Research article

The effect of economic complexity and relatedness on the economic growth of brazilian microregions♦

Felipe Micail da Silva **SMOLSKI**[1]
felipesmolski@hotmail.com | https://orcid.org/0000-0002-1982-3109

Janaina **RUFFONI**[1]
jruffoni@unisinos.br | https://orcid.org/0000-0002-7498-6437

Suelene **MASCARINI**[2]
smascarini@gmail.com | https://orcid.org/0000-0002-9926-7877

Abstract

This paper examines the effects of economic complexity and relatedness (proximity of productive knowledge) on the economic growth of Brazilian microregions. The perspective of economic complexity suggests that countries advancing in terms of productive sophistication and relatedness tend to achieve higher levels of income and economic growth. Additionally, regional-level analyses have consistently contributed to the growing body of evidence supporting this assertion. We employed dynamic spatial econometric models (SDM, SDGMM-SYS, and SDGMM-DIF) across 558 Brazilian microregions, utilizing employment data for 59 economic activities from 2002 to 2020. Our findings suggest that, firstly, an increased level of economic complexity in microregions is associated with a higher GDP per capita. Secondly, the density around industries (relatedness density) positively affects GDP per capita. Thirdly, and most importantly, our analysis reveals that economic complexity has unequal effects on the economic growth of microregions when examining distinct major Brazilian regions, particularly concerning per capita GDP. A lack of empirical evidence exists regarding the effects of relatedness and economic complexity on regional growth in Brazil, with issues of endogeneity and spatial dependence often overlooked. Our study addresses this gap by providing new insights into the effects of economic complexity and relatedness on regional growth, as well as the varying effects among Brazilian regions. Understanding these processes aids in formulating strategic interventions to promote economic complexity, relatedness, and regional economic growth.

Keywords

Economic complexity; Relatedness; GDP per capita; Microregions; Brazil.

[1] Universidade Unisinos, Escola de Negócios, Porto Alegre, RS, Brasil

[2] Universidade Estadual de Campinas (UNICAMP), Instituto de Economia, Campinas, SP, Brasil.

Recebido: 04/12/2023.
Revisado: 13/07/2024.
Aceito: 17/07/2024.
DOI: https://doi.org/10.1590/1980-53575442fjs

O efeito da complexidade econômica e *relatedness* sobre o crescimento econômico das microrregiões brasileiras

Resumo

O objetivo deste artigo é examinar os efeitos da complexidade econômica e *relatedness* (proximidade de conhecimentos produtivos) sobre o crescimento econômico das microrregiões brasileiras. A perspectiva da complexidade econômica sugere que os países que avançam em termos de sofisticação produtiva e *relatedness* tendem a alcançar altos níveis de renda e crescimento econômico. Além disso, estudos com análises em nível regionais tem consistentemente contribuído para o conjunto de evidências favoráveis a esta afirmação. Nós utilizamos modelos dinâmicos de econometria espacial (SDM, SDGMM-SYS e SDGMM-DIF) com as 558 microrregiões do Brasil, usando dados de emprego para as 59 atividades econômicas no período de 2002 a 2020. Nossos resultados sugerem que, em primeiro lugar, elevações nos níveis de complexidade econômica são associadas a elevações no PIB *per capita*. Em segundo lugar, a densidade das atividades (*relatedness density*) exibe uma correlação positiva com o PIB *per capita*. Em terceiro lugar, um achado importante é que existem efeitos desiguais da complexidade econômica sobre o crescimento econômico quando examinadas separadamente as grandes regiões no Brasil, principalmente em relação ao PIB *per capita*. Há uma falta de evidências com relação aos efeitos da *relatedness* e da complexidade econômica sobre o crescimento regional no Brasil e, as questões de endogeneidade e dependência espacial têm sido desconsideradas. Este estudo considera este *gap* relatando novas evidências dos efeitos da complexidade econômica e *relatedness*sobre o crescimento regional, bem como dos efeitos diferenciados com relação às regiões brasileiras. Um maior entendimento destes processos contribui para a formulação de intervenções estratégicas para promover complexidade econômica, *relatedness*e crescimento econômico regional.

Palavras-chave

Complexidade econômica; *relatedness*; PIB *per capita*; Microrregiões; Brasil.

Classification JEL

O47 ; R11; R11.

1. Introduction

The literature on economic complexity argues that creating more sophisticated and higher value-added products and services necessitates greater specific capabilities and productive knowledge. Consequently, countries endowed with these attributes excel in producing a wider array of products. This domain of research has engendered new indicators that encapsulate the benefits of specialization and productive diversification, endeavoring to illuminate the wealth-generation potential of nations (Hausmann and Hidalgo, 2011; Hidalgo and Hausmann, 2009; Simoes and Hidalgo, 2011).

The role of geographical location in the evolution of economic structures has garnered considerable attention in theoretical and empirical discourses, underscoring the geographically bound dynamics of place-based capabilities. This perspective has gained particular relevance in formulating smart specialization policies (Balland et al. 2018). Such an approach investigates how closely related activities within specific areas influence economic growth and the emergence of new industries, extending beyond the mere capacity to manufacture specific products. Recent studies have aimed to elucidate how increased economic complexity affects regional income, indicating that a stronger relatedness between activities promotes local economic growth, inequalities, employment creation, and the development of new industries (Balland et al., 2018; Chávez, Mosqueda, and Gómez-Zaldívar, 2017; Hartmann et al., 2020; Hausmann and Hidalgo, 2011; Mewes and Broekel, 2022; Queiroz, Romero, and Freitas, 2023; Vinci and Benzi, 2018).

Moreover, research efforts have been made to assess the effects of economic complexity on economic growth across various geographical scales, including nations and regions within economic blocs. For instance, Chávez et al. (2017) observed favorable effects of complexity on the income of Mexican states. Likewise, Gao and Zhou (2018) and Zhang et al. (2018) delved into the context of Chinese provinces and manufacturing sectors, corroborating the significant influence of complexity on regional growth.

In the Brazilian context, studies have identified a positive correlation between economic complexity, regional development, and diversification (Galetti et al., 2022; Teixeira, Missio, and Dathein, 2022). Furthermore, these studies have shown that increased relatedness increases de likelihood of entry into new sectors in more complex regions and dissects the analysis into various levels of regional complexity. Conversely, the complexity of a sector decreases the likelihood of its specialization in less complex regions while amplifying it in more complex regions (Freitas, Britto and Amaral, 2024; Queiroz, Romero, and Freitas, 2024).

This research, however, draws conclusions from methodologies that consider the simultaneity effects of the primary variables under investigation, a gap in the literature according to Teixeira, Missio, and Dathein (2022), and the influence of spatial heterogeneity and temporal dynamics. Hence, our study offers a significant opportunity to enhance the understanding of the relevance of complexity and relatedness to economic growth, making an original contribution to the regional literature on economic complexity. Although these investigations represent progress in the literature by concentrating on regional diversification, research remains scarce on relatedness, complexity, and regional growth, particularly in emerging countries.

Moreover, there is a considerable gap in our understanding of the variations and implications associated with the adoption of more complex activities in the country. It is crucial to discern the presence of unequal regional effects in Brazil to obtain deeper insights into the challenges and the continuity of income growth in historically disadvantaged areas. Simplifying the complex phenomena of economic development presents a significant challenge (Pereima 2020), making it necessary to devise novel policies to addressdeal with short- and long-term economic objectives (Dosi et al. 2020). In this regard, economic complexity methods of dimensionality reduction can contribute by handling data from thousands of economic activities and summarizing the geography of economic outputs (Hidalgo 2021; Hartmann et al. 2020).

In this context, the main objective of this study is to investigate the effects of economic complexity and relatedness on the GDP per capita growth of 558 microregions in Brazil from 2002 to 2020. This research contributes to the existing literature by providing new insights into the roles played by economic complexity and relatedness in driving economic growth in the microregions of an emerging country. Notably, this study utilizes employment data, departing from the conventional reliance on trade or export information to assess complexity, representing a significant advance in regional economic complexity studies.

Following this introduction, Section 2 elucidates the core concepts underpinning the framework of economic complexity and relatedness density, emphasizing their significance in the study of economic growth within a specific context. Section 3 offers a thorough description of the research methodology. Subsequently, the study presents the main findings and discussions, culminating in the concluding remarks.

2. Economic Complexity and Territory

The perspective of economic complexity has aimed to amplify its analytical framework, thereby facilitating a deeper insight into the structural transformations within nations and their respective territories. This has been achieved by introducing new empirical and longitudinal indicators (Hausmann and Hidalgo, 2011; Hidalgo and Hausmann, 2009; Simoes and Hidalgo, 2011). At its core, this approach evaluates the makeup of the local productive structure, often referred to as the *product space*. Therefore, the complexity and productive sophistication of a country or locality are quantified through methods derived from computer science, relying on the portfolio of local export production. This quantification indirectly assesses the technological sophistication embedded within the productive framework, generally called *economic complexity* (Gala 2017).

In light of this viewpoint, it is critical to consider that impoverished nations might predominantly export a single, scarce natural commodity, thereby impacting their economic complexity (Gala, 2017; Hidalgo and Hausmann, 2009). This consideration underscores the importance of two key indicators: non-ubiquity and diversity. Non-ubiquity rises in accordance with the level of complexity inherent in a product. Following the principles of economic complexity, not all products can be uniformly produced across all regions. Products that are more ubiquitous typically require lower productivity, knowledge, and technological capacities for their production. Conversely, products characterized by non-ubiquity demand a broader spectrum of expertise for their manufacture and, therefore, are exclusive to a select few regions. According to Gala (2017), the index depends on the combination of non-ubiquity and the diversity of distinct products that a location has the capacity to produce.

As a result, the ratio of non-ubiquitous exported products, economic activities, labor, or even research relative to their diversity indicates the inherent economic complexity within a locality. This highlights the crucial role of productive specialization in understanding the mechanism of wealth accumulation within a nation, potentially driving its economic development. Evidence indicates a negative correlation between the diversification of domestic production and the average ubiquity of exported products. Conversely, diversifying a nation's export portfolio positively correlates with the complexity of the manufactured goods. Essentially, more diversified and less ubiquitous economies tend to display higher economic complexity, indicative of more advanced production methodologies (Gala, 2017; Hausmann and Hidalgo, 2011).

Consequently, the paradigm of local specialization encompasses not only the composition of a nation's production and export portfolio in terms of diversification but also the complexity level of occupations and activities within regional areas, making it a pivotal factor in shaping economic growth and dynamism (Hausmann, Hwang, and Rodrik, 2007). Variations in regional income can thus be attributed to differences in territorial economic complexity, which aims to elucidate the localized capacity to create distinct economic production configurations (Hidalgo and Hausmann, 2009). The economic complexity perspective show that more intricate activities tend to cluster in larger urban centers due to their demand for profound knowledge and a division of labor. This clustering denotes a widespread network of individuals and public and private institutions equipped with complementary expertise required for engaging in complex services and industries (Balland et al. 2020).

Therefore, the local production structure depends on the degree of proximity between agents, known as the relatedness between activities, in which the ability to assimilate the knowledge essential for performing specific activities within a country and/or its regions is crucial. Achieving success in producing a particular good within a region or sector extends beyond just geographic or cultural proximity among activities; it also hinges on the degree of technological and cognitive proximity. Hidalgo (2021) articulated that the proximity between activities suggests that knowledge acquisition is encouraged through interactions among closely related sectors, striking a balance between similarity and differentiation, thereby fostering cooperation rather than competition. This conceptual framework has developed measures to assess the proximity between products, industries, sectors, occupations, and research fields, thereby understanding the dynamics, interrelations, and territorial capabilities required to produce specific goods.

The principle of relatedness, thus, implies that activities, research areas, sectors, or industries are closely related when they require similar types of knowledge or inputs for their execution, reflecting the various ways in which economies and organizations acquire knowledge (Hidalgo et al. 2018). Furthermore, it introduces the notion of agents' absorptive capacity, suggesting that assimilating new knowledge depends on the pre-existing knowledge associated with a specific activity within the region (Hidalgo 2021). As elucidated by Balland et al. (2018), regions marked by greater complexity and enhanced cognitive proximity between activities boost the potential for local economic advancement and the emergence of novel combinations. Therefore, the concept of relatedness holds significant importance within the literature on economic complexity as it underlines the interdependence of a country's ability to manufacture a specific product with its capacity to produce others, contingent upon their resemblance and the local potential to do so.

3. Methodology

3.1. Economic Complexity and Relatedness Indicators

We utilized employment data to calculate the Economic Complexity Index and the measure of relatedness, adhering to the methods outlined in previous studies for Brazilian microregions from 2002 to 2020 (Gao and Zhou, 2018; Queiroz, Romero, and Freitas, 2023; Zhang, Wu, and Wang, 2018). In accordance with the methodology developed by Hausmann and Hidalgo (2009; 2011), the economic complexity index juxtaposes the prevalence of specific product production within a particular location against the diversity of products available in that geographical area. It encapsulates the synergy of local productive knowledge or skills necessary for manufacturing goods that demand a high level of expertise or capabilities and thus is not commonly produced (non-ubiquity) with the breadth of productive specializations in a given locale (diversity).[1]

To determine the degree of local specialization in a particular activity, we used the revealed comparative advantage (RCA) indicator, initially proposed by Balassa (1965). Equation 1 illustrates the RCA:

$$RCA = \left(\frac{X_{cp}}{\Sigma_p X_{cp}}\right) \div \left(\frac{\Sigma_c X_{cp}}{\Sigma_{c,p} X_{cp}}\right) \tag{1}$$

[1] By examining the relationship between products and countries, their objective is to identify the unique capabilities or productive capacities inherent to specific regions. Through this analysis, it generates indicators of relative standings among countries/regions, shedding light on disparities in technological advancements and income levels over time.

X_{cp} represents the level of employment in economic activity p within microregion c, $\sum_c X_{cp}$ indicates the total employment in microregion c across all activities, $\sum_p X_{cp}$ represents the total employment in activity p for the microregion and $\sum_{c,p} X_{cp}$ represents the total employment across all activities in the country. A microregion is deemed to possess a comparative advantage in a specific activity when its RCA exceeds 1, denoting for $RCA_{c,t}(i) = 1$ each activity.

The matrix M, acting as the foundation for measuring diversity and ubiquity, is subsequently generated based on the RCA cutoff points (0,1) for each microregion and economic activity using Equations 2 and 3:

$$Diversity: k_{c,0} = \sum_p M_{cp} \tag{2}$$

$$Ubiquity: k_{p,0} = \sum_c M_{cp} \tag{3}$$

Where the diversity of a particular region ($k_{c,0}$) is determined by the count of specialized activities within the microregion, whereas the ubiquity of a specific economic activity is assessed by the number of microregions specializing in that particular activity ($k_{p,0}$).
From the averages of diversity for each microregion *c* and ubiquity of activities *p* produced by the locations in the analysis at hand, the economic complexity index (ECI) of the region (*kc*) is calculated with Equations 4 and 5 (Hidalgo, 2021; Hidalgo and Hausmann, 2009):

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{cp} k_{p,N-1} \tag{4}$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_p M_{cp} k_{c,N-1} \tag{5}$$

The reflections method calculates economic complexity because it creates "a symmetric set of variables for the two types of nodes in the network" (Hidalgo and Hausmann, 2009, p. 10572), microregions, and economic activities. This technique is iterative and involves the calculation of activity and microregion complexities using previous values of ubiquity and diversity, where N ≥ 1 indicates the number of iterations. It begins with the average values of $k_{p,0}$ and $k_{c,0}$. Subsequently, in the following rounds ($k_{p,^1}$, $k_{p,^2}$.. $k_{p,N}$ and $k_{c,^1}$, $k_{c,^2}$,.. $k_{c,N}$), previously calculated values are used until stability in the rankings of ECI and activity complexity index (ACI) complexities is achieved, without providing any new additional information. This is conducted using Equations 6 and 7:

$$k_c = \widetilde{M}_{cc'} k_{c'} \tag{6}$$

$$k_p = \widetilde{M}_{pp'} k_{p'} \tag{7}$$

In order to eliminate constant factors, the indices are standardize using Z score, Equation 8 (Hidalgo, 2021; Hidalgo and Hausmann, 2009):

$$ECI_c = \frac{\left(\vec{k}_c - mean(\vec{k}_c)\right)}{standard\_deviation(\vec{k}_c)} \tag{8}$$

where $\vec{k}$ is the eigenvector associated with the second-largest eigenvalue $\widetilde{M}_{cc'}$ of $\widetilde{M}_{pp'}$ and .

Our second variable of interest examines the effect of relevant microregional capabilities. Drawing from previous studies (Davies and Maré, 2021; Elekes, Baranowska-Rataj, and Eriksson, 2023; Galetti et al., 2022; Hane-Weijman, Eriksson, and Rigby, 2022; Hidalgo et al., 2007), we applied the concept of relatedness density. This metric quantifies the proximity of a new economic activity to the set of activities present in each location. Essentially, it assesses the probability of a location specializing in a new activity by considering the presence of activities with closely related capabilities within the same region. Therefore, the indicator for the density of activity *i* in a microregion *c* is derived from the relatedness of activity *i* to the activity for which the microregion holds an RCA, divided by the sum of the relatedness of activity *i* to all other activities in the country.

The first step involves identifying the proximity (ϕ) or relatedness between economic activities to compute the relatedness density. This is determined through the conditional probability that pairs of activities, denoted as *i*, are jointly produced in territory *j* (Equation 9):

$$\phi_{i,j,t} = \min\{P\left(RCAx_{i,t} \middle| RCAx_{j,t}\right), P\left(RCAx_{i,t} \middle| RCAx_{j,t}\right)\} \tag{9}$$

For this purpose, among the various established methods for measuring co-occurrence (Hidalgo, 2021; Van Eck and Waltman, 2009), we employed the cosine similarity, which is represented by Equation 10:

$$\phi_{i,j,t} = \frac{C_{i,j,t}}{\sqrt{S_{i,t} * S_{j,t}}} \tag{10}$$

Where C_{ij} represents the frequency with which two activities are observed together in period t; S_i and S_j denote the occurrences of activities i and j in time t. The relatedness density of a specific activity i in a microregion c is calculated from the relatedness between activity i and the ones for which the microregion has revealed RCA, divided by the sum of the relatedness between activity i and all other activities in the country. This calculation is illustrated by Equation 8, which assesses the proximity of a new product to the existing set of products in a given region:

$$RD_{i,c,t} = \frac{\sum_{j\in c, j\neq i} \phi_{i,j,t}}{\sum_{j\neq i} \phi_{i,j,t}}.100 \tag{11}$$

The relatedness density (RD) indicator ranges from 0 to 100%. A reading of 0% for microregion c indicates that no activity i is related to another activity i within microregion c. Conversely, a value of 100% suggests that all activities related to activity i are included in the portfolio of activities in microregion c, drawing from the framework of Hausmann and Hidalgo (2011; 2009) and Boschma et al. (2015).

3.2. Data and Variables

In this study, we utilized annual data for 558 microregions in Brazil (a spatial classification similar to the European Union's NUTS-3), covering the period from 2002 to 2020. The variables are presented in Table 1. Analyzing microregions is a common practice in the pertinent literature (Françoso, Boschma, and Vonortas, 2024; Galetti, Tessarin, and Morceiro, 2021; Queiroz, Romero, and Freitas, 2024). This approach helps control for unobserved factors derived from the administrative autonomy of different configurations (e.g., municipal units, states), thereby fostering a better understanding of the results and avoiding issues related to endogeneity and unobserved heterogeneity in panel data (Wooldridge 2010).

The dependent variable in this analysis is the economic growth of Brazil's microregions, denoted as $lgdppc$ (and its first difference, $\Delta lgdppc$). Economic growth is quantified using the logarithm of the GDP (Gross Domestic Product) per capita of the microregions. Data for this variable were sourced from the Brazilian Institute of Geography and Statistics (IBGE 2022).

Our main variables of interest are Economic Complexity, denoted as *ECIactiv*, and relatedness density, represented by *RDactiv*. Both indicators are derived from employment data of economic activities[2] from the RAIS (2022) database, compiled by the Brazilian Ministry of Labor and Employment. The database reveals detailed records of formal activities, quantity and salary, educational characteristics of the workforce employed, and data from establishments in Brazil across all public and private administrative structures, covering about 65% of the workforce and providing comprehensive information (Freitas, Britto, and Amaral, 2024; Galetti et al., 2022; Ulyssea, 2018). The RAIS database thus presents a faithful portrait of the formal sectors. However, it does leave out informal economic activities, which is seen as a limitation in this literature (Queiroz, Romero, and Freitas, 2024).

For robustness checks, we also created an alternative measure of economic complexity using the Fitness Complexity Method (*lfitness* – logarithm of fitness), presented in Appendix B (definition) and C (regressions). This method, introduced by Tacchella et al. (2012), defines the economic complexity index by adopting non-linear iterative equations that favor regions with diversified activities and penalize activities realized in many regions (Mariani et al. 2015). The fitness variable has also been employed in previous studies such as Gao and Zhou (2018), Fritz and Manduca (2021), Teixeira, Missio, and Dathein (2022), and Stojkoski, Koch, and Hidalgo (2023).

The variables *wECIactiv* and *wRDactiv* are indicators for adjacent regions, as defined by the spatial weight matrix, which will be elucidated in the subsequent section. It is important to note that *wECIactiv* and *wRDactiv* are indicative of their neighboring regions, as determined by the spatial weight matrix (Knn3).

Additionally, other control variables were also employed, namely $lgdppc_{t-1}$, *ldens*, *pempind*, *pexpcom*, *pempsup, and exppartbra.*, *pexpcom*, *pempsup*, *and exppartbra*. $lgdppc_{t-1}$ represents the logarithm of the GDP per capita from the previous period, *ldens* denotes the natural logarithm of the population density in the microregions and is derived from data provided by IBGE (2022). This variable was used to control for the potentially positive expected effects of population and urban agglomerations on economic growth, a practice employed in other studies (Mewes and Broekel, 2022; Balland et al., 2018; Frenken et al., 2007; Boschma and

[2] By employing data from the Annual Report of Social Information (2022) and using the National Classification of Economic Activities (v. 1.0), 59 activity groups (*activ*) were added.

Iammarino, 2009; Queiroz, Romero, and Freitas, 2024). Lastly, *pempind* pertains to the proportion of industrial employment relative to the overall employment within the microregion. This information is derived from RAIS (2022). Following Davies and Maré (2021), Galetti et al. (2022), and Hane-Weijman, Eriksson, and Rigby (2022), we expected a positive effect of industrial activities on economic growth.

Table 1 - Description of the variables and data sources used.

Variables	Description	Source	Signal	Reference
Dependent				
lgdppc	Logarithm of the GDP per capita	IBGE (2022)		Hidalgo and Hausmann (2009), Hausmann et al. (2011), Salles, Pinto, and Vasconcelos (2018), Gao and Zhou (2018), Zhang, Wu and Wang (2018)
Independent				
ECIactiv	Economic Complexity Index (activities)	RAIS (2022)*	(+)	Gao and Zhou (2018), Zhang, Wu and Wang (2018)
lfitness	Logarithm of fitness (activities)	RAIS (2022)*	(+)	Stojkoski, Koch, Hidalgo (2023)
RDactiv	Relatedness Density (activities)	RAIS (2022)*	(+)	Davies and Maré (2021), Galetti et al. (2022), Françoso, Boschma, Vonortas, (2024)
Controls				
ldens	Natural logarithm of the population density in the microregions	IBGE (2022)	(+)	Mewes and Broekel (2022), Balland et al. (2018), Frenken, Van Oort and Verburg (2007), Boschma and Iammarino (2009), Queiroz, Romero and Freitas, (2024), Stojkoski, Koch, Hidalgo (2023)
pempind	Proportion of industrial employment in relation to the overall employment within the microregion	RAIS (2022)	(+)	Davies and Maré (2021), Galetti et al. (2022), Hane-Weijman, Eriksson and Rigby (2022)
pexpcom	The proportion of primary products exported by the microregion	COMEXSTAT (2022), IBGE (2022)	(+)	Hausmann et al. (2011), Stojkoski, Utkovski and Kocarev (2016), Salles et at. (2018)
pempsup	Proportion of individuals with higher education levels relative to the total population of the microregion	RAIS (2022)	(+)	Zhu and Li (2017), Gao and Zhou (2018), Mewes and Broekel (2022), Galetti et al. (2022)
exppartbra	Microregional export share of Brazil's total exports	COMEXSTAT (2022)	(+)	Hausmann et al. (2011), Salles et al. (2018), Stojkoski, Utkovski and Kocarev (2016)
Spatial heterogeneity				
wlgdppc (*rho*) $wlgdppc_{t-1}$	GDP spatial autocorrelation with neighboring regions at time *t* and *t -1*		(+)	Verheij and De Oliveira (2020)
$wECIactiv_{t-1}$ $wRDactiv_{t-1}$	*ECIactiv* and *RDactiv* spatial autocorrelation with neighboring regions at time *t-1*		(+/-)	Fawaz and Rahnama-Moghadamm (2019), Verheij and De Oliveira (2020) Gómez-Zaldívar et al. (2020)

Note: *Index developed by the authors. Source: Developed by the authors.

pexpcom, which corresponds to the natural logarithm of the proportion of primary products exported by the microregion, using data from COMEXSTAT (2022). This variable assesses the effects of exporting natural products on regional income, considering the potential positive relationship of the commodity cycle in the Brazilian economy. Researchers such as Hausmann et al. (2011), Stojkoski et al. (2016), and Salles et al. (2018) have used this indicator in their studies. *pempsup* denotes the proportion of individuals with higher education relative to the total population of the microregion, utilizing data from RAIS (2022). This variable accounts for the potentially positive effect of human capital on economic growth, a factor considered in studies conducted by Zhu and Li (2017), Gao and Zhou (2018), Mewes and Broekel (2022), and Galetti et al. (2022). *exppartbra* represents the microregional export share of Brazil's total exports, using data from COMEXSTAT (2022). The economic complexity literature has demonstrated that international trade participation, particularly an increase in the share of exports of more technologically sophisticated products, is a hallmark of wealthier nations with greater potential for economic growth (Gala, 2017; Hausmann and Hidalgo, 2011; Hausmann, Hwang, and Rodrik, 2007; Hidalgo and Hausmann, 2009; Simoes and Hidalgo, 2011). *wlgdppc*(*rho*) and $wlgdppc_{t-1}$, which represent the effects of spatial autocorrelation with neighboring regions at time *t* and *t-1*, respectively, on the dependent variable *lgdppc*. This operationalization is in accordance with the methodology utilized by Verheij and De Oliveira (2020) and Martini (2020), based on the expectation of positive spillover effects of *lgdppc* growth on neighboring microregions.

Table 2 - Descriptive analysis of the variables utilized in the models.

Variable	N	Minimum	Mean	Median	Maximum	Standard deviation
lgdppc	10,602	5.24	8.78	8.82	12.12	1.06
RDactiv	10,602	0.00	26.55	27.00	70.00	11.19
ECIactiv	10,602	-3.17	0.00	-0.08	3.18	1.00
lfitness	10,602	-6.89	-0.99	-0.93	3.29	1.51
pexpcom	10,602	0.00	0.65	0.90	1.00	0.41
ldens	10,602	-2.85	1.22	1.11	6.82	1.55
pempind	10,602	0.00	0.16	0.14	0.72	0.12
pempsup	10,602	0.00	0.14	0.13	0.63	0.06
exppartbra	10,602	0.00	0.00	0.00	0.12	0.01

Source: Developed by the authors.

Table 2 offers descriptive statistics for the variables incorporated into the econometric models, as outlined in the methodology section. This dataset comprises 10,602 observations from 558 Brazilian microregions, thereby generating a balanced panel dataset covering a 19-year span from 2002 to 2020.

3.3. The Portrait of GDP per capita, Economic Complexity, and Relatedness Density in Brazilian Microregions

Figure 1 illustrates the evolution of GDP per capita at 2020 prices, economic complexity, and relatedness density from 2002 to 2020 in Brazilian microregions. In terms of GDP per capita, the legend denotes 5th-percentile intervals. Microregions in the lowest 20^{th} percentile - which, for example, had a GDP per capita ranging from BRL 243.00 to BRL 763.00 in 2002 - contrast with those in the highest 20th percentile, where the GDP per capita varied between BRL 2,657.00 and BRL 19,672.00 in 2002. High-income microregions are predominantly located in states such as São Paulo, Rio de Janeiro, Santa Catarina, Rio Grande do Sul, Mato Grosso, and Goiás throughout the periods analyzed.

By 2020, there was a significant increase in the GDP per capita of microregions within the state of Mato Grosso compared to the onset of the analysis period. The North and Northeast regions house the majority of microregions in the lowest percentile, indicating lower GDP per capita. Focusing on economic complexity, as depicted by *ECIactiv* (in Figure 1B, microregions with higher complexity (ECI > 1) are primarily found in the South and Southeast regions, which are recognized as dynamic economic centers in Brazil, with only a few exceptions in other regions. This trend has remained relatively consistent across the years examined.

As for relatedness density, indicated by *RDactiv* in Figure 1C, a pattern of microregion clustering is observed. However, it is important to highlight that there are no clear indications of specific cluster formations throughout the country.

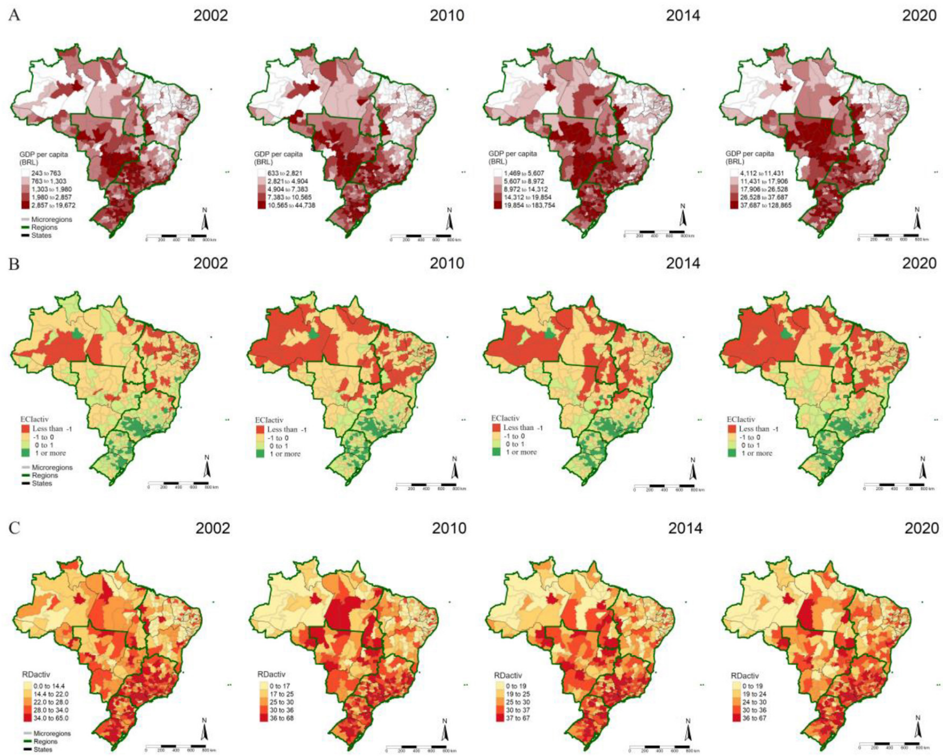

Figure 1 - Distribution of GDP per capita (A), economic complexity (B), and relatedness density (C) of Brazilian microregions in 2002, 2010, 2014, and 2020. Source: Developed by the authors.

To uncover patterns of local autocorrelation, Figure 2 illustrates the significance maps for the local indicator of spatial autocorrelation (LISA) for GDP per capita (A), economic complexity (B), and relatedness density (C). For an in-depth understanding of the LISA statistics, refer to Almeida (2012). In Figure 2A, we observe statistically significant clusters (p-value < 0.05) of high GDP per capita of the high-high type,[3] confirming the previously noted distribution pattern (concentration in the Southeast and South regions) while highlighting the emergence of new clusters in central-western Brazil.

[3] The LISA indicator, as detailed in Appendix A, identifies four statistically significant clusters within the spatial distribution of microregions: the high-high group (characterized by high values within the microregion and among neighboring regions), high-low (comprising high values within the microregion and low values in the neighboring microregions), low-high (encompassing low values within the microregion and high values in the neighboring microregions), and low-low (which includes low values both within the microregion and in the neighboring regions).

It is worth noting that the number of high-high clusters has decreased over the study period, declining from 81 in 2002 to 71 in 2020. Conversely, clusters featuring low *lgdppc* in the microregion and neighboring microregions (low-low) have formed, primarily in the Northeast region, and have expanded into the Northern region. The LISA Indicator analysis corroborates these findings, providing both visual and statistical evidence of the close relationship between *ECIactiv* and *lgdppc* over the period. The correlation between these variables is 0.40 (p-value = 0.01).

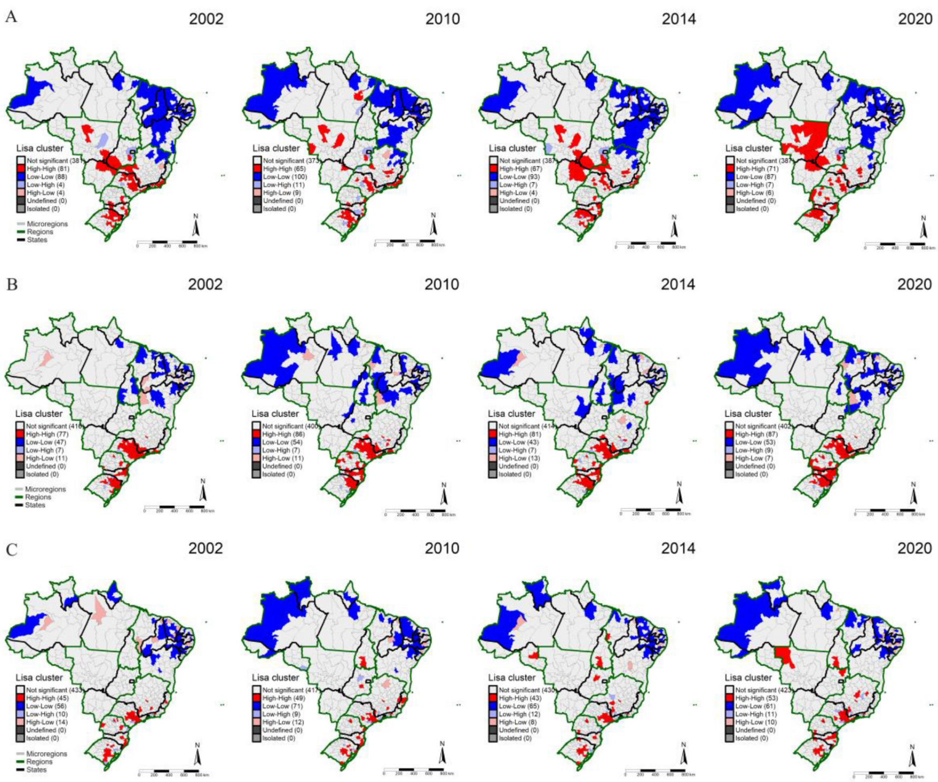

Figure 2 - LISA Significance (Knn3) for GDP per capita (A), economic complexity (B), and relatedness density (C) in 2002, 2010, 2014, and 2020. Source: Compiled by the authors.

These findings align with observations regarding the spatial distribution of the economic complexity index in the Brazilian context, as reported by Freitas and Andrade (2015). Moreover, within the economic complexity literature, it is well-established that regions hosting highly sophisticated activities are often regional economic hubs, leading to significant agglomerations (Balland et al. 2020). A recent study by Fritz and Manduca (2021) supports this association by examining US counties, where metropolitan areas demonstrate greater economic complexity, correspondingly resulting in higher income levels and enhanced economic growth.

Conversely, the northern and northeastern regions consistently exhibit the highest low-low values throughout all assessed periods. As previously established, higher *RDactiv* values indicate more favorable prospects for economic growth and the emergence of new activities. This further highlights the persistent challenges faced by these regions, which have historically been characterized by lower GDP per capita.

Over the duration of the analyzed period, regional economic complexity and local relatedness in these regions do not show significant signs of improvement in comparison to other regions across the country. However, it is important to emphasize that this study did not explore whether the spatial distance between these less economically complex regions and the more economically advanced regions, characterized by higher activity proximity (*RDactiv*), has decreased over time. This aspect warrants further examination in future studies, particularly in the context of convergence studies.

3.4. Econometric Models

The econometric analysis of spatial data significantly enhances our understanding of regional dynamics. Traditional econometrics often fails to address issues related to spatial autocorrelation and spatial heterogeneity, which are prevalent in many locations (Anselin 1988). Such analysis allows us to explore the interactions among various agents, the existence of externalities, and the potential for spillover effects (Golgher 2015). In our study, we employed spatial panel econometric regressions with time fixed effects, including Spatial Durbin Models (SDM), Spatial Durbin with Generalized Method of Moments in Differences (SDGMM-DIF), and Spatial Durbin with Dynamic Generalized Method of Moments (SDGMM-SYS).[4]

Explanatory and control variables are used at *t-1* in order to mitigate endogeneity concerns with the dependent variable. The fundamental model specification incorporates an array of explanatory variables denoted X_{it-1}, control variables, the spatial autocorrelation variable of the dependent variable ρWY_{it} and τY_{it-1}, which represents the fixed effects associated with unobserved characteristics that remain constant over time δ_t and the error term ε_{it} (Equation 12):

$$Y_{it} = \beta_0 + \rho WY_{it-1} + \tau Y_{it-1} + \beta X_{it-1} + \beta WX_{it-1} + \beta CONTROLS_{it-1} + \delta_t + \varepsilon_{it} \quad (12)$$

In the context of the fixed-effects spatial panel model over time, the variables are adjusted to eliminate the time-related effects among observations, following a process similar to that of a standard panel model. This involves subtracting the time-specific average value from each individual i, generating transformed observations ($\tilde{y}_{it-1}$ and $\tilde{x}_{it-1}$) for estimation. The formulation of the SDM panel model is described in Equation 13:

$$\tilde{y}_{it} = \rho W\tilde{y}_{it-1} + \tau\tilde{y}_{it-1} + \beta\tilde{x}_{it-1} + \beta W\tilde{x}_{it-1} + \beta \widetilde{CONTROLS}_{it-1} + \tilde{\varepsilon}_{it} \quad (13)$$

$W\tilde{y}_{it}$ represents the spatially lagged dependent variables, ρ stands for the spatial autoregressive coefficient, W denotes the spatial weights matrix, $\tilde{\varepsilon}_{it}$corresponds to the error term ($i.i.d.\sim N(0, \sigma 2)$). Furthermore $W\tilde{x}_{it}$, is added, representing the lagged independent variable, and θ is the spatial autoregressive coefficient of the independent variable.

A spatial dependency matrix *w*, often referred to as a neighborhood matrix or weight matrix, is utilized to identify regions in proximity and, consequently, those that are interrelated, i.e., neighbors. One of

[4] In the process of selecting the appropriate panel model considering the availability of data in a balanced panel format, we employed various tests, including the Hausmann, Chow, and LM (Lagrange multiplier) tests. Initially, we will determine the spatial regression models to be used based on the type of autocorrelation observed in the variables (spatial lag or errors). We will utilize both classical and robust LM tests, which are applied to the residuals from Ordinary Least Squares (OLS) models. The error LM test involves comparing the OLS model with the spatial error model, assuming λ=0, thereby testing the null hypothesis $H_0: \lambda = 0$. If H_0 is rejected, the Spatial Error Model (SEM) is preferred. On the other hand, the lag LM test compares the OLS model with a spatial lag model, assuming ρ=0, with the null hypothesis being H_0:ρ=0. If it is rejected, the lag model is considered more appropriate (Almeida 2012). The robust error tests do not make assumptions about ρ=0 and λ=0 but instead consider these values as unknown.

the most commonly employed matrices for this purpose is the contiguity matrix, wherein wij = 1 if two regions are contiguous (i.e., share a border) and wij = 0 if they are not (i.e., do not share a border) (Almeida 2012).[5] The Queen and Knn weight matrices, varying in orders from 1 to 4, were compared, and the one with the highest average over the study period was selected, which, in this instance, happened to be the Knn matrix of order 3.[6]

Additionally, we employed the SDGMM-DIF, since spatial model disturbances may potentially be correlated with time and exhibit heteroskedastic characteristics. The method of moments approach enables the derivation of consistent estimators for the variables while accounting for the disturbance process (Kapoor, Kelejian, and Prucha, 2007). In the SDGMM-DIF model, variables are first-differenced (Equation 14):

$$\begin{aligned} \Delta lgdppc_{it} = {} & \Delta\rho Wlgdppc_{it} + \Delta\beta lgdppc_{it-1} \\ & + \Delta\beta ECIactiv_{it-1} + \Delta\beta RDactiv_{it-1} \\ & + \Delta\theta WECIactiv_{it-1} + \Delta\theta WRDactiv_{it-1} \\ & + \Delta\beta CONTROLS_{it-1} + \delta_t + \eta_i + \varepsilon_{it} \end{aligned} \tag{14}$$

where ηi represents the region fixed effects.

An issue to consider in the regression estimation process is the potential endogeneity between economic complexity and relatedness indicators, as well as the economic growth indicators of the regions (Teixeira, Missio, and Dathein, 2022). To tackle this concern, Anderson and Hisao (1981) have suggested employing the lagged dependent variable as an instrumental variable. Fawaz and Rahnama-Moghaddam (2019) have utilized the spatially and temporally lagged dependent variable over one and two years. Similarly, Zhang et al. (2018) and Teixeira al. (2022) have used temporally lagged variables as instruments. Taking inspiration from these techniques and re-

[5] When considering all the first-order neighbors of a region, this process gives rise to a weight matrix known as the Queen's matrix, which can vary in order (1, 2, etc.). A different form of weight matrix employed is the *k*-nearest neighbors' matrix, where the *k*-neighbors (1, 2, 3, etc.) in closest proximity to a specific region, determined by centroid distances, are identified. Within this matrix, a value of wij = 1 indicates that the neighboring region is the closest to the reference region, whereaswhile wij = 0 denotes the other regions.

[6] Although different Moran's I indicators were tested for each year of the study, the Knn3 weight matrix was used in the regression models and in the LISA indicator' analyzes, due this matrix obtained the highest average of Moran's I during 2002-2020 comparing Queen and Knn matrices (orders 1 to 4).

search findings, this study adopted instrumental variables for the variables $lgdppc_{t-1}$, $wlgdppc_{t-1}$, $ECIactiv_{t-1}$, $RDactiv_{t-1}$, $wECIactiv_{t-1}$, and $wRDactiv_{t-1}$.

These variables are instrumented with their temporal lag (2nd difference). The underlying hypothesis is that the estimators are consistent and exogenous by employing these moment conditions. Sargan tests are then conducted to assess whether the utilized estimators are considered exogenous, indicating that the instruments are uncorrelated with the error term (Almeida 2012). Additionally, autocorrelation tests will be carried out, specifically AR (1) and AR (2). The estimated SDGMM-SYS model will be made with Equation 15:

$$\begin{aligned} lgdppc_{it} = \; & \rho Wlgdppc_{it} + \beta lgdppc_{it-1} \\ & + \beta ECIactiv_{it-1} + \beta RDactiv_{it-1} \\ & + \theta WECIactiv_{it-1} + \theta WRDactiv_{it-1} \\ & + \beta CONTROLS_{it-1} + \delta_t + \eta_i + \varepsilon_{it} \end{aligned} \tag{15}$$

4. Results and Discussion

Table 3 presents the outcomes of the SDM (Equations 1 and 2), SDGMM-SYS (Equation 5), and SDGMM-DIF (Equation 6) models. An initial examination shows a pronounced spatial dependence of the dependent variable (*rho*) across all models, as well as its first-time lag ($lgdppc_{t-1}$). This spatial dependence is crucial for the precise fitting of the proposed models, showcasing positive effects of growth on neighboring regions. Nevertheless, the influence on *lgdppc* is not statistically significant of surrounding microregions with a time lag ($wlgdppc_{t-1}$) in both the SDGMM-SYS and SDGMM-DIF models.

Table 3 – SDM, SDGMM-SYS, and SDGMM-DIF regressions, dependent variable *lgdppc*, Brazilian microregions from 2002 to 2020.

	SDM		SDGMM-SYS	SDGMM-DIF
	lgdppc	*Δlgdppc*	*lgdppc*	*Δlgdppc*
	(1)	(2)	(3)	(4)
rho	0.1067***	0.1058***	0.7743***	0.4437***
	(0.0125)	(0.0129)	(0.0689)	(0.1310)
$lgdppc_{t-1}$	0.6118***	-0.3942***	0.1127***	0.0814***
	(0.0078)	(0.0081)	(0.0241)	(0.0225)
$wlgdppc_{t-1}$	0.1533***	0.2623***	0.0207	0.0226
	(0.0144)	(0.0118)	(0.0331)	(0.0350)
$RDactiv_{t-1}$	0.0014***	0.0013***	0.0027	0.0021
	(0.0004)	(0.0004)	(0.0014)	(0.0015)
$ECIactiv_{t-1}$	0.0742***	0.074***	0.0665**	0.0633**
	(0.0057)	(0.0059)	(0.0205)	(0.0217)
$wRDactiv_{t-1}$	0.0005	0.0006	0.0021	0.0002
	(0.0006)	(0.0007)	(0.0023)	(0.0030)
$wECIactiv_{t-1}$	-0.0486***	-0.0478***	-0.0263	0.0098
	(0.0075)	(0.0077)	(0.0349)	(0.0412)
$pexpcom_{t-1}$	0.0502***	0.0496***	0.0401	-0.0188
	(0.0088)	(0.0091)	(0.0233)	(0.0184)
$ldens_{t-1}$	0.0013	0.0033	0.0186	0.1446
	(0.0028)	(0.0029)	(0.0149)	(0.2198)
$pempind_{t-1}$	0.1235***	0.1149***	0.0598	-0.0981
	(0.035)	(0.0362)	(0.1619)	(0.2043)
$pempsup_{t-1}$	-0.0279	-0.0893	0.1036	0.3026
	(0.0775)	(0.0795)	(0.1605)	(0.1856)
$exppartbra_{t-1}$	5.4578***	5.3348***	10.8440*	-4.8010*
	(0.6317)	(0.66)	(4.5958)	(1.9257)
N	9486	9486	10602	10602
R^2	0.89	0.21		
AIC	6553.3	6179.57		
Sargan: p-value			0.0302	0.4061
AR (1)			0.0000	0.0000
AR (2)			0.0422	0.0183
FE Microregion	No	No	Yes	Yes
FE Year	Yes	Yes	Yes	Yes

Note: Statistical significance: ***p < 0.001; **p < 0.01; *p < 0.05. Standard error in parentheses. AIC = Akaike Information Criterion. N = Number of observations. FE = Fixed Effects. Source: Developed by the authors.

Regarding the relationship between $ECIactiv_{t-1}$ and *lgdppc*, all regressions indicate that the coefficients are positive and significant. Notably, in the

SDGMM-DIF and SDGMM-SYS models, comprehensive testing confirmed whether the instrumental variables (temporal lags) produce exogenous regressors, as determined by the Sargan test. Additionally, tests were performed to detect any significant autocorrelation - AR (2), ensuring that the proposed models meet all test requirements ($p = 0.01$). As robustness checks presented in Appendix C, we also ran the same estimations using an alternative measure of economic complexity of microregions (*lfitness*) and the findings remained the same. Thus, our findings suggest that an increase in the economic complexity within microregions is associated with higher economic growth, namely, an increase in GDP per capita. These findings are consistent with previous evidence in other regional contexts, demonstrating a positive relationship between local productive sophistication and economic performance, even when considering spatial heterogeneity (Chávez, Mosqueda, and Gómez-Zaldívar, 2017; Fritz and Manduca, 2021; Hausmann and Hidalgo, 2011; Hidalgo and Hausmann, 2009; Mewes and Broekel, 2022; Salles, Pinto, and Vasconcelos, 2018; Stojkoski, Utkovski, and Kocarev, 2016; Zhang, Wu, and Wang, 2018).

Moreover, $wECIactiv_{t-1}$ coefficients are negative and statistically significant in all non-dynamic models (SDM), suggesting that the economic complexity of neighboring regions tends to adversely affect the economic growth of the microregion. Hence, it indicates the presence of a negative regional spillover, as an increase in the economic complexity of neighboring regions tends to reduce the GDP per capita growth of the microregion. Although unexpected, this result supports the findings of Gómez-Zaldívar et al. (2020), who observed negative effects of economic complexity in neighboring regions on economic growth in Mexican States. Accordingly, our results contribute additional insights into understanding the effects of complexity on Brazilian microregional growth.

The effects of the relatedness density — $RDactiv_{t-1}$ on *lgdppc* are positive and statistically significant in SDM Models 1 and 2. However, in the SDGMM-SYS and SDGMM-DIF models, the coefficients of $RDactiv_{t-1}$ are positive but not significant. These outcomes may suggest limited productive capabilities for knowledge-based specialization in the regions, coupled with economic growth policies that might not be specifically aimed at promoting proximity of economic activities, as highlighted by Davies and Maré (2021). Conversely, earlier studies have identified effects on diverse variables, such as comparative advantage in products and sectors (Alonso and Martín, 2019), job creation (Davies and Maré, 2021; Hane-Weijman, Eriksson, and

Rigby, 2022), and the emergence of new specialized occupations (Jefferson R B Galetti et al. 2022), which are indirect measures of economic growth.

Regarding the variable $wRDactiv_{t-1}$, representing the relatedness density of neighboring microregions, its coefficients are not significant in any model. These results may be attributed to two complementary explanations. Firstly, within the context of the relatedness indicator, the effects of neighboring regions are inherently captured, thereby already accounted for in the variable relatedness. Consequently, relatedness effects may predominantly manifest through their direct effects ($RDactiv_{t-1}$), rendering the spillovers or indirect effects not correlated with economic growth.

Control variables were integrated into various models to account for additional factors influencing regional GDP per capita growth, as recognized in existing literature. These factors significantly contribute to explaining variations in economic growth. Notably, the share of a microregion's exports in the country's total exports ($exppartbra_{t-1}$) showed statistical significance in explaining the growth of *lgdppc*, indicating significant effects across all models presented, although the effect became negative in Model 4. Conversely, the variable $pexpcom_{t-1}$ demonstrated positive and statistically significant effects solely in SDM models. This finding aligns with previous research (Hausmann et al., 2011; Salles, Pinto, and Vasconcelos, 2018; Stojkoski, Utkovski, and Kocarev, 2016), highlighting the positive effect of a microregion's export participation and the exportation of commodities on microregional GDP per capita growth.

The proportion of industrial employment in the microregion ($pempind_{t-1}$) revealed positive and statistically significant effects on *lgdppc* only in non-dynamic Models 1 and 2, whereas the proportion of employment with a higher education level ($pempsup_{t-1}$) has no statistical significance. Regarding the effect of agglomeration, the logarithm of the population density of the microregions ($ldens_{t-1}$) showed a positive effect but was not statistically significant in all models.

To assess potential variations in the relationship between economic growth, economic complexity, and relatedness density across different regions in Brazil, several SDM regressions were conducted for each of the country's five major regions. The findings, detailed in Table 4 (with robustness checks in Appendix C), consider both the dependent variables *lgdppc* (in levels) and its first differences ($\Delta lgdppc$).

In terms of the influence of $ECIactiv_{t-1,}$ across all regions, the effect is predominantly positive and statistically significant for both dependent variables - *lgdppc* and Δ*lgdppc*, except for Models 7 and 8 applied to southeastern Brazil. The magnitude of these effects varies across regions, with the northern region displaying the highest coefficients, values of 0.1539 and 0.1581 in Models 5 and 6, respectively. Southern Brazil, known for its high economic complexity and GDP per capita, follows closely, with coefficients of 0.1087 and 0.1025 in Models 9 and 10. The Central-West region exhibits relatively lower coefficients of 0.0964 and 0.103 in Models 1 and 2.

Despite the South's established economic complexity and elevated GDP per capita, the Central-West has seen a steady increase in microregions with high GDP per capita, gaining prominence in this indicator. In contrast, the North comprises numerous microregions characterized by lower economic complexity and GDP per capita. The positive relationship observed between economic complexity and GDP per capita growth suggests progression in northern Brazil's transition towards a more complex product structure, ultimately reflecting increased economic prosperity.

A comparative analysis between the South and Southeast, both characterized by a significant number of high GDP per capita microregions and historical dynamism, yields distinct outcomes. While the South demonstrates a noteworthy and statistically significant effect of economic complexity on *lgdppc*, the Southeast region shows non-statistically significant results. Studies such as Pinheiro et al. (2022) indicate the presence of a persistent negative reinforcement mechanism, where complexity continuously grows in already advanced areas, thereby increasing the spatial distance from less developed regions. These less developed regions maintain their production by diversifying into low-tech activities. This pattern suggests a decreasing significance of economic complexity in driving GDP per capita growth within the Southeast region, a pattern not observed in southern Brazil during the assessed.

Table 4 - SDM Regressions, dependent variable *lgdppc* and Δ*lgdppc*, separated by regions in Brazil from 2002 to 2020.

	Central-West		Northeast		North		Southeast		South	
	lgdppc	*Δlgdppc*	*lgdppc*	*Δlgdppc*	*lgdppc*	*Δlgdppc*	*lgdppc*	*Δlgdppc*	*lgdppc*	*Δlgdppc*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
rho	0.0055	0.0175	0.1007 ***	0.0986 ***	0.1095 ***	0.0872 **	0.0501 *	0.0563 *	0.016	-0.0005
	(0.0393)	(0.0434)	(0.0216)	(0.0224)	(0.0349)	(0.0374)	(0.0233)	(0.0243)	(0.0307)	(0.0319)
$wlgdppc_{t-1}$	0.5964 ***	-0.4048 ***	0.5953 ***	-0.4071 ***	0.4425 ***	-0.5729 ***	0.6233 ***	-0.3848 ***	0.3803 ***	-0.6302 ***
	(0.0256)	(0.027)	(0.0138)	(0.0141)	(0.0255)	(0.0268)	(0.0145)	(0.015)	(0.0229)	(0.0234)
$lgdppc_{t-1}$	0.0647	0.062	0.1171 ***	0.2206 ***	0.1675 ***	0.262 ***	0.1403 ***	0.1961 ***	-0.0017	0.0142
	(0.0503)	(0.048)	(0.0254)	(0.0218)	(0.0429)	(0.041)	(0.0272)	(0.0234)	(0.0419)	(0.0434)
$RDactiv_{t-1}$	-0.0059 ***	-0.0058 ***	0.0029 ***	0.0025 ***	0.0034 **	0.0033 *	0.0024 ***	0.0024 ***	-0.0032 ***	-0.0029 ***
	(0.0019)	(0.0019)	(0.0008)	(0.0008)	(0.0014)	(0.0014)	(0.0009)	(0.0009)	(0.001)	(0.001)
$ECIactiv_{t-1}$	0.0964 ***	0.103 ***	0.0626 ***	0.065 ***	0.1539 ***	0.1581 ***	0.019	0.0151	0.1087 ***	0.1025 ***
	(0.0272)	(0.0279)	(0.0086)	(0.0088)	(0.019)	(0.0197)	(0.0129)	(0.0133)	(0.016)	(0.0164)
$wRDactiv_{t-1}$	-0.0078 **	-0.0078 **	0.0004	0.0008	0.0032	0.0044 *	-0.0037 ***	-0.0043 ***	-0.0002	0.0003
	(0.0031)	(0.0033)	(0.0011)	(0.0012)	(0.002)	(0.002)	(0.0013)	(0.0013)	(0.0017)	(0.0018)
$wECIactiv_{t-1}$	-0.0249	-0.0278	-0.0718 ***	-0.0677 ***	-0.0315	-0.0337	0.0036	0.0087	0.0517 *	0.0479 *
	(0.0355)	(0.0361)	(0.0131)	(0.0135)	(0.0275)	(0.0282)	(0.0156)	(0.0161)	(0.0224)	(0.0231)
$pexpcom_{t-1}$	-0.0267	-0.0168	0.0497 ***	0.0515 ***	0.0761 ***	0.0754 ***	0.0622 ***	0.0545 ***	-0.0938 ***	-0.0946 ***
	(0.0397)	(0.0417)	(0.0127)	(0.0129)	(0.0267)	(0.0277)	(0.0195)	(0.0202)	(0.0273)	(0.028)
$ldens_{t-1}$	0.0386 **	0.0364 *	0.014 ***	0.0165 ***	-0.0159	-0.0141	0.0306 ***	0.0331 ***	-0.0413 ***	-0.0391 ***
	(0.0161)	(0.0169)	(0.0047)	(0.0048)	(0.0104)	(0.0107)	(0.0077)	(0.008)	(0.0101)	(0.0104)
$pempind_{t-1}$	0.3025 *	0.3065	0.1475 ***	0.104	-0.0147	-0.0366	0.2207 **	0.217 **	0.2981 ***	0.362 ***
	(0.1516)	(0.1577)	(0.0532)	(0.0547)	(0.1571)	(0.1723)	(0.0864)	(0.0896)	(0.0852)	(0.0877)
$pempsup_{t-1}$	-0.922 *	-1.026 *	-0.1494	-0.2029	0.4816 **	0.4174 *	0.6796 **	0.7103 **	0.7444 ***	0.7479 **
	(0.435)	(0.4498)	(0.1052)	(0.1075)	(0.2031)	(0.2064)	(0.272)	(0.2801)	(0.2849)	(0.294)
$exppartbra_{t-1}$	25.1841 ***	23.2609 ***	12.3528 ***	10.928 ***	14.0234 ***	13.5569 ***	2.3341 ***	2.1275 *	13.6557 ***	14.2539 ***
	(6.3993)	(6.5289)	(2.8013)	(2.8691)	(2.5562)	(2.5986)	(0.8861)	(0.936)	(1.949)	(2.0221)
R^2	0.862	0.229	0.891	0.217	0.868	0.308	0.862	0.213	0.88	0.338
AIC	731.32	695.26	1734.07	1603.67	852.55	808.76	2187.71	2100.12	700.23	632.65
EF Microregion	No	No	No	No	No	No	No	No	No	No
EF Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Statistical significance: ***p < 0.001; **p < 0.01; *p < 0.05. Standard error in parentheses. AIC = Akaike Information Criterion. FE = Fixed Effects. Source: Developed by the authors.

Concerning the variable $RDactiv_{t-1}$, the results demonstrate positive and statistically significant effects in the regressions performed for the Northeast, North, and Southeast (as presented in Equations 3, 4, 5, 6, 7, and 8 in Table 4). In particular, the effect of $RDactiv_{t-1}$ on both *lgdppc* and $\Delta gdppc$ is most notable in the North, suggesting that an increase in the proximity of activities within this region has a considerable influence on economic growth. This observation offers a unique perspective in the existing body of literature and may assist in the development of strategies aimed at enhancing local income, especially in Brazilian regions where the per capita GDP is traditionally lower than that in the southeastern and southern regions.

Conversely, in the South and Central-Wests, known for their roles as commodity exporters, the effects of $RDactiv_{t-1}$ are negative and statistically significant. This indicates that a lower relatedness density has a more pronounced effect on GDP per capita. Therefore, the influence of activity proximity on *lgdppc* in these regions diverges from the prevailing literature, suggesting that these areas benefit from activities that are not immediately proximate to their growth. Instead, they thrive through activities characterized by higher economic complexity, as observed earlier.

This evidence underscores the presence of different effects of economic complexity and relatedness on Brazilian microregions, offering supplementary insights to the conclusions reported by Davies and Maré (2021). They found that economic complexity and relatedness positively and synergistically contribute to the creation of employment in larger urban centers. However, this trend does not consistently apply to smaller cities in New Zealand, thereby underscoring the varied effects of economic complexity across different geographical contexts.

In contrast, Mewes and Broekel (2022), utilizing regional data from Europe, reveal that the positive influence of complexity on growth can be statistically significant, even in non-metropolitan areas. This provides evidence that regions outside metropolitan centers can also benefit from increased productive complexity. Nonetheless, our findings suggest that historically disadvantaged areas, marked by lower per capita incomes such as northern and northeastern Brazil, might witness more significant positive effects due to the evolution of economic complexity and relatedness density patterns. These effects may even exceed those in areas traditionally considered more developed, or that have undergone earlier industrialization processes, earlier, such as the country's South and Southeast.

As previously mentioned, the statistically significant positive effects of economic complexity on GDP per capita growth were apparent in numerous models. Still, the equally positive effects of relatedness density across various regions warrant further investigation. Future studies should explore the potential interaction effects between these variables. The results of this research underscore the existence of regional disparities in productive capabilities and economic performance across Brazilian microregions. These variances should be meticulously analyzed to aid in the crafting of targeted public policies designed to bolster economic growth.

5. Final Considerations and Implications

The literature on economic complexity has increasingly focused on how countries reshape their activities, engage in the production of more complex products, participate in innovative activities, and advance scientific research. Researchers have emphasized the importance of examining the effects of the proximity of capabilities, products, activities, technologies, and scientific research within the context of regional development, innovation, and economic growth (Balland et al., 2018; Chávez, Mosqueda, and Gómez-Zaldívar, 2017; Hartmann et al., 2020; Hausmann and Hidalgo, 2011; Mascarini et al., 2023; Mewes and Broekel, 2022; Queiroz, Romero, and Freitas, 2023; Vinci and Benzi, 2018).

In light of this context, our study aimed to investigate the effects of economic complexity and relatedness on the GDP per capita of Brazilian microregions, using data from 2002 to 2020. The findings provide empirical evidence that economic complexity has a positive effect on the economic growth of Brazilian microregions. These results were obtained while addressing concerns regarding spatial autocorrelation and endogeneity through the application of advanced econometric techniques and the inclusion of relevant control variables, alongside robustness checks with an alternative variable.

Our analysis reveals a positive, though somewhat modest, relationship between the density of knowledge proximate activities (relatedness) and the economic growth of microregions. When examining the effect of economic complexity and relatedness in Brazilian microregions, we found evidence that economic complexity has a positive influence on economic growth across all regions, albeit with variations in the magnitude of the effect. Thus, our findings suggest that the influence of economic complexity and relatedness on regions are unequal.

Furthermore, the results demonstrate that the northern, central-western, northeastern, and southern regions experienced a more significant effect on GDP per capita growth due to the evolution of economic complexity. Similarly, relatedness density, while closely associated with regions of greater economic complexity, had a more substantial effect on Northeast, Northern, and Southeast. These findings represent novel contributions to the existing literature, especially within the Brazilian context during the analyzed period, considering the historical challenges faced by these microregions.

The evidence presented in this manuscript offers valuable insights for enhancing economic growth while considering regional disparities. It suggests that to foster sustainable growth and prosperity, it is essential to encourage economic complexity. These insights could guide policies aimed at addressing the challenges faced by lagging regions, restructuring regional industries to increase the overall complexity of the production mix, elevating their economic complexity through technological sophistication, and moving towards high-quality activities to promote faster growth rates. Collectively, this may highlight the importance of economic complexity and human capital in driving economic growth and supporting the development of local labor markets into complex and related activities.

Given the spatial heterogeneity and historical context of Brazilian regions, policymakers could acknowledge limitations and promote specific regional capabilities for engaging in technological change, knowledge diffusion, and growth (Boschma 2017; Hidalgo et al. 2018; Hidalgo 2023). By focusing on developing the unique productive structures of each region, decision-makers could increase the diversity of economic activities, encourage productive change by leveraging existing capacities, and support activities in areas with dense interrelatedness. Additionally, they could explore deliberate efforts to enable regions to accumulate capabilities, facilitating product diversification.

Public policies should adopt a strategic approach that prioritizes economic activities closely related to the current structure (relatedness) while aiming for greater sophistication. This requires joint efforts from both the public and private sectors to uncover potential. Such efforts can prevent economic stagnation in less-developed regions by incentivizing the introduction of new products and technologies, and promoting the recombination of unrelated knowledge through policies that develop new paths of knowledge, as well consider the relevance of complex services on economic complexity.

Considering economic complexity as a critical factor for regional economic growth can fuel policy debates about region-specific policies, such as the smart specialization strategy. This could significantly enhance regional economic performance by leveraging local economic capabilities, promoting new specializations in more complex and specific activities, reshaping the landscape of local economic actors and institutions, and fostering meaningful spillover effects (Balland et al., 2018; Hartmann et al., 2017; Hidalgo and Hausmann, 2009). Such insights can focus on identifying potential strategies to enhance productive sophistication through the economic complexity transmission channel, leading to improved GDP per capita performance across regions.

The main limitation of this study is the lack of uncovering the factors and mechanisms behind regional disparities in the effects of economic complexity and relatedness on regional economies, and the absence of data on informal economic activities. Future research should determine whether regions that have historically lagged economically have experienced more pronounced growth in per capita production. It is also imperative to assess whether the disparities between economically disadvantaged and affluent regions have decreased over time.

Appendix A - Lisa indicator

Moran's Global I is an indicator utilized to assess the spatial autocorrelation of a given variable and its surroundings, based on a predetermined weight matrix. This tool thus measures the strength of this relationship (Equation 16):

$$I_W = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} \left(z_i - \bar{z}\right)}{\sum_i \left(z_i - \bar{z}\right)^2} \tag{16}$$

where $i \neq j$, w represents the assigned weight matrix, n the number of regions, z the variable of interest analyzed, i *and* j *the regions*. The null hypothesis of this relationship is that neighbors do not co-vary with the region, and for values of $I_W > 0$ there is a positive spatial autocorrelation. It is possible to obtain the statistical significance of this autocorrelation by calculating the threshold values for H_0 (Loonis and Bellefon, 2018).

$$E(I_W) = -\frac{1}{n-1} \tag{17}$$

$$\frac{I_W - E(I_W)}{\sqrt{Var(I_W)}} \sim N(0,1) \tag{18}$$

Moran's Global I therefore demonstrate the presence or absence of a pattern of spatial dependence using all data from the analyzed sample. However, Anselin (1995) developed a local indicator to analyze spatial associations, known as LISA (Local Indicator of Spatial Association), by which each observation can have its contribution analyzed. This approach enables the statistical analysis of significant clusters across locations. In this context, for each observation, it is possible to determine the extents of the spatial clusters in relation to similar values around the observation, and the sum of all LISA indicators from the observations proportionally represents the Global Moran indicator. Local Moran's I is:

$$I_i = (z_i - \bar{z}) \sum_j w_{ij} (z_j - \bar{z}) \tag{19}$$

Thus, values greater than zero of I_i indicate that there is clustering of similar values (above the average), and values less than zero indicate a mix of dissimilar values, for example, high values surrounded by low values (Loonis and Bellefon, 2018). For this clustering to be statistically significant, the combinations of similar or dissimilar values must differ from what would be observed in a randomized spatial distribution.

$$Z(I_i) = \frac{I - E(I_i)}{\sqrt{Var(I_i)}} \sim N(0,1) \tag{20}$$

Threshold values are defined to test the null hypothesis that the values are normally distributed. The LISA indicator then demonstrates the areas where local autocorrelation processes (Local Moran's I) surpass global autocorrelation processes (Global Moran's I) and where autocorrelation is absent (Loonis and Bellefon, 2018).

Appendix B - Fitness

The fitness complexity method (FCM) was introduced by Tacchella et al. (2012) with the objective of offering an alternative, non-linear, and iterative approach to measuring the economic complexity of countries/regions and their exported products or activities. Unlike the reflections method, FCM posits that "the complexity of a product cannot be defined as the average of the fitness of the countries producing it" (Tacchella et al., 2012, p. 1). Instead, a product or activity can only be deemed highly complex if it is produced in a highly competitive region or country. Tacchella et al. (2012) describe the iterative process of the FCM as follows:

$$\begin{cases} \tilde{F}_c^{(n)} = \sum_p M_{cp} Q_p^{(n-1)} \\ \tilde{Q}_p^{(n)} = \dfrac{1}{\sum_c M_{cp} \dfrac{1}{F_c^{(n-1)}}} \end{cases} \rightarrow \begin{cases} F_c^{(n)} = \dfrac{\tilde{F}_c^{(n)}}{\langle \tilde{F}_c^{(n)} \rangle_c} \\ Q_p^{(n)} = \dfrac{\tilde{Q}_p^{(n)}}{\langle \tilde{Q}_p^{(n)} \rangle_p} \end{cases} \tag{21}$$

Where F_c is the fitness of a region, Q_p is the complexity of an activity, obtained at a fixed point; M_{cp} ate the binary region-activity matrix M (see Section 3.1).

The iterative method is based on computing $\tilde{F}_c^{(n)}$ and $\tilde{Q}_p^{(n)}$ and normalizing the results to define the region's fitness and the complexity of activities for each (n) iteration, and the final value of fitness reflects the region's economic complexity (Gao and Zhou, 2018). The advantage of this iterative process is that, at each iteration, both variables are normalized, and the information is refined, as reported by Tacchella et al. (2012). In this study, we compute 20 iterations, and the final fitness variable was used in logarithm form (*lfitness*), as employed by Stojkoski, Koch, and Hidalgo (2023).

Appendix C - Robustness checks

Table 5 – SDM, SDGMM-SYS, and SDGMM-DIF regressions, dependent variable *lgdppc*, Brazilian microregions from 2002 to 2020 (economic complexity alternative and independent variables: *lfitness and wlfitness*).

	SDM		SDGMM-SYS	SDGMM-DIF
	lgdppc	*Δlgdppc*	*lgdppc*	*Δlgdppc*
	(1)	(2)	(3)	(4)
rho	0.1045***	0.1032***	0.7097***	0.4807***
	(0.0125)	(0.0129)	(0.0625)	(0.1347)
$lgdppc_{t-1}$	0.6227***	-0.3829***	0.1373***	0.0726**
	(0.0078)	(0.008)	(0.0241)	(0.0223)
$wlgdppc_{t-1}$	0.1463***	0.2524***	0.0002	0.0255
	(0.0142)	(0.0114)	(0.0351)	(0.0369)
$RDactiv_{t-1}$	0.001*	0.0009	0.0040**	0.0007
	(0.0005)	(0.0005)	(0.0015)	(0.0015)
$lfitness_{t-1}$	0.0359***	0.0352***	0.0600***	0.0221*
	(0.0034)	(0.0035)	(0.0079)	(0.0088)
$wRDactiv_{t-1}$	0.0005	0.0006	-0.0009	-0.0004
	(0.0007)	(0.0007)	(0.0025)	(0.0029)
$wlfitness_{t-1}$	-0.0224***	-0.021***	-0.0596***	0.0115
	(0.0053)	(0.0054)	(0.0149)	(0.0162)
$pexpcom_{t-1}$	0.0448***	0.0444***	0.0515*	-0.0273
	(0.0088)	(0.0091)	(0.0226)	(0.0182)
$ldens_{t-1}$	0.004	0.0061*	0.0280*	0.1080
	(0.0027)	(0.0028)	(0.0126)	(0.2125)
$pempind_{t-1}$	0.2009***	0.1925***	0.3551**	-0.0168
	(0.0332)	(0.0344)	(0.1161)	(0.1950)
$pempsup_{t-1}$	-0.0004	-0.0604	-0.0404	0.2839
	(0.0775)	(0.0795)	(0.1695)	(0.1883)
$exppartbra_{t-1}$	5.9014***	5.7785***	11.6200*	-3.8489
	(0.6321)	(0.6606)	(4.7336)	(2.0389)
N	9486	9486	10602	10602
R^2	0.89	0.21		
AIC	6621.21	6246.19		
Sargan: p-value			0.0027	0.6628
AR (1)			0.0000	0.0000
AR (2)			0.0147	0.0320
FE Microregion	No	No	Yes	Yes
FE Year	Yes	Yes	Yes	Yes

Note: Statistical significance: ***p < 0.001; **p < 0.01; *p < 0.05. Standard error in parentheses. AIC = Akaike Information Criterion. N = Number of observations. FE = Fixed Effects. Source: Developed by the authors.

Table 6 - SDM Regressions, dependent variable *lgdppc* and *Δlgdppc*, separated by regions in Brazil from 2002 to 2020 (economic complexity alternative and independent variables: *lfitness and wlfitness*)

	Central-West		Northeast		North		Southeast		South	
	lgdppc	*Δlgdppc*	*lgdppc*	*Δlgdppc*	*lgdppc*	*Δlgdppc*	*lgdppc*	*Δlgdppc*	*lgdppc*	*Δlgdppc*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
rho	0.0108	0.0193	0.0918***	0.0906***	0.1118***	0.0835*	0.049*	0.0553*	0.0344	0.0146
	(0.0393)	(0.0433)	(0.0217)	(0.0225)	(0.035)	(0.0374)	(0.0233)	(0.0243)	(0.0303)	(0.0314)
$wlgdppc_{t-1}$	0.6106***	-0.3899***	0.6084***	-0.3941***	0.47***	-0.5437***	0.623***	-0.3853***	0.3952***	-0.616***
	(0.025)	(0.0265)	(0.0136)	(0.014)	(0.0254)	(0.0267)	(0.0145)	(0.015)	(0.0227)	(0.0232)
$lgdppc_{t-1}$	0.0662	0.0665	0.0988***	0.1934***	0.173***	0.2612***	0.1376***	0.1931***	0.0393	0.0684
	(0.0482)	(0.0449)	(0.0255)	(0.0219)	(0.0432)	(0.0403)	(0.027)	(0.0229)	(0.0389)	(0.0378)
$RDactiv_{t-1}$	-0.0069***	-0.0069***	0.0024***	0.002**	0.0023	0.0021	0.0022**	0.0022*	-0.004***	-0.0037***
	(0.0019)	(0.002)	(0.0008)	(0.0008)	(0.0015)	(0.0015)	(0.0009)	(0.001)	(0.001)	(0.0011)
$lfitness_{t-1}$	0.0408***	0.0418***	0.0321***	0.0322***	0.0596***	0.0624***	0.011	0.0093	0.0518***	0.0494***
	(0.015)	(0.0155)	(0.0051)	(0.0052)	(0.0102)	(0.0105)	(0.008)	(0.0082)	(0.0098)	(0.0099)
$wRDactiv_{t-1}$	-0.007*	-0.007*	-0.0008	-0.0004	0.0028	0.0041	-0.0042***	-0.0048***	-0.0009	-0.0005
	(0.0032)	(0.0034)	(0.0013)	(0.0013)	(0.0021)	(0.0021)	(0.0014)	(0.0015)	(0.0019)	(0.0019)
$wlfitness_{t-1}$	-0.0303	-0.0292	-0.0177*	-0.0162	-0.0262	-0.0244	0.01	0.0131	0.0168	0.0167
	(0.02)	(0.0205)	(0.0085)	(0.0087)	(0.0154)	(0.0158)	(0.012)	(0.0124)	(0.0163)	(0.0166)
$pexpcom_{t-1}$	-0.014	-0.0026	0.0449***	0.0467***	0.0592*	0.0564*	0.063***	0.0558***	-0.0973***	-0.0969***
	(0.0395)	(0.0415)	(0.0127)	(0.013)	(0.027)	(0.028)	(0.0196)	(0.0203)	(0.0273)	(0.0281)
$ldens_{t-1}$	0.053***	0.053***	0.0141***	0.0172***	-0.0032	-0.0016	0.0317***	0.034***	-0.0348***	-0.0328***
	(0.0149)	(0.0156)	(0.0046)	(0.0047)	(0.0102)	(0.0104)	(0.0074)	(0.0076)	(0.0101)	(0.0104)
$pempind_{t-1}$	0.3703**	0.3693**	0.148***	0.1065*	0.2246	0.2563	0.2597***	0.2506***	0.5571***	0.6031***
	(0.1483)	(0.1542)	(0.0528)	(0.0543)	(0.1579)	(0.1732)	(0.0759)	(0.0786)	(0.0743)	(0.0765)
$pempsup_{t-1}$	-1.0557**	-1.1721***	-0.2001	-0.2528**	0.4176*	0.3806	0.7026***	0.7208***	0.9954***	0.9773***
	(0.4399)	(0.4547)	(0.105)	(0.1073)	(0.2068)	(0.2102)	(0.2673)	(0.2752)	(0.2764)	(0.2844)
$exppartbra_{t-1}$	27.6312***	26.1066***	13.3381***	12.0257***	16.9374***	16.3579***	2.3516***	2.1422*	15.2197***	15.6724***
	(6.3288)	(6.4603)	(2.8119)	(2.8801)	(2.5391)	(2.5865)	(0.8859)	(0.9357)	(19327)	(2.0067)
R^2	0.861	0.223	0.89	0.21	0.864	0.289	0.862	0.214	0.879	0.332
AIC	735.94	701.59	1762.8	1631.62	883.3	837.22	2187.1	2099.15	719.95	648.24
FE Microregion	No	No	No	No	No	No	No	No	No	No
FE Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Statistical significance: ***p < 0.001; **p < 0.01; *p < 0.05. Standard error in parentheses. AIC = Akaike Information Criterion. N = Number of observations. Source: Developed by the authors.

References

Almeida, Eduardo. 2012. *Econometria Espacial Aplicada*. Campinas, SP: Editora Alínea.

Alonso, Jose Antonio, and Víctor Martín. 2019. "Product relatedness and economic diversification at the regional level in two emerging economies: Mexico and Brazil". *Regional Studies* 53 (12): 1710–22.

Anderson, T. W., and Cheng Hsiao. 1981. "Estimation of dynamic models with error components". *Journal of the American Statistical Association* 76 (375): 598–606.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. *Kluwer Academic Publishers, Deodrecht*. Dorddrecht: Springer Science & Business Media.

———. 1995. "Local Indicators of Spatial Association—LISA". *Geographical Analysis* 27 (2): 93–115.

Balassa, Bela. 1965. "Trade Liberalisation and 'Revealed' Comparative Advantage". *Manchester* 33 (2): 99–123.

Balland, Pierre Alexandre, Ron Boschma, Joan Crespo, and David L Rigby. 2018. "Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification". *Regional Studies* 53 (9): 1252–68.

Balland, Pierre Alexandre, Cristian Jara-Figueroa, Sergio G. Petralia, Mathieu P.A. Steijn, David L. Rigby, and César A. Hidalgo. 2020. "Complex economic activities concentrate in large cities". *Nature Human Behaviour* 4 (3): 248–54.

Boschma, Ron. 2017. "Relatedness as driver of regional diversification: a research agenda". *Regional Studies* 51 (3): 351–64.

Boschma, Ron, Pierre Alexandre Balland, and Dieter Franz Kogler. 2015. "Relatedness and technological change in cities: The rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010". *Industrial and Corporate Change* 24 (1): 223–50.

Boschma, Ron, and Simona Iammarino. 2009. "Related Variety, Trade Linkages, and Regional Growth in Italy". *ECONOMIC GEOGRAPHY* 85 (3): 289–311.

Chávez, Juan Carlos, Marco T. Mosqueda, and Manuel Gómez-Zaldívar. 2017. "Economic complexity and regional growth performance: Evidence from the Mexican economy". *Review of Regional Studies* 47 (2): 201–19.

COMEXSTAT. 2022. "Plataforma de Consultas e Extrações de Dados Estatísticos do Comércio Exterior Brasileiro". 2022. Acessed July, 20, 2024. http://comexstat.mdic.gov.br/pt/geral.

Davies, Benjamin, and David C. Maré. 2021. "Relatedness, complexity and local growth". *Regional Studies* 55 (3): 479–94.

Dosi, G., M. C. Pereira, A. Roventini, and M. E. Virgillito. 2020. "The labour-augmented K + S model: A laboratory for the analysis of institutional and policy regimes". *EconomiA* 21 (2): 160–84.

Eck, Nées Jan Van, and Ludo Waltman. 2009. "How to normalize cooccurrence data? An analysis of some well-known similarity measures". *Journal of the American Society for Information Science and Technology* 60 (8): 1635–51.

Elekes, Zoltán, Anna Baranowska-Rataj, and Rikard Eriksson. 2023. "Regional diversification and labour market upgrading: Local access to skill-related high-income jobs helps workers escaping low-wage employment". *Cambridge Journal of Regions* XX (August): 1–14.

Fawaz, Fadi, and Masha Rahnama-Moghadamm. 2019. "Spatial dependence of global income inequality: The role of economic complexity". *International Trade Journal* 33 (6): 542–54.

Françoso, Mariane Santos, Ron Boschma, and Nicholas Vonortas. 2024. "Regional diversification in Brazil: The role of relatedness and complexity". *Growth and Change* 55 (1): e12702.

Freitas, Elton, Gustavo Britto, and Pedro Amaral. 2024. "Related industries, economic complexity, and regional diversification: An application for Brazilian microregions". *Papers in Regional Science* 103 (1): 100011.

Freitas, Elton Eduardo, and Emília Andrade Paiva. 2015. "Diversificação e sofisticação das exportações: uma aplicação do product space aos dados do Brasil". *Rev. Econ. NE* 46 (3): 79–98.

Frenken, Koen, Frank van Oort, and Thijs Verburg. 2007. "Related variety, unrelated variety and regional economic growth". *Regional Studies* 41 (5): 685–97.

Fritz, Benedikt S.L., and Robert A. Manduca. 2021. "The economic complexity of US metropolitan areas". *Regional Studies* 55 (7): 1299–1310.

Gala, Paulo. 2017. *Complexidade Econômica: uma nova perspectiva para entender a antiga questão da riqueza das nações*. 1º ed. Rio de Janeiro: Contraponto.

Galetti, Jefferson R B, Simone Tessarin, Paulo César Morceiro, and Milene Simone Tessarin. 2022. "Types of occupational relatedness and branching processes across Brazilian regions". *Area Development and Policy* 0 (0): 1–23.

Galetti, Jefferson Ricardo Bretas, Milene Simone Tessarin, and Paulo Cesar Morceiro. 2021. "Skill relatedness, structural change and heterogeneous regions: evidence from a developing country". *Papers in Regional Science* 100 (6): 1355–76.

Gao, Jian, and Tao Zhou. 2018. "Quantifying China's regional economic complexity". *Physica A: Statistical Mechanics and its Applications* 492:1591–1603.

Golgher, André Braz. 2015. *Introdução à Econometria Espacial*. Jundiaí: Paco Editorial.

Gómez-Zaldívar, Manuel, Felipe J. Fonseca, Marco T. Mosqueda, and Fernando Gómez-Zaldívar. 2020. "Spillover effects of economic complexity on the per capita gdp growth rates of mexican states, 1993-2013". *Estudios de Economia* 47 (2): 221–43.

Hane-Weijman, Emelie, Rikard H. Eriksson, and David Rigby. 2022. "How do occupational relatedness and complexity condition employment dynamics in periods of growth and recession?" *Regional Studies* 56 (7): 1176–89.

Hartmann, Dominik, Mayra Bezerra, Beatrice Lodolo, and Flávio L. Pinheiro. 2020. "International trade, development traps, and the core-periphery structure of income inequality". *EconomiA* 21 (2): 255–78.

Hausmann, Ricardo, and César A. Hidalgo. 2011. "The network structure of economic output". *Journal of Economic Growth* 16 (4): 309–42.

Hausmann, Ricardo, César A. Hidalgo, Sebastián Bustos, Michele Coscia, Sarah Chung, Juan Jimenez, Alexander Simões, and Muhammed A. Yıldırım. 2011. *The Atlas of economic complexity: mapping paths to prosperity*. Center for International Development, Harvard University.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. "What you export matters". *Journal of Economic Growth* 12 (1): 1–25.

Hidalgo, César A. 2021. "Economic complexity theory and applications". *Nature Reviews Physics* 3 (2): 92–113.

———. 2023. "The policy implications of economic complexity". *Research Policy* 52 (9).

Hidalgo, César A., Pierre Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldman, Koen Frenken, Edward Glaeser, et al. 2018. "The Principle of Relatedness". In *Springer Proceedings in Complexity*, Unifying Themes in Complex Systems IX. ICCS 2018, IX:451–57. Cham: Springer International Publishing.

Hidalgo, César A., and Ricardo Hausmann. 2009. "The building blocks of economic complexity". *PNAS* 119 (6166): 15.

Hidalgo, César. A., Bailey Winger, Albert-László Barabási, and Ricardo Hausmann. 2007. "The product space conditions the development of nations". *Science* 317 (5837): 482–87.

IBGE. 2022. "Sidra: sistema IBGE de recuperação automática". 2022. Acessed July, 20, 2022. https://sidra.ibge.gov.br/.

Kapoor, Mudit, Harry H. Kelejian, and Ingmar R. Prucha. 2007. "Panel data models with spatially correlated error components". *Journal of Econometrics* 140 (1): 97–130.

Loonis, Vincent, and Marie-Pierre de Bellefon. 2018. *Handbook of Spatial Analysis: Theory and Application with R*. Eurostat: INSEE.

Mariani, Manuel Sebastian, Alexandre Vidmer, Matsúš Medo, and Yi Cheng Zhang. 2015. "Measuring economic complexity of countries and products: which metric to use?" *European Physical Journal B* 88 (11): 1–9.

Mascarini, Suelene, Renato Garcia, Emerson Gomes dos Santos, Ariana Ribeiro Costa, and Veneziano Araujo. 2023. "Regional heterogeneity and the effects of the related and unrelated varieties on innovation". *Regional Science Policy and Practice*, nº August.

Mewes, Lars, and Tom Broekel. 2022. "Technological complexity and economic growth of regions". *Research Policy* 51 (8).

Pereima, João Basilio. 2020. "Economic development and complexity: Introduction to special issue". *EconomiA* 21 (2): 121–29.

Pinheiro, Flavio L, Pierre-Alexandre Balland, Ron Boschma, and Dominik Hartmann. 2022. "The Dark Side of the Geography of Innovation: Relatedness, Complexity, and Regional Inequality in Europe". 22.02. *Papers in Evolutionary Economic Geography*. Papers in Evolutionary Economic Geography. Utrecht: Utrecht University. Acessed July, 20, 2024. http://econ.geo.uu.nl/peeg/peeg2202.pdf.

Queiroz, Arthur Ribeiro, João Prates Romero, and Elton Freitas. 2023. "Economic complexity and employment in Brazilian states". *CEPAL Review* 139 (April).

Queiroz, Arthur Ribeiro, João Prates Romero, and Elton Eduardo Freitas. 2024. "Relatedness and regional economic complexity: Good news for some, bad news for others". *EconomiA*.

Salles, Fernanda Cimini, Elisa Pinto Rocha, and Felipe Lopes Vieira Vasconcelos. 2018. "A armadilha da baixa complexidade em Minas Gerais: o desafio da sofisticação econômica em um estado exportador de commodities". *Rev. Bras. Inov.* 17 (1): 33–62.

Simoes, Alexander J.G., and César A. Hidalgo. 2011. "The economic complexity observatory: An analytical tool for understanding the dynamics of economic development". *AAAI Workshop - Technical Report* WS-11-17 (January 2011): 39–42.

Stojkoski, Viktor, Philipp Koch, and César A Hidalgo. 2023. "Multidimensional economic complexity and inclusive green growth". *Communications earth & environment* 4 (130).

Stojkoski, Viktor, Zoran Utkovski, and Ljupco Kocarev. 2016. "The Impact of Services on Economic Complexity: Service Sophistication as Route for Economic Growth". *PLoS ONE* August (25).

Tacchella, Andrea, Matthieu Cristelli, Guido Caldarelli, Andrea Gabrielli, and Luciano Pietronero. 2012. "A new metrics for countries' fitness and products' complexity". *Scientific Reports* 2 (723).

Teixeira, Felipe Orsolin, Fabricio José Missio, and Ricardo Dathein. 2022. "Economic complexity, structural transformation and economic growth in a regional context: Evidence for Brazil". *PSL Quarterly Review* 75 (300): 47–79.

Ulyssea, Gabriel. 2018. "Firms, informality, and development: Theory and evidence from Brazil". *American Economic Review* 108 (8): 2015–47.

Verheij, Timo, and Heder de Oliveira. 2020. "Is economic complexity spatially dependent? A spatial analysis of interactions of economic complexity between municipalities in Brazil". *Revista Brasileira de Gestão e Desenvolvimento Regional* 16 (1): 318–38.

Vinci, Gianni Valerio, and Roberto Benzi. 2018. "Economic Complexity: Correlations between Gross Domestic Product and Fitness". *Entropy* 20 (10).

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. MIT Press.

Zhang, Qizi, Yejun Wu, and Lei Wang. 2018. "Implications of Increased Regional Economic Complexity". In *Transforming Economic Growth and China's Industrial Upgrading*, 153–78. Beijing, China: Springer.

Zhu, Shujin, and Renyu Li. 2017. "Economic complexity, human capital and economic growth: empirical research based on cross-country panel data". *Applied Economics* 49 (38): 3815–28.

♦ ACKNOWLEDGMENTS

This work was supported by CNPq under grants 313467/2023 and FAPESP under grants 2022/04734-8 and 2020/08751-9

CONFLITO DE INTERESSE

Os autores declaram não terem quaisquer conflitos de interesse.

EDITOR-CHEFE

Dante Mendes Aldrighi https://orcid.org/0000-0003-2285-5694

Professor - Department of Economics University of São Paulo (USP)

Fernando Salgueiro Perobelli https://orcid.org/0000-0003-2364-8865