrão de altas taxas de suicídio no sul e parte do centro-oeste. Na modelagem econométrica, para as variáveis explicativas, fez-se o uso de uma análise fatorial que condensou 17 variáveis em 4 fatores, que compreendem os temas pobreza, grau de ruralização, ensino superior e mulher. Seus resultados sugerem que a defasagem espacial significativa confirma o efeito contágio, a pobreza tem uma relação inversa com o suicídio, o grau de ruralização tem um efeito direto e as variáveis relativas ao sexo e à educação não parecem ter relevância na explicação.

Em outro estudo, também para todo o Brasil, mas agora considerando os municípios como unidades espaciais e o ano de 2010 como unidade de tempo, Dantas et al. (2018), a partir do uso da AEDE, examinam o efeito contágio por medidas de associação espacial e encontram um *I* de Moran que equivale à, aproximadamente, metade do encontrado por Gonçalves et al. (2011). Além disso, a partir de um *I* de Moran global e índice LISA bivariados, e mais regressões lineares bivariadas, verificou-se a relação entre o suicídio e 9 variáveis socioeconômicas apanhadas do Atlas de

Desenvolvimento Humano da PNUD, tais como IDHM, vulnerabilidade, razão de dependência e desemprego, concluindo que o suicídio tem uma fraca autocorrelação espacial, bem como uma pouco observada correlação espacial, e não espacial, com as variáveis socioeconômicas. Metodologicamente semelhante, ao explorar os conglomerados de suicídio, Santos e Barbosa (2017) fazem o mesmo para o Nordeste brasileiro, com exceção das regressões lineares bivariadas, e obtêm resultados parecidos, concluindo que o suicídio tem uma distribuição espacial aleatória e não tem relação espacial com variáveis socioeconômicas.

Não foram observados outros trabalhos que usam metodologias de análise espacial, apesar do efeito contágio ter sido observado também por Cutler et al. (2001), o qual percebe que esse efeito é pouco evidente em adultos e idosos, sendo uma provável justificativa para menor importância das relações sociais para esse grupo. Ao contrário dos suicídios ocorridos e tentados pelos mais jovens que, segundo casos específicos investigados pelo autor, o contágio é mais evidente.

Loureiro et al. (2010) não analisam o efeito contágio, mas buscam evidenciar a relação do suicídio com fatores econômicos. Eles observam que as taxas de suicídio têm uma maior relação com idades mais baixas e, ainda mais, variáveis econômicas têm maior impacto sobre o suicídio em indivíduos mais jovens. Esse resultado é encontrado com uma metodologia de dados em painel para os estados brasileiros, de 1981 a 2006, utilizada para verificar a influência de variáveis econômicas sobre o suicídio. Seus achados concordam com a tese verificada de que a felicidade está atrelada ao consumo relativo e não ao consumo absoluto, ao observar que a desigualdade tem uma relação direta e a pobreza tem uma relação inversa.

Com dados *cross-section*, Shikida et al. (2007) também avaliam o suicídio nos estados brasileiros, contudo partindo da teoria econômica do suicídio e estimando um modelo baseado em formulações teóricas estritamente estabelecidas. Eles concluem que a evidência empírica tem o ajuste esperado ao modelo teórico, verificando uma relação cúbica com a idade, *dummies* para a região sul e centro-oeste apresentaram grande relevância, bem como os gastos com saúde. Por outro lado, Schnitman et al. (2011), com uma análise de correlação e regressão clássica, não encontram muitas evidências de relações com fatores socioeconômicos e concluem que fatores biopsicológicos devem ser preponderantes na motivação do suicídio.

Em uma análise para os municípios do Brasil e do Rio Grande do Sul, especificamente este estado por ter os municípios com maiores taxas de suicídio, para o ano 2010, Fraga et al. (2016) utilizam ferramentas, como a Regressão de Poisson, Binomial e Quantílica, para verificar a influência de variáveis socioeconômicas sobre o suicídio. Foram observadas relações diretas com a pobreza, a cor/raça branca e indígena e, destacadamente, o nível de dependência agropecuária. Este último o autor associa às transformações recentes ocorridas na área rural dos municípios brasileiros e às questões econômicas, como a pobreza, a mudança nas formas de produção (máquinas e produtos químicos), o endividamento, a concentração da terra, o êxodo rural e o isolamento desse estrato social.

Dutra e Roehe (2013) realizaram pesquisa qualitativa sobre os cinco municípios com maiores taxas de suicídios entre agricultores do Rio Grande do Norte, a partir da aplicação de questionários a pessoas que vivem da agricultura nesses locais e são parentes de vítimas de suicídio. O produto de suas entrevistas comprova parte das conclusões tiradas por Fraga et al. (2016) acerca da condição do indivíduo no meio rural, o qual passa por situações degradantes do ponto de vista social e econômico. Para além disso, Dutra e Roehe (2013) interpretam as informações à luz da filosofia e psicologia, em uma abordagem existencial.

Por fim, interessa como recomendação para o aprofundamento empírico acerca do suicídio, sobretudo considerando aplicações em diversas partes do mundo, o artigo de revisão de Rehkopf e Buka (2006), que faz uma vasta análise da literatura empírica existente até então, acerca da associação do suicídio com características socioeconômicas de áreas geográficas. A revisão é feita tomando 221 análises encontradas, consideradas de alta qualidade, estratificando por características, como o tamanho da área agregada, as medidas socioeconômicas, a região de estudo e o desenho da modelagem e, por conseguinte, analisando o percentual de tipo de resultado, como positivo,

negativo e nulo. Como conclusão, há a confirmação da grande heterogeneidade de associações encontradas, conseqüente dos discrepantes resultados encontrados em diversos estudos, apesar de que tendências puderam ser observadas em conformidade com as estratificações usadas.

3. Metodologia

Monteiro et al. (2004) observam que, geralmente, a modelagem é precedida do que é comumente chamado de análise exploratória de dados espaciais (AEDE), a qual busca apresentar visualmente em gráficos e mapas os padrões espaciais e as relações de dependência. Neste trabalho, além de estatísticas descritivas e *box plots*, uma AEDE desenvolvida fará uso de diagramas de dispersão (*Moran scatterplot*), indicadores de autocorrelação espacial e análise de *clusters* espaciais. Os dois procedimentos mencionados inicialmente ainda não consideram os fatores espaciais, mas são capazes de fornecer uma ideia inicial de características importantes das observações. Seguindo Anselin (2005), os últimos são formas gráficas de observar medidas de tendência central e distribuição espacial dos dados e verificar a existência de *outliers* e pontos de alavancagem¹.

O *Moran scatterplot*, mais especificamente, é a primeira ferramenta utilizada, nesta pesquisa, que irá considerar o fator espaço, já que plota em um plano cartesiano a dispersão dos desvios da variável analisada contra sua defasagem espacial. Nos seus quadrantes, é possível identificar o padrão de associação em quatro categorias, alto-alto (AA), baixo-baixo (BB), alto-baixo (AB) e baixo-alto (BA), que indicam observações que têm valores altos (ou baixos) rodeados por valores baixos (ou altos) e vice-versa. Além de existirem *outliers* (nos quadrantes AB e BA) e pontos de alavancagem (nos quadrantes AA e BB), tipicamente categorizadas quando observações apresentam valores maiores que dois desvios-padrão (ANSELIN, 2005).

Anteriormente a isso, deve ser identificada a matriz de ponderação espacial W de dimensão $n \times n$, definida por Anselin e Rey (2014), como o componente mais importante da análise *cross-section* de dependência espacial, pois expressa a estrutura de vizinhança entre as observações. Apesar das diferentes formas possíveis de obter padrões de vizinhança, a construção de W realizada será baseada na relação binária expressa no padrão *queen* de primeira ordem, que considera vizinhas de uma unidade espacial todas as outras que compartilham com a primeira a mesma borda ou vértice, ou seja, regiões que fazem qualquer tipo de fronteira entre si. Assim, cada termo de peso espacial w_{ij} , da matriz W terá valor 1 sempre que i for vizinho de j e 0 caso contrário. Por fim, a matriz W considerada será a normalizada, em que o valor de cada w_{ij} será dividido pelo somatório de sua própria linha. Há ainda a possibilidade do uso de matrizes de ponderação que identificam padrões diferentes de vizinhança, tais como a matriz *root*, que considera como vizinhas apenas aquelas regiões que dividem o mesmo vértice de fronteira, as matrizes de distância, que avaliam como vizinhos aqueles mais próximos dentro de uma distância definida arbitrariamente ou, também, as matrizes de k -vizinhos, que consideram apenas como regiões vizinhas as k mais próximas, sendo k também arbitrário.

Após a obtenção da matriz de pesos espaciais, ainda segundo Anselin e Rey (2014), é possível obter o vetor das defasagens espaciais da variável objetivo, y , com o somatório do produto de suas observações pelos pesos espaciais, w_{ij} , para cada unidade espacial. Em suma, será obtido através de Wy e isso irá descrever, portanto, uma soma ou média ponderada dos valores vizinhos para essa variável. Isto posto, torna-se possível realizar todas as verificações necessárias neste trabalho que consideram o fator espaço.

¹ Torna-se necessário esse tipo de análise, dada a relevante influência que observações discrepantes podem exercer sobre o modelo, deturpando estimativas e testes. Os *outliers* espaciais são observações que destoam do resto em magnitude e demonstram relação inversa entre as vizinhanças, não seguindo o mesmo padrão de dependência espacial. São definidas, portanto, nas extremidades dos quadrantes de padrão alto-baixo e baixo-alto. Já os pontos de alavancagem seguem o mesmo padrão de associação espacial, mas exercem influência exacerbada sobre a determinação da medida de correlação espacial da amostra, por isso, são identificados nas extremidades dos quadrantes alto-alto e baixo-baixo.

Feito isso, será possível obter o nível de associação espacial dos dados. Como medida global, será considerada a estatística I de Moran, que mede o grau de autocorrelação da variável no espaço. Seu valor pode variar entre -1 e 1, quanto maior, mais positivamente autocorrelacionada será a variável; caso negativo, a variável terá relação inversa com suas observações vizinhas e valores próximos a zero indicarão pouca associação espacial, ou aleatoriedade. Seguindo Anselin (1995), e apenas transformando sua notação para a abordagem vetorial, o valor de I deve ser descrito pela Equação 1.

$$I = \frac{n}{S_0} \frac{z'Wz}{z'z} \quad (1)$$

Nessa equação, o número de observações, ou unidades espaciais, é dado por n . O valor de S_0 representa a soma de todas linhas e colunas de W , cujo resultado, neste caso, é um escalar igual a n , já que a matriz de pesos espaciais é normalizada nas linhas. O vetor z de dimensão n é identificado como os desvios da variável y em relação à média ($y_i - \bar{y}$). A validação estatística é baseada em um teste de pseudo-significância, que a partir de permutações aleatórias recalcula o índice I muitas vezes e gera uma distribuição de referência, dessa forma, o valor calculado pode ser comparado com a distribuição e computada uma pseudo-significância (ANSELIN, 2005).

Contudo, o índice global expõe apenas uma única medida para o conjunto de dados como um todo. Cabe, portanto, ressaltar a existência de padrões de associação local, que podem ter um padrão espacial estatisticamente significativo e diferente do exposto pelo indicador global. Para tanto, Anselin (1995) apresenta os indicadores LISA (*Local Indicator of Spatial Association*) que, a partir da decomposição dos indicadores globais, I de Moran neste caso, são capazes de fornecer informações acerca de *clusters* espaciais e seus padrões, fornecendo uma espécie de indicador para cada unidade espacial, conforme descrito na Equação 2.

$$I_i = z_i \sum_j w_{ij} z_j \quad (2)$$

Para obter um indicador de associação espacial entre duas variáveis, deve ser considerada a relação dos valores da segunda variável nos vizinhos com a primeira variável na unidade espacial em questão. Anselin et al. (2002) demonstram que para obter isso basta substituir na Equação 1 a defasagem espacial dos desvios de y (Wz , agora com z identificado pelos subscritos das duas variáveis) pela defasagem dos desvios da nova variável x (Wz_x), conforme a Equação 4. De maneira análoga, isso pode ser feito para os indicadores locais (Equação 4).

$$I_{yx} = \frac{n}{S_0} \frac{z_y' W z_x}{z_y' z_y} \quad (3)$$

$$I_{yx}^i = z_y^i \sum_j w_{ij} z_x^j \quad (4)$$

Dada a dificuldade de observar os resultados dos indicadores de associação local, sua interpretação fica facilitada com a plotagem do indicador de cada unidade espacial em um mapa. Portanto, a análise desses resultados é realizada, seguindo Anselin (2005), com o mapa de *clusters*, que identifica em cada unidade espacial o padrão de vizinhança, classificada conforme o observado em um diagrama de dispersão, com os quadrantes AA, BB, AB e BA, desde que tenha significância estatística. Ademais, também pode ser observado um mapa de significância que expõe o padrão de pseudoníveis de significância estatística gerado a partir de permutações, de forma análoga ao realizado ao I de Moran global.

A fim de realizar a modelagem para a variável estudada e verificar a relação entre essa e outras variáveis socioeconômicas, recorre-se ao modelo de regressão para dados *cross-section*, no caso, um modelo clássico estimado pelo método dos mínimos quadrados ordinários (MQO), representado como na Equação 5, na qual y é explicado por X , que representa uma matriz de covariáveis associada a um vetor de parâmetros β e mais um vetor de componentes aleatórios ε , normalmente distribuídos com média zero e variância constante $\sigma^2 I$.

$$y = X\beta + \varepsilon, \quad \text{onde: } \varepsilon \sim N(0, \sigma^2 I) \quad (5)$$

Como expõe LeSage e Pace (2009), para que esse modelo seja válido, é necessária a pressuposição observada sobre ε , a qual implica que todos os termos ε_i para cada unidade espacial i sejam independentes dos demais ($E(\varepsilon_i \varepsilon_j) = 0$) e não correlacionados com X ($E(\varepsilon|X) = 0$). No entanto, torna-se forçoso considerar essa simplificação caso o processo gerador dos dados analisados reflitam um contexto de dependência espacial, que poderia fazer com que os estimadores tivessem problemas de viés, inconsistência e/ou ineficiência. Consequentemente, cabe uma investigação sobre os erros da regressão para verificar a existência de processos desse tipo.

Para tal fim, pode ser aplicado um teste de Moran sobre os resíduos para verificar a existência de autocorrelação espacial. Se o I de Moran se mostrar estatisticamente significativo, deve ser definido um novo modelo de estimação que considere o processo de geração específico dos dados. O modelo a ser definido neste trabalho faz parte do que Monteiro et al. (2004) classificam como aqueles que “tratam a estrutura espacial de forma global”, isto é, com um único parâmetro, sobretudo SAR e SEM, e não com parâmetros que variam continuamente no espaço. Nas Equações 6, 7 e 8, podem ser observadas suas especificações.

$$y = \rho W y + X\beta + \varepsilon \quad (6)$$

Em um modelo de defasagem espacial autorregressivo (*Spatial Autoregressive Model - SAR*) observado na Equação 6, como é exposto por LeSage e Pace (2009), é modelado o efeito de dependência espacial inerente à variável dependente, portanto, é adicionado como variável explicativa o termo de defasagem espacial de y , representado pelo vetor $W y$, cujo efeito sobre a variável dependente é materializado pelo coeficiente ρ .

$$y = X\beta + \mu, \quad \text{onde: } \mu = \lambda W \mu + \varepsilon, \quad \text{logo:} \quad (7)$$

$$y = X\beta + (I_n - \lambda W)^{-1} \varepsilon$$

No modelo de erro espacial (*Spatial Error Model - SEM*), tenta-se modelar efeitos espaciais não incorporados ao modelo, como os de variáveis autocorrelacionadas espacialmente não consideradas, de forma que o efeito se apresenta como uma perturbação no termo de erro. Portanto, na Equação 7, na qual I_n é uma matriz identidade, é possível observar que o termo de erro μ , além de conter o erro aleatório ε , apresenta um termo autorregressivo μ cujo efeito é captado pelo coeficiente λ , que passa a ser modelado. Os processos observados nos modelos SAR e SEM também podem aparecer juntos e serem considerados em um único modelo (*spatial autoregressive confused - SAC* ou o *spatial auto-regressive and moving average - SARMA*), especificado na Equação 8.

$$y = \rho W y + X\beta + (I - \lambda W)^{-1} \varepsilon \quad (8)$$

Para definição do modelo espacial, será seguida a abordagem clássica do procedimento proposto por Florax et al. (2003) que consideram as seguintes etapas como critério de seleção: (1) estima-se o modelo clássico por MQO; (2) testa-se a hipótese de dependência espacial a partir dos

testes de Multiplicador de Lagrange (LM), descritos na Equação 9², na qual LM_ρ designa a estatística que tem como hipótese alternativa o uso da defasagem espacial de y no modelo, e LM_λ , que aponta o uso do modelo de erro espacial caso a hipótese nula seja rejeitada; (3) caso ambos os testes não sejam significantes, estima-se o modelo clássico por MQO; (4) senão, caso ambos sejam estatisticamente significantes, estima-se a especificação do que for apontado como mais significativo; (5) por outro lado, se apenas LM_ρ for significativa, estima-se o modelo SAR; (6) caso contrário, se nem isso acontecer, estima-se o modelo SEM. Além disso, se for conveniente, pode ser utilizada uma estratégia de especificação híbrida que incorpora as formas robustas do teste LM , sobretudo como critério de desempate na etapa 4.

$$LM_\rho = \frac{(\hat{\varepsilon}'Wy/\hat{\sigma}^2)^2}{nJ}, \quad LM_\lambda = \frac{(\hat{\varepsilon}'W\hat{\varepsilon}/\hat{\sigma}^2)^2}{T} \quad (9)$$

Como já foi afirmado, a estimação de modelos espaciais por MQO pode levar a estimadores viesados, inconsistentes e/ou ineficientes, devido a problemas de endogeneidade envolvendo as variáveis espaciais, Wy e $W\mu$. Como alternativa, a literatura largamente sugere o uso do método da máxima verossimilhança (MV), que é capaz de gerar estimadores ρ e λ com propriedades de consistência, eficiência e normalidade assintótica, apesar de também requerer a resguarda de alguns pressupostos, como a normalidade dos resíduos (LESAGE; PACE, 2009).

Após a estimação, a verificação de alguns testes é requerida. Para diagnosticar os resíduos, será aplicado o teste Jarque-Bera, cuja hipótese nula é a normalidade dos resíduos (para o MQO) e os testes Breusch-Pagan (também para MV) e Koenker-Bassett, com hipótese nula de homocedasticidade. Também será possível testar a especificação do modelo espacial, tendo como hipótese alternativa a preferência pelo parâmetro espacial. No caso do uso de variáveis instrumentais (VI), poderá ser verificado o esgotamento dos efeitos espaciais no modelo pelo teste Anselin-Kelejian, cuja hipótese nula é de que não há mais autocorrelação espacial nos resíduos do modelo (ANSELIN, 2005).

4. Base de dados

Para a análise realizada, foram utilizados dados *cross-section* colhidos da base de dados socioeconômicos do Atlas de Desenvolvimento Humano do Programa das Nações Unidas para o Desenvolvimento (PNUD), que faz uso do Censo Demográfico de 2010. Portanto, todas as variáveis extraídas dessa base serão referentes a esse período e serão utilizadas como covariáveis. Os dados de população relacionados à taxa de suicídio são também dessa fonte. Porém, os dados acerca do número de suicídios são extraídos da base do Datasus, que tem como fonte o Sistema de Informações de Mortalidade (SIM), referente aos anos de 2006 a 2015.

As outras variáveis utilizadas na análise, presentes no Quadro 1, foram selecionadas buscando captar características socioeconômicas dos municípios que tenham uma possível relação com a taxa de suicídio e em relevante parte tendo como embasamento a literatura teórica e empírica no processo de seleção.

Assim, como motivação para utilização, variáveis como o índice de desenvolvimento e de Gini, a renda e a proporção da população com ensino superior completo, que são vistos como elementos que levam à maior independência dos indivíduos, são de largo uso na literatura (GONÇALVES et al., 2011; DANTAS et al., 2018; SANTOS; BARBOSA, 2017; LOUREIRO et al., 2010; SHIKIDA et al., 2007; SCHNITMAN et al., 2011; FRAGA et al., 2016) a urbanização atende a um critério de viés bem mais teórico, como é elucidado em Durkheim (2000); bem como a fecundidade, vista por Durkheim (2000) como um elemento que fortalece os laços familiares, e o decil dos ricos, não

² No qual $\hat{\sigma}^2$ e $\hat{\varepsilon}$ são, respectivamente, a variância e os resíduos estimados, T é o traço da matriz $((W'+W)W)$ e J é um escalar que capta relações entre W , X , $\hat{\sigma}^2$ e os estimadores de MQO.

encontrado em outros trabalhos empíricos, mas que quantifica o quão alto é o nível de riqueza das classes mais abastadas enfatizadas por Durkheim (2000) como mais propensas ao suicídio. A taxa de desemprego (a qual esperava-se ter uma relação inversa com a taxa de suicídio) também é verificada por Dantas et al. (2018), Santos e Barbosa (2017), Loureiro et al. (2010) e Fraga et al. (2016). Ademais, como alternativa ao Gini, a desigualdade é dada pelo índice de Theil (usado por Schnitman et al. (2011)), razão de dependência (usado por Schnitman et al. (2011), Dantas et al. (2018) e Santos e Barbosa (2017)) e pobreza (usado por Schnitman et al. (2011), Loureiro et al. (2010), Gonçalves et al. (2011), Fraga et al. (2016), Dantas et al. (2017) e Santos e Barbosa (2017)).

Quadro 1 – Descrição das variáveis

Variável	Descrição
Taxa de Suicídio	Taxa da média anual de óbitos ocorridos entre 2006 e 2015 para cada 100.000 habitantes (classificação do óbito pela CID-10: X60-X84).
Taxa de urbanização	Razão entre a população urbana e a população total.
IDH-M	Índice de Desenvolvimento Humano Municipal.
Renda <i>per capita</i>	Renda <i>per capita</i> média de residentes em domicílios particulares permanentes.
Taxa de desocupação	Taxa de desocupação da população de 18 anos ou mais de idade.
Taxa de fecundidade	Número médio de filhos que uma mulher deverá ter ao terminar o período reprodutivo.
Índice de Gini	Grau de desigualdade na distribuição de indivíduos segundo a renda <i>per capita</i> .
Ensino superior	Proporção da população de 25 anos ou mais com ensino superior completo.
Índice de Theil	Desigualdade na distribuição de indivíduos segundo a renda domiciliar <i>per capita</i> .
Razão de dependência	Razão entre os <15 e >65 anos em relação aos >15 e <64 anos.
Pobreza	% vulneráveis à pobreza (menos de 1/2 salário mínimo de renda domiciliar <i>per capita</i>).
Decil dos ricos	Renda <i>per capita</i> mínima do décimo mais rico.

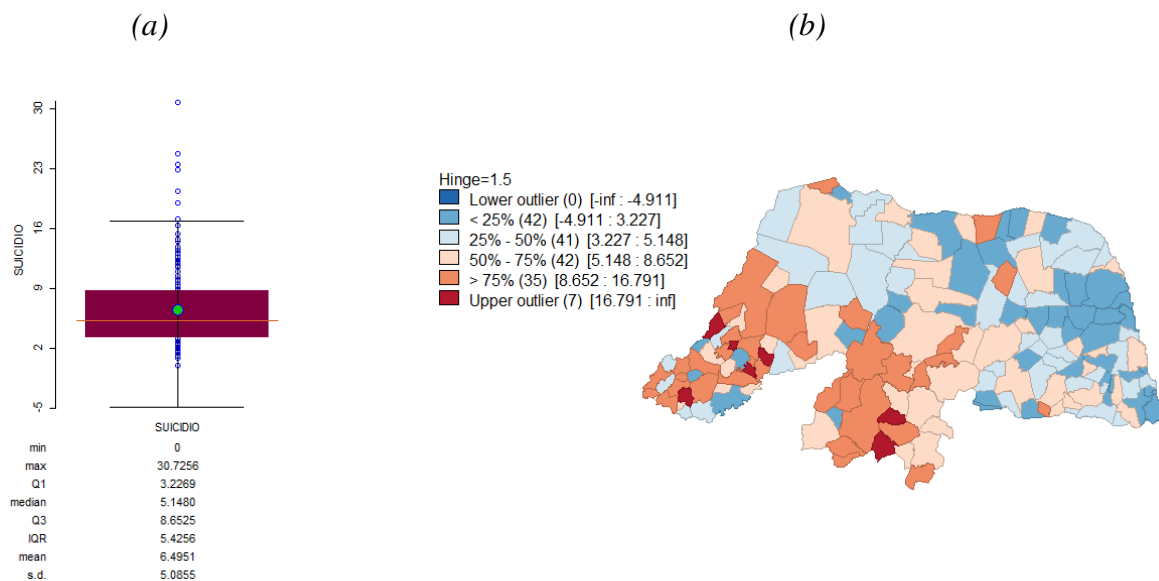
Fonte: Elaboração própria, com base nas descrições da PNUD e Datasus.

No *box plot* da taxa de suicídio, apresentado na Figura 2, é possível notar visualmente o grau de dispersão da variável. A caixa, que contém 50% das observações mais prováveis, ou seja, 83 municípios, está delimitada pelas taxas de suicídio nos valores de 3,22 (1º quartil - Q1) e 8,65 (3º quartil - Q3). A linha laranja representa o município de Serra Negra do Norte, que indica o valor mediano 5,14, ou seja, que divide a amostra da variável ao meio. A média observada para os municípios é de 6,49 suicídios para cada 100.000. O valor mínimo foi observado em 11 municípios, que não relataram nenhum caso de suicídio, e o valor máximo foi observado no município de São José do Seridó, com uma taxa de 30,72.

Além disso, acima da linha superior do *box plot*, podem ser constatadas 7 observações discrepantes caracterizadas como *outliers* altos, nomeadamente: José da Penha (17,04), Frutuoso Gomes (18,90), Rodolfo Fernandes (20,37), Rafael Godeiro (22,85), Ouro Branco (23,41), Viçosa (24,72) e São José do Seridó (30,72). O último foi verificado também por Santos e Barbosa (2017) como a maior taxa de suicídio do Nordeste. Esses municípios podem ter suas localizações identificadas pelos municípios de cor marrom escura do *box map* na Figura 2, que apresenta informações semelhantes ao *box plot*, mas localizadas geograficamente.

Quanto à espacialidade das informações, os dados caracterizam cada um dos municípios do Rio Grande do Norte como observações e unidades espaciais, se tratando, portanto, de um estudo ecológico, em que a unidade de análise é uma população ou grupo de indivíduos, e não o indivíduo propriamente dito. Para esse tipo, Gonçalves et al. (2011) mencionam a possibilidade de se incorrer na “falácia ecológica”, na qual a generalização pode ser indevida, eles alertam que não necessariamente as relações refletirão a situação dos indivíduos especificamente. Contudo, além de esses serem os dados disponíveis, são também utilizados para formulação de políticas públicas na área da saúde, embasando seu potencial de relevância.

Figura 2 – Plotagem e distribuição da taxa de suicídio dos municípios em quartís: *Box Plot* (a) e *Box Map* (b)



Fonte: Elaboração própria.

A descrição mais específica de cada variável é fornecida no Quadro 1, mais especificamente com relação a taxa de suicídio, o procedimento adotado foi, inicialmente, a realização de uma média anual do valor absoluto de suicídios³ para cada município, ocorridos nos dez anos circundantes a 2010, isto é, entre 2006 e 2015. Assim, a taxa foi obtida pela razão entre esta média e a população municipal referente a 2010 multiplicada por 100.000.

Desta forma, a taxa de suicídio deve ser interpretada como a quantidade média anual de suicídios ocorridos a cada 100.000 habitantes para cada município. Com esse procedimento, de uma média que abarca dez anos, procurou-se amenizar os problemas de grandes flutuações anuais e de eventos raros em populações pequenas mencionados por Fraga et al. (2016), já que poucos municípios possuem uma população maior do que 100.000 habitantes.

Para um melhor esclarecimento das análises dos mapas expostos, serão muitas vezes mencionadas as mesorregiões e microrregiões do Rio Grande do Norte. Assim, no caso das quatro mesorregiões (e respectivas microrregiões em parênteses), essas podem ser identificadas como Oeste Potiguar (Mossoró, Chapada do Apodi, Médio Oeste, Vale do Açu, Serra de São Miguel, Pau dos Ferros, Umarizal), que inclui toda porção oeste do estado até parte do centro e final do norte do litoral; Central Potiguar (Macau, Angicos, Serra de Santana, Seridó Ocidental, Seridó Oriental), toda a região que se inicia ao sul do estado que é ladeado por porções da Paraíba e vai até o lado norte, onde encontra o litoral; Leste Potiguar (Litoral Nordeste, Macaíba, Natal, Litoral Sul), localizado em toda região litorânea leste; e, por fim, Agreste Potiguar (Baixa Verde, Borborema Potiguar, Agreste Potiguar), que fica entre o Central Potiguar e Leste Potiguar, sendo a única região sem contato com o mar.

Observando o *box map*, fica perceptível a predominância de baixos valores nas mesorregiões do Leste Potiguar e do Agreste Potiguar, com essas contendo boa parte dos valores dos dois primeiros quartís, isto é, com a maioria de suas observações com taxas de suicídio abaixo de 5,14. De forma oposta, valores mais altos predominam nas mesorregiões Central Potiguar e Oeste Potiguar, sobretudo nas microrregiões da Serra de São Miguel, Pau dos Ferros, Umarizal, Médio Oeste, Chapada do Apodi, Seridó Ocidental e Oriental e Serra de Santana. Em suma, visualmente, é

³ Óbitos decorrentes de lesões autoprovocadas intencionalmente categorizadas a partir da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-10, X60-84).

bastante perceptível o aumento das taxas de suicídio espacialmente do litoral em direção ao interior do estado.

Também serão analisados de forma mais específica, na Análise Exploratória de Dados Espaciais (AEDE), o Índice de Desenvolvimento Humano Municipal (IDHM) e a taxa de urbanização, buscando verificar a hipótese de relação entre estas variáveis e a taxa de suicídio.

5. Resultados

Nas próximas subseções, serão explorados os resultados das inferências realizadas.⁴ Em um primeiro momento, serão observadas as medidas de associação espacial, buscando investigar o efeito contágio relacionado aos casos de suicídio nos municípios do Rio Grande do Norte e as relações espaciais deste com os níveis de desenvolvimento e urbanização. Em seguida, proceder-se-á com a modelagem econométrica-espacial, buscando verificar o efeito de variáveis socioeconômicas, conforme a melhor estratégia de estimação for constatada.

5.1. Análise Exploratória de Dados Espaciais

Para uma análise do efeito contágio do suicídio, a primeira estatística a ser analisada é o I de Moran global, pois ele fornece um nível de associação geral entre as unidades espaciais para a variável, considerando a matriz *queen* de ponderação espacial. Dessa forma, foi constatado um valor I igual a 0,2686 para a taxa de suicídio entre os municípios do Rio Grande do Norte, aproximadamente igual ao encontrado por Dantas et al. (2018) e Santos e Barbosa (2017) para os municípios do Brasil e do Nordeste, respectivamente. Uma autocorrelação espacial considerada baixa, contudo, estatisticamente significativa com um pseudo p -valor de 0,001 gerado a partir de 999 permutações em que o I foi recalculado. O valor positivo da autocorrelação espacial, mesmo que baixo, evidencia a existência de um efeito contágio do suicídio entre os municípios, cujos padrões locais devem ser investigados.

Na Figura 3, é observado o Moran *scatterplot*, ou gráfico de dispersão de Moran, em que no eixo X apresentam-se os desvios da taxa de suicídio contra a defasagem espacial desses desvios no eixo Y, conforme a matriz de vizinhança *queen*, escolhida baseando-se nos trabalhos de Gonçalves et al. (2011), Dantas et al. (2018) e Santos e Barbosa (2017). No gráfico central, é possível constatar no primeiro quadrante seis pontos de alavancagem (José da Penha, Frutuoso Gomes, Rafael Godeiro, Ouro Branco, Viçosa e São José do Seridó), que são observações com altas taxas de suicídio discrepantes, rodeadas por uma vizinhança que também tem altos valores, considerando a média. No quarto quadrante, é observado um *outlier* espacial, o município de Rodolfo Fernandes, que tem uma alta taxa de suicídio discrepante, rodeado por municípios com baixas taxas.

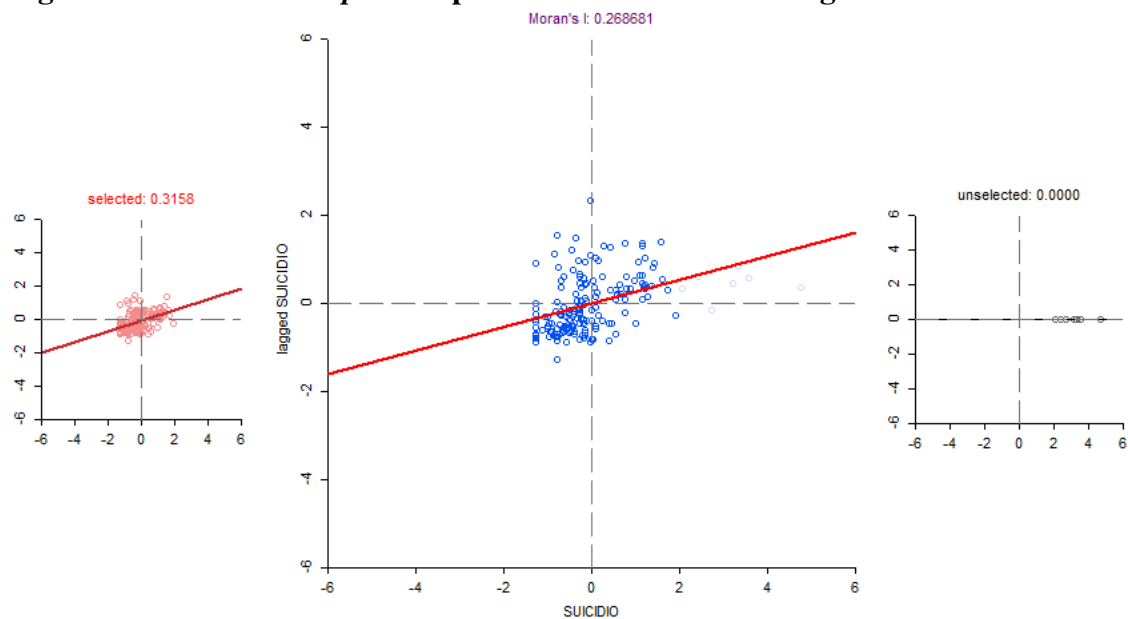
Juntos, esses sete municípios equivalem aos *outliers* globais encontrados no *box plot* e *box map* da Figura 2. Se fossem excluídos da amostra, como no gráfico com pontos vermelhos à esquerda do gráfico central da Figura 3, seria obtido um I de Moran global no valor de 0,3158. A reta vermelha presente nos gráficos deve ser entendida como uma reta de ajuste descrita por $z = \beta Wz + \varepsilon$, no qual $I = \hat{\beta}$, logo o I de Moran de 0,2686 do gráfico central será a inclinação dessa reta de regressão (ANSELIN et al., 2002).

Além disso, também foi avaliado o I de Moran considerando diferentes matrizes de ponderação. Como exemplo, a matriz *rook* rendeu um I significativo de 0,28, bem próximo do gerado usando a matriz *queen*. As matrizes que consideram vizinhos os municípios em raios de 60 km, 100 km e 150 km resultaram em correlações espaciais de, respectivamente, 0,29, 0,22 e 0,17, todos significantes e expondo um resultado óbvio: quanto maior a distância, menor o efeito espacial. Por último, a matriz cujos vizinhos considerados são sempre apenas os dez mais próximos rendeu um I de Moran de 0,36. Contudo, dada a arbitrariedade na especificação dessas outras

⁴ Os resultados da Análise Exploratória de Dados Espaciais (AEDE) foram obtidos usando o *software* GeoDa e os referentes a modelagem econométrica com o GeoDaSpace.

matrizes de ponderação, a matriz *queen*, amplamente já utilizada, será a escolhida para todo o resto das verificações.

Figura 3 – Moran scatterplot: dispersão dos desvios e defasagens da taxa de suicídio



Fonte: Elaboração própria.

A análise de *clusters* da taxa de suicídio confirma, em grande parte, aquilo que já havia sido observado no *box map*. Na Figura 4(a), segundo o indicador local de autocorrelação espacial (LISA), verificam-se *clusters* com padrão de baixas taxas de suicídio com municípios vizinhos semelhantes, identificados em azul escuro, em quase toda região metropolitana de Natal e na maior parte do litoral leste do estado. Esse é o padrão com mais municípios, 28 ao todo.

Outros 9 municípios participam de conglomerados de alta taxa de suicídio nas microrregiões do Seridó, de Pau dos Ferros e de Umarizal. A maioria dos municípios, 117 ao todo, não teve indicadores LISA estatisticamente significantes ao nível de 5%. Assim, pode ser afirmado que altas taxas de suicídio estão mais presentes no interior do estado, enquanto grandes conglomerados de baixas taxas de suicídio se encontram próximos ao litoral, sobretudo no entorno de Natal e da região metropolitana. A maior parte desses *clusters* identificados ressaltam, portanto, o efeito contágio examinado por Durkheim (2000).

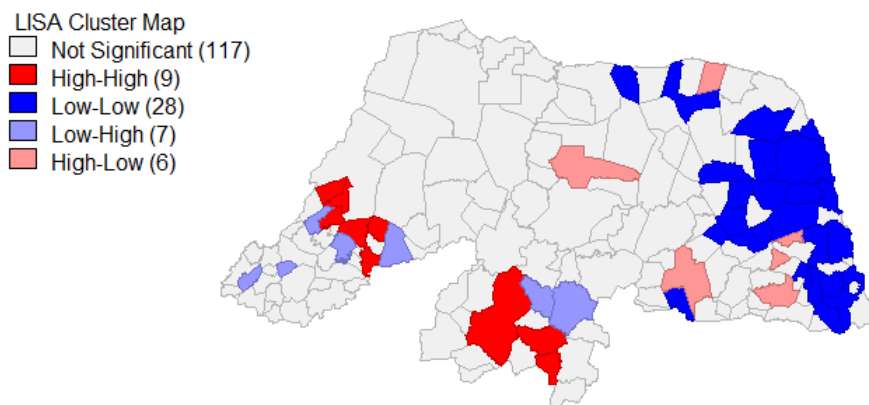
Com intuito de verificar relações espaciais da taxa de suicídio com outras variáveis-chave, foi realizada também uma análise de *clusters* a partir de indicadores LISA bivariados. Na Figura 4(b), é exposto o nível de associação do suicídio com a defasagem espacial do IDHM nos municípios vizinhos. O índice de desenvolvimento humano, usado em praticamente toda a literatura nacional, foi escolhido haja vista que fornece um resumo de informações importantes acerca da renda, longevidade e educação, cuja generalidade pode ser importante para explicar o suicídio não apenas no próprio município.

Assim, foram detectados *clusters* (21 municípios) de baixa taxa de suicídio rodeados por baixos níveis de desenvolvimento municipal nas microrregiões do Litoral Nordeste e da Baixa Verde. Em uma relação semelhante, no Seridó, foram observados municípios (9) onde há altas taxas de suicídio e a vizinhança apresenta altos níveis de desenvolvimento. Examinando a situação de maneira ampla, o *I* de Moran global bivariado calculado para essa relação foi de 0,2022, baixo, mas significativo, ressaltando uma relação direta. Esses resultados podem estar apenas refletindo a relação positiva dos níveis de desenvolvimento com a independência alcançada pelos indivíduos na região conglomerada que, segundo Durkheim (2000), é um fator direto de influência nos atos de suicídio. Apesar de que, foram detectados 12 municípios cujo padrão é de baixas taxas de suicídio

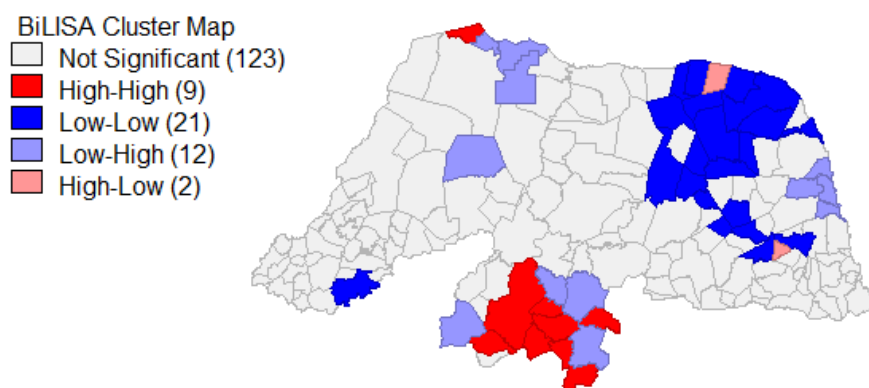
associadas a altos níveis relativos de desenvolvimento. Ao todo, 123 municípios não tiveram indicadores LISA significativos estatisticamente.

Figura 4 – LISA univariado e bivariado: análise de *clusters* da taxa de suicídio, IDHM e taxa de urbanização

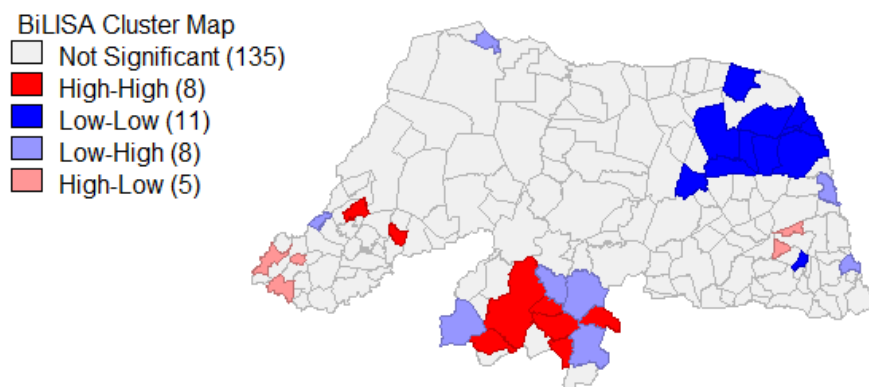
(a) *Taxa de suicídio (univariado)*



(b) *Suicídio x IDHM*



(c) *Suicídio x Urbanização*



Fonte: Elaboração própria.

O mesmo foi realizado para a taxa de urbanização na Figura 4(c), embasando-se na ideia de que esse fenômeno tem um caráter, em grande parte, eminentemente espacial, no qual dificilmente uma cidade se torna altamente urbanizada sem que suas vizinhas sofram influência. No entanto, o I de Moran global bivariado constatado foi ainda menor do que os anteriores, 0,1645, ainda que significativo, com um pseudo p -valor de 0,001, mas 135 municípios não alcançaram estatísticas LISA significantes. Os padrões encontrados foram semelhantes aos anteriores, sobretudo ressaltando relações diretas.

5.2. Modelo econométrico

Inicialmente, para a modelagem, foram escolhidas variáveis socioeconômicas consideradas relevantes para a explicação das taxas de suicídio, em conformidade com as recomendações da literatura empírica e teórica. À vista disso, foram determinados para a matriz de variáveis X : o índice de desenvolvimento, a renda e os níveis superiores de educação, que são vistos como elementos que levam à maior independência dos indivíduos e são utilizadas em praticamente todos os modelos da literatura empírica revisada, bem como também a desigualdade (GONÇALVES et al., 2011; DANTAS et al., 2018; SANTOS; BARBOSA, 2017; LOUREIRO et al., 2010; SHIKIDA et al., 2007; SCHNITMAN et al., 2011; FRAGA et al., 2016). A urbanização já explorada na AEDE, teorizada e verificada empiricamente por Durkheim (2000) e abordada também por Gonçalves et al. (2011), deverá constar. O desemprego, que espera-se ter uma relação inversa com a taxa de suicídio, também é verificado por Dantas et al. (2018), Santos e Barbosa (2017), Loureiro et al. (2010) e Fraga et al. (2016), e verificou-se a fecundidade, vista por Durkheim (2000) como um elemento que fortalece os laços familiares e que, portanto, tem relação inversa com a taxa de suicídio. No Quadro 1, pode ser observada a descrição das variáveis.

Com essas variáveis, foi estimado um modelo clássico por MQO com a intenção prévia de verificar os testes de detecção de relação espacial no modelo, haja vista que a literatura e a AEDE já indicam a presença desse fenômeno. Contudo, os testes sobre os resíduos rejeitaram as hipóteses de normalidade e homocedasticidade. Assim, foi adicionada uma variável *dummy* para caracterizar os municípios identificados como *outliers*. Os resultados dos testes desses modelos iniciais estão presentes na Tabela A1, no apêndice. Nota-se que a presença da *dummy* melhorou bastante o resultado dos testes de heterocedasticidade, mas não foi suficiente para não rejeitar a hipótese de normalidade.

Ainda assim, é dado continuidade à interpretação dos testes do modelo com a variável *dummy*, seguindo o procedimento proposto por Florax et al. (2003). O teste I de Moran para os resíduos foi significativo, demonstrando a existência de relações espaciais no modelo. Os testes LM para os termos de defasagem espacial do suicídio (ρ) e de erro espacial (λ) foram todos significantes, com exceção do teste robusto para λ , sendo o valor da estatística do teste LM_ρ sempre mais forte do que o LM_λ .

Por conseguinte, foi estimado o modelo SAR por dois métodos diferentes, o da Máxima Verossimilhança (MV) e o de Mínimos Quadrados em Dois Estágios (MQ2E) utilizando as defasagens espaciais das variáveis explicativas como Variáveis Instrumentais (VI), considerando que no MV exige-se o pressuposto de normalidade assintótica dos resíduos, enquanto que o método por VI essa hipótese não precisa ser satisfeita. O resultado de ambos encontra-se, respectivamente, nas Tabelas A2 e A3, no apêndice.

Ao nível de 5% de significância estatística, no modelo por MV, o índice de Gini e o IDHM foram significantes, ambos com relações diretas, isto é, quanto maior o desenvolvimento e a desigualdade, mais elevadas serão as taxas de suicídio. Por outro lado, a renda per capita apresentou uma relação significativa e inversa, na qual maiores níveis de renda podem levar a uma redução da taxa de suicídios, apesar da constatação de um baixo coeficiente em termos absolutos. No modelo por VI, considerado mais adequado por não supor normalidade, dos parâmetros do vetor $\hat{\beta}$, apenas o índice de Gini foi significativo e também positivo. Em ambos, o parâmetro de autocorrelação ρ foi

positivo e significativo. O teste Breuch-Pagan para o primeiro confirmou a homocedasticidade dos resíduos e o teste LR (Razão de Verossimilhança) ressaltou a significância do ρ . No modelo por VI, o teste Anselin-Kelejian comprovou que não há mais autocorrelação espacial nos resíduos do modelo.

Tabela 1 – Testes estatísticos do modelo estimado por MQO

Teste	Valor	Prob.
Jarque-Bera	4.8248	0.0896
Breusch-Pagan	14.6127	0.1021
Koenker-Bassett	17.6980	0.0388
<i>I</i> de Moran (ϵ)	2.4131 (I=0.1039)	0.0158
LM_ρ	10.6743	0.0011
LM_ρ * (robusto)	7.3737	0.0066
LM_λ	4.1379	0.0419
LM_λ * (robusto)	0.8373	0.3602
$LM(\rho + \lambda)$	11.5115	0.0032

Fonte: Elaboração própria.

Contudo, tendo em vista os problemas envolvidos nessa estimação inicial, sobretudo relacionados à normalidade dos resíduos, buscou-se alterar a especificação do modelo com a mudança de variáveis explicativas, visando ao alcance de melhores diagnósticos. Assim, a nova especificação trocou o índice de desigualdade de Gini pelo de Theil (usado por Schnitman et al. (2011)), as taxas de fecundidade e de desocupação foram excluídas e foram incluídas as variáveis razão de dependência (usado por Schnitman et al. (2011), Dantas *et al.* (2018) e Santos e Barbosa (2017)), pobreza (usado por Schnitman *et al.* (2011), Loureiro *et al.* (2010), Gonçalves et al. (2011), Fraga et al. (2016), Dantas et al. (2018) e Santos e Barbosa (2017)) e a o decil dos ricos, que ao determinar a renda mínima dos 10% mais ricos resalta quão abastados são os ricos de uma sociedade, expondo também, certamente, o grau de independência desses indivíduos.

Os testes do novo modelo estimado por MQO estão presentes na Tabela 1. A nova especificação gerou testes que não rejeitam as hipóteses de normalidade e homocedasticidade dos resíduos (exceção do Koenker- Bassett). Um adendo deve ser feito também para a possibilidade de algum nível de multicolinearidade devido à semelhança de algumas variáveis, caráter bastante comum atribuído a esses tipos de dados socioeconômicos. Quanto aos testes de caráter espacial, o diagnóstico foi igual ao do modelo anterior, apenas com valores de estatísticas diferentes. Consequentemente, estimou-se um modelo SAR por Máxima Verossimilhança. Os resultados são observados na Tabela 2.

Os resultados da Tabela 2, possivelmente o modelo estimado mais adequado, constataram um ρ estatisticamente significativo ao nível de 5%, ressaltado ainda pelo teste *LR* que confirmou o modelo SAR como a especificação mais apropriada, comparada a um modelo sem defasagem espacial da variável dependente. Assim, pode ser afirmado, segundo o modelo da Tabela 2, com 5% de significância, que uma elevação no valor de uma unidade da taxa de suicídio na média das taxas das cidades vizinhas que dividem alguma fronteira (ou no aumento da taxa de cada uma dessas cidades em uma unidade) deve elevar, em média, a taxa de suicídio no município em questão.

Quanto às variáveis explicativas da matriz *X*, foi observado que o ensino superior tem uma relação direta, com significância aceita na margem, confirmando as formulações de Durkheim (2000). Além disso, a nova variável incorporada, razão de dependência, obteve significância e relação inversa com o suicídio, ressaltando que os suicídios ocorrem mais na população economicamente ativa antes de aposentar.

Tabela 2 – Resultados do modelo SAR

Variável	Coefficiente	Erro-padrão	Estatística Z	Prob.
Constante	4.1626	14.5161	0.2868	0.7743
Defasagem espacial (ρ)	0.2546	0.0841	3.0284	0.0025
Taxa de urbanização	-0.3095	1.6485	-0.1877	0.8511
IDH-M	21.8730	17.2905	1.2650	0.2059
Ensino superior	0.4733	0.2508	1.8873	0.0591
Renda <i>per capita</i>	0.0004	0.0137	0.0300	0.9761
Índice de Theil	8.0818	4.0591	1.9910	0.0465
Razão de dependência	-0.2215	0.0740	-2.9914	0.0028
Pobreza	-0.0049	0.0894	-0.0549	0.9562
Decil dos ricos	-0.0111	0.0048	-2.3205	0.0203
<i>Dummy outliers</i>	15.4698	1.2925	11.9685	0.0000

Teste de Breusch-Pagan: 15.4142 (Prob.= 0.08017); Pseudo- $R^2 = 0.5967$;

Teste de Razão de Verossimilhança (ρ): 9.2751 (Prob.= 0.0023).

Fonte: Elaboração própria.

A variável incorporada decil dos ricos apresentou significância estatística, mas com uma relação inesperada (como com a renda no primeiro modelo), já que se esperava que o maior nível de riqueza fosse levar a maiores taxas de suicídio, segundo Durkheim (2000). Já o índice de Theil tem uma relação direta e significativa com o suicídio, evidenciando que quanto maior a desigualdade de renda observada, maior serão os níveis de suicídio. Esses dois resultados devem estar em congruência com a confirmação da tese verificada e confirmada por Loureiro et al. (2010), de que a felicidade está atrelada ao consumo relativo e não ao consumo absoluto, ao observar que a desigualdade tem uma relação direta e a pobreza (e possivelmente a riqueza) têm uma relação inversa com o suicídio.

A urbanização não mostrou relevância em nenhum dos modelos estimados, provavelmente ressaltando que morar no campo ou em zonas urbanas não importa para as taxas de suicídio no Rio Grande do Norte. O IDHM, apesar de ter sido relevante no primeiro modelo, deixou de ser significativo na nova especificação. As influências da renda e da pobreza sobre o suicídio também não podem ser consideradas estatisticamente diferentes de zero. O grande destaque é para a defasagem espacial do suicídio, que se mostrou relevante em todos os modelos para explicar o suicídio, ressaltando o efeito contágio.

6. Conclusão

Esta pesquisa buscou verificar padrões espaciais da taxa de suicídio e suas relações com fatores socioeconômicos, partindo da tese enfatizada por Durkheim (2000), de que parte dos suicídios ocorrem devido a uma espécie de imitação, na qual indivíduos propensos a cometer suicídio teriam uma maior possibilidade de realizá-lo após o conhecimento de casos próximos. Dessa forma, o principal objetivo foi testar a hipótese para os municípios do Rio Grande do Norte para o que ficou caracterizado como um efeito contágio do suicídio.

Assim, a partir da análise espacial realizada, ficou constatado o efeito contágio entre os municípios do RN, apesar de indicadores de autocorrelação espacial serem considerados baixos e ressaltados em microrregiões específicas, mas foram sempre significantes, seja na Análise Exploratória de Dados Espaciais ou na modelagem econométrica-espacial.

A análise econométrica, embora tenha verificado relações espaciais sempre evidentes e significantes, não demonstrou essa constância e asserção com as variáveis socioeconômicas. A análise das significâncias estatísticas e a direção dos sinais se tornaram maleáveis e muito sensíveis às mudanças de especificação do modelo, certamente comprovando as conclusões de Rehkopf e

Buka (2006), que avaliam uma grande heterogeneidade de comportamentos dessas variáveis em estudos desse tipo. Além de ocorrer a inexistência de um modelo teórico fechado sobre o assunto, provavelmente por sua impossibilidade de formulação.

Dessa forma, uma análise econométrica mais profunda, considerando outras especificações, áreas geográficas, métodos e tipos de variáveis, embasando-se em uma vasta revisão da literatura, seria recomendada, ficando como sugestão para futuras pesquisas. Como exemplo, as relações possíveis envolvendo o interior do estado, local no qual ficou constatado que municípios mais distantes do litoral apresentam maiores taxas de suicídio (relações com a incidência solar, produção agropecuária, pluviometria, gastos com saúde etc.).

Portanto, as inferências feitas a partir das variáveis socioeconômicas têm como resultado poucas conclusões, mas ainda permitem que sejam propostas políticas de prevenção ao suicídio. A teoria indica que a população com alto grau de independência socioeconômica é mais propensa ao suicídio e os resultados mostram que isso ocorre sobretudo com aqueles com maior nível de escolaridade, nomeadamente, com ensino superior completo. Portanto, parcela das políticas de prevenção ao suicídio poderiam ser aplicadas sobre essa parte da população. Também há uma indicação, contrariamente ao que se poderia pensar, que suicídios podem ser mais frequentes sobre a população economicamente ativa relativamente aos dependentes mais idosos.

Além disso, a relação direta entre taxa de suicídio e desigualdade de renda mostra que a redução do suicídio pode ser um subproduto de políticas de redistribuição de renda. Com relação aos resultados puramente espaciais, sua interpretação pode levar a sugestão de políticas localmente focadas, direcionadas para áreas interioranas e rurais, como nas microrregiões do Seridó, de Pau dos Ferros e de Umarizal que apresentam padrões de altas taxas de suicídio com evidentes efeitos de contágio.


Referências

- ANSELIN, L. Local indicators of spatial association - LISA. *Geographical Analysis*, v. 27, n. 2, p. 93-115, 1995.
- ANSELIN, L. Exploring spatial data with GeoDaTM: a workbook. *Center for Spatially Integrated Social Science*. Urbana, IL. 2005.
- ANSELIN, L.; REY, S. J. *Modern spatial econometrics in practice: a guide to GeoDa, GeoDaSpace and PySAL*. GeoDa Press LLC, 2014.
- ANSELIN, L.; SYABRI, I.; SMIRNOV, O. *Visualizing multivariate spatial correlation with dynamically linked windows*. In: Proceedings of the CSISS Workshop on New Tools for Spatial Data Analysis. Santa Barbara, CA. 2002.
- ASSUNÇÃO, R. M. Estatística espacial com aplicações em epidemiologia, economia e sociologia. In: Congresso da Associação Brasileira de Estatística. *Anais...* São Carlos: Associação Brasileira de Estatística, 2001.
- CRUZ, B. D. O.; FURTADO, B. A.; MONASTERIO, L.; JÚNIOR, R. *Economia regional e urbana: teorias e métodos com ênfase no Brasil*. Brasília: IPEA, 2011.
- CUTLER, David M.; GLAESER, E. L.; NORBERG, K. E. Explaining the rise in youth suicide. In: GRUBER, J. (Ed.) *Risky behavior among youths: an economic analysis*. University of Chicago Press, 2001. p. 219-270.
- DANTAS, A. P.; AZEVEDO, U. N. D.; NUNES, A. D.; AMADOR, A. E.; MARQUES, M. V.; BARBOSA, I. R. Analysis of suicide mortality in Brazil: spatial distribution and socioeconomic context. *Revista Brasileira de Psiquiatria*, v. 40, n. 1, p. 12-18, 2018.

- DURKHEIM, E. O. *O suicídio: estudo de sociologia*. Tradução de Mônica Stahel. São Paulo: Martins Fontes, 2000.
- DUTRA, E.; ROEHE, M. V. Suicídio de agricultores no Rio Grande do Norte: compreensão fenomenológica preliminar dos aspectos existenciais. *Psicologia em Revista*, v. 19, n. 1, p. 108-118, 2013.
- FLORAX, R. J. G. M.; FOLMER, H.; REY, S. J. Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics*, v. 33, n. 5, p. 557-579, 2003.
- FRAGA, W. S.; MASSUQUETTI, A.; GODOY, M. R. Determinantes socioeconômicos do suicídio: um estudo para os municípios do Brasil e do Rio Grande do Sul. *Revista Econômica*, v. 18, n. 2, p. 1-37, 2016.
- GONÇALVES, L. R. C.; GONÇALVES, E.; OLIVEIRA JÚNIOR, L. B. Determinantes espaciais e socioeconômicos do suicídio no Brasil: uma abordagem regional. *Nova Economia*, v. 21, n. 2, p. 281-316, 2011.
- LESAGE, J. P.; PACE, R. K. *Introduction to spatial econometrics*. Statistics: textbooks and monographs. New York: CRC Press, 2009.
- LOUREIRO, P. R. A.; MENDONÇA, M. J. C.; SACHSIDA, A. *Os determinantes econômicos do suicídio: um estudo para o Brasil*. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada (IPEA), 2010. (Texto para Discussão, n. 1487).
- MONTEIRO, A. M. V., CÂMARA, G.; CARVALHO, M. S.; DRUCK, S. *Análise espacial de dados geográficos*. Planaltina: Embrapa, 2004.
- REHKOPF, D. H.; BUKA, S. L. The association between suicide and the socio-economic characteristics of geographical areas: a systematic review. *Psychological Medicine*, v. 36, n. 2, p. 145-157, 2006.
- SANTOS, E. G. O; BARBOSA, I. R. Conglomerados espaciais da mortalidade por suicídio no nordeste do Brasil e sua relação com indicadores socioeconômicos. *Cadernos Saúde Coletiva*, v. 25, n. 3, p. 371-378, 2017.
- SHIKIDA, C. D.; ARAUJO JR, A. F.; GAZZI, R. A. V. Teoria econômica do suicídio: estudo empírico para o Brasil. *Análise Econômica*, v. 25, n. 48, p. 123-147, 2007.
- SCHNITMAN, G.; KITAOKA, E. G.; AROUCA, G. S. D. S.; LIRA, A. L. D. S.; NOGUEIRA, D.; DUARTE, M. B. Taxa de mortalidade por suicídio e indicadores socioeconômicos nas capitais brasileiras. *Revista Baiana de Saúde Pública*, v. 34, n. 1, p. 44-59, 2011.
- SGOBIN, S. M. T. *Custos diretos e indiretos de tentativas de suicídio de alta letalidade internadas em hospital geral*. Dissertação (Mestrado em Saúde Pública). Faculdade de Ciências Médicas da Universidade Estadual de Campinas (UNICAMP). Campinas, 2013.

ORCID

Stefany Silva Amaral  <https://orcid.org/0000-0002-6420-4964>

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

Apêndice

Tabela A1 – Testes estatísticos do modelo inicial estimado por MQO, com e sem *dummy*

Teste	Modelo inicial		Modelo c/ <i>Dummy</i>	
	Valor	Prob.	Valor	Prob.
Jarque-Bera	89.7920	0.0000	8.6616	0.0132
Breusch-Pagan	39.6330	0.0000	9.8858	0.2731
Koenker-Bassett	17.2520	0.0158	11.1090	0.1956
<i>I</i> de Moran (ϵ)	1.9820 (I=0.082)	0.0475	3.1814 (I=0.1424)	0.0015
LM_ρ	7.7340	0.0054	17.3620	0.0000
LM_ρ * (robusto)	12.9810	0.0003	11.0117	0.0009
LM_λ	2.6310	0.1048	7.7669	0.0053
LM_λ * (robusto)	7.8780	0.0050	1.4165	0.2340
$LM(\rho + \lambda)$	15.6120	0.0004	18.7786	0.0001

Fonte: Elaboração própria.

Tabela A2 – Resultados do modelo SAR inicial estimado por Máxima Verossimilhança

Variável	Coefficiente	Erro-padrão	Estatística Z	Prob.
Constante	-17.8946	10.3855	-1.7230	0.0849
Defasagem espacial (ρ)	0.3254	0.0838	3.8826	0.0001
Taxa de urbanização	0.8466	1.7310	0.4891	0.6248
IDH-M	33.6046	16.3913	2.0502	0.0404
Renda <i>per capita</i>	-0.0147	0.0070	-2.0915	0.0365
Taxa de desocupação	-0.0236	0.0703	-0.3356	0.7372
Taxa de fecundidade	-1.1036	0.8272	-1.3341	0.1822
Índice de Gini	14.5413	6.5692	2.2136	0.0269
Ensino superior	0.1587	0.2402	0.6608	0.5087
<i>Dummy outliers</i>	15.5278	1.3435	11.5576	0.0000

Teste de Breusch-Pagan: 11.4714 (Prob.= 0.1763); Pseudo- $R^2 = 0.5663$;Teste de Razão de Verossimilhança (ρ): 14.7594 (Prob.= 0.0001).

Fonte: Elaboração própria.

Tabela A3 – Resultados do modelo SAR inicial estimado por MQ2E (defasagem espacial das variáveis explicativas (*WX*) como variáveis instrumentais de *Wy*)

Variável	Coefficiente	Erro-padrão	Estatística Z	Prob.
Constante	-17.2005	10.3142	-1.6677	0.0954
Defasagem espacial (ρ)	0.5100	0.1243	4.1037	0.0000
Taxa de urbanização	0.4040	1.5605	0.2589	0.7957
IDH-M	26.0741	15.3502	1.6986	0.0894
Renda <i>per capita</i>	-0.0105	0.0073	-1.4467	0.1480
Taxa de desocupação	0.0009	0.0796	0.0118	0.9906
Taxa de fecundidade	-0.5617	0.8179	-0.6867	0.4922
Índice de Gini	15.3549	7.2821	2.1086	0.0350
Ensino superior	0.1149	0.2694	0.4264	0.6698
<i>Dummy outliers</i>	15.5125	1.5668	9.9009	0.0000

Teste de Anselin-Kelejian: 3.3970 (Prob.= 0.0653); Pseudo- $R^2 = 0.5683$.

Fonte: Elaboração própria.