MASTER PROGRAMME
APPLIED ECONOMETRICS AND FORECASTING

MASTER´S FINAL WORK
DISSERTATION

INEQUALITY OF OPPORTUNITIES: A STUDY BASED ON THE SPATIAL ECONOMETRICS METHODOLOGY

FLÁVIO RAMOS TERTO JÚNIOR

OCTOBER, 2018

LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER PROGRAMME
APPLIED ECONOMETRICS AND FORECASTING

MASTER´S FINAL WORK
DISSERTATION

INEQUALITY OF OPPORTUNITIES: A STUDY BASED ON THE SPATIAL ECONOMETRICS METHODOLOGY

FLÁVIO RAMOS TERTO JÚNIOR

SUPERVISION:
PROF. Dr.ª. ISABEL MARIA DIAS PROENÇA

OCTOBER, 2018

Flávio Ramos Terto Júnior[1]

Advisor: Prof. Drª. Isabel Maria Dias Proença[2]

Abstract

The child contact to basic goods and services can play an important role in adult life, and a just society allows all its individuals to have access without distinction due to the person's circumstantial characteristics. Through The Theory of Justice of Rawls (1971) and the work of Roemer (1996), the success of an individual can be understood through two components: the effort of total responsibility of the individual and the innate characteristics that are not changeable.

The literature argues that inequalities caused by innate characteristics are socially unfair, therefore characteristics such as sex, ethnicity, per capita income of the family, among others, should not determine the inequality of opportunities. This work intends to estimate the probability that the individual will have access to a good or service taking into account the spatial dynamics. The results point to an almost universalization in access to electric lighting between Brazilian states, however, there is still discrimination in access to channelled water, basic sanitation, the probability of completing the studies in the correct age and accessibility to Information and technology. The study also finds spatial patterns between Brazilian states, as well as a north-south polarization in access to goods and services.

Keywords: Theory of justice, equal opportunity, inequality, spatial econometrics

[1] Contact: flavioterto@gmail.com

[2] ISEG Lisbon School of Economics & Management - Department of Mathematics and CEMAPRE. Contact: isabelp@iseg.ulisboa.pt

Flávio Ramos Terto Júnior[1]

Orientação: Prof. Drª. Isabel Maria Dias Proença[2]

Resumo

O acesso em criança a bens e serviços básicos pode desempenhar um importante papel na vida adulta, sendo que uma sociedade justa permite que todos os seus indivíduos tenham o acesso sem distinção devido às características circunstanciais da pessoa. Através da Teoria da Justiça de Rawls (1971) e do trabalho de Roemer (1996) pode entender-se o sucesso de um indivíduo através de duas componentes: esforço - de total responsabilidade do indivíduo - e características inatas - não alterável. A literatura defende que as desigualdades provocadas por características inatas são socialmente injustas, portanto características como género, etnia, renda per capita da família, entre outras não devem determinar a desigualdade de oportunidades. O presente trabalho tem o objetivo de estimar a probabilidade de um indivíduo ter acesso a um bem ou serviço tendo em conta a dinâmica espacial. Os resultados apontam numa quase universalização no acesso a iluminação elétrica entre os estados brasileiros, entretanto ainda há discriminação no acesso a água canalizada, acesso a saneamento básico, probabilidade de completar os estudos na idade correta, acesso a informação e tecnologia. O estudo também encontra padrões espaciais entre os estados brasileiros, bem como uma polarização Norte-Sul no acesso as bens e serviços.

Palavras-chaves: Teoria da Justiça, Igualdade de oportunidades, Desigualdade, Econometria Espacial

[1] Contact: flavioterto@gmail.com

[2] ISEG Lisbon School of Economics & Management- Department of Mathematics and CEMAPRE. Contact: isabelp@iseg.ulisboa.pt

Dedicated to my beloved parents

Acknowledgments

I want to thank God for making me healthy and capable; for being my safe port in my despair hours during my master degree. To my parents, Flávio Terto and Etiene Terto, who definitely are the responsible for all my conquers; for investing with very sacrifice in my studies all my life, for the great support and example given. To my younger brother, for considered me an example, and motivated me to search and do the right thing. To the rest of my family, especially to my grandparents with their advices, and my aunts for the incentives of all kinds. To Vanessa Sousa for the support at the end of this journey and for hearing me talking about this theme. To Patricia Barros for the support when I most needed. To Sr. Mendes do Vale for the support.

Thank you.

I want to thank the teacher Drª Isabel Proença, for accepting the supervision of my dissertation and being an excellent counselor, thank you for the patient and great ability to teach.

Thank you to all the docents of the master Applied Econometrics and Forecast of the Lisbon School of Economics and Management for raising my knowledge through their great teaching and knowledge in the area.

Thank you to the professors who incentivized my application to this master programme, Drª Ana Paula Amazonas and Drº Luís Campos.

I want to thank the friends I had made during my course, who contributed to this journey and for the good moments lived in this past years.

Finally, thank you to the whole institution and collaborators who make the University of Lisbon and the Lisbon School of Economics and Management.

Contents

Acronyms

2SLS two-stage least squares

GMM General Method of Moments

HDI Human Development Index

IBGE Brazilian Institute of Geography and Statistics

HOI Human Opportunity Index

LISA Local indicators of spatial association

OLS Ordinary Least Squares

PNAD National of Domains Samples Survey

SAC Spatial Lag Combined

SAR Spatial Autoregressive Model

SEM Spatial Error Model

Chapter 1

Introduction

The poorest half of the global population own less than 1% of total wealth. In sharp contrast, the richest decile holds 88% of the world's wealth, and the top percentile alone account for 50% of global assets, says the Credit Suisse Global Wealth Databook (2017). This number reveals a strong inequality of results in the world. There are two ways of interpreting and addressing this contrast. One, is based on the theories of stratification, where the most influential of which is given by Marx, who argued that all stratification systems are determined by the distribution of economic resources.

In contrast to stratification and using the conception of just society also linked to the concept of social equality, Rawls (1971) spoke about a kind of inequality: inequality of opportunity. Roemer (1996, 1998) worked on this concept and admitted that the difference between personal advantages is guided by two distinct but related components: one, individual innate characteristics, which cannot be altered, such as ethnicity, sex, and the other the effort exerted by each one.

Innate characteristics, although have an impact on inequality, they are not the responsibility of the individual, while the effort depends on herself/himself. Being the inequality caused by the former, which skips the individual decision, is considered socially unfair.

In the constitution of democratic countries, a number of basic rights are guaranteed to their citizens, regardless of their personal standards, including health, basic education, housing, security and protection in motherhood and childhood. As an example, the Brazilian Constitution:

> **art 6.** The health, food, work, housing, transport, leisure, safety, security, protection of motherhood and childhood, and assistance to the destitute, in accordance with this Constitution.
>
> In Brazilian Constitution (1988), art. 6º.[1]

These rights are called by Rawls (1971) and by Roemer (1996) as primary goods, goods that aid in the personal development and growth of a socially just society. Access to these goods cannot be differentiated, that is, a person cannot be privileged because of innate personal characteristics, as is defended in the Portuguese Constitution:

> **art. 13º** can be privileged, benefited, prejudiced, deprived of any right or meaning of reason of ancestry, sex, race, language, source of religion, religion, belief or ideology, education, economic situation, social condition or sexual.
>
> In Portuguese Constitution (1974), art. 13º.[2]

In this context, based on data from Latin America and the Caribbean, Barros et al. (2009) develop an indicator to measure how socially fair a society is, the Human Opportunity Index (HOI) for access to essential goods and services, such as sanitation, water, electricity, and school.

In an earlier study, Barros et al. (2000), pointed to the instability of the Brazilian economy in the 1980s and 1990s, which resulted in high levels of income inequality, reflecting a greater number of people living in conditions of

---

[1] translated by Flávio Terto
[2] translated by Flávio Terto

extreme poverty. However, since the government of Fernando Henrique Cardoso, income transfer policies have been developed, which makes it interesting to analyze the impact on the poverty in the following years.

From Barros et al. (2009) several articles were published using HOI. The present study, intends to estimate the probability that the individual will have access to a good or service through its spatial dynamics, using the microdata of the National of Domains Samples Survey IBGE (2015) in Brazil and with the aid of spatial econometric techniques. That is, this work aims to investigate if nearby observations influence the outcome of the neighbor. Spatial analysis contributes to the formulation of public policies, through a better design when a local variable influences the same (or another characteristic) in a geographically close locality.

The central problem of this thesis lies in the identification of which factors influence the inequality of opportunities. In this paper equality is considered as follows: two individuals located in the same relative position in two distinct distributions must have the same probability of access to the primary goods. Consequently, any factor that is not controlled by the individual and has an impact on equality will be considered socially unfair.

In view of the above, the general objective of this study is to identify circumstantial factors that may lead to inequality of opportunities and how these interact with neighboring agents.

The specific objectives are:

• To quantify the characteristics of individuals and their relation to access to public goods;

• To identify the relevance of the spatial association between the locations where individuals live in determining inequality.

The thesis is structured in five chapters. The first one is the Introduction. The second chapter contains a brief review on the Opportunity Index and presents the definitions of inequality. The third chapter reviews the methodology used, as well as its advantages. The fourth chapter describes the steps for building this work, the data, as how they were analyzed to accomplish the intentions of this work and analyzes the results generated by the work. The conclusions of the study, together with its limitations and suggestions for future work are presented in the fifth chapter.

Chapter 2

Review of Literature

2.1 Equality Viewed from Social Justice

The notion of a just society is closely related to the concept of social equality and justice according to Rawls (1971). For the author, justice is the first virtue of social institutions, as truth is for systems of thought.

An important distinction in the study of social equality is the difference between distributive inequality and equal opportunity or meritocracy. According to Habibis and Walter (2015), equal opportunities focus on the extent to which people have a similar chance for rewards, with less attention to the effects this has on the distribution of wealth, income, and influence.

The theoretical approach used in this work will be based on The Theory of Justice by Rawls (1971), The Equality Theory systematized by Dworkin (1981) and Theories of Distributive Justice by Roemer (1996). Dworking (1981) stated that justice requires equality of something, not only results, but also resources, which we can consider access to resources.

In economics, Roemer (1996) develops a concept in which effort and innate characteristics are included as determinants of inequality, that is, there are two components that can determine inequality: (i) Inborn characteristics: Inequality is not controlled by the individual, an example is sex and ethnicity, and ii) the effort: which depends exclusively on the individual.

According to Peragine (2004), inequalities arising from factors over which individuals have no control are unfair and must be compensated by society. Therefore, if the variables derived from the circumstance component exert influence on the inequality between people, we have a socially unjust society.

2.2 Human Opportunities Index

As seen in the work by Ramos and Van de Gaer (2016), the elements that define the inequality of opportunities are often called in the literature variables of responsibility (effort) and variables of non-responsibility (components innate). Despite the different denomination, the idea, presented by Dill and Gonçalves (2013) is very similar. For these authors, factors for which the individual is not responsible, such as gender or parental education, should not exert influence on the inequality of results while factors resulting for the individual choice, such as the effort exerted by him or her, can affect the inequality of results and often do it.

Institutions are fair when no arbitrary distinctions are made between people in the assignment of basic rights and duties (Rawls, 1971). In the search for the measurement of how these distinctions occur and how this contributes to the formation of inequalities in access to opportunities, Barros et al. (2009) developed the Human Opportunity Index (HOI). The index is a measure of how much the access to certain basic goods and services available in a society is allocated based on the principle of equal opportunities, Vega et al. (2010).

The HOI allows to evaluate how non-accountable variables Influence the access to a given good or service. The work of Barros et al. (2009) uses a sample of individuals with 16 years or less, as Dill and Gonçalves (2013) argue that the

HOI calculated considering only individuals of this age group reflect the inequality due exclusively to the innate condition variable. In fact, individuals of this age are not able to choose their effort and therefore the differences observed in their access to basic goods and services derive from their personal characteristics and environment. In addition, for Vega et al. (2010) interventions to reduce inequality for agents of this age are less expensive and more effective than in adulthood.

As explained by Dill and Gonçalves (2013), the index is defined by combining two elements: The first is the coverage rate of a given good or service, which reveals the proportion of the population that has access to a given opportunity. The second is the inequality in opportunity, given by the index of dissimilarity. This index was elaborated by Duncan and Duncan (1955) and can be derived from the Lorenz curve measuring the segregation between different groups. The HOI assumes values from 0 to 100. The closer it is to 100, the greater the equality in access between people with different innate characteristics is.

To build the Opportunity Index, Barros et al. (2009) estimate the probability of an individual having access to a good or service according to their innate characteristics through a logistic regression. The latter explains the access or not to a good or service as a function of circumstantial variables.

Consider the probability of individual i to have access to opportunity or service A given a set of circumstances $(x_{1i} \dots x_{mi})$, given by,

$$(1) \qquad P_{iA} = P(A_i = 1 | x_{1i} \dots x_{mi}) = \frac{e^{\beta_0 + \sum_{k=1}^{m} x_{ki}\beta_k}}{1 + e^{\beta_0 + \sum_{k=1}^{m} x_{ki}\beta_k}} , \qquad i = 1, \dots, N \ and \quad k = 1, \dots, m$$

As stated in Barros et al. (2008), HOI estimation could have been obtained through a variety of parametric, non-parametric or semi-parametric procedures.

The present study chose to follow the same methodology found in most of the studies using HOI.

After estimating the probability of all individuals, we estimate the coverage rate for opportunity or service A in location S, $C_{A,S}$, according to,

(2) $$C_{A,S} = \frac{1}{N}\sum_{i=1}^{N} P_{iA,S} \ , \quad S = 1,\ldots,27$$

with $P_{iA,S}$ the probability of individual i to have access to opportunity or service A in location S. S identify the Brazilian States plus the federal district.

According to Dill and Gonçalves (2013), although the coverage rate is important, does not reveal the degree to which the opportunities are distributed across distinct individual. Therefore, the dissimilarity index that measures the inequality in the coverage rate for groups with comparable circumstances is calculated, providing a measure of inequality, and is given by,

(3) $$D_{A,S} = \frac{1}{2C_{A,S}}\sum_{i=1}^{N} \frac{1}{N} | P_{iA,S} - C_{A,S}|$$

according to Barros et al. (2009), its value lies from 0 to 1 and the closer to 1 it is the more unfair the society in study is.

The Human Opportunity Index is calculated from the combination of the coverage rate ($C_{A,S}$) and the dissimilarity index ($D_{A,S}$), according to,

(4) $$IOH_S = C_{A,S} * (1 - D_{A,S})$$

where (1-$D_{A,S}$) is the fraction to achieve equality (Vega et al., 2010).

Therefore, we can conclude that if there is an equal distribution of the opportunity among individuals, the dissimilarity index will be 0, and the HOI will be equal to the coverage rate.

However, there is a strong inequality in the supply of intra and interregional public goods and services in Brazil (for example, only about 50% of Brazilian municipalities have a sewage collection system based on PNAD data (2014)). The results showed that the states with the worst proportions in the number of municipalities with adequate collecting networks are in the North and Northeast regions and the best proportions are in the South and Southeast regions, indicating spatial association. In this perspective, it is necessary to incorporate the spatial component to explain the inequality among individuals from different regions.

Since the work of Fujita et al. (1999) there was a revolution in geography economic studies, although with a greater focus on agglomerations of firms. The referred work provides a basis for explaining why some regions grow and others not, that is, for explaining the inequality in growth between regions.

According to Golgher (2012), phenomena with spatial interactions, diffusion processes and spatial hierarchies imply that the location and distance between observations should be incorporated into studies that address these issues. There are many studies about inequality of opportunity, but the number of works that incorporate the spatial dimension are incipient, therefore this work using spatial econometrics to explain HOI gives a rich contribution to the literature on the subject.

Chapter 3

Spatial Econometrics: Methodology

3.1 Spatial dependence

The first law of geography dictates that everything is related to everything else, but things close are more related than distant things (Tobler, 1970). This is the founding principle on which corrective measures for spatial autocorrelation are based. That is, we must question the spatial independence of the collected data set and use Spatial Econometrics to avoid incorrect statistical inference which may be encountered when using Classical Econometrics.

What characterizes Spatial Econometrics, according to Lesage (1998), is spatial dependence and spatial heterogeneity. Spatial dependence or spatial autocorrelation indicates that observation associated with a location (i) is influenced by other observations of a different location, that is,

(5) $$y_i = f(y_j)\,, \qquad i = 1,2 \ldots n, \qquad j \neq i$$

where i indexes observations collected at i = 1,…,n points

Spatial heterogeneity occurs when the causality relationship among y_i and y_j change according to the location. In the more general case, consider that we can expect a different relationship for each observation in space, leading to,

(6) $$y_i = X_i \beta_i + \varepsilon_i$$

Where X_i represents a (1 x k) vector of explanatory variables with an associated set of parameters β_i, y_i is the dependent variable at observation (or location) i and ε_i denotes a stochastic error in the linear relationship. This approach will not be used, because modelling spatial heterogeneity it is quite complex and requires a Bayesian approach.

On the other hand, modeling spatial dependence is less complex. Moreover, works that consider spatial dependence between observations may describe problems in a better way that are not effectively addressed by standard econometrics. According to Lesage (1999), the presence of spatial dependence between observations, or spatial heterogeneity in the modeled relations, invalidates the basic assumptions of Gauss-Markov, traditionally used in linear regression models.

According to Vieira (2009), the problem of spatial autocorrelation bears some resemblance to temporal autocorrelation. In fact, if regions were aligned, that is, if the neighbor of the "front" and the "neighbor" behind existed, the econometric treatment would be identical to that of the time series. In fact, the spatial dependence is not so simple to model. Moreover, according to Vieira (2009), the spatial autocorrelation can be modelled in three ways, the autoregressive spatial lag model, the spatial error model and a combination of the two.

Spatial autocorrelation presents consequences similar to temporal autocorrelation. If the errors are autocorrelated with each other, the OLS is inefficient and the variance estimators will be biased. Anselin and Bera (1998) formally described the spatial autocorrelation as follows,

$$Cov(y_i, y_j) = E(y_i, y_j) - E(y_i)E(y_j) \neq 0\,, \qquad i \neq j \tag{7}$$

To specify a Spatial Econometric Model additionally to the condition above, it is necessary to introduce an intuitive logical pattern based on the spatial structure that gives the spatial correlation.

3.2 Spatial Contiguity Matrix W

According to Anselin (2003), the variance-covariance matrix of the errors shows a spatial structure when there exists covariances that are different from zero, that is, $Cov(\varepsilon_i, \varepsilon_j) \neq 0, i \neq j$. There are two ways to define the spatial pattern of the variance-covariance matrix. One, specifies the covariance as a function of the distance separating two pairs of locations and the other as a function of the contiguity status. The first, requires a function that decreases with the distance and a given set of values for the spatial parameter to guarantee that the resulting covariance matrix is define positive. The second, specifies a stochastic process that connects the value of a random variable in a locality to the values of that variable in neighboring localities. Thus, the neighbors of each locality are determined by the Contiguity Matrix W, instead of a function that decreases with distance.

Concerning the approach based on the distance, observe that the notion of proximity is relative, since proximity does not necessarily have to be related to the Euclidean distance between two points. Distinct criteria can be considered, such as economic, social and political distances (see, for example, the discussion in Vieira, 2009).

Measures of contiguity rely on a knowledge of the size and shape of the observational units depicted on a map. From this representation, we can determine which units are neighbors or represent observational units in reasonable proximity to each other (Lesage, 1998).

Spatial dependence should conform to the fundamental theorem of regional science: distance matters. Observations located nearest reflect a greater degree of spatial dependence than those that are more distant. The principle is embedded in the contiguity matrix W, whether it is defined based in the criterion known as "tower" or the one known as "queen". These procedures make a neighborhood pair according to the movements in chess of the tower (rook) and the queen, respectively. The contiguity matrix W may be defined as having the element (i, j) equal to 1 if i and j are neighbors and 0 otherwise, as in Magalhães et al. (2005).

The status of neighbor is attributed according to the contiguity criteria tower when observation share a lateral border. On the other hand, according to the criterion queen observation are considered neighbors, if they share lateral borders and vertices.

As an example assume a spatial distribution of regions according to the figure below.

Figure 1: Example of a Spatial distribution of regions

1	2	3
4	5	6

Using the criterion queen, the connectivity matrix of order 6x6 would be represented as follows,

$$W = \begin{Bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{Bmatrix}$$

The matrix W is symmetric and by convention the main diagonal is equal to zero. Lesage (1999) recommends normalizing the matrix W by the rows, that is, in such a way that the sum of the elements in each row equals 1. The normalized version of the matrix above is,

$$W = \begin{Bmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{Bmatrix}$$

In this work we will use the contiguity matrix row normalized and based on the queen criterion, which takes into account the boundaries and vertices between the localities. As stated, the matrix W gives the spatial correlation structure presented by the data. Although in the literature of Spatial Econometrics there is no consensus on a specific structure for the matrix W, other than to follow the assumption that the closest localities have a greater correlation than the more distant localities. For example, Campos (2004) indicates that the matrix can also be assembled with the inverse of the distance between cities or the inverse of the distance squared, to capture the effect of overflow in the case of clusters.

3.3 Spatial Linear Regression Models

Lesage (1999) presents the most general spatial autoregressive model (SAC), which can be described as follows,

$$\begin{aligned} y &= \rho W_1 y + X\beta + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + \mu\,, \qquad \mu \sim N(0, \sigma^2 I_n) \end{aligned} \tag{8}$$

where y is the vector of the dependent variable $(Nx1)$, X is the matrix of explanatory variables (Nxk), ε the vector of $(Nx1)$ error term, μ a vector with white noise variables, ρ and λ are scalar unknown parameters, β $Kx1$ is a vector of unknown coefficients and W_1 and W_2 are known matrices containing the spatial weights.

From this model it is possible to derive distinct models by imposing constraints on the parameters. For example, supposing $X = 0$ and $W_2 = 0$, we have the pure first-order autoregressive Spatial Model, where the variation of y is a function of the variation of the y of neighboring units, without any explanatory variable,

$$y = \rho W_1 y + \varepsilon\,, \qquad \varepsilon \sim N(0, \sigma^2 I_n) \tag{9}$$

Whit only $W_2 = 0$, the model becomes the general Spatial Autoregressive Model (SAR) (also known as the Spatial Lag Model),

$$y = \rho W_1 y + X\beta + \varepsilon\,, \qquad \varepsilon \sim N(0, \sigma^2 I_n) \tag{10}$$

The spatial lag term $\rho W_1 y$ represents the weighted average value of the neighbor observations, that is, the effect of the neighbor is modelled endogenously. The parameter ρ quantifies this effect, that is, the effect of the neighbors on y.

It should be noted that the reduced form of model (10) shows a non-zero error correlation, according to,

(11) $$y = (I - \rho W_1)X\beta + (I - \rho W_1)\varepsilon$$

Moreover, the spatial lag for a given observation i is not only correlated with the error error term i, but with the error term in all observations. Therefore, Anselin (1988) indicates that the OLS estimator of model (10) will be biased and inconsistent, proposing a Maximum Likelihood method to estimate the parameters.

The imposition of $W_1 = 0$ leads to the Spatial Error Model (SEM),

(12) $$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W_2\varepsilon + \mu\ , \qquad \mu \sim N(0, \sigma^2 I_n) \end{aligned}$$

Note that the coefficient λ gives the spatial autocorrelation coefficient of the errors. Through the reduced form of (12) we realize that spatial dependency may be seen as omitted variables of the model, according to,

(13) $$y = X\beta + (I - \lambda W_2)^{-1}\mu$$

OLS estimator for the SEM model is biased and inconsistent due to correlation between all the locations. Therefore, estimation should be performed by the Maximum Likelihood method. However, Conley (1999) introduces an estimation procedure based the General Method of Moments (GMM), which is consistent with cross-sectional dependence (where spatial dependence is a particular case).

Kelejian and Prucha (1997, 1998, 1999) developed GMM procedures based on the previous work that are designed specifically for the estimation of spatial models. Essentially, their procedure is based on a three-step algorithm. In the first step the regression SAR model in (10) is estimated by two-stage least squares (2SLS), where the instruments are the model variables themselves. This estimator is related to the computationally simple "pseudo" ML estimator.

In the second step the autoregressive parameter ρ is estimated in terms of the residuals obtained via the first step and the generalized moments procedure suggested in Kelejian and Prucha (1995). Then the regression model in (10) is estimated by 2SLS after transforming the model via a Cochrane-Orcutt type transformation to account for the spatial correlation.

The estimation procedure is valid if the following assumptions concerning the spatial model are verified, (see Kelejian and Prucha, 1998),

Assumption 1: All diagonal elements of the spatial weighting matrix W are zero.

Assumption 2: The matrix $(I - \rho W)$ is nonsingular with $|\rho| < 1$.

Assumption 3: The row and column sums of the matrices W_1 and $(\mathrm{I} - \rho W_1)$ are bounded uniformly in absolute value.

Assumption 4: The matrix X_n has full column rank (for N large enough). Furthermore, the elements of the matrices X_n are uniformly bounded in absolute value.

Assumption 5: The innovations $\{\varepsilon_{i,N}: 1 \leq i \leq N, N \geq 1\}$ are distributed identically. Further, the innovations $\{\varepsilon_{i,N}: 1 \leq i \leq N\}$ are for each n distributed (jointly) independently with $E(\varepsilon_{i,N}) = 0$, $E(\varepsilon_{i,N}^2) = \sigma_\varepsilon^2$ where $0 < \sigma_\varepsilon^2 < b$ with $b < \infty$. Additionally, the innovations are assumed to possess finite fourth moments.

Assumption 6: The instrument matrices have full column rank $p \geq k + 1$.

Assumption 7: The estimators are well defined asymptotically.

Assumption 8: The autoregressive parameter is uniquely identifiable.

3.4 Dependence Indicators

One of the most used tests to diagnose presence of spatial autocorrelation is the Moran test index proposed by Moran (1950). Moran Global Index is a statistical operator able to detect the possible spatial autocorrelation in a given variable and can be computed as follows (Anselin, 1995),

(14) $$I = \frac{N}{W}\left(\frac{\sum_i \sum_j w_{ij} h_i h_j}{\sum_i h_i^2}\right)$$

with N the number of observations, w_{ij} the elements of the spatial matrix; W is the sum of the elements of the matrix; h_i e h_j are values of the variable, measured in deviation around the mean; and the indices i and j refer to different locations. Using a standardized W matrix, Lesage (1998) calculates I according to,

(15) $$I = \left(\frac{e'We}{e'e}\right)$$

where e represent the residuals from the SAR model. Cliff and Ord (1981) prove that the asymptotic distribution for Moran's I on least-squares residuals is a standard normal distribution under the null hypothesis of spatial independence. The interpretation of index I is analogous to the Pearson's correlation coefficient. These values are, in most cases, restricted to the interval between -1 and 1. Negative values of I support an inverse ratio, while positive values support a positive ratio. Values close to 0, support that there is not spatial dependence.

Ferreira et al. (2012) use the Moran (I) Local Index to address the possible existence of spatial association regimes, focusing on the identification of local agglomeration. The authors aim to identify possible spatial clusters for the Opportunity Index. This indicator may be calculated as follows,

$$(16) \quad I_i(d) = \frac{(z_i - z)}{s^2} \sum_j w_{ij}(d)(z_j - z)$$

where w_{ij} is the weighting factor in the spatial matrix W for the region i and j; d is the distance measurement; z_i and z_j are values observed at position i and neighbours j; z is the sample mean; and s^2 is the sample variance.

Ferreira et al. (2012) explain that significant and high values would indicate with high probabilities that there are clusters of spatial association in polygons with high associated values, as well as in polygons with low associated values, due to the fact that the index is the product of deviations from the mean.

In this work will be presented through the Moran Diagram, a two-dimensional chart comparing normalized values of an area with the mean of its neighbors, divided into four quadrants, which can be interpreted as Ferreira et al. (2012),

- Quadrant 1 (Q1 – right and above): positive values, "positive" means "high-high" are accompanied by high values in adjacent units and quadrant 2 (Q2 – left and below): negative values, low means "low-low" mutually adjacent with low values; evidence of a positive spatial autocorrelation or spatial clustering.
- Quadrant 3 (Q3 – right and below): positive values, low mean-low-high and quadrant 4 (Q4 – left and above): negative values, low-high positive mean; evidence of negative spatial autocorrelation or spatial outliers, where high values surrounded by low values or low values surrounded by high values.

As an alternative to Moran statistics, this work used a proposed test in Anselin (1995), LISA – Local Indicators of Spatial Association. According to Xavier (2014), the LISA consists in testing spatial correlation between a variable of a locality and the average of the same variable of the neighboring localities.

Other alternative is the local measure of spatial autocorrelation introduced by Geary (1954). This is an interesting alternative, because it is not limited to linear associations and is based on a quadratic difference, Anselin (2017). According to Anselin (2017), the Geary c statistic can be expressed equivalently as a ratio of two sums of squares, that is, the quadratic difference between the observations in i and j in the numerator and the sum of the squared deviations of the mean in the denominator:

$$(17) \quad c = \frac{\sum_i \sum_j w_{ij}(x_i - x_j)^2 / 2S_0}{\sum_i (x_i - \bar{x})^2 / (n-1)}$$

The term S_0 corresponds to the sum of all the weights ($\sum_i \sum_j w_{ij}$); x_i and x_j are values observed at position i and neighbours j.

According to Anselin (2017), the factor 2 is included for making the statistical expected value under the null hypothesis of non-spatial autocorrelation close to the value of 1 (not zero). Statistics smaller than one indicate a small difference between an observation and its neighbors, suggesting positive spatial autocorrelation, while statistics larger than one suggest negative spatial autocorrelation (due to large differences between an observation and its neighbors). Anselin (2017) indicates an empirical reference distribution that represents a computational approach to obtain the distribution of the statistic under the null.

Although they are good spatial correlation identifiers, the Moran and the Geary tests are not able to distinguish which spatial dependence structure is present in the model. That is, although the null hypothesis is defined as the inexistence of any form of spatial dependence, the alternative is broadly specified (just the negation of the null). In the other hand, the LISA distinguishes the spatial dependence structure.

Chapter 4

Empirical Analysis

4.1 Overview

This empirical application is centered on the analysis of the factors that influence the spatial dynamics of access to services and public goods, using the methods introduced in chapter 3. The focus is on the occurrence of spatial dependence in the Human Opportunity Index, allowing the identification of the characteristics of the Brazilian states that may lead to inequality between states and the formation of clusters. There is a vast literature on economic growth that advocates that different public service offerings bring about different outcomes in human capital, Hall and Jones (1999) call these components of social infrastructure. With great inequality between states and regions in Brazil, there is a need to investigate this "infrastructure" through spatial econometrics.

4.2 Data

From the National of Domains Samples Survey IBGE (2015), seven public goods and services were chosen to estimate the probability of access of an individual to these goods. The choice was made in similarity to Dill and Gonçalves (2013). According to the authors, goods and basic services whose access can play an important role in adult life are:

- Access to piped water
- Access to Lighting
- Access to basic sanitation

- Access to the correct disposal of garbage
- Access to Internet
- Access to Mobile
- Chance of completing the 6th grade at the correct age

The National of Domains Samples Survey IBGE (2015) inquired 356.924 individuals in the 26 Brazilian states plus the federal district. Variables from PNAD used to estimate the seven Human Opportunity Indices – HOI can be seen in table A.1.

Access to piped water, access to lighting, access to basic sanitation and correct disposal of garbage are associated with housing characteristics and, according to Dill and Gonçalves, are related to quality of life. The probability of completing the 6th grade at the correct age is related to access to basic education, also calculated by Barros et al. (2009)

In addition, according to the World Bank (2006), clean water, health systems and basic sanitation are the most important determinants of life expectancy at birth. Terto et al. (2017) estimate how these assets determine the likelihood of stillbirths.

The opportunity of access to the Internet and access to mobile phone, have the objective of capturing the computerization and digitization of the individual. Besides being the internet a space that democratizes information and functions as a social space, Poster (1995).

The specification of the probabilities of access to each good includes the explanatory variables similar to the work of Barros et al. (2009) and Dill and Gonçalves (2013). The definition of these variables is included in Table A.1.

Binaries for gender and ethnicity of the individual were considered, to capture the effects of direct discrimination. The area of residence was included with the aim of confronting disparities between urban and rural areas. The gender of the reference person, that is, the person declared as responsible for the domicile when conducting the survey was added to capture indirect discrimination, while the presence of the mother and the number of people living at home were considered to control for aspects of family structure.

The education of the reference person of the household was considered as a proxy for family background. The logarithm for the monthly household income per capita was added to capture the effect of the available resources that the individual has access to.

To construct a dummy for race, the five categories presented in the PNAD were aggregated in only two, white and non-white following the same criteria of Bourguignon et al. (2007), which consider in the first category white and yellow individuals and in the second black, brown and indigenous.

The level of effort is relevant to determine the probability of access to the goods. However, this variable is not observed and its omission may lead to endogeneity problems in the estimation of the referred probability. To overcome this problem a cut in the sample was made. As a result, the study considered only individuals aged between 7 and 16 years, since according to Barros et al. (2009) individuals with such age are not able to choose their effort, and therefore, the differences observed in their access to basic goods and services stem from their innate characteristics. The resulting sample includes 56,439 individuals. Descriptive statistics of the variables can be seen in table A.2.

After estimating, the individual probabilities to access each of the seven goods described before by a Logit model having as explanatory variables those mentioned previously which one included in table A.1., the Opportunity Index was calculated following the methodology presented in section 2.2 for each state and each good. Then a spatial model was specified, for each of the estimated Opportunity Indices, with explanatory variables the respective spatial components and an observed variable, the Human Development Index (HDI) of the state having as observations the 26 Brazilian states plus the Federal District.

The HDI is a comparative measure to classify regions in developed or not, and simultaneous give the degree of their development. It is presented in the report of the United Nations Development Programme (1990). The HDI is calculated from three dimensions: life expectancy, education index and income index. United Nations Development Programme (2013) began using a new method of calculating the HDI, following three indices,

i) Life Expectancy Index (LEI)

(18) $$LEI = \frac{LE - 20}{85 - 20}$$

LEI is equal to 1 when Life expectancy at birth is 85 and 0 when Life expectancy at birth is 20.

ii) Education Index (EI)

(19) $$EI = \frac{MYSI+EYSI}{2}, \quad MYSI = \frac{MYS}{15}, \quad EYSI = \frac{EYS}{18}$$

MYSI is a mean years of schooling index, 15 is the projected maximum of this indicator for 2025. EYSI is an expected years of schooling index, 18 is equivalent to achieving a master's degree in most countries.

iii) Income Index (II)

(20) $$II = \frac{\ln(GNIpc) - \ln(100)}{\ln(75.000) - \ln(100)}$$

GNI is the gross national income. II is equal to 1 when GNI per capita is $75,000 and 0 when GNI per capita is $100.

The HDI is the geometric mean of the previous three dimensions,

(21) $$HDI = \sqrt[3]{LEI\ x\ EI\ x\ II}$$

Therefore, it is a variable that captures the effect of several components of development. The values used here are from the Atlas of Human Development in Brazil (2013) database and are in table A.III.

4.3 Results

To perform the calculation of each Opportunity Index, a code for STATA was developed and it is shown in the appendix C. The estimated results for the coefficients of the variables used in the estimation of the opportunity indexes, as well as their mean values and the p-value of the t-test for the statistical significance of the estimated parameters can be found in appendix A.

In table A.V, we find the estimated HOI for each Brazilian state plus the federal district and each good. Figures B.1 to B.7 show the distribution of the values of each opportunity index respectively and figures B.8 to B.14 show the representation of the measurements of the local spatial indices, namely, LISA (see section 3.4).

To test the presence of the spatial dependence, the Moran and Geary indexes were calculated and shown in the table A.VI. Both tests support the hypothesis of spatial dependence, for the HOI of access to piped water, access to lighting, access to internet, mobile phone access and access to basic sanitation. However, for the probability of completing the 6th year at the correct age a hypothesis of spatial dependence is weakly supported.

After the analysis of the spatial dependence, the spatial models were estimated, as described in the previous section. The results of the estimation of the spatial models are found in table A.VIII (SAC model) and table A.IX (SAR model).

4.4 Discussion of the Results

The results obtained from the calculation of the indices, presented in the table A.V, support the thesis of polarization in the access of public goods and services between the Brazilian regions. In general, the greatest inequality in access is verified in the states located to the North and Northeast of Brazil, in contrast, the highest values are found in South and Southeast states.

In relation to the indices associated to quality of life, that is, the characteristics of house, lighting is the public good that is better distributed, having an average value of 99.44. This result, like that found by Dill and Gonçalves (2013), denotes the universalization of access to this good, which is an important issue, since access to lighting allows access to other assets, such as access to television and information through the computer and allows study at night time.

By contrast, access to correct treatment of sewage is the one with the lowest mean 51,68, in addition to presenting the highest standard deviation compared to all calculated HOI (table A.V). In fact, for example, social inequality is reflected in the lack of basic sanitation, although it is a serious problem throughout Brazil, the proportion of families living in irregular areas affected by the lack of sewage treatment is much greater, since according to Instituto Trata Brasil (2016), 90% of sewage from irregular areas is neither collected nor treated.

The index (table A.V) referring to the correct destination of the trash also presents a great variability between the states. It is noticeable the strong polarization of this index in figures B.3 and B.10, with the lowest value found in the state of Maranhão (59,13) and the largest in the state of São Paulo. A major global trend is a concern in large metropolises in dealing with the correct disposal of waste, see European Green Capital, European Commission (2018), as well as increasingly demanding legislation on recycling.

In similarity to access to lighting, access to piped water does not differ greatly from other accesses analyzed. Although these accesses are a measure of the quality of life and in disease prevention, the high values in the HOI of the channeled water of a state do not guarantee the quality of the respective water nor its frequency in an acceptable regime.

Regarding the indices (table A.V) referring to education and access to information, there is an alarming scenario. Although there is little variability in the probability of completing the sixth year at the correct age, this index presents the lowest average among all indexes. In addition of having little probability of access to this good, the country in general presents a large value for this index of inequality of opportunity, that is, the variables of circumstances exert a great influence on the probability of completing the sixth year at the correct age. According to table A.V, the results indicate that, in the best of scenarios, an individual has in average a 29.23% chance of completing the sixth grade at the correct age.

In terms of internet and mobile access, the numbers are low. This result turns out to be a contradiction compared to the one expected and concluded by Dill and Gonçalves (2013), since given the high indexes of the access to the lighting, high rates of access to the internet and mobile phone were expected as well. The results thus corroborate the polarization present in the access to goods between the North / Northeast and South / Southeast regions.

Concerning the results included in Table A.IV obtained from the estimation of the probability of access to each good, the factors of home environment, mother present at home, family income per capita, number of people in the residence, gender and years of studies of the reference person presented similar results to those in the works that used these variables as controls, as in Assis et al. (2007).

The estimated coefficient of the variable that measures the influence of the presence of the mother in the residence on the probability of the individual having access to the opportunities analyzed was significant and positive for practically all HOIs (at a level of significance of up to 10%), except for the correct destination of garbage and lighting.

On the other hand, the data show that per capita income increases the probability that the individual has access to all goods and services. It is observed that, the logarithm of income is significant at the 1% level, except for the index of opportunity for access to piped water where it is significant at 10%.

In contrast, the coefficient of the number of people in the residence is negative, pointing to an inverse relation with the number of people and the probability of the individual having access to the basic goods and services. It should be noted that this variable is significant at the 1% level for all estimated HOIs, that is, regardless of the opportunity, a higher number of people at home decreases the chance of the individual having a good quality of life.

Note that for individuals whose reference person in their household is female they are less likely to have access to piped water (significant at 5%), access to basic sanitation (significant at 5%), mobile phone access 1%) and internet access (significant at 1%). That is, less access to information and good sanitary conditions.

It should be noted, however, that the reference person's years of study are significant at 1% and positive for access to piped water, the probability of completing the 6th year at the correct age, Internet access and cell phone access. The data indicate also that the higher the reference person's years of education the greater the likelihood that individuals will have access to basic goods and services during childhood and adolescence.

The coefficient of the variable relative to the locality of the individual, if it is of urban zone or rural zone, displays the greater average effects. The data indicate that individuals living in urban areas have a greater probability of access to the goods and services analyzed. The disparity between inequality of opportunities in urban and rural areas in Brazil has been addressed by Santana (2014) and these results are in line with his work.

The characteristics of the individual, gender and ethnicity should not be decisive for access to basic goods and services, however, the results point to the influence of these circumstances on the likelihood of access to some of the services as it was registered in the literature (see for example Henriques (2001), Barros and Mendonça (1995) and Quadros (2004)).

The results point to racial discrimination in the access almost to all the goods and services analyzed, except for the ability of the individual to be likely to complete the 6th grade at the correct age. It should be noted that this variable is significant at 1% in the estimation of these HOI, with individuals considered white and yellow being more likely to have access to basic services compared to those declared in other ethnicities.

For gender analysis, men are more likely to have access to light. In turn, women are more likely to complete the 6th grade at the correct age, access to mobile phones and access to the internet. That is, women are more likely to get more information via technology and years of studies at the correct age.

In addition to the components of the individual's innate characteristics and their residences, the study analyzed the influence of the individual's location on the probability of access to goods and services.

Concerning now the spatial analysis of the HOI it should be noted that the data does not support the null hypothesis of no spatial dependence for Moran and Geray tests (see table A.VI), rejecting this hypothesis at the significance level of 1% for all the HOI except for the one related to the probability of completing the 6th year at the correct age, which shows a p-value of 0.1 for the test statistic of Moran and 0.02 for the test statistic of Geray.

In figures B.1 to B.7, the distribution of HOI values in the Brazilian states were presented. According to Marques (2010), the simple analysis of the association indices for each area considered is not sufficient to identify clearly the presence of Clusters and spatial dependence. Therefore, it is through the Moran Diagram that we can determine if the HOI concentration is spatially significant and if the observations are similar agents, as is shown in figures B.15 to B.21, giving evidence of spatial association.

In addition to the dispersion diagram, the Local Index Spatial Association (LISA) is presented. According to Xavier (2014), the LISA identifies a set of clusters in which the relation in space is significant, since it produces a value for each object, identifying similar or atypical groupings and allowing for the determination of a spatial autocorrelation index. In figures B.8 to B.14 the LISA is shown for each of HOI, respectively.

With the evidence of spatial dependence, a report of the local dependence was made using the LISA method and the Moran diagram, both described in section 4.

The red part of the images represents the states and neighbors where the concentration of HOI is high and significant for the total HOI values of the states. The blue part, occurring opposite to the red, correspond to regions of absence of concentration or low importance of HOI. The figures can identify the Clusters for the indexes of the analyzed opportunities. Except for the probability of completing the 6th year at the correct age, for all opportunities the state of São Paulo belongs to the Cluster of access highest to opportunities while the state of Maranhão, in contrast indicates the zone with smaller space opportunities.

To summarize, the results shown above give evidence of spatial correlation. However, the questions that arise are: what are the consequences of spatial correlation and which factors can contribute to its advancement. To approach these questions, the SAC regressions and the two-step SAR model, described in the previous chapter, were estimated and the results presented in table A.IX.

The SAC model, is more general among the two since it considers the spatial relationship between errors as well as considering the autoregressive spatial component. The estimation results show that in no case the coefficient capturing the spatial dependence or the omitted variables of the model is standing statistical signification. Therefore, this model was discarded for spatial analysis.

The SAR model was estimated with instruments the variables themselves and correcting for heteroscedasticity in the estimation of the covariance matrix. Results are in Table A.IX, it is verified that the estimators of the coefficients of the spatial lag are high except for the probability of obtaining the 6th year at the correct age. They indicate strong influence of the states in the determination of the indexes of their neighbors.

Again, except for the probability of completing the 6th year at the correct age, the HDI was significant in all models at 1%. The coefficient of this variable is positive, indicating that the higher the HDI of the state, the greater the likelihood that individuals will have access to basic goods and services.

For the probability of the individual completing the 6th grade at the correct age, there is an alarming problem with very low values for the respective opportunity index, indicating that the variables of circumstances influence this probability a lot and do not present any spatial pattern.

Chapter 5

Final Remarks

5.1 Conclusion

Since the first studies on inequalities in the nineteenth century, the number of studies that address social inequality among individuals has increased using different methodologies and different focuses. Part of this work agrees that the inequality of opportunity in access to basic goods and services can play a key role in the adult life of the individual. Therefore, reducing the inequality in access to these goods should be a focus of public policies that aim to promote an environment conducive to the development of people and a more just society.

This work intended to contribute to the field of study on inequality of opportunities and to help the decision-making of public policies concerning development programmes. The study found that the innate variables of the individuals, that is the components of no responsibility of the individual, influence the inequality of opportunity in the access to basic goods and services. It is therefore imperative that discrimination in that access should be tackled gradually so that there is a fair society approach and people can compete on an equal basis.

The contextualization raised on the objective studied shows the importance of this theme not only for the Brazilian states, but for most of the developing countries. The estimation results point to possible key variables to promote a lower inequality of opportunity, ensuring that the public power fulfills the charter of most of the western states, where characteristics of circumstance such as gender, ethnicity, per capita income does not determine in large access to basic goods and services.

Moreover, the spatial econometrics approach allowed us to identify patterns of inequality of opportunity among Brazilian states, and to measure the relation of the access to public goods in each locality with neighbors, finding a polarization among the Brazilian regions in the access to public services.

5.2 Further Research and Limitations

In order to deepen knowledge about equal opportunities and the development of a more just society, the paper poses as suggestions for future work: to estimate the evolution of inequality of opportunity over time, to use more sophisticated econometric models that allow to consider spatial dependence in explanatory variables as well and estimating the HOI through information from smaller federative units, such as municipalities instead of states.

In fact, the use of states as spatial unit was one of the limitations of the study, since when using PNAD data, it is known that most of the research is carried out in large capitals, which makes it difficult to more accurately capture the presence of clusters in the state boundaries, due to the difference between capitals and municipalities that are furthest from capitals.

Despite the high explanatory power of the variable of the Human Development Index in the HOI of the states, it is interesting to investigate other characteristics of the states that are also determinants of equality of opportunity.

Bibliography

Anselin, L. (2017). *A Local Indicator of Multivariate Spatial Association: Extending Geary's C*. [pdf]. Available at: http://priede.bf.lu.lv/ftp/pub/TIS/datu_analiize/GeoDa/ LA_multivariateGeary1.pdf.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2), pp. 153–166.

Anselin, L. and Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: A. Ullah and D.E.A. Giles, Eds. *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, pp. 237–89.

Anselin, L. and Florax, R. (1995). *New Directions in Spatial Econometrics*. Berlin: Springer-Verlag.

Anselin, L. (1995). Local indicators of spatial association – LISA. Geographical Analysis, 27(2), pp. 93-115.

Anselin, L. (1988). *Spatial Econometrics: methods and models*. Norwell: Kluer.

Assis, A. *et al.* (2007). Desigualdade, pobreza e condições de saúde e nutrição na infância no Nordeste brasileiro. *Cadernos de Saúde Pública*, 23(10), pp. 2337–2350.

Atlas Brasil, (2013). *Atlas do Desenvolvimento Humano no Brasil 2013*. Available at: http://atlasbrasil.org.br/2013/pt/consulta/. Access in: 14 Fev. 2018.

Barros, R., Ferreira, F., VEGA, J. and CHANDUVI, J. (2009). *Measuring inequality of opportunities in Latin America and the Caribbean*. Washington, DC: Palgrave Macmillan and the World Bank, pp. 222.

Barros, R., Molinas, J. and Saavedra, J. (2008). Measuring inequality of opportunities for children. Working paper, Unpublished, The World Bank. Available at: http://siteresources.worldbank.org/INTLACREGTOPPOVANA/Resources/IneqChildrenPaesdeBarrosMolinasSaavedra.pdf.

Barros, R., Henriques, R., and Mendonça, R. (2000). *A estabilidade inaceitável: desigualdade e pobreza no Brasil*. Desigualdade e pobreza no Brasil. Rio de Janeiro: IPEA.

Barros, R., Mendonça, R., (1995). *Os determinantes da desigualdade no Brasil*. Seminário de pesquisa econômica. Rio de Janeiro: FGV.

Bourguignon, F., Ferreira, F., and Menéndez, M. (2007). Inequality of opportunity in Brazil. *Review of Income and Wealth*, Series, 53, n. 4, pp. 585-618.

Brazilian Constitution (1988). Constituição da República Federativa do Brasil. *Senado Federal*, pp. 7.

Campos, L. (2004). *O processo de aglomeração produtiva em Pernambuco*. PhD. Universidade Federal de Pernambuco.

Cliff, A and Ord, J. (1981). *Spatial processes: models and applications*. London: Pior Limited.

Conley, T. (1999). GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, 92(1), pp. 1–45.

Credit Suisse (2017). Global Wealth Databook 2017. *Credit Suisse Research Institute*, (129).

Dill, H. and Gonçalves, F. (2013). Igualdade de oportunidade entre os estados brasileiros: uma análise microeconométrica com base nos dados da PNAD 2009. *Nova Economia*, *23*(2), pp. 307-328.

Duncan, O., and Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review, 20*(2), pp. 210-217.

Dworkin, R. (1981). What is Equality? Part 2: Equality of Resources. *Philosophy & Public Affairs*, 10(4), pp. 283–345.

European Commission (2018). European Green Capital Award, Available at: http://ec.europa.eu/environment/europeangreencapital/index_en.htm.

Ferreira, D., Oliveira, V., and Gonçalves, F. (2012), Igualdade de oportunidades: uma análise espacial para os municípios brasileiros a partir do censo 2010. In: *Anais 40º Encontro Nacional de Economia*. Porto de Galinhas: ANPEC.

Fujita, M., Krugman, P., and Venables, A. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge and London: MIT Press.

Geary, R. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 5, pp. 115–145.

Golgher, A. (2012). *Introduction to Spatial Econometrics*. [pdf]. Available at: http://rri.wvu.edu/wp-content/uploads/2012/12/TD1_pronto.pdf.

Habibis, D. and Walter, M. (2015). *Social Inequality in Australia: Discourses, Realities and Futures*. Oxford: Oxford University Press.

Hall, R. and Jones, C. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others?. *Quarterly Journal of Economics*, 114(1), pp. 83–116.

Henriques, R. (2001). *Desigualdade Racial No Brasil: Evolução Das Condições de Vida Na Década de 90*. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada.

IBGE, (2015). Pesquisa Nacional por Amostra de Domicílios 2015. Available at: https://ww2.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2015/default.shtm. Access in: 15 out. 2017.

IBGE, (2014). Pesquisa Nacional por Amostra de Domicílios 2014. Available at: https://ww2.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2014/default.shtm. Access in: 15 jun. 2017.

Instituto Trata Brasil, (2016). Pesquisa Saneamento Básico em Áreas Irregulares – Relatório Brasil. Available at: http://www.tratabrasil.org.br/datafiles/estudos/areas-irregulares/volume2/relatorio-completo-areas-irregulares.pdf. Access in: 15 jan. 2018.

Kelejian, H. and Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40, pp. 509–533.

Kelejian, H. and Prucha, I. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17, pp. 99–121.

Kelejian, H. and Prucha, I. (1997). Estimation of spatial regression models with autoregressive errors by two-stage least squares procedures: a serious problem, *International Regional Science Review*, 20, pp. 103–111.

Kelejian, H. and Prucha, I. (1995). A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model, *International Economic Review,* Department of Economics, University of Maryland, Working Paper 95-03 (forthcoming in International Economic Review).

LeSage, J. (1999). *The Theory and Practice of Spatial Econometrics*. Ohio: University of Toledo.

LeSage, J. (1998). *Spatial econometrics*. Ohio: University of Toledo.

Magalhães, A., Hewings, G. and Azzoni, C. (2005). Spatial Dependence and Regional Convergence in Brazil. *Investigaciones Regionales*, 6, pp. 5–20.

Marques, A. (2010). Análise Exploratória de dados de área para índices de furto na mesorregião de presidente Prudente-SP. In: *Anais III Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*. Recife: SIMGEO.

Moran, P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2), pp. 17-23.

Peragine, V. (2004). Ranking of income distributions according to equality of opportunity. *Journal of Income Inequality*, 2, pp. 11-30.

Portuguese Constitution (2005). Constituição da República Portuguesa. *Assembleia da República Portuguesa*, pp. 91.

Poster, M. (1995). CyberDemocracy: Internet and the Public Sphere. *University of California*, [online]. Available at: http://faculty.humanities.uci.edu/poster/writings/democ.html. Access in: 20 maio 2018.

Quadros, W. (2004). Social Brasileira Recente. *Estudos Avançados*, 18(50), pp. 95–117.

Ramos, X. and Van de Gaer, D. (2016). Approaches to Inequality of Opportunity: Principles, Measures and Evidence. *Journal of Economic Surveys*, 30(5), pp. 855–883.

Rawls, J. (1971). *A Theory of Justice.* Cambridge: Harvard University Press.

Roemer, J. (1998). *Equality of opportunity*. Cambridge: Harvard University Press.

Roemer, J. (1996). *Theories of Distributive Justice*. Cambridge: Harvard University Press.

Santana, L. (2014). *Desigualde de Oportunidades no meio rural do estado de Pernambuco*. Msc. Universidade Federal Rural de Pernambuco.

Terto, F., Silva, R., Albuquerque, J. and Freitas, M. (2017). O Impacto de Fatores Evitáveis na Natimortalidade Brasileira: Uma Análise A Partir de Modelos de Variáveis Limitadas. In: *Anais 22ª Congresso Brasileiro de Economia*, Belo Horizonte: CO.

Tobler, W. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46, pp. 234-240.

United Nations Development Programme (2013). *Human Development Report 2013*. Available at: http://hdr.undp.org/sites/default/files/hdr_2013_en_technotes.pdf

United Nations Development Programme (1990). *Human Development Report 1990*. Available at: http://hdr.undp.org/sites/default/files/reports/219/hdr_1990_en_

Vega, J., Barros, R., Saavedra, J. and Giugale, M. (2010). *Do our children have a chance? The 2010 Human Opportunity Report for Latin America and the Caribbean*. Washington, DC: World Bank, 176 p.

Vieira, R. (2009). *Crescimento Econômico no estado de São Paulo: uma análise espacial*. São Paulo: UNESP.

World Bank (2006). *World Development Report 2006: Equity and development.* Washington, DC.

Xavier, M. (2014). *A Localização da Confecção em Pernambuco: o mapa da integração entre os municípios do agreste para a produção de roupas*. Msc. Fundação Joaquim Nabuco.

Appendices

Appendices A

Tables

Table A.I

Variables for individual HOI specification

	Variables	Variable specification	Speciation
Dependent variables	Opportunity Index for access to piped water	ioh_aguacanaliza	Binary: 1 – Access 0 – No access
	Opportunity Index for access to illumination	ioh_iluminacao	Binary: 1 – Access 0 – No access
	Opportunity Index for Internet access	ioh_utilizouinter	Binary: 1 – Access 0 – No access
	Opportunity Index for cellphone access	ioh_telefo	Binary: 1 – Access 0 – No access
	Opportunity Index for access to basic sanitation	ioh_esgoto	Binary: 1 – Access 0 – No access
	Opportunity Index for correct garbage destination	ioh_destinolixo	Binary: 1 – Access 0 – No access
	Opportunity Index for access to finishing the 6th grade at the right age	ioh_sixth_comp	Binary: 1 – Access 0 – No access
Explanatory variables		idade_13y	Binary: 1 – person with 13 years 0 – otherwise
		idade_14y	Binary: 1 – person with 14 years 0 – otherwise
		idade_15y	Binary: 1 – person with 15 years 0 – otherwise
		idade_16y	Binary: 1 – person with 16 years 0 – otherwise
	Gender	feminino	Binary: 1 – female 0 – male
	Mother´s presence	maedom	Binary: 1 – is present 0 – isn't present
	Number of people at home	numpessoas	Linear
	Municipal area	codareacens	Binary: 1 – urban 0 – rural
	Referenced person´s home education	anosetudo_ref	Quadratic
	Ethnicity	branco	Binary: 1 – white 0 – not white
	Monthly household income per capita	lnrend	Logarithm natural
	Gender of the home reference person	fem_ref	Binary: 1 – female 0 – male

Source: Self elaboration

Table A.II

Descriptive Statistics for individual HOI specification

Variable	Mean	Std. Dev.	Min	Max
Opportunity Index for access to piped water	0.9296	0.0344	0.8406	0.9735
Opportunity Index for access to illumination	0.9951	0.0037	0.9861	0.9994
Opportunity Index for Internet access	0.6148	0.0719	0.4501	0.7154
Opportunity Index for cellphone access	0.6152	0.0591	0.4886	0.6965
Opportunity Index for access to basic sanitation	0.5359	0.0849	0.3576	0.6589
Opportunity Index for correct garbage destination	0.7870	0.0774	0.5911	0.8932
Opportunity Index for access to finishing the 6th grade at the right age	0.2727	0.0088	0.2501	0.2923
Gender	0.4856	0.4997	0	1
Mother´s presence	0.8741	0.3317	0	1
Number of people at home	4.5748	1.6496	1	18
Municipal area	0.8202	0.3839	0	1
Referenced person´s home education	4.4390	3.7764	1	17
Ethnicity	0.3762	0.4844	0	1
Monthly household income per capita	6.3912	2.9057	1.9459	27.6310
Gender of the home reference person	0.1023	0.3031	0	1

Source: Self elaboration

Table A.III

Human Development Index of Brazilian states

State	IDH
Acre	0,663
Alagoas	0,631
Amapá	0,708
Amazonas	0,674
Bahia	0,66
Ceará	0,682
Distrito Federal	0,824
Espírito Santo	0,74
Goiás	0,735
Maranhão	0,639
Mato Grosso	0,725
Mato Grosso do Sul	0,729
Minas Gerais	0,731
Pará	0,646
Paraíba	0,658
Paraná	0,749
Pernambuco	0,673
Piauí	0,646
Rio de Janeiro	0,761
Rio Grande do Norte	0,684
Rio Grande do Sul	0,746
Rondônia	0,69
Roraima	0,707
Santa Catarina	0,774
São Paulo	0,783
Sergipe	0,665
Tocantins	0,699

Source: Atlas of Human Development in Brazil (2013)

Table A.IV

Estimated coefficients and APES obtained by logistic regression

	IOH – Access to piped water		IOH - Access to illumination		IOH – Probability of complete the 6th grade at the right age		IOH – Access to Internet		IOH – Access to cellphone		IOH – Access to basic sanitation		IOH Access to the right garbage treatment	
	Coefficient	APE	Coefficient	APE	Coefficient	APE	Coefficient	APE	Coefficient	APE	Coefficient	APE	Coefficient	APE
constante	1,4213* (0,2388)	-		-	-2,3051* (0,1034)	-	-2,1494* (0,1245)	-	-0,3175* (0,1338)	-	-2,5975* (0,2442)	-	-0,9786* (0,1850)	-
idade_13y	-	-	-	-	3,2381* (0,0948)	0,58560	-	-	-	-	-	-	-	-
idade_14y	-	-	-	-	3,2138* (0,0948)	0,58340	-	-	-	-	-	-	-	-
idade_15y	-	-	-	-	1,8850* (0,1328)	0,30630	-	-	-	-	-	-	-	-
idade_16y	-	-	-	-	0,8279* (0,1428)	0,11790	-	-	-	-	-	-	-	-
maedom	0,2406*** (0,1361)	0,0132	-	-	0,1243** (0,0509)	0,01650	0,5238* (0,0395)	0,10560	0,0913** (0,0401)	0,0192	0,2745* (0,0603)	0,0556	-	-
branco	0,6946* (0,1177)	0,0337	0,8164* (0,2354)	0,003	-	-	0,6068* 0,0673)	0,11770	0,4603* (0,0531)	0,0963	0,5762* (0,1250)	0,1152	0,5637- (0,1211)	0,0546
feminino	-	-	-0,1132** (0,0571)	-0,0005	0,1082* (0,0248)	0,01460	0,0650* (0,0158)	0,01260	0,3219 (0,0251)	0,0675	-	-	-	-
lnrenda	0,0330*** (0,0199)	0,0017	0,3794* (0,1223)	0,0017	-	-	0,1041* (0,0198)	0,02010	0,0797* (0,0128)	0,0166	0,0599* (0,0107)	0,0119	0,0237* (0,0087)	0,0023
numpessoas	-0,2171* (0,0199)	-0,0114	-0,2974* (0,0664)	-0,0013	-0,0475* (0,0078)	-0,00640	-0,0792* (0,0170)	-0,01530	-0,2214* (0,0134)	-0,0463	-0,1093* (0,0203)	-0,0217	-0,1076* (0,0342)	-0,0106
codareacens	2,6557* (0,1333)	0,219	4,5592* (1,0227)	0,1774	0,1108* (0,0394)	0,01480	1,8000* (0,0900)	0,39280	1,0463* (0,0808)	0,2369	2,8539* (0,1446)	0,5318	3,4288* (0,2231)	0,6521
fem_ref	-0,3217** (0,1458)	-0,0181	-	-	-	-	-0,3909* (0,0579)	-0,07790	-0,3017* (0,0586)	-0,0644	-0,4404** (0.1809)	-0,0902	-	-
anosestudo_ref	0,0042* (0,2388)	0,0002	-	-	0,0006* (0,0002)	0,00008	0,0065* (0,0007)	0,00120	0,0029** (0,0005)	0,0006	-	-	-	-
Pseudo R2	0,2616		0,3148		0,2634		0,1424		0,0898		0,1578		0,3384	

(*) significative values at 1%. (**) significative values at 5%. (***) significative values at 10%
Source: Self elaboration

Table A.V

Human Oportunity Index by Brazilian states

Region	State	HOI – Access to piped water	HOI – Access to illumination	HOI – Probability to complete the 6th grade at the right age	HOI – Access to Internet	HOI – Access to cellphone	HOI – Access to basic sanitation	HOI – Access to the right garbage treatment
North	Acre	87,28	98,67	25,26	51,39	51,73	41,57	67,98
	Amapá	90,98	99,17	25,96	59,34	55,93	47,46	77,04
	Amazonas	90,50	99,07	26,36	57,65	54,28	45,70	74,56
	Pará	89,26	99,03	26,51	55,66	55,62	43,24	70,07
	Rondônia	92,85	99,61	27,33	60,25	60,69	47,54	74,04
	Roraima	93,03	99,53	27,52	61,90	59,92	49,75	78,69
	Tocantins	91,24	99,46	27,07	57,18	58,99	49,35	74,59
Northwest	Alagoas	88,52	99,07	26,12	50,51	54,72	44,13	68,57
	Bahia	91,09	99,36	28,18	56,03	57,95	48,48	74,57
	Ceará	91,26	99,36	27,19	56,06	57,74	48,74	74,72
	Maranhão	83,89	98,60	25,02	44,46	48,59	35,79	59,13
	Paraíba	91,10	99,42	27,27	56,08	57,55	49,67	75,00
	Pernambuco	92,92	99,46	27,33	59,80	60,49	52,67	79,30
	Piauí	87,33	99,03	26,41	48,42	52,73	41,44	65,76
	Rio Grande do Norte	90,43	99,31	27,03	55,82	58,44	47,99	72,01
	Sergipe	87,19	98,87	28,18	50,69	53,58	42,55	66,53
Center Southwest	Distrito Federal	96,30	99,84	29,23	68,28	66,52	61,60	86,79
	Goiás	95,69	99,85	28,10	64,72	64,83	59,75	85,07
	Mato Grosso	92,14	99,61	25,88	57,82	60,22	51,51	75,87
	Mato Grosso do Sul	95,23	99,82	27,07	64,06	64,13	58,74	83,78
Southeast	Espírito Santo	94,36	99,74	27,07	63,71	64,39	54,93	79,84
	Minas Gerais	93,99	99,68	27,74	63,09	62,85	55,38	80,37
	Rio de Janeiro	97,34	99,94	28,34	70,51	67,94	62,65	89,08
	São Paulo	97,38	99,94	27,44	71,75	69,47	65,93	89,33
South	Paraná	95,58	99,82	26,61	67,45	67,29	61,78	84,11
	Rio Grande do Sul	96,51	99,84	28,07	71,10	69,71	64,32	86,26
	Santa Catarina	95,95	99,85	26,87	70,34	69,81	62,62	83,44

Source: Atlas Brasil (2013). Available at: http://atlasbrasil.org.br/2013/pt/consulta/

Table A.VI

Diagnosis of Spatial Dependence. Moran's and Geary's Test

Opportunity Index	**Moran**		**Geary**	
	Index	p-value	Index	p-value
Access to piped water	0,5584	0	0,3295	0
Access to illumination	0,5245	0	0,3766	0
Probability to complete the 6th grade at the right age	0,1986	0,1	0,6426	0,02
Access to Internet	0,6189	0	0,2791	0
Access to cellphone	0,6599	0	0,2577	0
Access to basic sanitation	0,6924	0	0,236	0
Access to the right garbage treatment	0,5503	0	0,3414	0

Source: Self elaboration

Table A.VII

Variables for states HOI specification

Variables	**Variable specification**	**Speciation**
Human Oportunity Index	IDH	Linear

Table A.VIII

HOI for states calculated according to equation 8 – SAC Model

	IOH – Access to piped water		IOH - Access to illumination		IOH – Probability of complete the 6th grade at the right age		IOH – Access to Internet		IOH – Access to cellphone		IOH – Access to basic sanitation		IOH Access to the right garbage treatment	
	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p
Coefficient	0,3032 (0,1127)	0,0007	0,6741 (0,1840)	0,0002	0,1772 (0,1567)	0,2582	-0,3332 (0,0713)	0,0000	-0,1911 (0,0502)	0,0000	-0,4209 (0,0801)	0,0000	-0,2325 (0,0888)	0,0088
IDH	0,5156 (0,0979)	0,0000	0,0549 (0,0115)	0,0000	0,1002 (0,0517)	0,0528	1,0020 (0,1774)	0,0000	0,6955 (0,1355)	0,0000	0,9487 (0,1656)	0,0000	1,1447 (0,2157)	0,0000
Rho	0,2776 (0,1790)	0,1210	0,2831 (0,1916)	0,1395	0,0844 (0,6751)	0,9004	0,3783 (0,1504)	0,0119	0,5052 (0,1348)	0,0001	0,5222 (0,1233)	0,0000	0,2548 (0,1843)	0,1668
Lambda	-0,3029 (0,2985)	0,3102	-0,2726 (0,3101)	0,3792	0,0977 (0,6468)	0,8798	-0,1960 (0,2913)	0,5009	-0,3295 (0,3159)	0,2969	-0,0062 (0,2942)	0,9831	-0,1752 (0,2946)	0,5519

Source: Self elaboration

Table A.IX

HOI for states calculated according to equation 10 – SAR Model

	IOH – Access to piped water		IOH - Access to illumination		IOH – Probability of complete the 6th grade at the right age		IOH – Access to Internet		IOH – Access to cellphone		IOH – Access to basic sanitation		IOH Access to the right garbage treatment	
	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p	Coefficient	value-p
Coeficient	0,1512 (0,1322)	0,2526	0,4645 (0,2020)	0,0215	0,2356 (0,5328)	0,6583	-0,3306 (0,0939)	0,0004	-0,2032 (0,0685)	0,0030	-0,4049 (0,0954)	0,0000	-0,2590 (0,1157)	0,0251
IDH	0,4126 (0,1179)	0,0004	0,0448 (0,0145)	0,0020	0,1090 (0,1081)	0,3131	0,8928 (0,2232)	0,0000	0,6294 (0,1648)	0,0001	0,867 (0,1939)	0,0000	0,9693 (0,2536)	0,0001
Rho	0,5204 (0,2036)	0,0105	0,5010 (0,2115)	0,0178	-0,1538 (2,2309)	0,945	0,5018 (0,1730)	0,0037	0,6017 (0,1520)	0,0000	0,6028 (0,1341)	0,0000	0,4492 (0,2255)	0,0463
R2	0,7837		0,7185		0,2406		0,8454		0,8510		0,8865		0,7895	

Source: Self elaboration

Appendices B

Figures

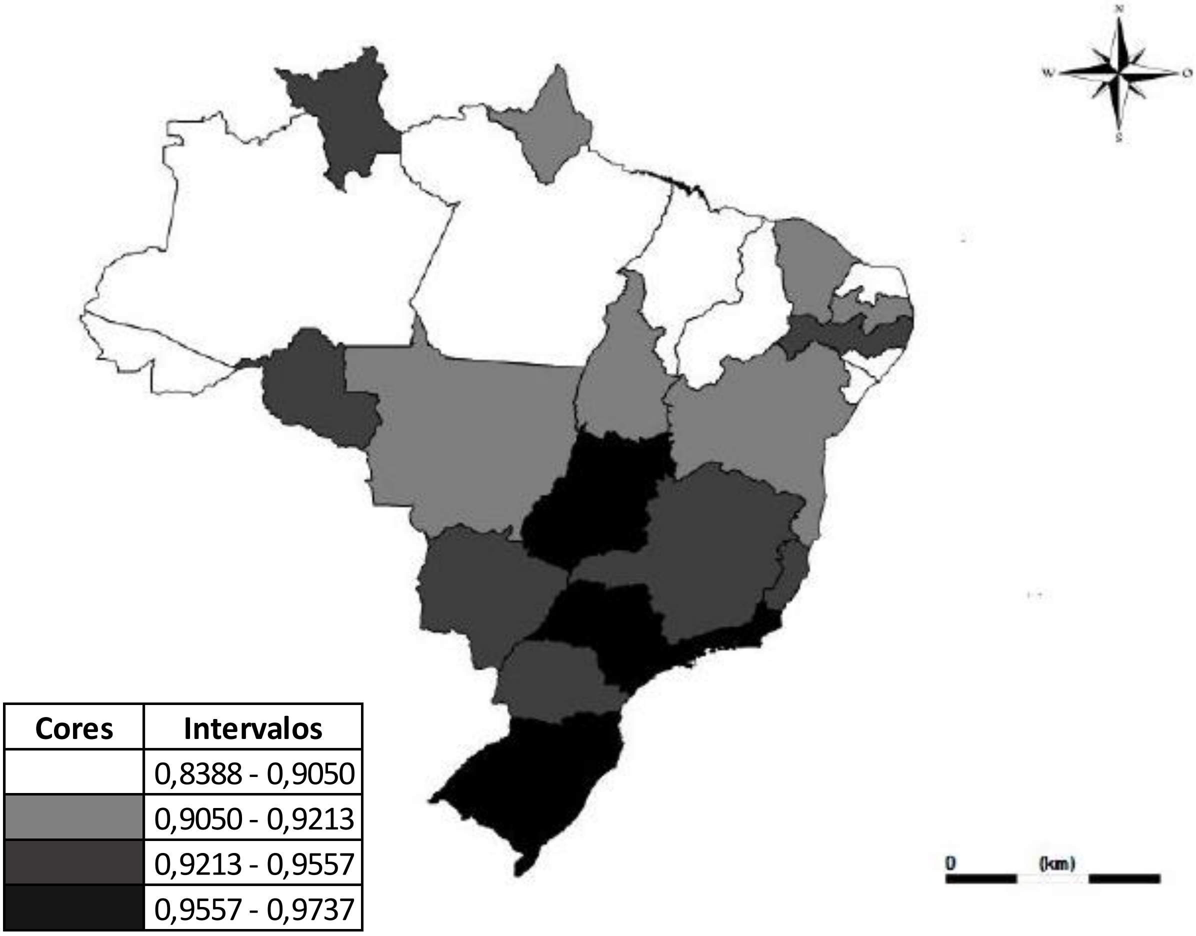
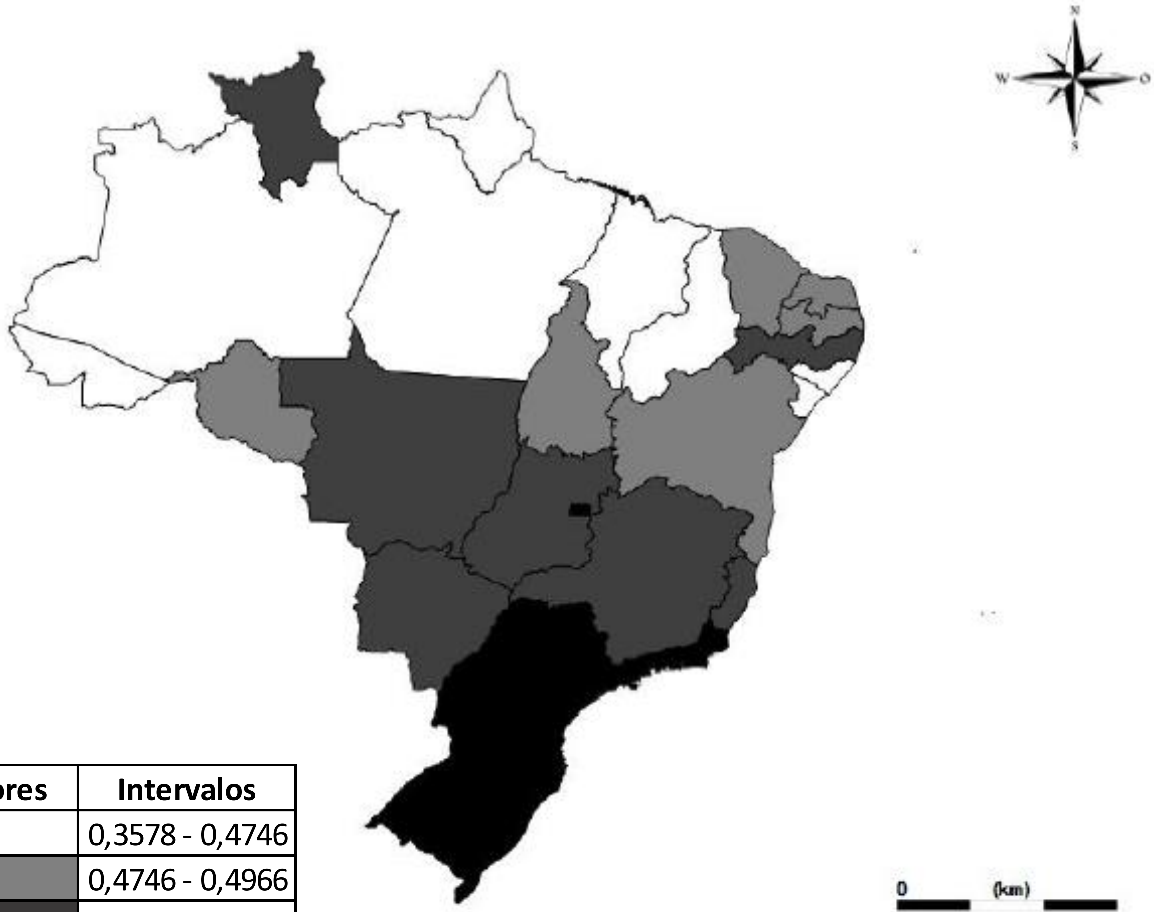


Figure B.1 – Distribution of the Opportunity Index for access to piped water



Cores	Intervalos
	0,3578 - 0,4746
	0,4746 - 0,4966
	0,4966 - 0,5974
	0,5974 - 0,6593

Figure B.2 – Distribution of the Opportunity Index relative to basic sanitation

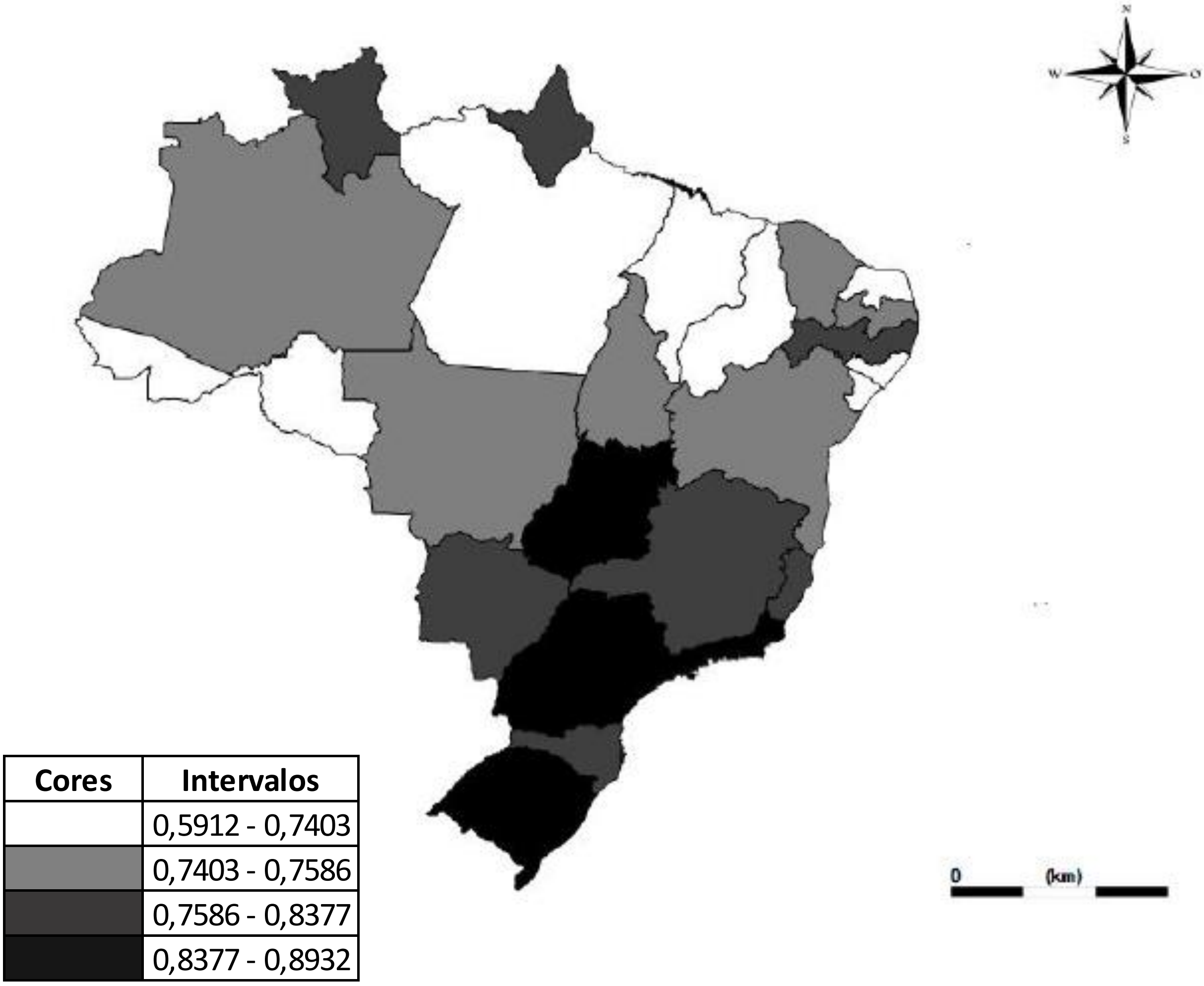


Figure B.3 – Distribution of the Opportunity Index relative to correct garbage treatment

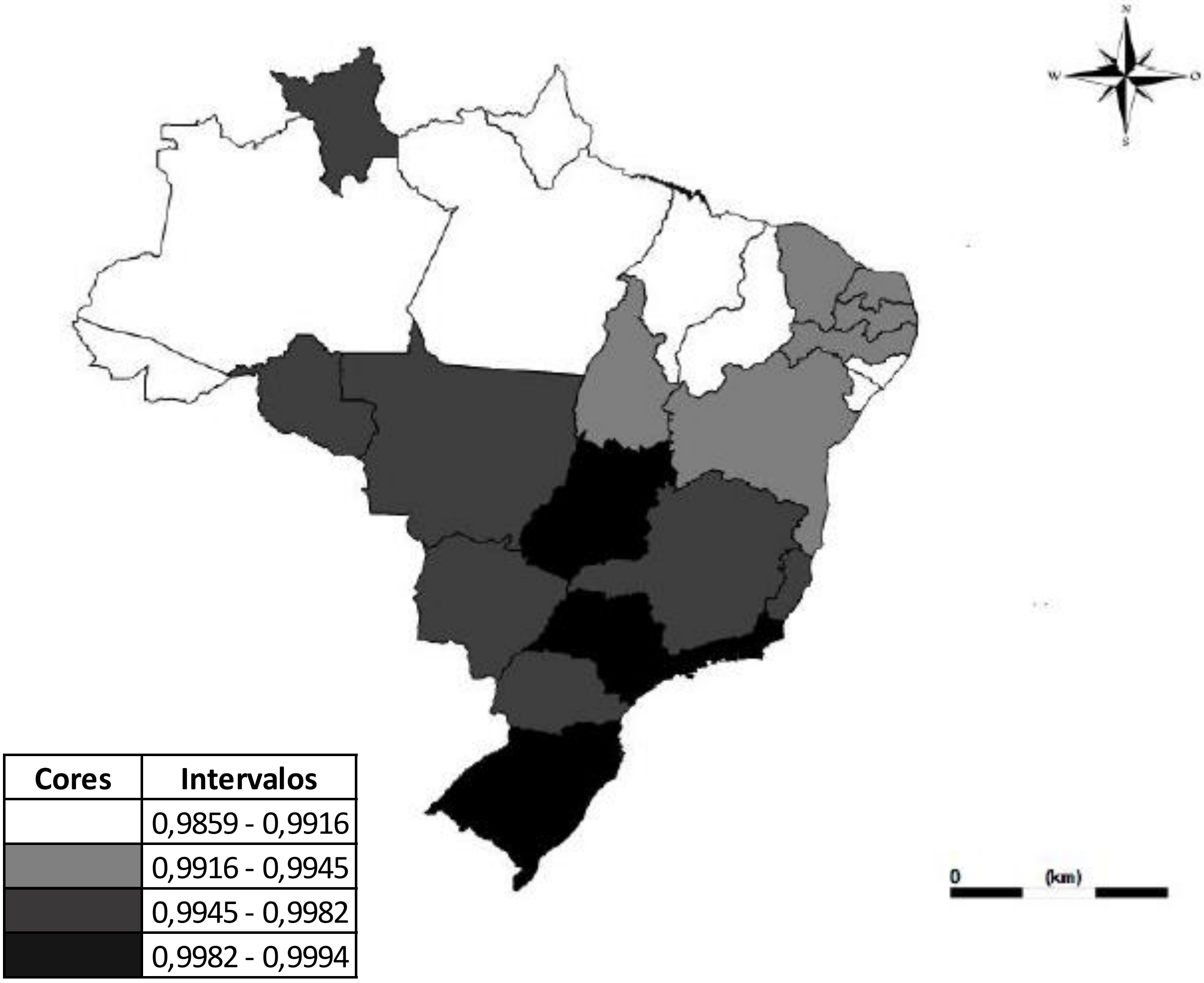


Figure B.4 – Distribution of the Opportunity Index relative to illumination

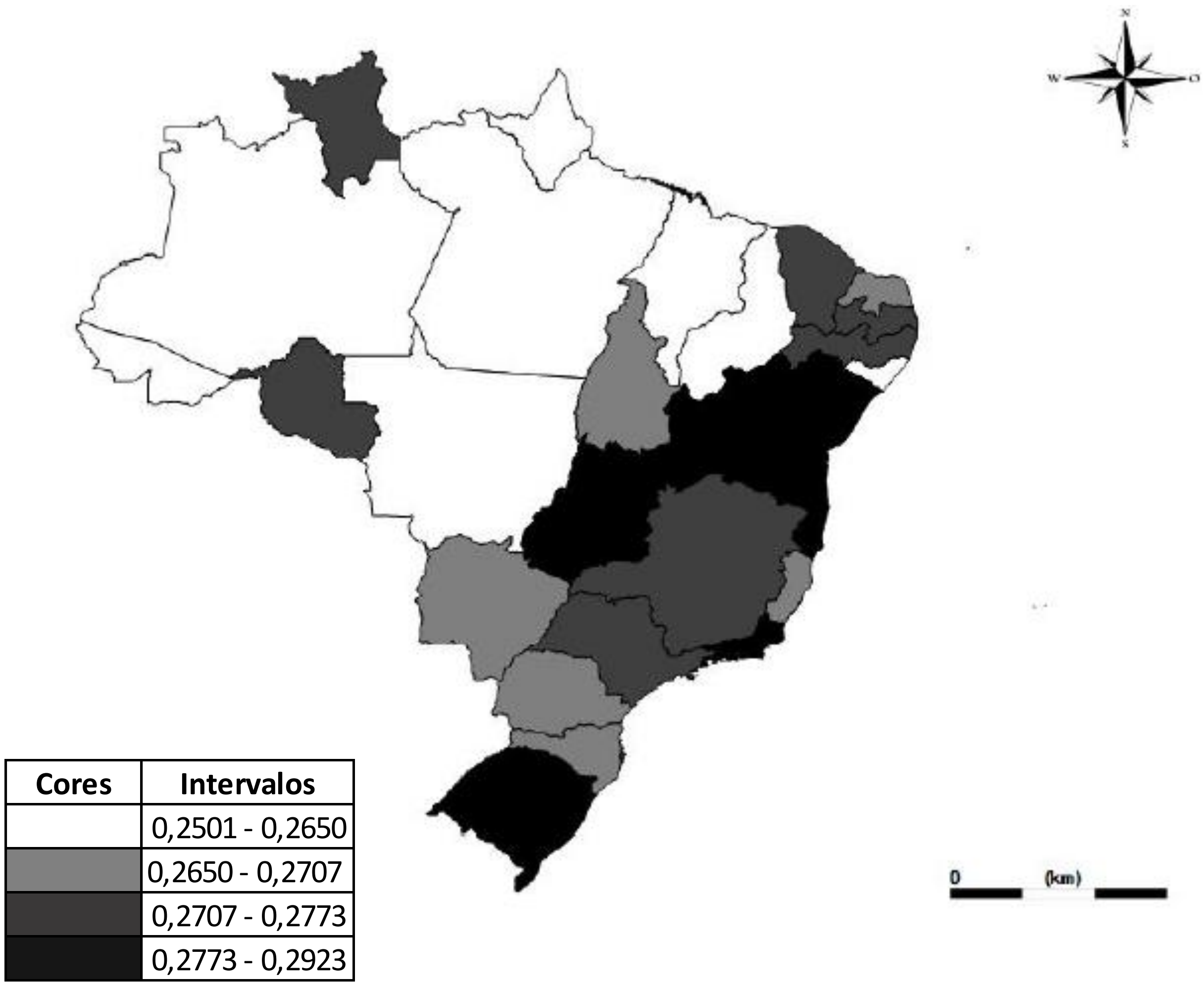


Figure B.5 – Distribution of the Opportunity Index relative to the probability of finishing the 6^{th} grade at the right age

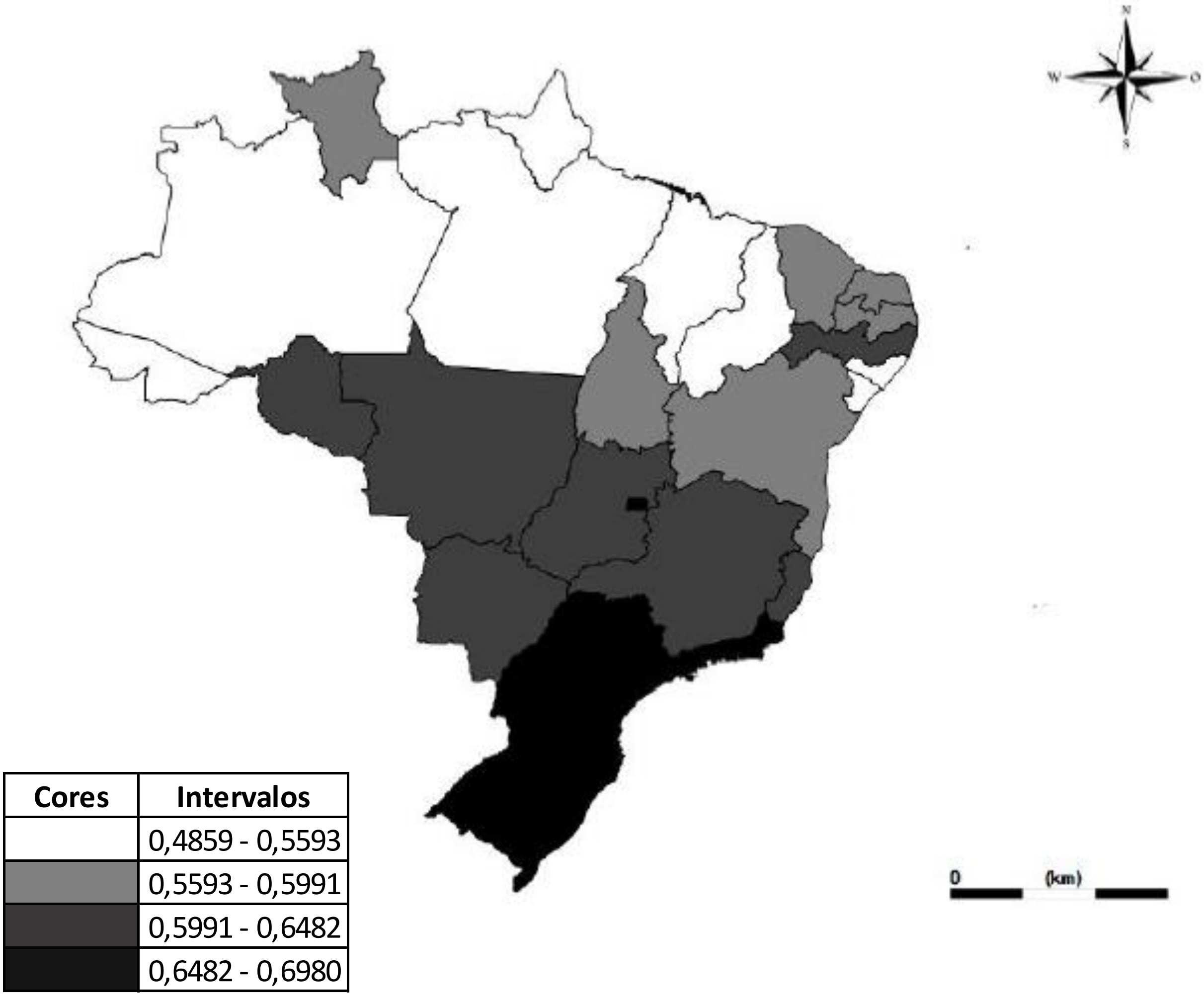


Figure B.6 – Distribution of the Opportunity Index relative to cellphone access

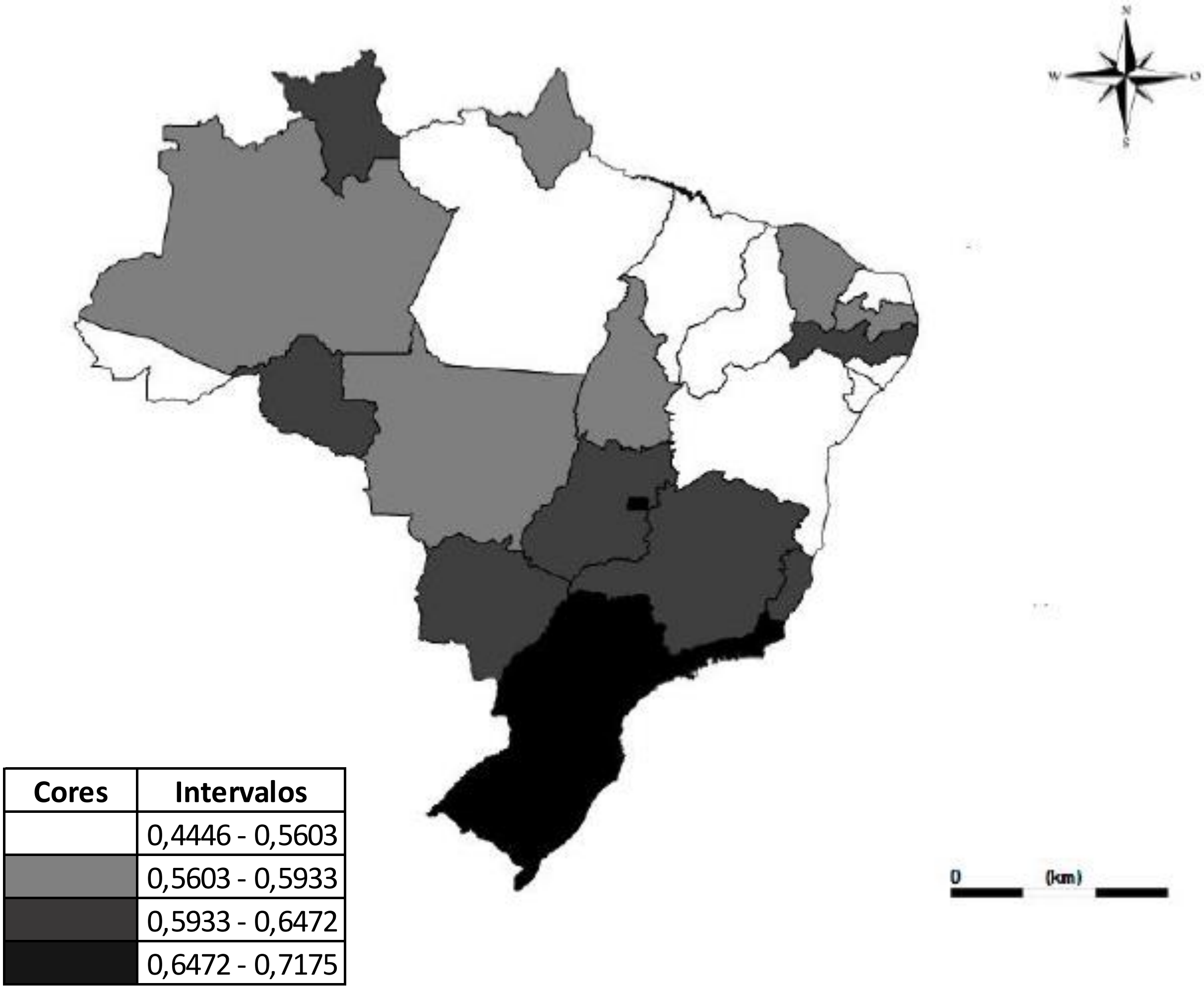
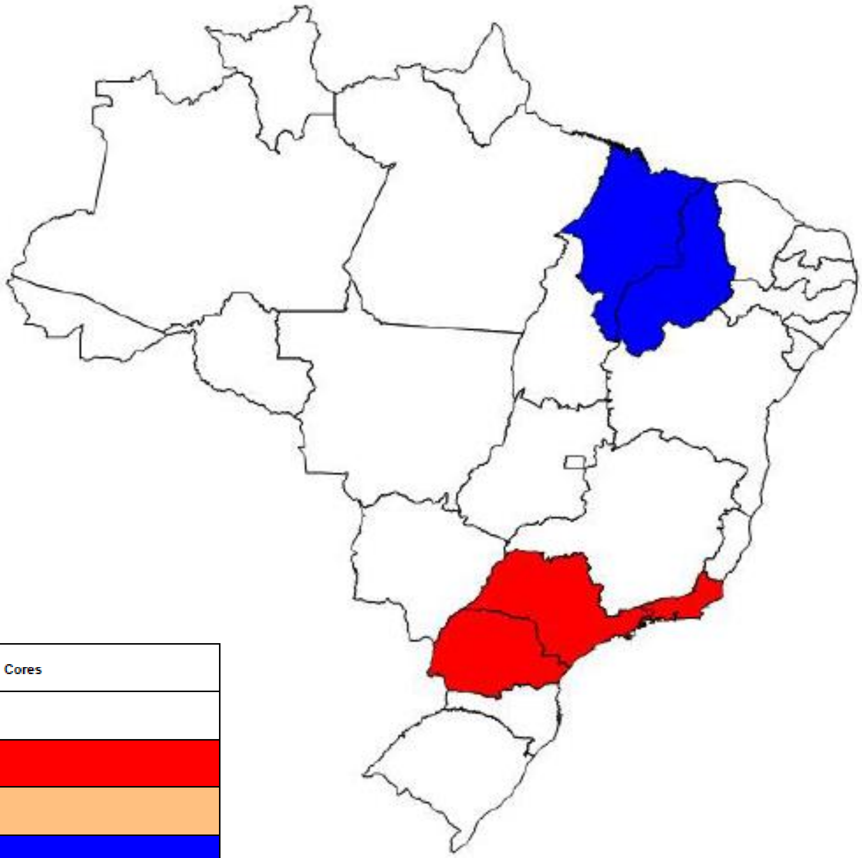


Figure B.7 – Distribution of the Opportunity Index relative to Internet access



Classificação	Cores
Não significativo	
Alto-Alto	
Alto-Baixo	
Baixo-Baixo	
Baixo-Alto	

Figure B.8 – LISA of the Opportunity Index relative to water access

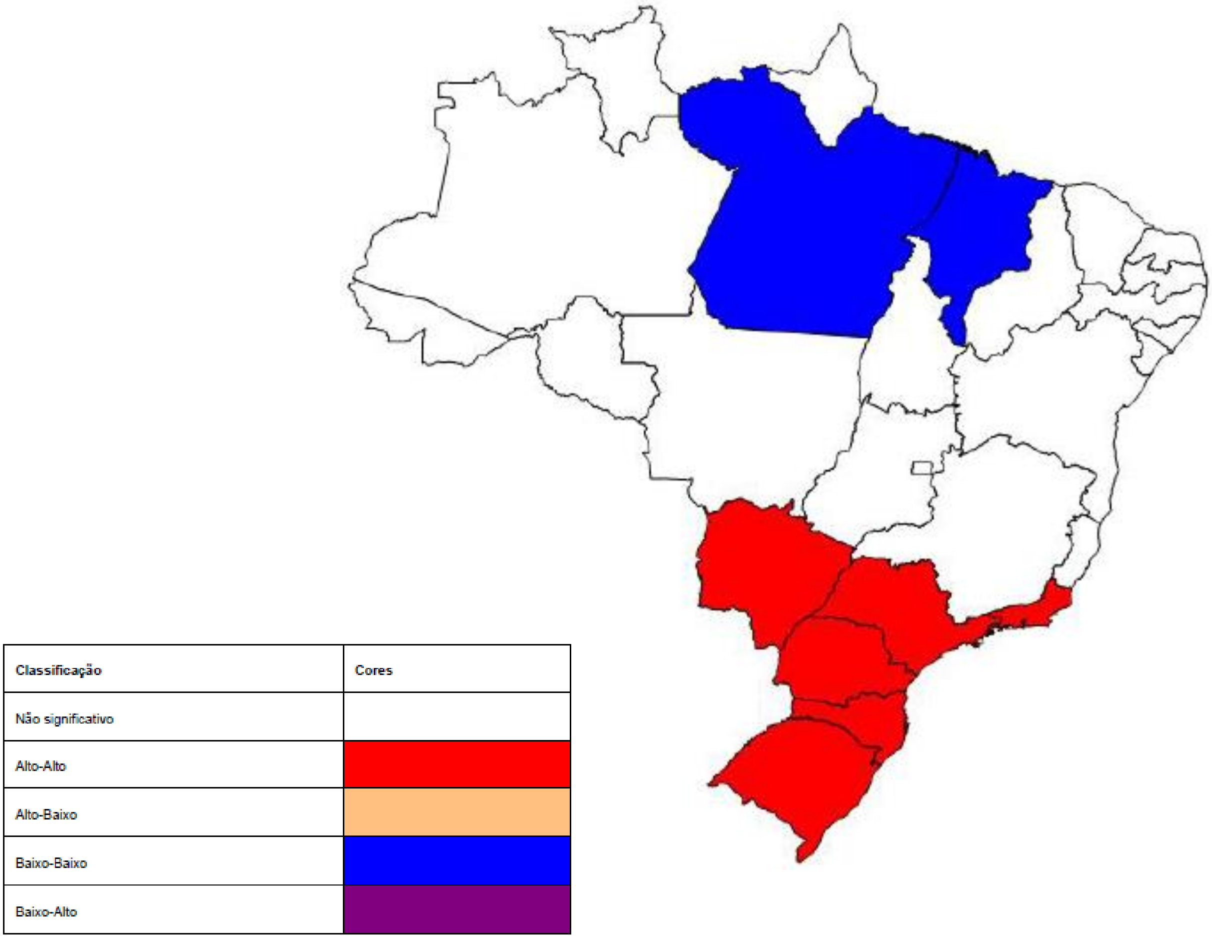


Figure B.9 – LISA of the Opportunity Index relative to basic sanitation

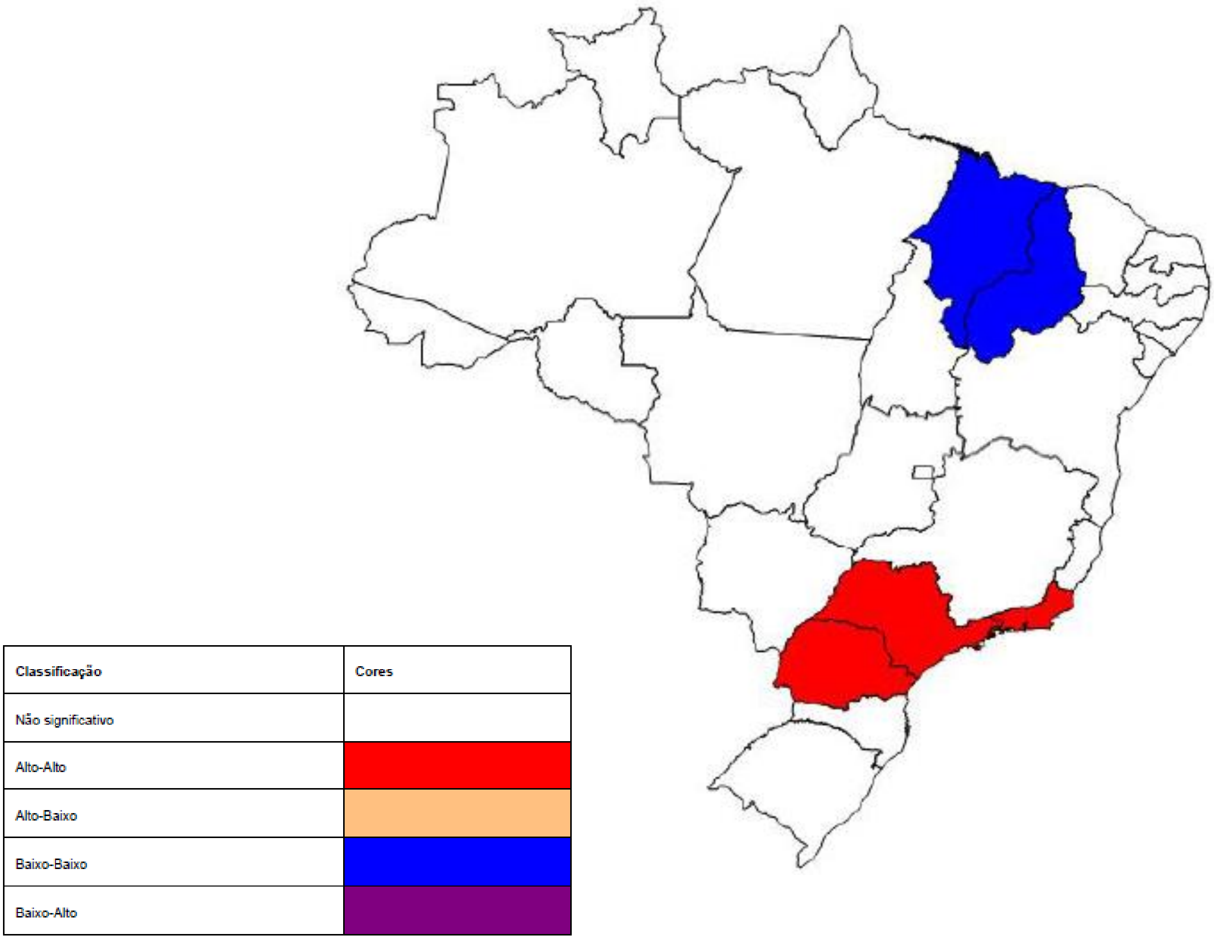
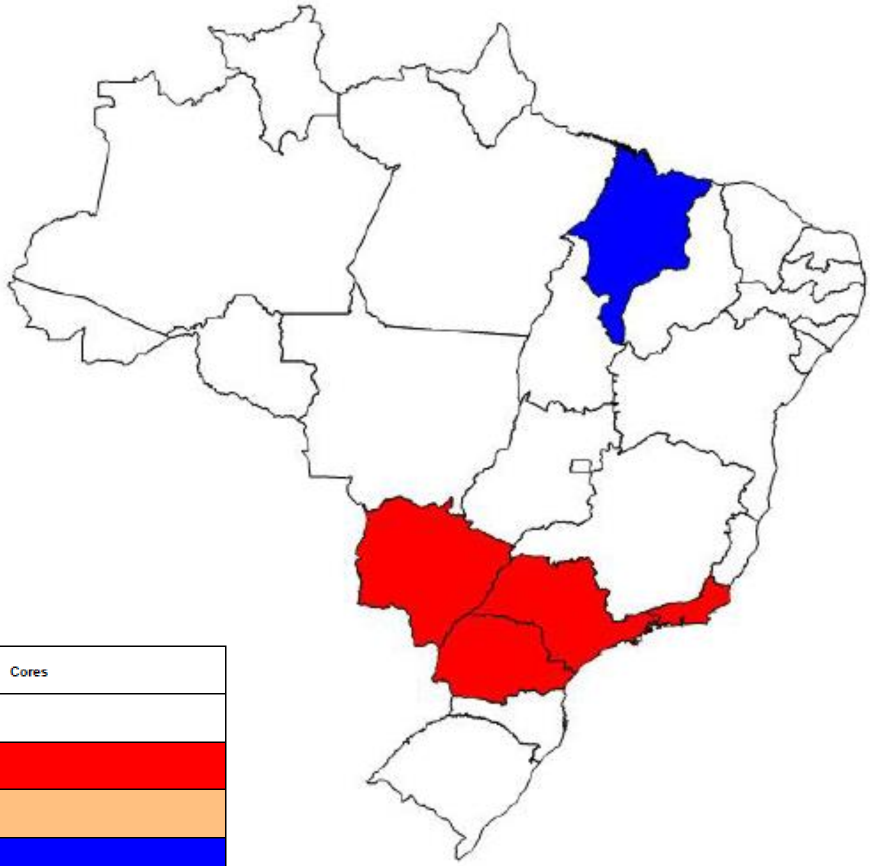


Figure B.10 – LISA of the Opportunity Index relative to correct garbage treatment



Classificação	Cores
Não significativo	
Alto-Alto	
Alto-Baixo	
Baixo-Baixo	
Baixo-Alto	

Figure B.11 – LISA of the Opportunity Index relative to illumination access

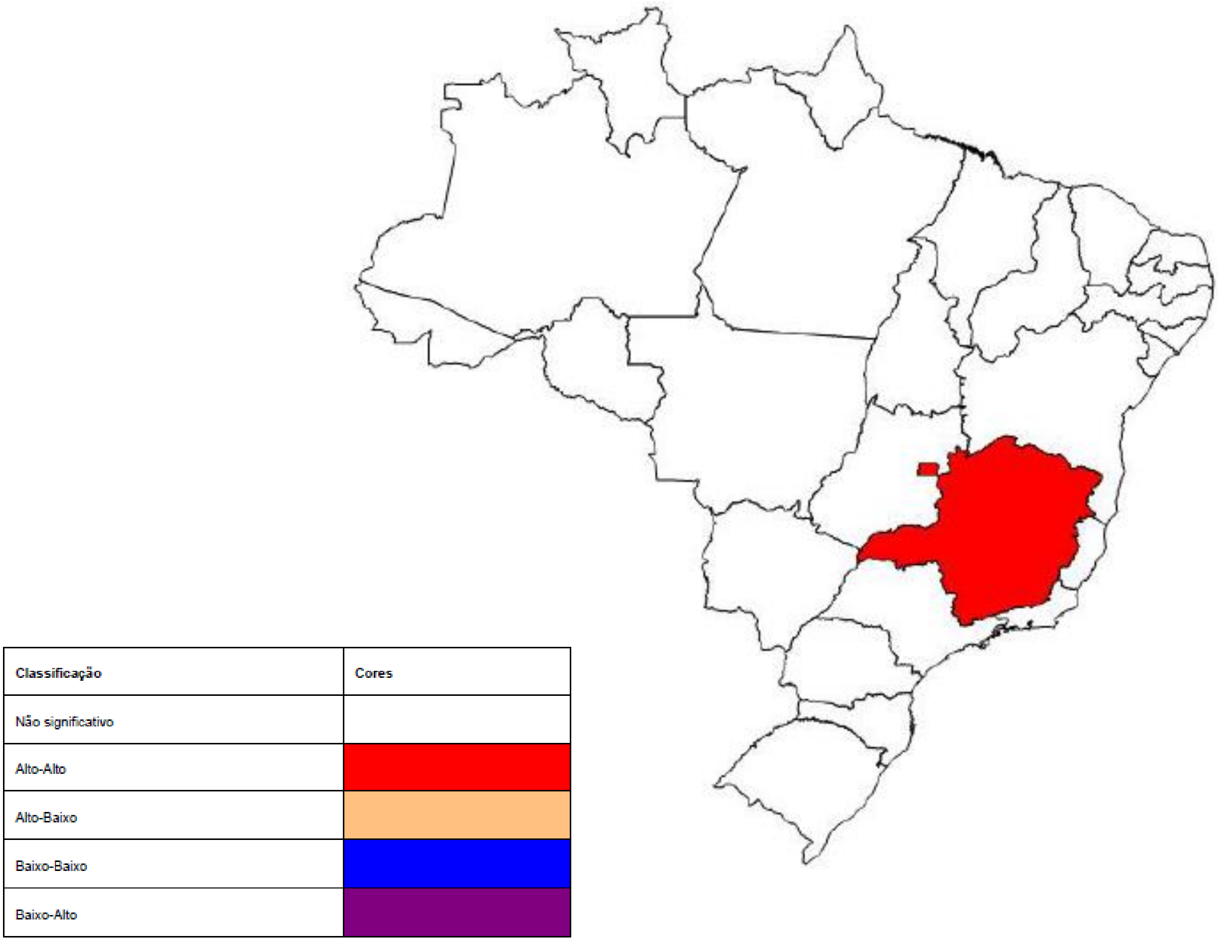


Figure B.12 – LISA of the Opportunity Index relative to the probability of finishing the 6th grade at the right age

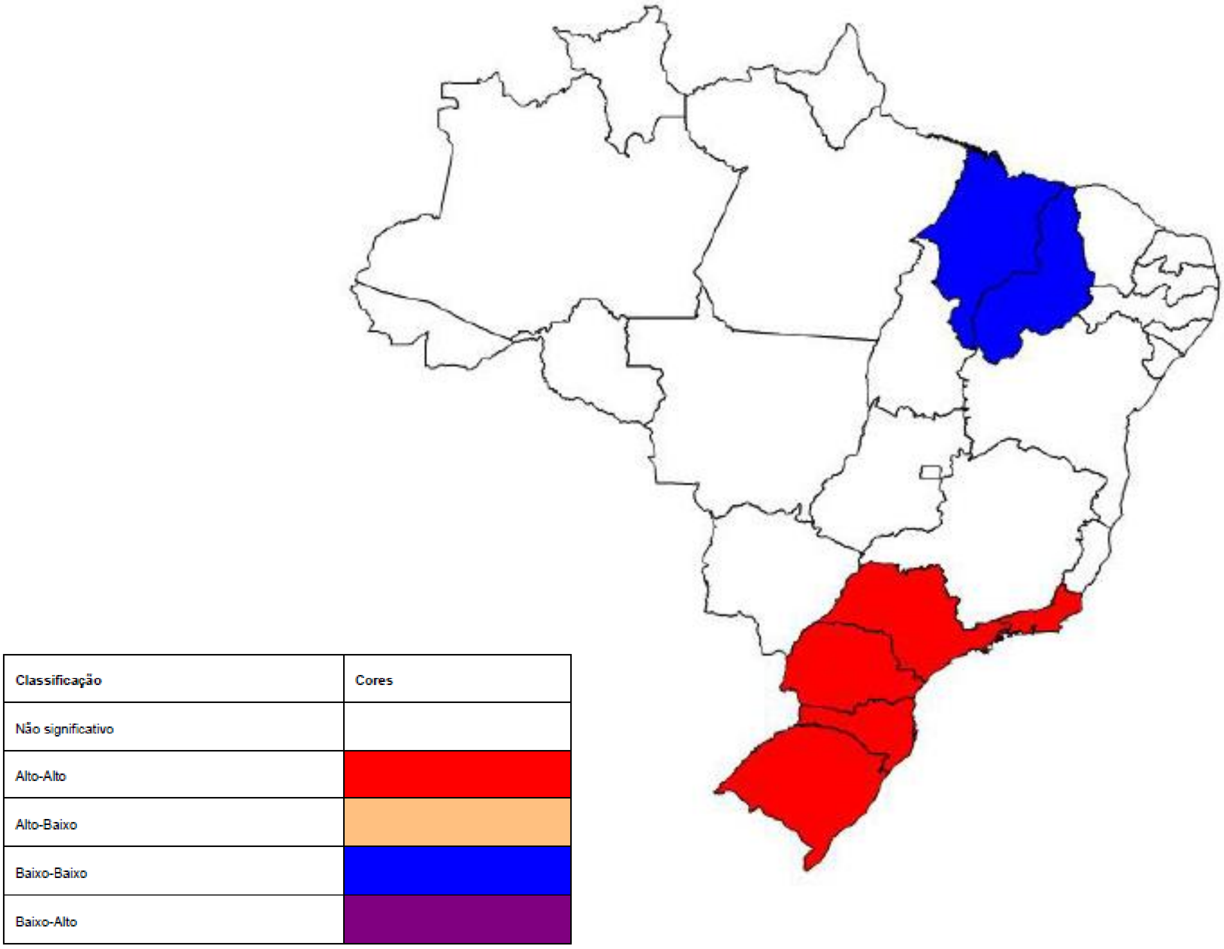
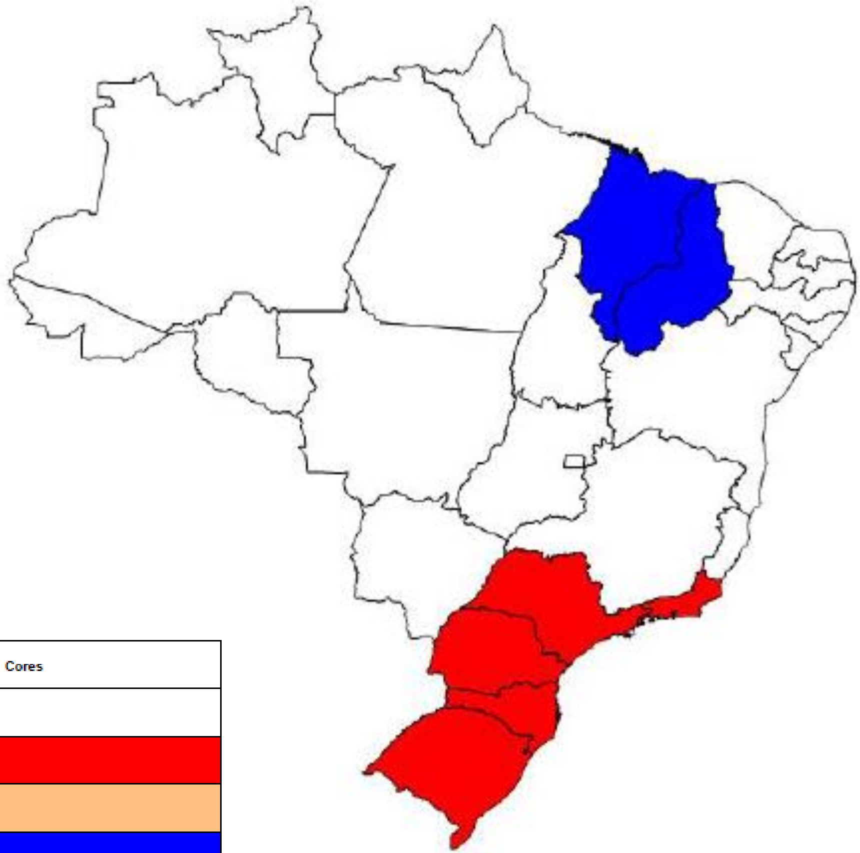


Figure B.13 – LISA of the Opportunity Index relative to cellphone access



Classificação	Cores
Não significativo	
Alto-Alto	
Alto-Baixo	
Baixo-Baixo	
Baixo-Alto	

Figure B.14 – LISA of the Opportunity Index relative to Internet access

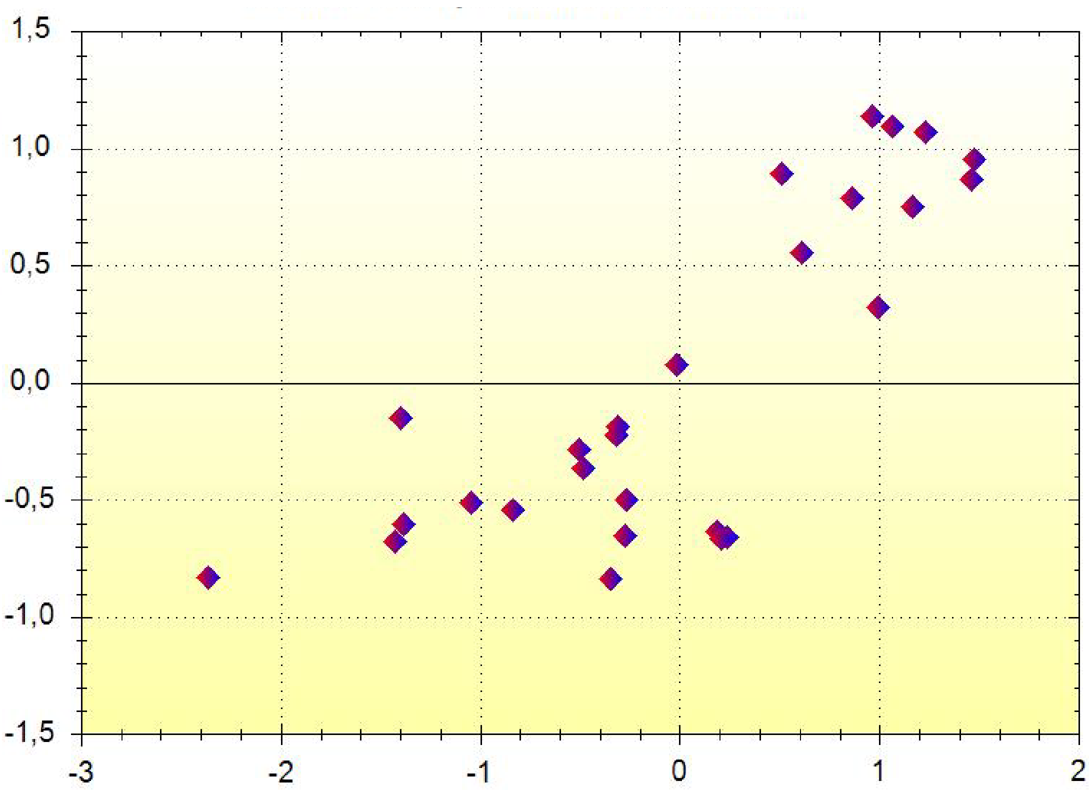


Figure B.15 – Moran Diagram for the Opportunity Index for access to piped water

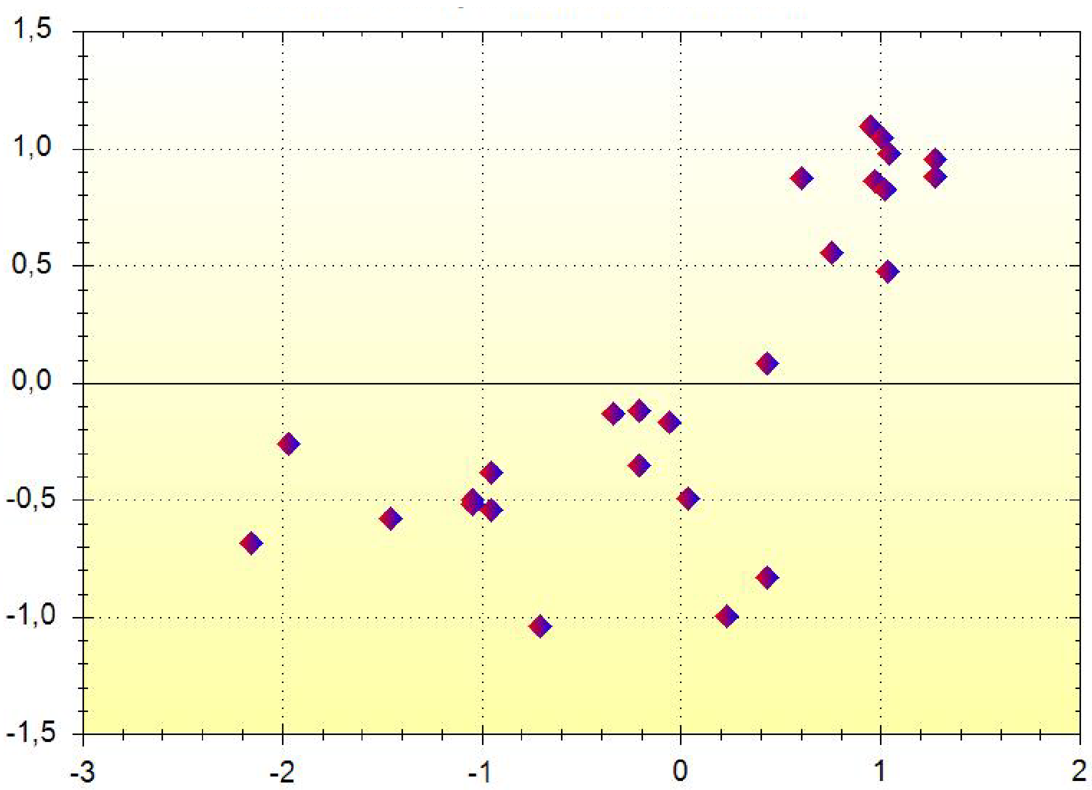


Figure B.16 – Moran Diagram for the Opportunity Index relative to illumination access

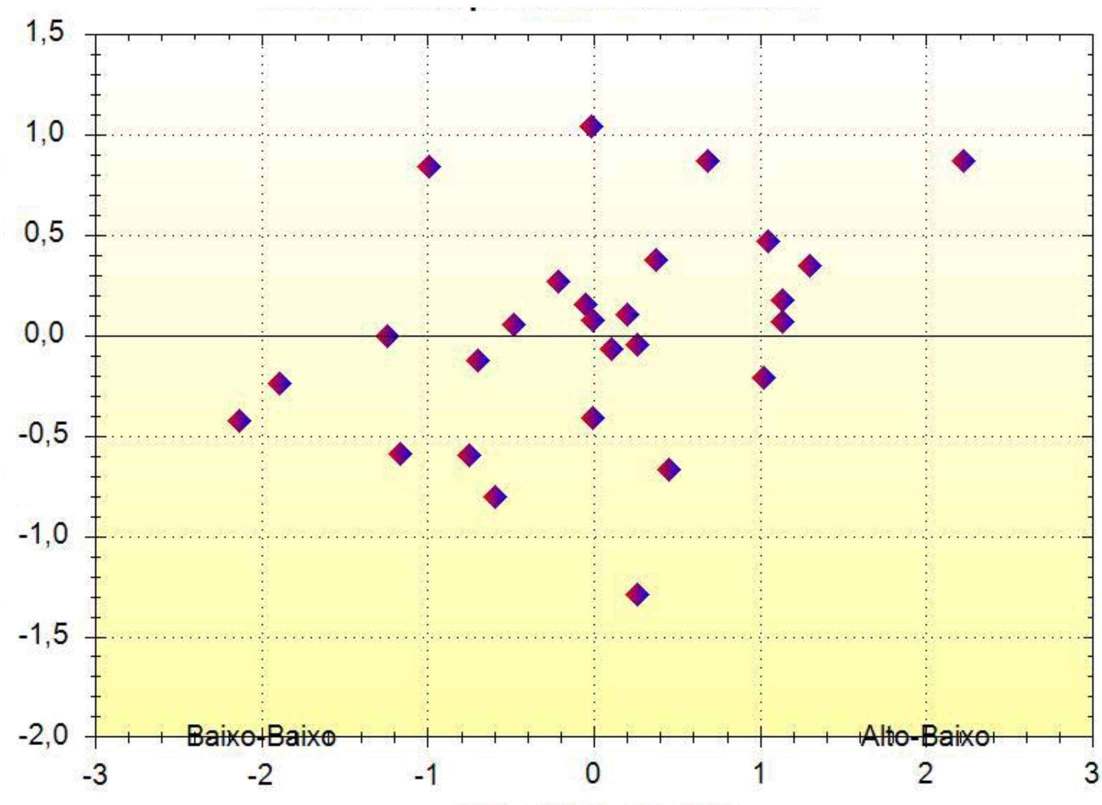


Figure B.17 – Moran Diagram for the Opportunity Index relative to the probability of finishing the 6th grade at the correct age

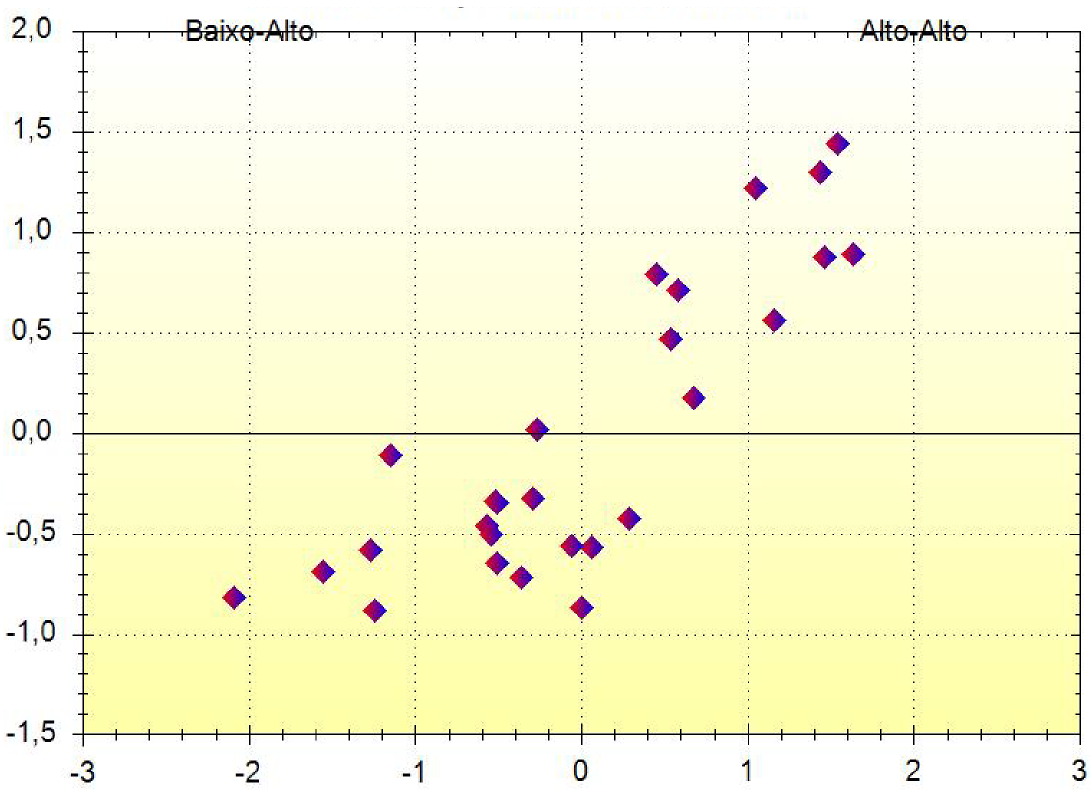


Figure B.18 – Moran Diagram for the Opportunity Index relative to Internet access

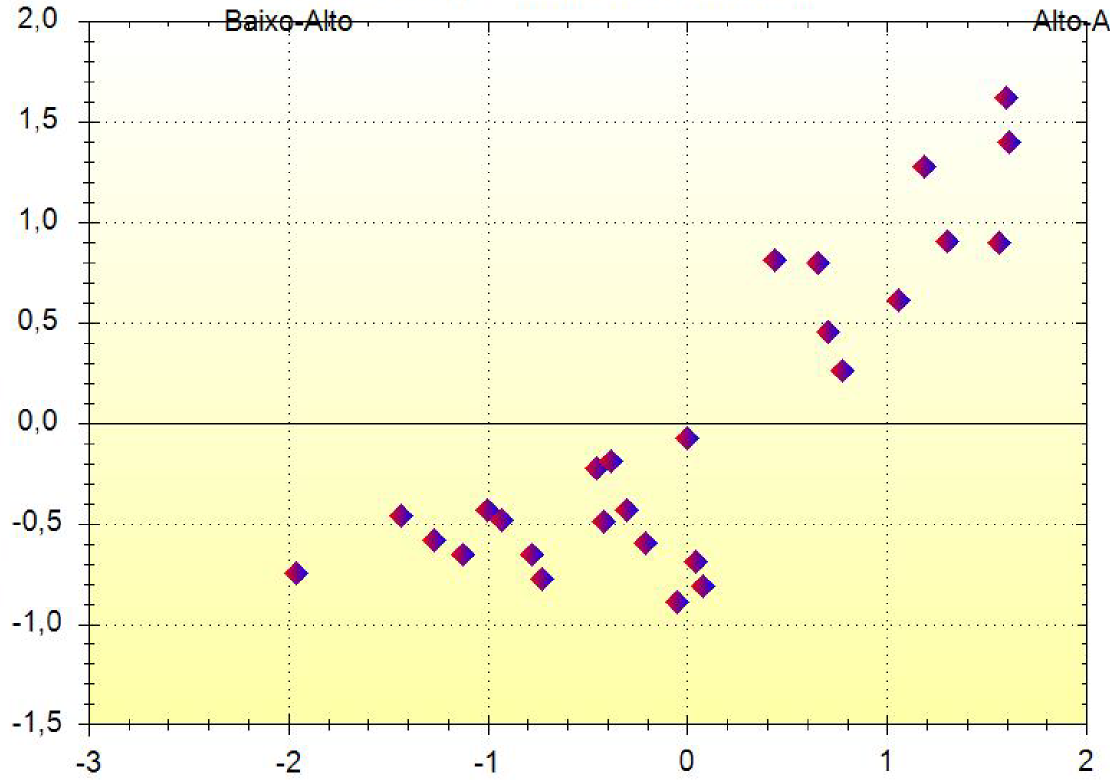


Figure B.19 – Moran Diagram for the Opportunity Index relative to cellphone access

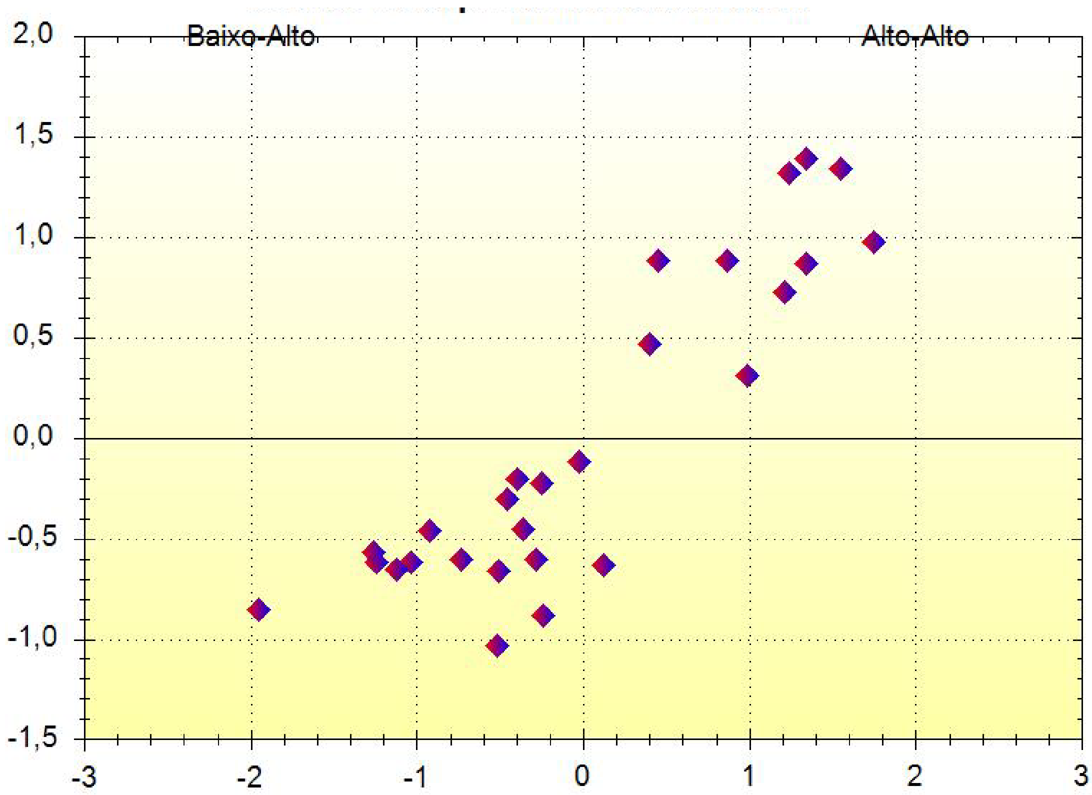


Figure B.20 – Moran Diagram for the Opportunity Index relative to basic sanitation

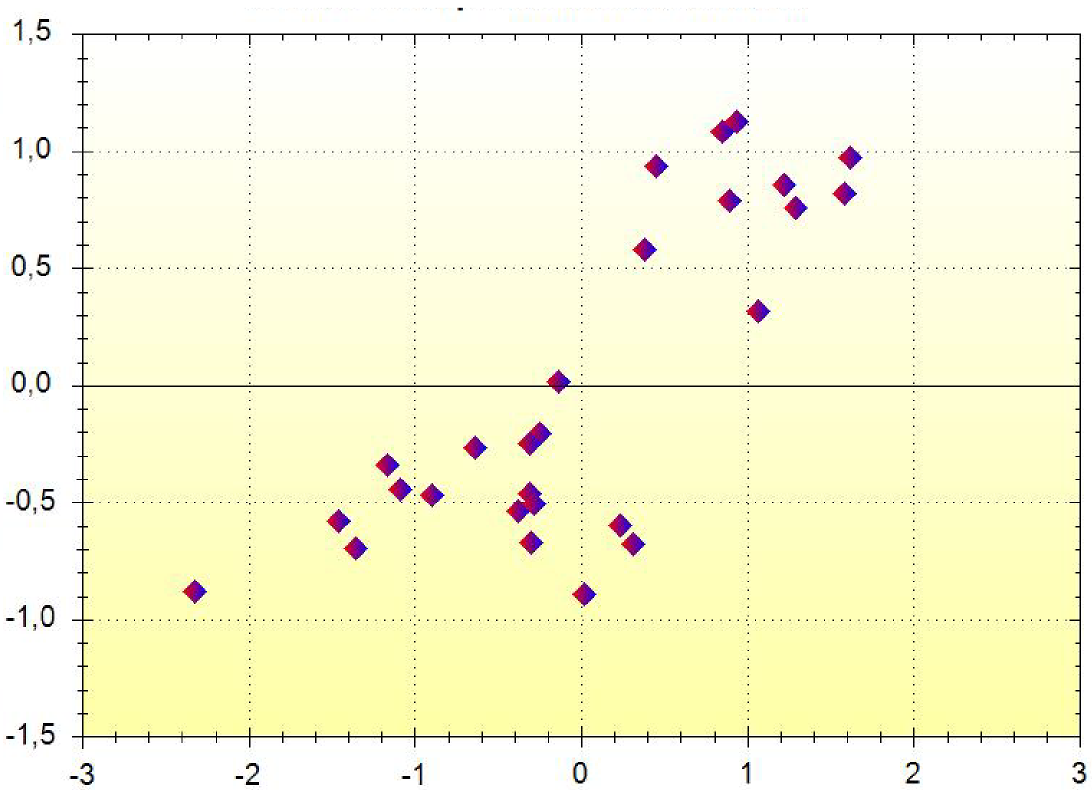


Figure B.21 – Moran Diagram for the Opportunity Index relative to the correct garbage destination

Appendices C

Stata code and Data set

This link[3] redirects to an OneDrive shared folder that contains the data used in this dissertation, and the *do file* with the Stata code used to develop the empirical application.

The file *txt* data that contains the microdata of PNAD (2015) extracted from the IBGE can be found in the "Data" Folder. In the same folder is a *xlsm file*(IDH) with the IDH of each state extracted from Atlas Brazil.

The *do file* used to estimate the HOI, can be found at "PNAD_do".

[3] https://phdisegutl-my.sharepoint.com/personal/terto_aln_iseg_ulisboa_pt/Documents/Forms/All.aspx?slrid=09d0b39e-c015-7000-9fdc-4a00d74ec3b8&RootFolder=%2Fpersonal%2Fterto_aln_iseg_ulisboa_pt%2FDocuments%2FDisserta%C3%A7%C3%A3o%2FAppendices%20C&FolderCTID=0x0120001216A46D97E2404CA421B09D3D828CFF