

Knowledge Economy in Brazil: Analysis of Sectoral Concentration and Production by Region

José Antonio de França¹ & Wilfredo Sosa Sandoval²

¹ PhD in Accountancy and Economics, University of Brasilia (UnB), Brazil

² PhD in Mathematics, Catholic University of Brasilia (UCB), Brazil

Correspondence: José Antonio de França, PhD in Accountancy and Economics, University of Brasilia (UnB), Brazil.

Received: August 29, 2021

Accepted: October 10, 2021

Online Published: October 12, 2021

doi:10.5539/ijef.v13n11p57

URL: <https://doi.org/10.5539/ijef.v13n11p57>

Abstract

The research presented in this article investigates and analyzes the concentration of knowledge production in Brazil, in the context of a public policy, at postgraduate level, by using the spectral methods grounded on the LQ (location quotient) and CI (concentration index) indicators, in three dimensions, from 2013 to 2018. The dimensions are economics, geography, and time. Economics is represented by Fields and Major Fields of knowledge production. Geography corresponds to the regions identified by each Federation unit (FU). Time is a chronological unit of the timeline in which knowledge is produced. The research then evaluates knowledge concentration in the income performance of the families by FU. The results are robust and indicate significant evidence that sectorial knowledge production in Brazil is regionally unequal and impacts on family incomes, but those family incomes evolve regardless of the knowledge concentration level produced. The research contributions are relevant to assist public policy regulators and monitoring managers, as well as to encourage future discoveries in regional economics applications.

Keywords: knowledge economy, knowledge Concentration Index (CI), LQ Regional Matrix (LQRM), local and global percentages

1. Introduction

This article investigates and analyzes the concentration of knowledge produced in Brazil at postgraduate level, in the context of agglomerations, grounded on the quantitative data of master theses and doctoral dissertations, by Field and Major Field of knowledge, by federation unit (FU), within the timeline from 2013 to 2018, and investigates the contribution of knowledge concentration produced in the income performance of families.

Field is the aggregate of all knowledge produced within the geography represented by a FU. Major Field aggregates the knowledge produced by all related Fields, in the same timeline unit. Knowledge concentration is a non-negative neutral measure that, when orbiting near zero, suggests that a FU is developed. Thus, there are three variables to evaluate the concentration of the knowledge economy: FU, which is equivalent to a geographical territory unit, named Region; Field and Major Field, which correspond to the sectors in which the economy produces knowledge; and time, which is a chronological unit of the timeline in which knowledge is produced. With these variables, a local and global analysis of knowledge concentration is produced.

To investigate the degree of knowledge concentration, the article introduces the concentration index (CI), which uses spectral analysis metrics, in the context of linear algebra theory, applied to the study of agglomerations, guided by the location quotient (LQ) model in three dimensions: economics, geography and time. These metrics indicate how equal or unequal knowledge production is in terms of regional development, and how concentrated or dispersed the knowledge produced by each region is. To evaluate knowledge concentration performance in the average household income (AHI) of families and the statistical significance of this performance, the Data Envelopment Analysis (DEA) deterministic model and the quantile linear regression model are used.

In the LQ specification for the three dimensions, the S sectors (k) of the economy represented by the Fields are identified, totaling 81; by Major Fields, totaling 9; by geographic R regions (i) identified by each FU, totaling 27; and by the T timeline composed of time units (t), totaling 6, divided in periods p of four P units (t), so that $P=4t$ and each period p is defined by $T-P+1$ T timeline unit. Thus, an economic variable V is defined, observed in the

V format (i,k,t) of regional production. Thus, from the LQ in the three dimensions, LQ regional matrices (LQRM) are produced, associated to the covariance matrices (COVM) whose characteristic polynomial eigenvalues (λ) generate the CI calculation defined by the vector norm model.

The DEA deterministic model is specified by having each FU as DMU (Decision Maker Analysis), as an input to the CI variable, with the product being the AHI variable, whose combination of both variables, in the production function, produces the efficiency score as a first stage. To obtain statistical significance in the quantile model, the DEA efficiency score is regressed for each FU, against the respective CI, segregating the result in percentiles.

The metrics for the three models, LQ, CI, and DEA, contain the following characteristics: (a) LQ assumes a non-negative quantum from zero to plus infinite $[0, +\infty]$, and for any quanta higher/lower than 1, the indication is of disproportionality of knowledge production by region, equivalent to regional inequality; otherwise, it is said that the region produces knowledge proportionally equivalent to the aggregate of the Field or Major Field. (b) the IC quantum is also non-negative, varying from zero to plus infinite $[0, +\infty]$ and any quanta higher than zero indicate a non-concentrated knowledge production and, consequently, is drawn away from the *optimal status* of development. (c) the DEA score is also limited to the interval from 0 to 1 $[0, 1]$ because, being 1, it corresponds to the higher relation between product/input that determines efficiency. Lastly, the statistical significance of the efficiency score is acceptable between 0.10 and 0.01, by percentile, calculated by the quantile regression model.

In the context of the premises presented, the research investigates how proportional, concentrated and efficient the knowledge produced in Brazil is, at a postgraduate level, considering the number of master theses and doctoral dissertations, separated by Field, Major Field, region, and period, within the timeline defined from 2013 to 2018.

The actions necessary for the production and success of the research are as follows: to retrieve the quantitative data of master's thesis and doctoral dissertations defended, by FU and Field, from CAPES' thesis and dissertations database, for each time unit, and group them by Major Field; to retrieve the AHI values of IBGE statistics, by FU and time unit; calculate the LQ, the local and global percentages, and the CI, by FU, for each period of the timeline; to evaluate the fulfillment of targets defined by the regulator in market conditions (MC); and to evaluate the AHI's performance according to knowledge concentration, by FU, Field and Major Field, in each period of the timeline, as well as to calculate and analyze the statistical significance of the quantile estimators of the DEA score.

MC must be arbitrated by the regulator as a reference for the effectiveness of public policy, however, public policy must be understood not only as an action of the State, but as a set of actions that seek to build a real future (Boneti, 2017).

The results produced by the research allow managers and regulators to correct possible deviations in achieving the goals during the implementation of the public policy of knowledge production so that the planned objectives are met.

The article is structure in four other sections. Section 2 addresses the evolution of the agglomeration and knowledge production binomial, highlighting the relevant contributions of the literature from the end of the 19th century to the early 21st century. Section 3 presents the evolution of the methodology and the research model. Section 4 analyzes knowledge economy in Brazil. Section 5 summarizes the main research contributions as well as the main limitations.

2. Agglomeration and Knowledge Production

The agglomeration and knowledge production binomial was studied by Feser, Renski, and Goldstein (2008), who evaluated subregional concentration of the research and development industry in 18 universities and other institutions in the Appalachian region, United States, as of the 1990s, with the objective of investigating its relationship with employment. They developed knowledge infrastructure indicators for the region, using the location quotient (LQ) model in two dimensions and other methods. The evidence from the study, as stated by the authors, suggests that agglomeration is associated with the creation of new businesses and not with employment growth.

In the context of knowledge as the foundation of a society's transformation, Hájková and Hájek (2014) used the DEA model to evaluate the impact of the knowledge bases of European cities in population and economic growth in the period from 2004 to 2009. The authors' evaluation indicates that the cities of the new member-states transform their knowledge bases in urban growth more efficiently than those from EU-15, due to their low initial level of creation and systematic knowledge transfer, because of the economic convergence and internal knowledge transfer, which were supported by sales.

Knowledge, as a grouping engine of the industry, promotes the appearance of urban agglomerations, an increase in income levels and, consequently, in social welfare, as addressed by Marshal (1890; IV,X,3), Gabe and Abel (2010). In this context, knowledge boosts the development of a region and/or a nation, as it builds the capacity and the skills of human capital (Romer, 1996) to increase and improve productivity, whether by the proper use of technology or by the rational use of natural resources, or even by the improvement of processes, as addressed by Quatraro (2010), Cooke et al. (1997) and Antonelli (2008). It may even be by interacting with and impacting other sectors of the economy, as suggested by the findings of Zabala-Iturriagoitia, Voigt, Gutiérrez-Gracia, and Jiménez-Sáez (2007) in a similar study in the European Union.

In Brazil, the knowledge production industry, segregated by Major Field and Field, is spread out in the five macro-regions or geopolitical regions of interest of the education policy. This distribution gains emphasis in the Major Field and Field according to the direction of the local economic activity, where a region stands out in comparison with the others, creating a regional (local) economic characteristic. Examples of such characteristic include the prominence of agribusiness in the Central-West region; of tourism in the Northeast region; and of technology and mining in the Southeast and South regions.

3. Model Specifications

This section retrieves from the literature the development of the agglomeration measurement model and defines the research model.

3.1 Location Quotient (LQ)

LQ is a social indicator traditionally used in the literature of regional economics to measure the spatial effects of agglomerations. Seminal studies such as the one attributed to Haig (1920) and the one produced by Florence (1929) are evidence of that use. However, these studies present the LQ as being only two-dimensional, represented by economics and geography, which is a limitation to the use of this indicator in a matrix application.

Later studies, such as those of Alexander (1954), Tiebout (1962, p. 9), Richardson (1985), Stimson, Stough, and Roberts (2006, p. 107), Gabe and Abel (2010), and Thulin (2014), have also measured the effects of each sector in the spatial activity of the economy through the LQ in two dimensions, but other studies, such as those by North (1955) and Gilmer and Keil (1989), criticize the effectiveness of the results obtained through the analytical LQ model with the seminal specification.

The seminal conception of the two-dimensional LQ model is one-off and specific. Thus these two dimensions are the restriction that limits the dynamic application of the model with a matrix-based use. The model attributed to Haig (1920) by the Division of Research and Statistics of New York (June 2017) is specified as follows:

$$LQ = \frac{\text{Industry } X\text{'s \% Share of jobs in the Regional Economy}}{\text{Industry } X\text{'s \% Share of jobs in the U.S. Economy}}$$

and it considers:

$$LQ = \begin{cases} > 1 \Rightarrow \text{exports the excess} \\ = 1 \Rightarrow \text{in balance} \\ < 1 \Rightarrow \text{imports what is lacking} \end{cases}$$

The model proposed by Florence (1929) is specified as follows:

$$\text{Local concentration} = \frac{A}{C} * \left(\frac{B}{D}\right)^{-1}$$

and it considers A to be a location with a specified industry; C to be a location that gathers the entire specified industry; B to be all the locations in which there are specified industries; and D to be all locations of all specified industries.

On the same line of the seminal concept of the two-dimensional **LQ**, Stimson, Stough & Roberts (2006, p. 107) also rewrote the model as follows:

$$LQ_{ir} = \frac{E_{ir}}{E_r} * \left(\frac{E_{iN}}{E_N}\right)^{-1}$$

and they consider E_{ir} to be employment in sector i in region r ; E_r to be total employment in region r ; E_{iN} to be employment in sector i in the reference area (N = national reference); and E_N to be total employment in the area of national reference.

3.2 Research Model

Realizing the limitation imposed upon the LQ in only two dimensions, De França (2020) and De França and Sosa (2020) proposed a matrix-based approach. This approach consists in adding a third variable: time. With the introduction of a new three-dimensional LQ, the context of theory of linear algebra and spectral analysis is also introduced. With the three dimensions, the specification of the LQ removes the restriction present since the seminal studies and constitutes a consistent non-parametric social indicator that signals proportional equalities/inequalities between regions, as well as indicating a concentration/dispersal in regional development, within a timeline, with vast application in the economy. The model is conceived based on a set of equations, as follows:

1) General three-dimensional LQ model and introduction of economic variable V

$$LQ(i, k, t) = \frac{\left[\frac{V(i,k,t)}{\sum_{j=1}^S V(i,j,t)} \right]}{\left[\frac{\sum_{l=1}^R V(l,k,t)}{\sum_{l=1}^R (\sum_{j=1}^S V(l,j,t))} \right]} \tag{1}$$

the specification considers V to be the variable that brings together the three dimensions of the economy; k and j are the sectors of the economy; i and l are the regions of geography; t is the unit of the timeline.

If the production by FU is proportionally equivalent, one expects $LQ(i, k, t) \equiv 1$.

2) Aggregation of the economy according to the V economic variable

a) Aggregation of the production by k sector in each t unit of time

$$ASV_{(k,t)} = \sum_{l=1}^R V(l,k,t)$$

b) Aggregation of the production by i region in each t unit of time

$$ARV_{(i,t)} = \sum_{j=1}^S V(i,j,t)$$

Total aggregation of the production in each t unit of time

$$TA_{(t)} = \sum_{l=1}^R (\sum_{j=1}^S V(l,j,t))$$

3) Definition of the global (GP) and local (LP) percentages according to the V economic variable

$$GP(k, t) = \frac{ASV_{(k,t)}}{AT_{(t)}} \tag{2}$$

$$LP(i, k, t) = \frac{V(i,k,t)}{ARV_{(i,t)}} \tag{3}$$

The GP and the LP allow Regulators and Public Policy Managers to follow up and monitor the fulfillment of the target defined in the market condition (MC), within the timeline for which the target was established.

a) LQ restrictions according to V based on GP and LP

The LQ calculated by the model of Equation 1 is rewritten by Equation 4. This is necessary to show that, when the mathematical limitation of the division by zero takes place, the LQ is defined as zero. This occurs when a sector does not produce in any region.

$$LQ(i, k, t) = \begin{cases} 0 & \text{if } V(i, k, t) = 0 \\ \frac{LP(i,k,t)}{GP(k,t)} & \text{if } V(i, k, t) \neq 0 \end{cases} \tag{4}$$

b) LQ (LQRM) and Covariance (COVM) matrices

The elements of each LQRM are defined based on the LQ indexes in the three dimensions. A $LQRM_i^p$ sized SxP is defined for each region i and period p and, from each of those matrices, a $COVM_i^p$ sized PxP is defined. For both matrices, S represents the sectors of the economy (Field and Major Field), and P represents the units of time of each p period, which is defined by P intervals of units of time. Thus, Equation 5 defines the elements of each $LQRM_i^p$ and Equation 6 defines the elements of each $COVM_i^p$.

$$LQRM_i^p = \overline{LQ}_i^p(t) = \frac{1}{S} \sum_{j=1}^S LQ(i, j, t) \tag{5}$$

$$COVM_i^p = \frac{1}{S} (LQRM_i^p)^T (LQRM_i^p) \tag{6}$$

c) Eigenvalues (λ) and CI

Each CI is defined by a vector of eigenvalues (λ) from a $COVM_i^p$. The CI is a real neutral number, sized $R \times (T-P+1)$, which evaluates the consistency of regional development aiming at the equitable offer of opportunities, calculated through Equations 7 and 8.

$$Det (COVM_i^p - \lambda I) = Pol(\lambda) \quad (7)$$

$$CI(i, p) = |AV(i, p)| = \sqrt{\sum_{w=1}^P \lambda_w^2} \quad (8)$$

If the economy is concentrated, $CI(i, p) \equiv 0$ and the opportunities are distributed.

4) The efficiency of knowledge concentration performance

The DEA performance scores are evaluated by the production function (Equation 9) as per models CCR and BCC introduced by Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984).

$$Score(h) = f\left(\frac{1}{AHI}, CI\right) \quad (9)$$

CCR model:

$$\text{Min } h_0 = \left[\frac{\sum_{i=1}^r v_i X_{i0}}{\sum_{j=1}^s u_j Y_{j0}} \right]$$

Subject to:

$$\left[\frac{\sum_{i=1}^r v_i X_{ik}}{\sum_{j=1}^s u_j Y_{jk}} \right] \geq 1, \forall k$$

$$v_i, u_j, \forall j, i$$

BCC model:

$$\text{Max } h_0$$

Subject to:

$$X_{i0} - \sum_{k=1}^n X_{ik} \lambda_k \geq 0, \forall i$$

$$-h_0 Y_{j0} + \sum_{k=1}^n Y_{jk} \lambda_k \geq 0, \forall j$$

$$\sum_{k=1}^n \lambda_k = 1, \lambda_k \geq 0, \forall k$$

In which $X = (CI)$; $Y = (1/RDM)$; $i =$ less efficient FU_i (DMU); $j =$ more efficient FU_j (DMU); $k = (i,j)$; $\lambda =$ higher productivity of DMU; $u, v =$ weights.

5) The statistical significance of knowledge concentration performance

The quantile model is recommended by the literature because it enables the evaluation of the consistency of the coefficients in a dynamic manner, in the timeline, in quartiles and percentiles, and also because it dispenses with some of the properties demanded by the traditional models, such as homoscedasticity, as discussed by Li (2015) and Hinojosa (2017), and as done by Henriques (2019).

One of the model specifications is presented by Koenker (2005, pp. 123-125), which defines the rate of convergence as

$$Q_y(\tau x) = X^T \beta_\tau$$

and argues, in a contextualized manner, that, in order to explore the asymptotic behavior of the regression classification process, one must consider the linear model in scale of location, defined by

$$Q_{y_i}(\tau | x_i) = x_i^T \beta + x_i^T \gamma F_u^{-1}(\tau)$$

Arising from the **iid** error model

$$Y_i = x_i^T \beta + (x_i^T) u_i$$

In which, for the application of the research: β is the asymptotic estimator; Y is the DEA performance score; X is the CI of the eigenvalues vector; F is the conditional distribution function.

The purpose of introducing this model into the research is to evaluate the signal (direction) and the statistical significance of the coefficients.

4. An Analysis of Knowledge Economy in Brazil

The primary data that make up the sample gather 471,308 master's theses and doctoral dissertations defenses in the 81 Fields of the 9 Major Fields of knowledge, from 2013 to 2018. The programs approved for those fields totaled 4,654 specialties at the master and doctorate level, both academic and professional, offered by 6,877 courses in the 27 FUs (Table A4) in the geography. On average, more than 76,300 documents for master's theses and doctoral dissertations were produced per year in the last five years. The choice for the production of knowledge, at a graduate level (master's and doctorate), is adequate due to the coverage of the entire national territory, the reliability of the data, and the organization of the document production system by CAPES (Coordination for the Improvement of Higher Education Personnel) and by CNPq (National Council for Scientific and Technological Development), the regulatory agencies for knowledge production in Brazil. CAPES and CNPq were both created in 1951 with the mission of regulating knowledge production at a master's and PhD level and foster scientific and technological production, respectively.

4.1 Dimensions of Knowledge Economy

The knowledge economy is made up of the regions (i) represented by each of the 27 FUs; by the sectors (k) represented by each of the 81 Fields and 9 Major Fields; and by the units of time (t) distributed over the 6 years of the 2013 to 2018 timeline, making up 3 p periods of P units of time each ($P=4t$), so that ($T-P+1=3=p$), with the following distribution: $p_1=P$ (2013 to 2016); $p_2=P$ (2014 to 2017); and $p_3=P$ (2015 to 2018). Thus, the economic variable V, defined in the model, contains the database with all the required attributes.

4.2 Analysis of the Descriptive Statistics of Knowledge Production

The statistical estimators of the primary distribution data, totaled by FU and level of knowledge, are shown in Table A1. The estimators reveal two relevant characteristics: one is the proximity and position change between mean and median, where the mean is sometimes above and sometimes below the mean. The other is the reduced dispersion measured by the Coefficient of Variation (CV). At the master's degree level, the two largest dispersions orbit around $\frac{1}{2}$ standard deviation of the mean (0.442 and 0.542), attributed to the FUs AP and AC, and the two smallest ones (0.048 and 0.053) are attributed to FUs MS and RJ. These magnitudes of the CV suggest a sector concentration of the production of knowledge around the regional mean. At the doctorate level, the elasticity of the dispersion is higher, with 2 FUs, AC and RR, showing a CV around 2 standard deviations from the median (2.082 and 1.922), while the lowest two, SP and DF, disperse with 0.099 and 0.118. The most elastic amplitude of the CV reflects an insufficiency of regional production, such as in FUs AC and RR, with high inequality.

Table A2 shows the estimators at the master's and doctorate levels. One characteristic of this distribution consists in the positioning of the estimators of the mean at the upper half of the median throughout the entire timeline. Another characteristic is the opposite direction of the dispersion (CV), which increases at the master's level and decreases at the doctorate level. The first characteristic produces an effect of asymmetry between the FUs, and the second indicates an increase in inequality at the master's level and a reduction of inequality at the doctorate level.

4.3 Analysis of the Dynamics of the Global Percentages by Major Field (GPMF)

The global percentages shown in Table 1 were calculated according to the model defined in equation (2) and represent the execution of the public policy of knowledge production in Brazil. The quantum of each GPMF, in each column, is the participation of the MF in the global knowledge economy. In this context, it can be observed that MF6 (applied social sciences) leads with the highest production of knowledge in five of the six units of time, at the master's level, and, at the doctorate level, throughout the entire timeline, while the lowest production of knowledge is that of MF2 (biological sciences) in five of the six units of time, at the master's level, and, at the doctorate level, it is that of MF8 (language studies and arts) throughout the entire timeline.

These results supply robust evidence that biological sciences, MF2, (studies of the origin of life), at the master's level, shows a low priority, with a participation ranging between 5.08% and 7.15% of all knowledge produced, although it places higher at the doctorate level, between 8.75% and 10.5%, being, however, less than half of the most productive MF (applied social sciences). To evaluate the fulfillment of the target set by the regulator, we compare the MC to the percentages of the execution of each MF, but the aforementioned MF is not observed. In this condition, indirectly, the comparison is made by observing the CI quanta of each MF, which is developed in the subsection.

Table 1. Global percentages by major field (GPMF) of the knowledge produced at the master's and doctorate levels – 2013 to 2018

MF	GPMF of knowledge at the master's level						GPMF of knowledge at the doctorate level					
	M2013	M2014	M2015	M2016	M2017	M2018	D2013	D2014	D2015	D2016	D2017	D2018
MF1	0.1038	0.1033	0.0936	0.0913	0.0877	0.0861	0.1325	0.1321	0.1276	0.1319	0.1211	0.1235
MF2	0.0715	0.6064	0.0594	0.0542	0.0551	0.0508	0.0994	0.1050	0.0991	0.0949	0.0919	0.0875
MF3	0.0953	0.0902	0.0887	0.0882	0.0831	0.0814	0.0951	0.0970	0.0952	0.0934	0.0970	0.0901
MF4	0.1608	0.1583	0.1612	0.1605	0.1630	0.1598	0.1697	0.1671	0.1709	0.1663	0.1692	0.1618
MF5	0.1211	0.1259	0.1251	0.1291	0.1339	0.1381	0.0744	0.0787	0.0758	0.0862	0.0864	0.0852
MF6	0.1689	0.1697	0.1676	0.1657	0.1602	0.1635	0.2000	0.1920	0.1908	0.1906	0.1941	0.2006
MF7	0.1077	0.1089	0.1091	0.1125	0.1153	0.1130	0.1011	0.0963	0.1032	0.0984	0.0917	0.0940
MF8	0.0600	0.0612	0.0691	0.0641	0.0575	0.0608	0.0628	0.0649	0.0643	0.0592	0.0593	0.0595
MF9	0.1109	0.1184	0.1263	0.1345	0.1440	0.1465	0.0651	0.0670	0.0730	0.0792	0.0893	0.0978

Source: the authors. MF identifies the major field of knowledge; M identifies the master's level; D identifies the doctorate level. MF1: agricultural sciences; MF2: biological sciences; MF3: exact and Earth sciences; MF4: human sciences; MF5: health sciences; MF6: applied social sciences; MF7: engineering; MF8: linguistics, language studies and arts; MF9: multidisciplinary.

4.4 Analysis of the CI Matrices of the Execution of the Public Policy of Knowledge Production

The IC matrices represented in Table 2 were calculated according to the model specified in Equation (8). Each line corresponds to a FU associated to the knowledge produced in a Field (F) and to the total knowledge produced in all of the Fields (MF), at the master's and doctorate levels, in each p period, in the T timeline from 2013 to 2018. The total of p periods dimensioned by the P intervals of 4 units of time is determined thus: $p=T-P+1=3$.

Through the metrics of the spectral model, an LQ index approaching 1 implies a CI quantum approaching zero. The left portion of Table 2, Knowledge concentration (MF), shows that only RS approaches this metric, around 0.04 at the master's level, and ranging from 0.05 to 0.09 at the doctorate level. To the right, Knowledge concentration (F) does not show signs of convergence. Apart from this evidence, the CI quanta suggest that, by MF , some FUs are less distant from the convergence than by F , seeing as 7/27 (CE, DF, MG, PE, PR, RJ, SP) show a CI lower than 1 in all CIMMF and CIDMF periods. In Knowledge concentration (F), the CI quanta are all higher than 1. This suggests that the view of the economy regulator in Brazil is focused on knowledge production by Major Field, at the expense of the Field.

Thus, out of the findings of the research, two are significantly relevant: (1) in knowledge by MF , the reality of 6 of the 27 FUs (AC, AP, RN, RO, RR, TO) is very different from that of the others in the production of knowledge, indicating a high level of regional inequality; (2) in knowledge by F , FUs of the south and southeast regions (MG, RJ, PR, SC, SP) show less inequality, despite being far from the convergence.

Considering that the Market Condition (MC) is not observed, it is possible to safely infer that the regulator's target, if there is one, was not met, because such a fulfillment would necessarily imply a CI quantum close to zero, a situation that is only present in UF **RS**. Under these circumstances, there is no evidence of monitoring of the implementation of the public policy for knowledge production, or, if there was such a monitoring, it was not effective.

Table 2. Concentration indexes (CI) of the knowledge produced by Major Field (MF) and Field (F) in each FU-Brazil from 2013 to 2018

FU	Knowledge concentration (MF)						Knowledge concentration (F)					
	CIMMF			CIDMF			CIMF			CIDF		
	p1	p2	p3	p1	p2	p3	p1	p2	p3	p1	p2	p3
AC	10.0744	6.1605	5.4980	35.6535	6.7353	13.2142	11.4649	9.6299	7.7852	4,268.187	4.3766	8.0042
AL	0.4327	0.4987	0.3342	15.1212	9.5183	5.6702	10.1385	8.4423	6.8032	12.6220	10.8111	9.7492
AM	1.6599	1.5948	1.4336	8.3739	6.4440	5.7757	13.0074	13.3033	13.1925	24.0904	21.4690	18.9440
AP	9.0390	6.9998	6.2402	24.2343	21.696	27.122	47.2588	47.8487	55.9755	863.4173	157.3701	156.4332
BA	0.9435	0.8535	0.6457	1.2877	1.1213	1.0323	2.9947	3.2639	3.6160	5.3960	5.6828	5.1948
CE	0.1134	0.1071	0.1102	0.2164	0.2726	0.2979	5.2700	5.8436	6.1164	7.6071	6.6686	6.8980
DF	0.4127	0.5491	0.7917	0.7435	0.7543	0.6966	6.4884	6.4594	6.1473	9.1721	9.0369	8.3190

FU	Knowledge concentration (MF)						Knowledge concentration (F)					
	CIMMF		CIMMF		CIDMF		CIMF		CIDF		CIDF	
	p1	p2	P3	p1	p2	p3	p1	p2	p3	p1	p2	p3
ES	0.2323	0.1808	0.2141	1.8420	1.5907	1.2198	4.6936	5.0681	5.0667	23.6314	18.2317	15.5213
GO	0.6277	0.6004	0.5863	1.5219	1.5243	1.5712	4.3531	4.7343	4.2545	9.2611	8.3062	6.2982
MA	1.1300	0.9694	0.8186	7.4482	10.5419	13.279	10.8323	12.5010	12.8083	65.3155	54.6575	48.2809
MG	0.2906	0.2727	0.2767	0.9445	0.8077	0.7833	1.8639	1.7243	1.6748	4.3786	3.8759	3.5953
MS	1.1833	1.0696	1.0839	4.7011	4.0072	3.3657	4.1416	4.1063	4.0700	15.8626	16.6014	16.0415
MT	1.2362	1.4502	1.4719	14.0411	9.7252	9.2579	11.4781	11.7680	12.7817	53.7161	26.8797	24.6546
PA	0.7241	0.8339	0.8699	2.0334	1.6454	1.6653	17.3799	15.9237	15.8336	20.0373	18.8655	22.8586
PB	0.3436	0.3196	0.2655	1.2998	1.1296	1.2639	3.2059	2.9317	2.5896	9.3363	8.0001	7.8407
PE	0.4262	0.4565	0.4112	0.2730	0.3137	0.3292	22.0088	30.9578	41.5005	25.4482	19.0283	11.6025
PI	1.5191	1.7565	1.8920	6.0433	6.8507	5.9878	31.8226	44.0378	49.2489	62.2206	63.7700	59.4250
PR	0.1915	0.2004	0.2121	0.2912	0.2647	0.2557	2.2182	1.7704	1.4966	5.5549	5.0487	4.3537
RJ	0.6096	0.5861	0.5471	0.2338	0.2431	0.2471	6.2275	5.9810	5.8562	6.2904	7.2004	8.4851
RN	5.4742	6.0439	8.2240	4.2519	3.0658	2.7915	8.8235	8.9600	9.3359	17.4340	14.9417	13.7853
RO	4.6602	4.6024	5.5005	35.6535	27.390	23.413	111.295	96.9868	86.7089	4,268.187	2,920.981	1,858.2584
RR	5.4309	5.2555	4.6748	35.6535	6.7353	13.214	11.1337	12.2021	12.7947	4,268.187	4.3766	8.5811
RS	0.0402	0.0412	0.0419	0.0994	0.0790	0.0596	1.4317	1.3857	1.3865	1.8530	1.8146	1.7812
SC	0.3990	0.3873	0.3916	1.0271	0.8575	0.7816	3.5209	3.3040	3.1661	20.1596	24.7655	30.8693
SE	0.3537	0.2896	0.2410	3.1284	2.9679	2.4189	77.2995	56.1673	45.4750	14.0933	50.4631	99.2052
SP	0.1284	0.1293	0.1317	0.1168	0.1136	0.1145	1.6451	1.6660	1.7355	1.1485	1.1735	1.2341
TO	4.2182	4.7406	4.5765	19.2853	17.4029	15.055	34.5271	34.7619	32.8040	77.7294	51.6991	20.3451

Source: the authors. CIMF: CI of the master's level by field; CIDF: CI of the doctorate level by field; CIMMF: CI of the master's level by major field; CIDMF: CI of the doctorate level by major field; FU: Federation Unit.

4.5 Analysis of the Performance of Knowledge Concentration

The performance status of the concentration of knowledge in the improvement of family income is analyzed by the DEA score, in a first stage. This entails the use of the production function defined by Equation (9). The status defines the classification of performance as fully efficient, partially efficient, and approaching the efficiency frontier, as shown in Table 3.

Full efficiency is met when the FU/DMU is benchmark with all scores equal to 1; in this article, the total of scores is equal to 12. Partial efficiency happens when the FU/DMU obtains at least one score =1 among all possible. An approach to the efficiency frontier is shown when the FU/DMU does not satisfy the first two conditions, but has at least one DEA score above the 0.75 cutoff defined in this research. The criterion for computing the quantum of the score in each FU/DMU considers the whole of the dimensions Field or Major Field, period, level of knowledge, and the fixed and variable effects, input and output, CCR and BCC, attributes of the DEA model. In this criterion, only 16 of the 27 FU/DMU obtained a score higher/equal to the cutoff.

1) Status of classification of performance efficiency

The performance efficiency results were obtained with the use of the open-source *Gretl* statistical package. Full efficiency was only achieved by FU RS, which scored 1 in the 12 possible scores, indicating that knowledge concentration improved the regional family income, as highlighted in the first line of Table 3.

The status of partial efficiency was achieved by 8 of the 27 FU/DMU (AL, BA, CE, MA, MG, PA, PB, SP), which scored between 2 and 8 out of the 12 possible scores. This score shows that the FUs with the best performance are SP and MG, followed by AL and MA, and, with the worst performance, CE and PA.

Approaching the Efficiency Frontier are 7 of the 27 UF/DMU (AM, PI, AC, SE, PE, PR, RN), which obtained from 2 to 5 scores at least equal to the cutoff. This score suggests that a high CI contributes little to the improvement off family income.

Thus, the model's responses for the three efficient performance statuses suggest that the *optimal* knowledge concentration impacts the AHI. However, the data from Table A4 reveal that the highest AHI, attributed to FU FD, did not even score the minimum of the cutoff. In light of this robust evidence, the inference is that, while a CI approaching zero improves family income, this income can be improved regardless of the *optimal status* of knowledge production, as in the findings of Zabala-Iturriagoitia, Voigt, Gutiérrez-Gracia, and Jiménez-Sáez

(2007); Susiluoto (2003); Cooke, Uranga, and Etxebarria (1997); Antonelli (2008), and Quatraro (2010), who did not find evidence that a higher income is related to efficiency.

In spite of the robust evidence presented, it is necessary to consider a relevant restriction to the model in a first stage. The CI and the AHI are associated to a same FU in the timeline. However, due to social mobility, knowledge produced in one FU may impact income in another FU. If this mobility takes place, performance converted into efficiency may present an upward or downward bias and impact the classification status.

Table 3. Performance of knowledge concentration in household income measured by DEA between 0.75 and 1 per FU - Brazil from 2013 to 2018

UF	Full Efficiency		Partial Efficiency		Neighborhood Efficiency Frontier	
	F	MF	F	MF	F	MF
RS	12/12	12/12				
SP			8/12			
MG			8/12			
AL			6/12	6/12		
MA			6/12	6/12		
BA			4/12			
PB			2/12			
CE			2/12	2/12		
PA				2/12		
PR					4/12	
AM					3/12	3/12
AC					3/12	3/12
PI					2/12	3/12
SE					2/12	3/12
PE					2/12	5/12
RN						2/12

F: Field; MF: Major Field. Full efficiency all scores DEA=1; Partial efficiency part of DEA scores = 1; Neighborhood of the efficiency frontier $0.75 \leq \text{Scores DEA} < 1$.

4.6 Analysis of the Statistical Significance of the Concentration of Knowledge

Knowing the impact of the degree of knowledge concentration on family income, which defined the efficiency status, what is under analysis now is the statistical significance of this concentration and the direction of the coefficients, in a second stage, through the quantile model of linear regression, as defined in 3.2(4), using the DEA score and the degree of knowledge concentration measured by the CI. This model was used by Brown and Scott (2012) to explain the migration of human capital and agglomerations in Canada.

For each of the 16 FUs with performances placed between full efficiency and approaching the efficiency frontier (Table A3), 24 regressions were conducted (three by period/Field and three by period/Major Field, in percentiles 0.25; 0.50; 0.75; and 0.99) and in all coefficients the direction is positive (+). The choice was made to use only the CI of the knowledge produced at the master's level (CIM) because the production base is larger, with a broader coverage of observations.

The coefficients of FUs RS and SP, in all p periods, are statistically significant with 99% reliability. FUs CE and PR also show, in all p periods, 95% and 99% reliability. The other FUs present a statistical significance that varies from 90% to 99%, in aggregates Field, Major Field, and p periods.

The coefficients of all p periods, in percentiles 0.75 and 0.99, are statistically significant in all FUs, by Field and Major Field, such that, in percentile 0.99, all coefficients are statistically significant at 99% reliability.

These results show robust evidence that the efficiency scores produced by the DEA model, in the first stage, are relevant to indicate, in a consistent manner, that the production of knowledge in Brazil impacts family income, but that family income evolves independently from the proportionality of the knowledge produced in each region.

5. Conclusion

This article presented the results of the research that investigated and analyzed the concentration of knowledge produced in Brazil at postgraduate level, based on master's theses and doctoral dissertations from 2013 to 2018,

by Field and Major Field of knowledge production within the regional geography.

Market conditions with the metrics defined by the regulator for the fulfillment of goals are not observable. With the absence of these metrics, one infers the achievement of the goal with a **CI** near zero, which defines the *optimal status* of a region's development. In the geography of all 27 FUs, only **FU RS** indicates the fulfillment of the goal with a **CI** bearing near zero for the knowledge produced by Major Field. For the knowledge produced by Field, all FUs present divergence from the fulfillment of the goal.

The Major Fields with highest level of quantitative participation in total knowledge production are Social Sciences and Human Sciences, while Biological Sciences and Multidisciplinary present the lowest levels of participation.

Efficiency of knowledge concentration performance, in the improvement of family income, was only observed in FU RS with a DEA score equal to 1 in all time periods. Only 16 out of the 27 FUs had a performance score between 0.75 and 1. The statistical significance of these performance scores is consistent with 10%, 5%, and 1%.

The evidence is robust when indicating that knowledge production in Brazil is not regionally developed, and the absence of a market condition with targets setting does not allow for the fulfillment of the market regulator's objectives. To mitigate this gap, the present research referred to the spectral methodology grounded on LQ regional matrices (LQRM), covariance matrices (COVM), whose characteristic polynomial eigenvalues (λ) generate the CI calculation, which indicated that knowledge concentration improves family income. Nevertheless, family income evolves regardless because social mobility may interfere in the region's income. However, this hypothesis has not been assessed.

Lastly, the results presented encourage future research in the field of knowledge production, as well as assist researchers, regulators and managers in applications of the regional economics.

References

- Alexander, J. W. (1954). The Basic-Nonbasic Concept of Urban Economic Functions. *Economic Geography*, 30(3), 246-261. <https://doi.org/10.2307/141870>
- Antonelli, C. (2008). *Localised Technological Change. Towards the Economics of Complexity*. Routledge, New York. <https://doi.org/10.4324/9780203932001>
- Boneti, L. (2017). Pol íficas p úblicas por dentro. São Paulo, Mercado de Letras.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078-109. <https://doi.org/10.1287/mnsc.30.9.1078>
- CAPES. Brasil. Retrieved from <http://www.capes.gov.br>
- IBGE. Brasil. Retrieved from <https://www.ibge.gov.br>
- Brown, W. M., & Scott, D. M. (2012). Human capital location choice: Accounting for amenities and thick labor markets. *Journal of Regional Science*, 52(5), 787-808. <https://doi.org/10.1111/j.1467-9787.2012.00772.x>
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2, 429-444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
- Cooke, P., Uranga, M. G., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organizational dimensions. *Research Policy*, 26, 475-491. [https://doi.org/10.1016/S0048-7333\(97\)00025-5](https://doi.org/10.1016/S0048-7333(97)00025-5)
- De França, J. A. (2020). *Análise da Concentração de Setores Produtivos: Teoria e aplicações*. Tese (Doutorado em Economia) Universidade Católica de Brasília. Brasília (Junho/2020).
- De França, J. A., & Sosa, W. (2020). *Concentration Analysis of Productive Sectors Regional Studies*. Center de Recerca Matem àtica (CRM). Preprint N.1250 (November 4).
- Division of Research and statistics of New York. State of Opportunity (Jun, 2017).
- Feser, E., Renski, H., & Goldstein, H. (2008). Clusters and Economic Development Outcomes Analysis of the Link Between Clustering and Industry Growth. *Economic Development Quarterly*, 22(4). <https://doi.org/10.1177/0891242408325419>
- Florence, P. S. (1929). *The Statistical Method in Economics and Political Science: A treatise on the quantitative and institutional approach to social and industrial problems*. Harcourt, New York.
- Gabe, T., & Abel, J. R. (2010). Agglomeration of Knowledge. *Urban Studies Journal*.

<https://doi.org/10.1177/0042098010371988>

- Gilmer, W., & Keil, R. (1989). *The location quotient and central place theory*. Federal Reserve Bank of Dallas. Retrieved from <https://www.dallasfed.org/~media/documents/research/papers/1989/wp8916.pdf>
- Haig, R. M. (1928). *Location Quotients: A statewide and Regional Analysis Division of Research and Statistics of New York State* (2017). Retrieved from <https://labor.ny.gov/stats/PDFs/Location-Quotients-A-Statewide-and-Regional-Analysis.pdf>
- Hájek, V., & Hájek, P. (2014). Efficiency of knowledge bases in urban population and economic growth – Evidence from European cities. *Cities*, 40(Part A), 11-22. <https://doi.org/10.1016/j.cities.2014.04.001>
- Henriques, I. C. (2019). *Eficiência do setor bancário brasileiro: modelo DEA dois etágios com regressão truncada bootstrapped*. Dissertação (Mestrado em Administração) Universidade de Brasília. Brasília, 2019.
- Hinostroza, A. A. A. (2017). *Regressão quantílica bayesiana em modelos de fronteira de produção estocástica*. Dissertação (mestrado em estatística). Universidade Federal do Rio de Janeiro, 2017.
- Koenker, R. (2005). *Quantile regression*. Cambridge. <https://doi.org/10.1017/CBO9780511754098>
- Mingxiang, L. (2015). Moving Beyond the Linear Regression Model: Advantages of the Quantile Regression Model. *Journal of Management*, 41(1), 71-98. <https://doi.org/10.1177/0149206314551963>.
- Marshall, A. (1920). *Principles of Economics* (6th ed.). MacMillan, Toronto.
- North, D. C. (1955). Location Theory and Regional Economic Growth. *Journal of Political Economy*, 63(3), 243-258. <https://doi.org/10.1086/257668>
- Quatraro, F. (2010). Knowledge Coherence, Variety and economic Growth: Manufacturing Evidence from Italian Regions. *Research Policy*, 39, 1289-1302. <https://doi.org/10.1016/j.respol.2010.09.005>
- Richardson, H. W. (1985). Input-Output and Economic Base Multipliers: Looking Backward and Forward. *Journal of Regional Science*, 25(4), 607-661. <https://doi.org/10.1111/j.1467-9787.1985.tb00325.x>
- Romer, D. (1996). *Advanced Macroeconomics* (4th ed.). McGraw-Hill, U.S.A.
- Silva, J. de S., Ferreira, M. de O., & Leite, J. R. F. (2017). Eficiência técnica dos produtores de mandala do Vale do São Francisco. *REA*, 15(1). <https://doi.org/10.25070/rea.v15i1.367>
- Stimson, R. J., Stough, R. R., & Roberts, B. H. (2006). *Regional Economic Development: Analysis and Planning Strategy* (2nd ed.). Verlag Berlin Heidelberg, Springer.
- Susiluoto, I. (2003). Effects of ICT on Regional Economic Efficiency. *Helsinki City Urban Facts Office Web Publications*, (16).
- Thulin, P. (2014). Local multiplier and economic base analysis. *Working Papers 29 Series from Swedish Entrepreneurship Forum*. Retrieved from https://entreprenorskapsforum.se/wp-content/uploads/2014/11/WP_29.pdf
- Tiebout, C. M. (1962). The Community Economic Base Study. *Supplementary Paper No. 16*. Committee for Economic Development.
- Zabala-Iturriagoitia, J. M., Voigt, P., Gutiérrez-Gracia, A., & Jiménez-Sáez, F. (2007). Regional Innovation Systems: How to Assess Performance. *Regional Studies*, 41(5), 661-672. <https://doi.org/10.1080/00343400601120270>.

Appendix

Table A1. Descriptive statistics of knowledge production by FU - Master's and Doctoral levels – from 2013 to 2018

FU	Master's Dissertations							Doctoral Theses						
	Mean	Median	SD	CV	Min	Max	NP	Mean	Median	SD	CV	Min	Max	NP
AC	92	76	50	0,542	47	171	549	3	-	5	2,082	0	13	6
AL	400	379	67	0,168	338	515	2399	58	63	23	0,408	23	84	6
AM	631	622	71	0,112	521	711	3785	113	121	28	0,250	73	149	6
AP	79	62	35	0,442	57	147	475	9	10	4	0,478	3	14	6
BA	2.148	2.160	164	0,076	1898	2379	12890	577	566	99	0,172	465	717	6
CE	1.761	1.709	182	0,103	1562	2055	10563	481	509	123	0,256	278	607	6
DF	1.780	1.769	206	0,116	1557	2072	10681	603	598	71	0,118	530	716	6

FU	Master's Dissertations							Doctoral Theses						
	Mean	Median	SD	CV	Min	Max	NP	Mean	Median	SD	CV	Min	Max	NP
ES	1.106	1.119	141	0,127	902	1293	6637	138	132	35	0,250	86	184	6
GO	1.352	1.422	189	0,140	1022	1519	8109	290	297	52	0,179	217	346	6
MA	431	430	73	0,170	336	556	2586	84	81	55	0,655	19	145	6
MG	6.190	6.176	624	0,101	5469	6947	37140	1.898	1.903	246	0,130	1590	2232	6
MS	800	809	38	0,048	740	836	4801	138	134	59	0,427	69	231	6
MT	651	667	64	0,098	565	740	3904	75	68	32	0,433	33	129	6
PA	1.299	1.301	203	0,156	1024	1547	7792	267	282	57	0,213	170	328	6
PB	1.441	1.451	125	0,087	1231	1585	8646	413	414	34	0,083	372	455	6
PE	2.166	2.152	156	0,072	1996	2352	12995	755	751	114	0,151	585	900	6
PI	469	470	34	0,072	430	514	2816	40	37	12	0,297	27	59	6
PR	4.239	4.257	550	0,130	3623	4948	25435	1.137	1.150	227	0,199	816	1422	6
RJ	7.654	7.716	409	0,053	7148	8167	45922	2.568	2.589	305	0,119	2178	2989	6
RN	1.567	1.667	412	0,263	1036	2007	9402	464	487	87	0,186	329	560	6
RO	160	168	36	0,226	114	196	958	8	7	4	0,564	3	14	6
RR	117	113	32	0,275	68	158	703	1	-	2	1,922	0	4	6
RS	5.580	5.658	419	0,075	4974	6109	33481	1.963	2.047	347	0,177	1465	2350	6
SC	2.403	2.394	210	0,087	2166	2660	14415	732	739	173	0,236	496	924	6
SE	587	579	49	0,083	535	655	3523	91	87	36	0,391	56	151	6
SP	13.460	13.675	730	0,054	12499	14192	80761	6.860	7.124	676	0,099	5850	7433	6
TO	208	219	54	0,260	132	272	1248	18	18	10	0,551	5	29	6

NP= number of production – from 2013 to 2018.

Table A2. Descriptive statistics of knowledge production in Brazil - master's and doctoral levels - from 2013 to 2018

Estimators	Master's Level						Doctoral Level					
	2013	2014	2015	2016	2017	2018	2013	2014	2015	2016	2017	2018
Mean	647	665	712	757	770	802	195	215	237	261	271	286
Median	421	446	481	497	509	526	143	168	182	185	213	220
C.V	1,06	1,07	1,09	1,14	1,15	1,17	1,20	1,16	1,13	1,12	1,11	1,11
Min	0	0	3	19	11	10	0	0	0	0	0	0
Max	3167	3481	3893	4454	4700	4887	1495	1573	1648	1749	1785	1988
Production	52408	53870	57646	61357	62402	64933	15801	17387	19227	21163	21958	23156
No. Field	81	81	81	81	81	81	81	81	81	81	81	81

Source: Authors.

Table A3. Statistical significance estimators (p-value) of the 16 FUs classified in the three statuses of efficiency and performance in knowledge production by percentile - Brazil - 2013 to 2018

UF	IC	p-value of percentiles in p1				p-value of percentiles in p2				p-value of percentiles in p3			
		P(0.25)	P(0.5)	P(0.75)	P(0.99)	P(0.25)	P(0.5)	P(0.75)	P(0.99)	P(0.25)	P(0.5)	P(0.75)	P(0.99)
SP	F	***	***	***	***	***	***	***	***	***	***	***	***
SP	MF	***	***	***	***	***	***	***	***	***	***	***	***
RS	F	***	***	***	***	***	***	***	***	***	***	***	***
RS	MF	***	***	***	***	***	***	***	***	***	***	***	***
AL	F	0,214	**	**	***	0,452	*	**	***	*	**	***	***
AL	MF	0,452	*	**	***	0,484	*	**	***	0,279	**	**	***
MG	F	***	***	***	***	***	***	***	***	***	***	***	***
MG	MF	0,122	0,122	***	***	0,257	0,107	***	***	0,264	0,114	***	***
MA	F	0,196	**	**	***	0,311	**	**	***	0,305	**	**	***
MA	MF	0,764	*	**	***	0,712	*	**	***	0,639	*	**	***
BA	F	***	***	***	***	***	***	***	***	***	***	***	***
BA	MF	***	0,42	***	***	**	0,279	***	***	0,545	0,104	*	***
PB	F	***	***	***	***	***	***	***	***	***	***	***	***
PB	MF	***	0,146	***	***	***	0,136	***	***	***	*	***	***
CE	F	**	***	***	***	**	***	***	***	***	**	**	***
CE	MF	***	***	***	***	***	***	***	***	***	***	***	***

UF	IC	p-value of percentiles in p1				p-value of percentiles in p2				p-value of percentiles in p3			
		P(0.25)	P(0.5)	P(0.75)	P(0.99)	P(0.25)	P(0.5)	P(0.75)	P(0.99)	P(0.25)	P(0.5)	P(0.75)	P(0.99)
PA	F	0,123	*	**	***	0,365	*	**	***	0,451	0,215	***	***
PA	MF	0,623	*	*	***	0,721	*	*	***	**	0,435	***	***
AM	F	**	*	***	***	***	0,133	***	***	0,364	0,103	***	***
AM	MF	0,843	0,596	***	***	***	0,716	***	***	0,635	0,19	**	***
PI	F	***	0,459	***	***	**	0,497	***	***	0,016	0,567	***	***
PI	MF	**	0,655	***	***	0,528	0,259	**	***	0,554	0,286	**	***
AC	F	***	*	***	***	**	***	***	***	***	**	***	***
AC	MF	0,975	0,952	***	***	0,539	0,529	***	***	***	0,899	***	***
SE	F	*	0,115	***	***	0,811	0,811	***	***	***	0,678	***	***
SE	MF	***	0,151	***	***	***	*	***	***	***	**	***	***
PE	F	***	0,298	***	***	**	0,231	***	***	0,305	*	***	***
PE	MF	***	***	***	***	**	**	**	***	0,129	**	**	***
RN	F	***	*	***	***	***	**	***	***	***	*	***	***
RN	MF	0,865	0,799	***	***	0,854	0,66	***	***	0,89	0,656	***	***
PR	F	***	***	***	***	***	***	***	***	***	***	***	***
PR	MF	***	***	***	***	**	***	***	***	***	**	***	***

UF: Federation Unit; CI: Concentration Index; p1:(2013 a 2016); p2:(2014 a 2017); p3:(2015 a 2018); *:p-value 0.10; **:p-value 0.05; ***:p-value 0.01; P: Percentile.

Table A4. Average household income (AHI) by FU – Brazil – From 2013 to 2018 - in BRL 1

FU	FU name	Time Units (t)					
		2013	2014	2015	2016	2017	2018
AC	Acre	665	670	752	761	769	909
AL	Alagoas	557	604	598	662	658	714
AM	Amazonas	735	739	753	739	850	791
AP	Amapá	797	753	840	881	936	857
BA	Bahia	734	697	736	773	862	841
CE	Ceará	612	616	681	751	824	855
DF	Distrito Federal	2.034	2.055	2254	2351	2548	2460
ES	Espírito Santo	1.018	1.052	1074	1157	1205	1295
GO	Goiás	1.083	1.031	1078	1140	1277	1323
MA	Maranhão	571	461	509	575	597	605
MG	Minas Gerais	1.047	1.049	1128	1168	1224	1322
MS	Mato Grosso do Sul	1.196	1.053	1044	1283	1291	1439
MT	Mato Grosso	1.109	1.032	1043	1139	1247	1386
PA	Pará	627	631	671	708	715	863
PB	Paraíba	682	682	774	790	928	898
PE	Pernambuco	667	802	825	872	852	871
PI	Piauí	649	659	728	747	750	817
PR	Paraná	1.246	1.210	1241	1398	1472	1607
RJ	Rio de Janeiro	1.303	1.139	1284	1429	1445	1689
RN	Rio Grande do Norte	794	695	819	919	845	956
RO	Rondônia	834	762	823	901	957	1113
RR	Roraima	837	871	1008	1068	1006	1204
RS	Rio Grande do Sul	1.293	1.318	1434	1554	1635	1705
SC	Santa Catarina	1.357	1.245	1368	1458	1597	1660
SE	Sergipe	787	758	782	878	834	906
SP	São Paulo	1.357	1.432	1482	1723	1712	1898
TO	Tocantins	793	765	816	863	937	1045

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).