The Population Distribution Dynamics of Brazilian MCAs: An Spatial Evaluation for 1970-2010 Period

Diego Firmino Costa da Silva
Instituto de Pesquisa Econômica Aplicada (Assistente de Pesquisa - Bolsista). Endereço: Rua Dr. Odilon Lima, 44/02, 50850-010, Afogados, Recife, PE, Brazil. E-mail: diegofirmino@gmail.com, Phone: +55 (81) 32711196

Raul da Mota Silveira Neto
Universidade Federal de Pernambuco (Professor). Address: Av. Prof. Moraes Rego, 1235, Recife, PE, 50670-901, Brazil. E-mail: rau.silveira@uol.com.br, Phones: +55 (81) 21268381 r. 224, + 55 (81) 91381357

Resumo

Este artigo tem como foco principal a dinâmica populacional brasileira entre 1970 e 2010. Neste sentido, nosso objetivo é explorar o comportamento da distribuição populacional, utilizando tanto a abordagem tradicional de *rank* quanto as cadeias de Markov. A fim de obter informações mais precisas sobre a dinâmica e a evolução da distribuição populacional, a dependência espacial é introduzida através da análise de LISA Markov e *Spatial Markov Chains*. O formato da distribuição indica que a divergência no tamanho populacional das Áreas Mínimas Comparáveis (AMC) é decrescente. A estimação da lei de Zipf traz evidências de que a distribuição populacional está, a cada década, de distanciando da distribuição de Pareto. A abordagem utilizando as cadeias de Markov traz como principais evidências a alta persistência das AMCs permanecerem nas suas classes iniciais com o passar das décadas e o fenômeno que diferentes contextos espaciais tem efeitos diferentes sobre a transição das localidades.

Palavras-Chave: Distribuição populacional, Lei de Zipf, Cadeias de Markov, Dependência espacial

Abstract

This paper focuses on the Brazilian population dynamics between 1970 and 2010. In this sense, our objective is to explore the behavior of the Brazilian population distribution, revisiting the traditional rank-size rule and Markov chain approaches. In order to bring up more accurate information on the dynamics and evolution of the population distribution, the spatial dependence is introduced through the analysis of LISA Markov and Spatial Markov Chains. The distribution shape may indicate that divergence in population size of Minimum Comparable Areas (MCAs) is decreasing. The Zipf's law estimation indicates that the population distribution is, every decade, moving away from Pareto law. Markov chain approach brings as main evidence the high-persistence of MCAs to stay in their own class size from one decade to another over the whole period, and different spatial contexts have different effects on transition for regions.

Keywords: Population distribution, Zipf's Law, Markov Chains, Spatial dependence

JEL: C14, C21, C23, O18

Área 1 - Economia Regional

The Population Distribution Dynamics of Brazilian MCAs: An Spatial Evaluation for 1970-2010 Period

Resumo

Este artigo tem como foco principal a dinâmica populacional brasileira entre 1970 e 2010. Neste sentido, nosso objetivo é explorar o comportamento da distribuição populacional, utilizando tanto a abordagem tradicional de *rank* quanto as cadeias de Markov. A fim de obter informações mais precisas sobre a dinâmica e a evolução da distribuição populacional, a dependência espacial é introduzida através da análise de LISA Markov e *Spatial Markov Chains.* O formato da distribuição indica que a divergência no tamanho populacional das Áreas Mínimas Comparáveis (AMC) é decrescente. A estimação da lei de Zipf traz evidências de que a distribuição populacional está, a cada década, de distanciando da distribuição de Pareto. A abordagem utilizando as cadeias de Markov traz como principais evidências a alta persistência das AMCs permanecerem nas suas classes iniciais com o passar das décadas e o fenômeno que diferentes contextos espaciais tem efeitos diferentes sobre a transição das localidades.

Palavras-Chave: Distribuição populacional, Lei de Zipf, Cadeias de Markov, Dependência espacial

Abstract

This paper focuses on the Brazilian population dynamics between 1970 and 2010. In this sense, our objective is to explore the behavior of the Brazilian population distribution, revisiting the traditional rank-size rule and Markov chain approaches. In order to bring up more accurate information on the dynamics and evolution of the population distribution, the spatial dependence is introduced through the analysis of LISA Markov and Spatial Markov Chains. The distribution shape may indicate that divergence in population size of Minimum Comparable Areas (MCAs) is decreasing. The Zipf's law estimation indicates that the population distribution is, every decade, moving away from Pareto law. Markov chain approach brings as main evidence the high-persistence of MCAs to stay in their own class size from one decade to another over the whole period, and different spatial contexts have different effects on transition for regions.

Keywords: Population distribution, Zipf's Law, Markov Chains, Spatial dependence

JEL: C14, C21, C23, O18

Área 1 - Economia Regional

1. Introduction

According to UN Population Division data, the percentage of people living in urban centers in Brazil went from 55.8% in 1970 to 87% in 2010. Therefore, it is noteworthy that Brazil has become an increasingly urban country in the last 40 years and consequently, the study of cities becomes increasingly important to achieve the understanding of the national urban system. In this sense, these simple evidences awakens the need for research regarding population dynamics (growth and size distribution) of Brazilian cities, a topic that has been studied since the 1970s for the United States and the European countries cases, but which is still scarce for the Brazilian case.

Population growth of cities is also associated with the population movement between municipalities. To illustrate, according to data from the Brazilian Institute of Geography and Statistics (IBGE) in 1970 there were 66 cities in Brazil with more than 100 000 inhabitants, 105 cities in the early 1980s, 398 in 2000, and this number increased to 488 in 2010. Meanwhile, in the last decade the number of cities with over 1 million inhabitants went only from 13 to 14. In this sense, the concern is not only for the scale of urbanization, but also for the distribution of population across the urban hierarchy that becomes a challenge for policy makers to establish strategies for cities of different sizes. This observation raises some questions: How cities of different sizes grow during the process of development and transformation of a country? Is the degree of cities-size mobility slow or fast in the last 40 years? Are the movements within the distribution affected by spatial dependence?

Regarding this context, some authors investigated the behavior of cities's size distribution (Dobkins and Ioannides, 1999; Black and Henderson, 2003; Gabaix and Ioannides, 2004; Gallo and Chasco, 2007). About the Brazilian case, there are few studies on the behaviour of population distribution. Oliveira (2004a), when analyzing the evolution of city-size distribution in Brazil between 1936 and 2000, found evidence that smaller cities grew less than large ones until the 1990s. Trindade and Sartoris (2009) examined the evolution of size distribution of cities in Brazil between the 1920-2000 period and the results show evidence of divergence, similarly to Oliveira (2004a). Justo (2012) finds evidence of low interclass mobility and high persistence in the population distribution behavior of 431 minimum comparable areas between 1910 and 2010. Moro and Santos (2013) also found low mobility for the period of 1970-2010, but they only as sample the municipalities that existed in 1970, not covering all Brazilian territory.

For studies on the characterization and evolution of population distribution, the national literature available does not include information in their databases beyond the year 2000 (e.g.: Oliveira, 2004a; Trindade and Sartoris, 2009), do not go further in the investigation of spatial effects (e.g.: Justo, 2012) or do not cover all Brazilian territory (e.g.: Moro and Santos, 2013). In this sense, the first focus of this thesis is to assess the behavior of the population size distribution of Brazilian Minimum Comparable Areas (MCA) covering all Brazilian territory between 1970 and 2010. [1] Furthermore, the intention of the first part of the study is to advance in providing more accurate evidences, taking into account the possibility of spatial dependence in population size distribution by using current spatial techniques. In analyzing the evolution of Brazilian minimum comparable areas (MCA) population distribution, we begin by revisiting the traditional rank-size rule and Pareto distribution approaches. Therefore, in order to bring up information on the dynamics and evolution of the population distribution, we lead with the estimation of transition probability matrices associated with discrete Markov chains (Kemeny and Snell, 1976). And then, the spatial

[1] A MCA is a municipality or an aggregation of municipalities necessary to enable consistent spatial analyses over time. More details follow in section 4.2.

dependence is introduced through the analysis of LISA Markov and Spatial Markov Chains, both developed by Rey (2001).

In addition to these introductory notes, in the section 2 we present the foundations for distributions dynamics of population and spatial dependence through the literature review and presentation of Zipf's law and Markov Chain approaches. Section 3 describes the dataset and reports the results of the empirical analysis regarding distributional dynamics. The conclusions are presented in section 4.

2. Distribution Dynamics of Population and Spatial Dependence

In this section, we present the literature review and methodologies used to analyze the behavior of the MCAs's population distribution covering the entire Brazilian territory in the period 1970-2010. The first methodology deals with the traditional estimation of Zipf's Law, which verifies that the distribution of city sizes follows the Pareto distribution. This technique only provides some information about static distribution at each point in time. As suggested by Duranton (2006), even though it may be only a rough first approximation, Zipf's law nonetheless remains a useful benchmark to think about the distribution of city sizes. Therefore, in order to bring up information on the dynamics and evolution of the population distribution, i.e., to access information on the movements of the localities within the distribution, techniques based on the Markov Chain will be explored in the remaining topics in this section.

2.1 Literature Review

Based on the integration of spatial statistic with modeling using Markov chains, Rey (2001) studies the evolution of the regional distribution of income taking into account the transitions of both the individual economies and those of their respective geographic neighbours within a distribution of income. Using data from 48 U.S. states for the period 1929-1994, the main result found by the author is that the rates of upward or downward mobility of the states within the distribution was sensitive to the position of its neighbour in the same distribution. And a possible implication in terms of policy is that, for example, a policy to reduce regional disparities could be more effective when the receiving state of the policy is surrounded by less disparate states.

Black and Henderson (2003) examine the evolution of the US city size distribution by applying Zipf 's Law to the 20th century US city size distribution. They then turn to a more general approach to analyze the evolution and trends of the size distribution of cities by modeling the transition process of cities directly. In relation to the Zipf's Law, the Pareto parameter estimated for the whole sample lending modest support to the view of increasing urban concentration in recent decades. For the top one-third of cities, the rise in the Pareto parameter would suggest decreasing urban concentration in the US over time. But the fact is that this difference between the estimated parameters suggests that the relationship between rank and city size is not log-linear. In relation to city size distribution, the authors conclude that existing cities tend to move up the size distribution 'fairly quickly', but to move down extremely slowly. Additionally, there is some tendency in the USA towards increasing urban concentration, with a greater proportion of cities in large relative size categories.

Gallo and Chasco (2007) analyzed the evolution of population growth for a group of 722 Spanish municipalities during the period 1900-2001. Attentive to the fact that the omission of spatial autocorrelation could cause a bias to the OLS estimator, the authors followed the strategy suggested by Anselin (1988) and estimated a spatial model SUR for Zipf's Law's (size distribution of cities follows a Pareto Law) and two main phases are found: divergence (1900-1980) and convergence (1980-2001). Furthermore, the authors estimated transition matrices associated with discrete Markov chains to obtain information concerning the movements of the urban groups within the population distribution. In this case, the results indicated that the municipalities located on the ends of the distribution would be more persistent in staying in those positions in the ranking, while medium-sized cities were more likely to move into smaller categories. The authors, however, do not explore the approach suggested by Rey (2001) who proposes the use of a spatial Markov matrix.

The international literature that refers to the behavior of population distribution is much wider than the one discussed in this thesis. Dobkins and Ioannides (1999) using the data from U.S. Census and cover metropolitan areas between 1900 and 1990, the authors found evidences of divergent growth, if spatial evolution is ignored, and convergent growth in the presence of very significant regional effects. Lalanne (2013) investigate the hierarchical structure of the Canadian urban system. Some papers discuss theoretical issues on city size distributions (Gabaix, 1999; Gabaix and Ioannides, 2003; Duranton, 2006; Gan et al., 2006). The diversity of international studies is very large, but this is not the case of the national literature. Below we list some work featured in national literature regarding distributional population behavior.

Oliveira (2004a), when studying the evolution of size distribution in Brazilian cities and testing the validity of Zipf's Law, estimates Pareto coefficient for Brazil between 1936 and 2000. The obtained results do not allow the conclusion that the rule of order and size applies to Brazil. Only in 1960 and 1970 this rule is true, but represents a transition period, since the coefficient decreased constantly over the period studied. This reduction represents an increase in inequality in the size of Brazilian cities. In this study, the author does not take into account the spatial factors that could influence the results.

Trindade and Sartoris (2009) examine the evolution of the relationship between the size of Brazilian cities and their population distribution for the period between 1920 and 2000. Using models based on Zipf's Law, Markov chain and taking spatial effects into account, the authors find in their results that there is a persistent population concentration in a small number of areas. As Gallo and Chasco (2007), the authors do not use the approach of spatial Markov matrix proposed by Rey (2001), which would have clearer information about the spatial relation existing in size distribution of cities.

Monastério (2009) analyzes the changes in the spatial distribution of population and manufacturing employment in Brazil between 1872 and 1920. To this end, the used tools, which combine the spatial analysis techniques Exploratory Spatial Data Analysis (ESDA) of the Markov chain, as, suggested by Rey (2001). The sample consists of minimum comparable areas in the Northeast, Southeast and South, and the state of Goiás. That is, the other states of the Midwest and the northern region are outside the sample used. The analysis revealed differences in the trajectories of the areas within states, the role of space in the dynamics and the tendency to increase in concentration during the studied period, especially with regard to manufacturing occupation. The analysis using the Markov matrix spatially conditioned indicated that the neighborhood was essential to the destinations of AMCs. Localities with little dense neighbors tended to approach the low-density profile of its neighbors.

Justo (2012) seeks to identify the dynamics of population growth for a group of 431 minimum comparable areas in Brazil between 1910 and 2010. For this, the author estimated spatial models for Zipf's law achieving results that point to the divergence, which has been losing strength in recent decades. Furthermore, through the estimation of functions of non-parametric densities, the author attempts to characterize the population distribution and through a process stationary first order Markov Chain shown the growth process of Brazilian cities. The results point to a low interclass mobility and high persistence. The probability of remaining cities on the class itself between a decade and another over the last hundred years is high. Despite being a very recent paper, as Trindade and Sartoris (2009), the author does not use the approach of spatial Markov matrix proposed by Rey (2001).

Moro and Santos (2013) test the Zipf's Law in order to describe the spatial distribution of the Brazilian cities and Markov Chains analysis to examine the dynamics of the cities within the urban system. Additionally, the authors introduce spatial dependence in both Zipf's law estimation and Markov. To estimate the Zipf's law equation, they used the full sample of municipalities between 1970 and 2010. The results point that the Pareto coefficient is much smaller than 1, featuring a polarized and asymmetric urban structure. Regarding the spatial Markov approach, the results show strong evidence that the probability of urban growth of a municipality depends on the surrounding urban context, and there is a low mobility for the period 1970-2010. However, in the Markov chain analysis, they use as sampling only municipalities that existed in 1970, not covering all Brazilian territory in the following decades. In this way, territory and population of new municipalities (created from the subdivision of former municipalities) will be excluded from the sample, skewing the results with selection bias.

2.2 Zipf's Law

The evolution of the size distribution of cities is explored through the law of Zipf, or rank-size rule. Zipf (1949) stated that the size distribution of cities follows a Pareto law (Pareto, 1896) by claiming that:

$$R = a \cdot S^{-b} \tag{1}$$

where R is the classification order of the city in the size distribution of population, S is the city's population, a and b are parameters, b is the Pareto exponent. Formally, the size distribution of cities depends on the value of b parameter. In the limit, if b tends to infinity, all cities will have the same size. The smaller the value of b, the greater the inequality in the size distribution of cities.

In terms of the Pareto distribution, this means that the probability of city size be greater than some S is proportional to $1/S$: $P(Size > S) = \alpha / S^{b}$, the statement of Zipf's Law implies a Pareto exponent of unity, $b=1$. According to this law, populations of cities within any group of cities at any point in time are inversely proportional to the ranking of their populations in this group. According to Gabaix (1999), one proposed explanation for Zipf's Law is if cities grow randomly, with the same expected growth rate and the same standart deviation, the limit distribution will converge to Zipf's law.

At this moment to is interesting to point the differences between Zipf's law and rank-size rule. Using Gabaix (1999) words, Zipf's law states that the probability that a city has a size greater than S decreases as *1/S*. The rank-size rule states that we should expect the size S_i of a city of rank *I* to follow a power law: the size of the city of rank *I* varies as *1/i* , and the ratio of the second largest city to the largest city should be ½, the ratio of city 3 and city 2, 2/3, and so on. These size ratios are often used to compare actual urban patterns with "ideal" (Zipf) patterns. In fact, even if Zipf's law is verified exactly, the rank-size rule will be verified only approximately, if our probabilistic interpretation of Zipf's law is correct.

The b parameter can be interpreted as an indication of inequatility. More precisely, the high value of b represents a greater possibility of mobility. That is, as the inequality in the localities's size is small, the possibility of mobility is higher in the rank. Greater dispersion of population among the cities implies increasing convergence of cities's sizes and a greater number of cities with a population close to average size (the smaller the size variance). Empirically, the logarithms are taken on both sides of equation (1) and the linear expression for each city each year is estimated:

$$\ln R_{it} = \ln a_t - b_t \cdot \ln S_{it} + \varepsilon_{it} \tag{2}$$

Unfortunately, it is not possible to have information on the dynamics of the distribution only from the estimation of equation (2). According to Gallo and Chasco (2007), Zipf's law allows the characterization of the overall evolution of the size distribution of cities, but gives no information on the movements of the cities within the distribution. It is not to possible to answer, for example, why is it that some cities are present in certain positions of the distribution over time. Another limitation of Zipf's law to study the population distribution of cities is that in addition to not realize movements within the distribution, it does not take into account the possibility that these movements are affected by spatial dependence. To clarify these issues, in the next topics we lead with the estimation of transition probability matrices associated with discrete Markov chains (Kemeny and Snell, 1976), which will make it possible to follow the progress of each group of cities of a certain size in time. And then, the spatial dependence is introduced through the analysis of LISA Markov and Spatial Markov Chains, both developed by Rey (2001).

2.3 Markov Chains and Spatial Dependence

The study of distributive population dynamics according to the position of the cities and the trend configuration of population distribution over time is a method aimed at describing the law of motion driving the evolution of the distribution as a Markovian stochastic process. Once estimated the motion up or down probabilities in the population hierarchy during a transition period of a given length, the law is used to calculate a limiting population distribution characterizing a stochastic steady-state income distribution to which the system converges over time. Through modeling of the transition process of the minimum comparable areas directly, we can examine the evolution and trends in the MCAs size distribution. Compared to continuous stochastic kernels, for example, one of advantages of using this method listed by Gallo and Chasco (2007) is that discrete probability distribution and transition matrices are easier to interpret: various descriptive indices and the long-run or ergodic distribution are easier to compute. The Zipf's law, as density functions, allows the characterization of the evolution of the global distribution, but Gallo and Chasco do not provide information about the movements of the localities within the distribution. Specifically, they do not say if the locations that were in a region of the distribution at the beginning remain or not in the same region of the distribution at the end of period.

We denote F_t as the distribution of the cross-section population of the minimum comparable areas at time *t* related to an average in the country. Defining a set of *K* different size classes, we discretize the population distribution in K relevant classes. To proceed with the estimation, first we need to assume that the distribution frequency follows a first order stationary process of Markov. This assumption requires transition probabilities, p_{ij}, of order 1, which means independence of the classes at the beginning periods (*t-2, t-3, ...*). If the order is higher, the transition matrix will not be clearly specified. That is, we only have part of the necessary information to describe the true evolution of the population distribution. Following this assumption, the evolution of a size distribution is represented by a transition probability matrix, *M*, in which each element (*i, j*) indicates which is the probability that a city in class *i* at time *t* will be in the class *j* in the following period. The (*K, 1*) vector indicating the frequency of cities in each size class at time *t*, is described by:

$$F_{t+1} = MF_t \tag{3}$$

where M is the matrix of transition probability ($k \times k$) representing the transition between two distributions:

$$M = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1K} \\ p_{21} & p_{22} & \cdots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1} & p_{K2} & \cdots & p_{KK} \end{bmatrix} \tag{4}$$

where each element $p_{ij} \geq 0$ represents the probability that the cities of a particular size class *k* at time *t-1* will be in the class *j* at time *t* and $\sum_{j=1}^{K} p_{ij} = 1$.

The elements of the matrix *M* can be estimated by the frequency of changes from a size class to another. According to Amemiya (1985) or Hamilton (1994), the maximum likelihood estimator of p_{ij} is:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i} \tag{5}$$

where n_{ij} is the total number of cities moving from class *i* in the decade *t-1* to *j* in decade *t* and n_i is the total number of municipalities which remains in *i* for all *T-1* transitions.

If the transition probabilities are stationary, in other words, if the probabilities of two classes do not vary over time, then:

$$F_{t+s} = F_t M^s \tag{6}$$

Thus, we can define the steady state distribution (also known as ergodic distribution) of F_t, which is characterized when *s* tends to infinity in the equation (6), since the changes represented by *M* are repeated an arbitrary number of times. Such distribution of steady state exists if the Markov chain is regular, which means, if and only if, for an *m*, M^m has no inputs with a value equal to zero. In this case, the matrix of transition probabilities converges to M^* of rank 1. Then the existence of a steady state distribution, F^*, is characterized by:

$$MF^* = F^* \tag{7}$$

The vector F^* describes the future distribution of cities' sizes, if the movements observed in the sample period are repeated *ad infinitum*. Each row of M^t tends to the limit of the distribution when $t \to \infty$.

To get a sense of speed with which the urban localities move within the distribution, it is possible to calculate the matrix of mean first passage times M_P, where one element M_{Pij} indicates the expected time for a unit of observation to move from class *i* to class *j* for the first time. For a regular Markov chain, M_P is defined as

$$M_P = (I_K - Z + ee'Z_{dg})D \tag{8}$$

where I_K is the identity matrix of order K, Z is the fundamental matrix: $Z = (I_K - M + M^*)^{-1}$, M^* is the limiting matrix, *e* is the unit vector, Z_{dg} results from Z setting off-diagonal entries to 0, and *D* is the diagonal matrix with diagonal elements $d_{ii} = 1/\alpha_i$, given that $\alpha = \{\alpha_1, \alpha_2, ..., \alpha_r\}$ is the limiting probability vector for *M* (Kemeny and Snell, 1976).

A consideration about limitations of traditional Markov chains to study the dynamics of cities is that they do not capture the spatial dependence that may exist between the studied observational units. This spatial dependence can arise from measurement errors, such as boundary mismatches between the administrative data and market processes (Rey, 2001), and may reflect amenities and knowledge spillovers or trade and migration flows. In this sense, in order to bring up information on the spatial dependence in the population distribution, techniques developed by Anselin (1995) and Rey (2001) will be explored in the remainder of this section.

Local Indicators of Spatial Association

Anselin (1995) suggest a class of Local Indicators if Spatial Association (LISA) for the analysis of spatial clustering and hot spots. These local statistics can provide more detailed insights on the location specific nature of spatial dependence. The local Moran statistic is given by

$$I_i = \frac{z_i}{\sum_i z_i^2} \times z_i^{\circ} \qquad (9)$$

where z_i expresses the observation for region *I* on a variable as a deviation from the mean, and z_i° is the spatial lag for location *i*

$$z_i^{\circ} = \sum_{j=1}^{n} w_{ij} z_j \qquad (10)$$

As pointed by Rey (2001), the local Moran can be used to map an observation's location in absolute space into a relative space that consider not only its point in an a-spatial variate distribution, but also the location of its neighbours in the same distribution. The position of neighbours is summarized by the spatial lag from z_i°. The LISA implementation, then, divides the observations in four classes according to their local Moran statistic, as summarized in table below. Additionally, several geographical aspects can be viewed in a Moran scatter plot, consisting of pairs of local Moran and its spatial lag local Moran for each location.

Table 1 – LISA Classifications

Class	Own Value (z_i)	Neighbours Value (z_i°)
HH	Above Average	Above Average
HL	Above Average	Below Average
LH	Below Average	Above Average
LL	Below Average	Below Average

Source: Rey (2001)

Rey (2001) suggest an extension of LISA approach that integrates the local indicators of spatial integration into a dynamic framework based on Markov chain. The main motivation for this extension is the fact that Moran scatter plot only brings spatial information about the locational distribution of a given variable at a point in time.

In each period, the local Moran statistic for each observation can be classified into four mutually exclusive classes. Thus, there are twelve possible transitions a local Moran may experience over two or more periods. Moreover, These twelve transitions can be divided in three groups: Type 0 – The region-neighbours pair remains at same level; Type I – Only the region moves, but its neighbours were in the same category; Type II – Involves a transition of only the neighbours in relative space, but the region in question remains in the previous state; and Type III – Involves a transition of both a locality and its neighbours. The Type III can be broken down into two subgroups: IIIA – which occurs when both state and neighbours move in the same direction in the distribution; and IIIB - occurs when locality and neighbours move in opposite directions.

Rey (2001) also suggests two interesting measures that can be obtained using the frequency of each type of transition between two periods. One is a flow measure, which can be understood as a measure of instability in the short-term spatial dynamics. A measure of instability or flux of the short run can be given by

$$\text{Flux}_t = 1 - \frac{F_{0,t}}{n} \tag{11}$$

Where $F_{I,t}$ is defined as the number of observation that experience a transition of type *I* in the period $t \rightarrow t+1$, and $n = F_{0,t} + F_{I,t} + F_{II,t} + F_{III,t}$. This flux measure varies between 0 and 1, where 1 indicates a high instability.

Since the relationship between the locality and its neighbours remains cohesive under Type 0 and Type IIIA, a measure of spatial cohesion is given by

$$\text{cohesion}_t = \frac{F_{IIIA,t} + F_{0,t}}{n} \tag{12}$$

This cohesion measure varies between 0 and 1, where 1 indicates a high cohesion. It is the percentage of locations that move in the same direction of its spatial lag or locality-neighbours pair that remains in the same class from the previous period.

In the original implementation of LISA, developed by Anselin, a bifurcation of high and low value relative to the mean was used. This correspond to discretize the distribution in k=2 classes. According to Rey (2001), with respect to Markov chains, such classification may too aggregate and darken some of transitional dynamics in the income distribution.

Spatial Markov Chains

Rey (2001) suggests a modification in the traditional Markov matrix, conditioning the transition probability (p_{ij}) to the *j* initial class of the spatial lag of the variable in question. Here, this conditioning concerns the population size class of the spatial lag in the initial period. This combination of traditional Markov matrix with the spatial autocorrelation is called spatial Markov matrix. Conditioned on the class of spatial lag in the initial period, this matrix can be constructed by dividing the traditional matrix ($k \times k$) in *k* conditional matrices of dimension (*k,k*), this is, the traditional matrix ($k \times k$)is decomposed into a system ($k \times k \times k$). In other words, an explicit test of adherence or propulsive influence of neighbours of an economy can be based on the comparison between the different states transitions conditioned to the initial state of its spatial lag (Rey, 2001).

For the *k-th* matrix conditional, an element $p_{ij|k}$ is the probability of a region in class *i* at time *t* convert the class *j* in the next moment on the understanding that its spatial lag was in class *k* at time *t*. This matrix is shown in Table 1 below where *k = 4.*

Table 2 – Spatial Markov Transition Probability Matrix

Spatial Lag	t_i	$t_i + 1$			
		1	2	3	4
1	1	$p_{11\|1}$	$p_{12\|1}$	$p_{13\|1}$	$p_{14\|1}$
	2	$p_{21\|1}$	$p_{22\|1}$	$p_{23\|1}$	$p_{24\|1}$
	3	$p_{31\|1}$	$p_{32\|1}$	$p_{33\|1}$	$p_{34\|1}$
	4	$p_{41\|1}$	$p_{41\|1}$	$p_{43\|1}$	$p_{44\|1}$
2	1	$p_{11\|2}$	$p_{12\|2}$	$p_{13\|2}$	$p_{14\|2}$
	2	$p_{21\|2}$	$p_{22\|2}$	$p_{23\|2}$	$p_{24\|2}$
	3	$p_{31\|2}$	$p_{32\|2}$	$p_{33\|2}$	$p_{34\|2}$
	4	$p_{41\|2}$	$p_{41\|2}$	$p_{43\|2}$	$p_{44\|2}$
3	1	$p_{11\|3}$	$p_{12\|3}$	$p_{13\|3}$	$p_{14\|3}$
	2	$p_{21\|3}$	$p_{22\|3}$	$p_{23\|3}$	$p_{24\|3}$
	3	$p_{31\|3}$	$p_{32\|3}$	$p_{33\|3}$	$p_{34\|3}$
	4	$p_{41\|3}$	$p_{41\|3}$	$p_{43\|3}$	$p_{44\|3}$
4	1	$p_{11\|4}$	$p_{12\|4}$	$p_{13\|4}$	$p_{14\|4}$
	2	$p_{21\|4}$	$p_{22\|4}$	$p_{23\|4}$	$p_{24\|4}$
	3	$p_{31\|4}$	$p_{32\|4}$	$p_{33\|4}$	$p_{34\|4}$
	4	$p_{41\|4}$	$p_{41\|4}$	$p_{43\|4}$	$p_{44\|4}$

Notes: Elaboration by the Author based on Rey (2001)

Table 2 can be used to test the negative or positive influence of geographic neighbours in a region. In this case, dividing the cities in four size classes (small, medium-small, medium-large and large), for example. If we want to know the effect of medium-large sized neighbours on the transition to move up or down of a city, we analyze the matrix elements in the third conditional, where the spatial lag is medium-large. Per instance, the $p_{34|3}$ element stands for the possibility of a region in the medium-large class to move upwards given that its neighbours are in medium-large class.

Furthermore, it is possible to know the influence of spatial dependence on the transition probability comparing the elements of the traditional transition matrix with the elements of the spatial Markov matrix. For example, if $p_{34} > p_{34|3}$, then the probability of an upward movement in the classification of a city in the medium-large class is higher than the probability of one in the medium-large class with neighbours in the same class. Generally speaking, if the neighbourhood has no effect on the probability of transition, then the conditional probability is equal to the probability of the traditional Markov matrix

$$p_{ij|1} = p_{ij|2} = \ldots = p_{ij|k} = p_{ij} \quad \forall ij \qquad (13)$$

The main gain in analyzing the dynamics of the spatial conditioning is capturing the influence of the location and thus the influence of the dimensions of the neighbors about the possibilities for mobility of minimum comparable areas within the populational hierarchy. Beyond to providing a more detailed view of the geographic dimension of population distribution, some interesting questions concerning the characteristics of population mobility can be formulated, in analogy with the questions that Rey(2001) has brought forward for the income distribution theme. Some of these: Is MCA's probability of moving up or down the distribution related to the current, or past movements of its neighbors? Is this form of spatial dependence of a similar magnitude for upward as opposed to downward moves in the distribution? These are some of the questions that can be answered by using of a spatial Markov matrix.

3 Results

In order to examine the behavior of the population distribution between the minimum comparable areas covering the entire Brazilian territory, a series of empirical evidences is presented according to the characteristics of the methodologies applied. Besides the data implementation, the following two subsections deal with the estimation of density functions and Zipf's law. Then, techniques based on Markov Chain will be explored, in order to bring up information on the dynamics and evolution, as well as the possibility of spatial dependence on the behavior of population distribution.

3.1 Data Implementation

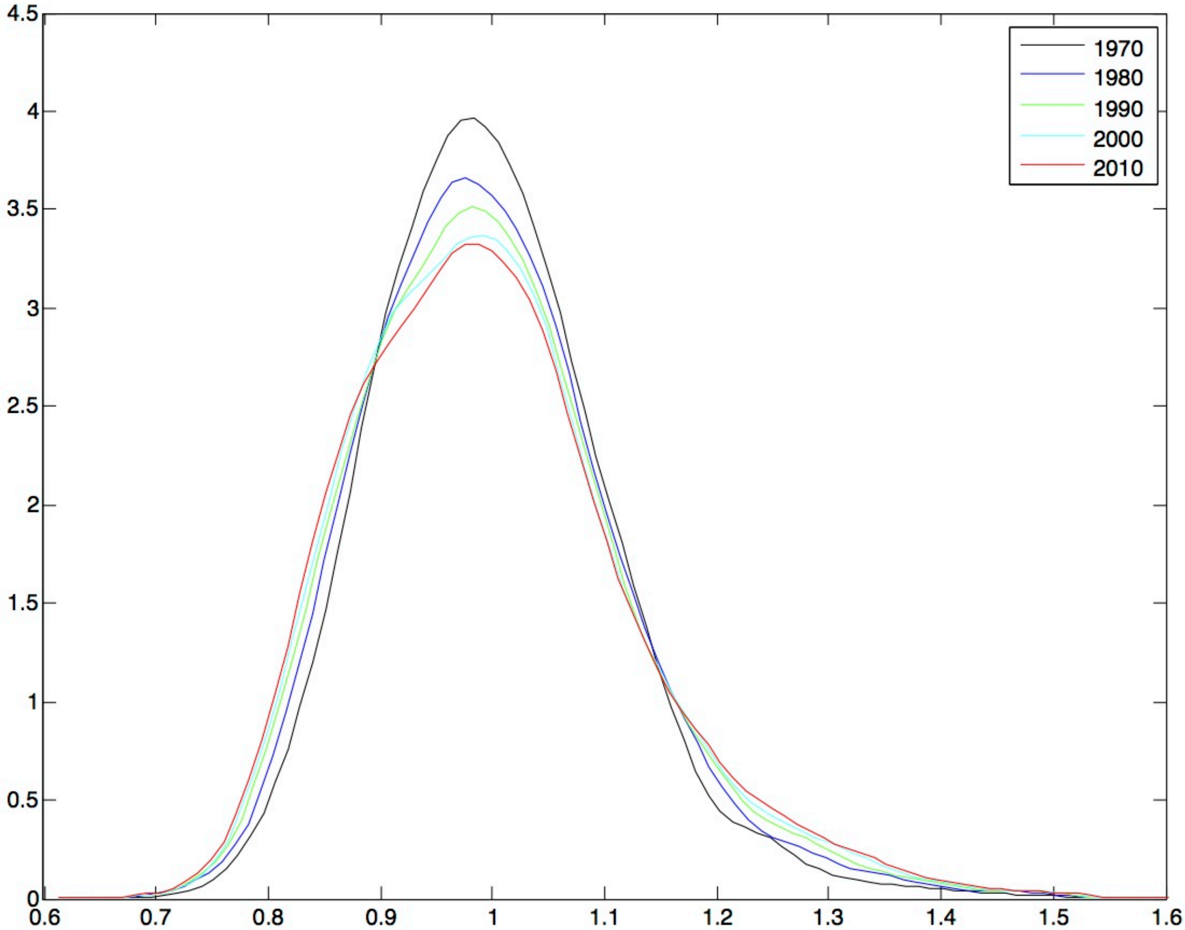
The main source of data is the Brazilian Demographic Census for the years 1970, 1980, 1991, 2000 and 2010 conducted by the Brazilian Institute of Geography and Statistics (IBGE). Although the municipality constitute the smallest unit of observation in political and administrative terms to which is possible to obtain economic and demographic data with coverage of entire Brazilian territory for various periods of time, the intertemporal comparisons in a strictly municipal geographic level become inconsistent with changes in the number, area, and border of municipalities that occurred over the decades. Specifically, over the period 1970-2010, the number of municipalities increases from 3952 to 5565. Therefore, to allow consistent comparisons over time, it is necessary to aggregate these municipalities into broader geographical areas, called Minimum Comparable Areas (MCA). Based on the aggregation of municipalities developed by IPEA (Reis et al., 2010), this study has 3659 MCAs relating to the aggregation of all Brazilian municipalities for each census from 1970 to 2010, covering all territory and avoiding selection bias problem.

3.2 Evolution of Brazilian MCAs size distribution

To investigate the evolution of the Brazilian population distribution shape for 1970-2010 period, a non-parametric normal kernel density with bandwidth value of 0.0245 was estimated for the urban population distribution for each decade[2]. Following Gallo and Chasco (2007), relative population size are considered and the Figure 1, below, shows the distributions of the relative log of population size in 1970, 1980, 1991, 2000 and 2010. The Kernel density plot may be interpreted as the continuous equivalent of a histogram in which the number of intervals has been set to infinity. Adopting a similar strategy to interpretation of Gallo and Chasco, 1 on the horizontal axis indicates Brazilian average MCA size, 1.5 indicates 50% higher than the average, and so on.

[2] The largest bandwidth among the optimum values calculated for each decade was chosen. The optimum values for each decade were calculated using the Matlab function ksdensity.

Figure 1 – Normal Kernel Density Functions for the Population Distribution of MCAs, 1970-2010



Notes: Elaboration by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

Observing the figure, it is remarkable that over the decades there is a loss of concentration of AMCs around the mean. However, this presents a deconcentration rhythm regressive each decade, i.e., the distance between the lines is decreasing. The behavior in the distribution shape may indicate that divergence in population size of MCAs is decreasing, or polarized sizes distribution. In other words, the size of the localities are not converging to the same level, but diverging at a diminishing rate.

Table 3 below shows a statistical summary information, clarifying the Figure 1. Through observation of the columns, obviously the mean is equal to 1 since the values are normalized, the median value decreases over the decades while the standard deviation increases. Clearly, the 2010 distribution is more dispersed around the mean, and this seems to be the trend between 1970 and 2010. Specifically, the distribution became more dispersed in approximately 20% between 1970 and 2010. However, as already pointed out, this divergence is decreasing, between 1970 and 1980 and between 2000 and 2010 the standard deviation growth was 9% to 2.1%, respectively. Still observing the values of the table, with the reduction of the median over the decades we note that the greater dispersion arises mainly from the increased presence of cities above the mean.

Table 3 – Summary Statistics – Relative Population of Brazilian MCAs, 1970 – 2010

Year	Mean	Median	Standart Deviation
1970	1	0.9914	0.1044
1980	1	0.9905	0.1137
1991	1	0.9902	0.1194
2000	1	0.9885	0.1231
2010	1	0.9880	0.1258

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

The evidence of decreasing divergence is similar to others found in the literature, as in Justo (2012) for 431 brazilians MCAs between 1910 and 2010. On the other hand, Trindade and Sartoris (2009) found evidence that the behavior of the Brazilian population distribution already shows a trend of increasing number of municipalities with population below average between 1920 and 2000.

Tables 4a and 4b (below) allow us to contextualize such regional changes in population distribution emphasized in the preceding paragraphs. While the numbers remain fairly stable over the decades, the North region had the largest growth in the number of MCAs above the national median from 2.27% in 1970 to 3.03% in 2010. In this period, as can be seen in Table 4a the North Region's share in the total population increased from 4.43% to 8.32%. On the other hand, the southern region had the highest percentage reduction of MCAs above the median from 9.51% in 1970 to 7.76% in 2010. The participation of the southern region in the total population decreased from 17.71% to 14.36% between 1970 and 2010.

Table 4a – Relative Population Above the Median per Region, 1970 – 2010

Region	Above Median (%)				
	1970	1980	1991	2000	2010
North	2.27	2.71	2.84	2.92	3.03
Northeast	19.24	19.90	19.98	19.98	19.70
Southeast	16.64	16.40	16.67	16.64	16.81
South	9.51	8.72	8.03	7.82	7.76
Midwest	2.32	2.27	2.46	2.62	2.71

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

Table 4b – Participation on the Total Population per Region, 1970 – 2010

Region	Percentage of Population (%)				
	1970	1980	1991	2000	2010
North	4.43	5.56	6.83	7.60	8.32
Northeast	30.18	29.25	28.94	28.12	27.82
Southeast	42.79	43.47	42.73	42.65	42.13
South	17.71	15.99	15.07	14.79	14.36
Midwest	4.89	5.72	6.42	6.85	7.37
Total	93,134,846	119,011,052	146,825,475	169,799,170	190,747,731

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

Table 5 provides information about what percentage of MCAs within each region was higher than the median of the country, which allows us to observe how these changes occurred in the population distribution within each region. As can be seen, the biggest changes in population distribution occurred within the North and Midwest regions. For example, from 143 minimum comparable areas of the North, 58% were above the median in 1970 and this percentage increased to 77.62% in 2010. In the South, there is a reduction of 58.6% to 47.81% in the percentage of MCAs with population above the national median.

Table 5 – MCAs with Population Above the Median per Region, 1970 – 2010

Region	Above Median (%)					
	1970	1980	1991	2000	2010	Total
North	58.04	69.23	72.73	74.83	77.62	143
Northeast	54.24	56.09	56.32	56.32	55.55	1298
Southeast	43.47	42.83	43.54	43.47	43.90	1401
South	58.59	53.70	49.49	48.15	47.81	594
Midwest	38.12	37.22	40.36	43.05	44.39	223

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

The non-parametric normal kernel density functions estimates, as well as the descriptive tables foregoing, have as main role to illustrate the city size distribution patterns. One of the limitations of this information is that it does not permit us to make more precise statements about the size distribution of cities. In this sense, the next topics will bring evidences obtained through the Zipf's law approach, which allows the characterization of the overall evolution of the size distribution of cities, and, in order to bring up information on the dynamics and evolution of the population distribution, techniques based on Markov Chain will also be explored.

3.3 The Rank Size Rule for Brazilian MCAs

Table 6 presents the estimation of rank-size equation (2) for all Brazilian MCAs in each decade using an OLS estimator. In the 1970s, the estimated Pareto coefficient approaches the Zipf's law, with an estimated value of 0.95. According to this rule, city populations among any group of cities at any time are proportional to the inverse of the ranking of their populations in that group (Gallo and Chasco, 2007). In the following decades, this coefficient deviates increasingly from the unit value, reaching 0.77 in 2010. This parametric analysis is consistent with the previously obtained evidence of the decreasing distances between the size population distributions of MCAs seen in Figure 1, acquired through the non-parametric normal kernel density analysis. What seems natural, for a more equitable distribution of population between locations (with less variance; 1970 in relation to 2010, for example), the change in position between the cities become easier due to the fact that most localities have closer sizes.

Table 6 – Classic Rank-Size Equation for log(rank) as dependent variable, 1970 – 2010

Explanatory Variables	OLS									
	1970		1980		1991		2000		2010	
Variable										
Intercept	16.1272	**	15.4354	**	15.1158	**	14.8880	**	14.7113	**
ln Population	-0.9517	**	-0.8698	**	-0.8251	**	-0.7950	**	-0.7713	**
No. Obs.	3659		3659		3659		3659		3659	
R-squared	0.913		0.920		0.925		0.929		0.928	
Log Likelihood	-2458		-2311		-2220		-2110		-2150	
JB stat	3302	**	4373	**	6572	**	9942	**	12381	**

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

This result, from the classical equation of rank-size, is qualitatively consistent with Trindade and Sartoris (2009), Justo (2012) and Moro and Santos (2013). Of these, what the first two have in common with our work is the use of the entire Brazilian territory and both rural and urban populations of the observed units, which allows a comparison with our analysis. Moreover, unlike our analysis, both two studies use a very high level of aggregation, 920 MCAs between 1920 and 2000, and 431 observational units between 1910 and 2010, respectively. In 1970 their estimated coefficients were 0.794 and 0.77, respectively. The only difference between these two estimates and our work is the level of aggregation, but 3659 observational units is much closer to reality for 1970. Therefore, we can see that the high aggregation level used by these two studies lead to results that indicate a higher population concentration than the reality, making their evidence inaccurate[3]. Although Moro and Santos (2013) use municipality as observational unit (more disaggregated than MCAs), they only take the urban population into account, which makes our results quantitatively incomparable.

3.4 Brazilian Population Distributional Dynamics

Unfortunately, it is not possible to have information on the dynamics of the distribution estimating the Zipf's law equations. The approach of the last topic gives no information on the movements of the cities within the distribution. Apart from this, it does not take into account the possibility that these movements are affected by spatial dependence. To assess these empirical issues on the size distribution of Brazilian minimum comparable areas, in the next topics we lead with the estimation of transition probability matrices associated with discrete Markov chains (Kemeny and Snell, 1976), which will make it possible to follow the progress of each group of Brazilian MCAs in time. And then, the spatial dependence is introduced through the analysis of LISA Markov and Spatial Markov Chains, both developed by Rey (2001).

Traditional Markov Chains

In order to observe the behaviour of transition from the relative population levels over time, Table 7 shows the traditional Markov transition probability matrix for four classes of relative population according to quartiles for each decade between 1970 and 2010. The class with the MCAs with smaller populations relative is represented by the first quartile. Therefore, if a MCA is in the first quartile class means that it is among the 25% smaller in terms of relative population, and if the MCA is inserted in the fourth quartile means that 25% is among the largest in terms of relative population.

Table 7 – Markov Transition Probability Matrix for Brazilian MCAs population, 1970-2010

t_i	t_i+1			
	1	2	3	4
1	0.9208	0.0781	0.0005	0.0005
2	0.0787	0.8249	0.0951	0.0014
3	0.0005	0.0971	0.8356	0.0667
4	0.0000	0.0000	0.0686	0.9314

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

From Table 7, several points can be observed. Firstly, the transitions probabilities on the main diagonal are relatively high. If the MCA is in the ith class, the probability of being in

[3] In 1970 the Brazilian territory was divided into 3952 municipalities.

the same class the decade after is at least 82.49% and at most up 93.14%. Specifically, the probability that the MCA in the second quartile remain in this class in the next period is 82.49%. The high probabilities on the main diagonal show a low interclass mobility, a high persistence of MCAs to stay in their own class from one decade to another over the whole period. However, since these probabilities are not exactly equal to 1, we have the possibility to analyse how the MCAs in each cell move to other cells. Secondly, the probability to continue in the initial state, given by the diagonal elements, is higher in the two extreme classes. In particular, the largest and smallest MCAs have less probability of moving to another categories, i.e., these localities have less interclass mobility than the medium-size cities. Since the elements of main diagonal do not assume the value 1, so there is no possibility of parallel or uniform growth between MCAs. This result is an evidence that population distribution structure suffered changes during the period 1970-2010.

Continuing with the reading of Table 7, we realize that the non-diagonal elements are extremely smaller than elements in the main diagonal. Nevertheless, during 1970 to 2010, the medium classes (2 and 3) have more probability of inter-class mobility than extreme classes (1 and 4), the biggest transition probability among different classes is 9.71% which occurs from third to second quartile; next is the second to third class moving, 9.51%. That is, the largest flows occur between the MCAs that are in the second and third classes. This evidence, together with high persistence of both largest and smallest MCAs to stay in the initial class, highlights the major role of medium-size localities in the processes of urban agglomeration that occurred in Brazil during the last 40 years. This evidence is in agreement with Andrade and Serra (2001), as they assert that Brazilian population is undergoing a process of polarization reversal, in which the medium-sized cities play a decisive role in an automatic decentralization of economic activities. In addition, the probabilities of MCAs move up or down more than two steps are extremely small.

The first mean passage time indicates the expected time for a locality to move from class i to class j for the first time. To determine the speed with which the urban municipalities move within the distribution, Table 8 displays the mean first passage time matrix for relative population based on equation (8)[4].

Table 8 – First Mean Passage Time Matrix in Decades for Brazilian MCAs population, 1970-2010

t_i	t_i+1			
	1	2	3	4
1	4.00	13.00	33.16	75.36
2	37.47	4.00	20.72	63.31
3	57.47	20.28	4.00	43.73
4	72.05	34.86	14.58	4.00

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

On average, the number of years to reach any class other than the original is relatively high: the shortest time passage is 13 decades and the longest is 75.36 decades. As expected, more distant classes take longer to be reached. For example, for a MCA that was originally in class 1 achieve the class 3, it takes on average 33.6 decades. The faster declines in the 3 and 4 classes (20.28 and 14.58 decades, respectively) may indicate that localities in these classes are more likely to lose relative population. This evidence suggests a general progressive suburbanization process in which big cities stop to grow, favouring the progressive appearance of smaller population cores (Gallo and Chasco, 2007).

4 $M_P = (I_K - Z + ee'Z_{dg})D$

The ergodic distribution can be interpreted as the long-run equilibrium in the distribution of relative population of the MCAs. As stated by Gallo and Chasco (2007), given a regular transition matrix, with the passage of many periods, there will be a time when the distribution of urban municipalities will not change any more: that is the ergodic or limit distribution. As the population relative discretization was made from the quartiles, the ergodic distribution naturally will be similar to the initial distribution of classes (25% of MCAs in each class) and does not bring interesting results to be interpreted.

A consideration about limitations of traditional Markov chains to study the dynamics of cities is that they do not capture the spatial dependence that may exist between the studied observational units. In order to take into account the possibility of spatial dependence in population distribution dynamics of Brazilian minimum comparable areas, we introduce in the following topics the spatial dependence through the analysis of LISA Markov, that integrates the local indicators of spatial association into a dynamic framework based on Markov chains, and Spatial Markov Chains, that extends the transition probabilities from traditional Markov chains to be conditioned on the initial relative population class of its spatial lag. Both approaches were developed by Rey (2001).

LISA Markov

Table 9 below summarizes the spatial transitions using the same classification system proposed by Rey (2001). The LISA Markov matrix was estimated for decennial time interval: between 1970 and 1980, 1980 and 1991, 1991 and 2000, and between 2000 and 2010. The main motivation is the distinction between the decades over the last 40 years of changes in the population configuration of the country. Performing the estimation with LISA decennial intervals give us an idea of how the changes on the behavior of MCA's population distribution in relation to their respective spatial lags have occurred in every period. As can be seen, the transition probability of Type 0 is high; there is a higher probability of minimum comparable areas and their spatial lag to remain in the same classification. Additionally, this probability increases each decade indicating a low mobility between the classes. Specifically, between 1970 and 1980 the probability of transition Type 0 was 86.89%, and between 2000 and 2010 this probability increases to 96.91%. This evidence of stability in the population distribution behavior over time corroborate with the normal non-parametric kernel density functions estimates (Figure 1) and with the summary statistics (Table 3).

Table 9 – LISA Spatial Transitions (Decennial) for Brazilian MCAs population, 1970 – 2010

Interval	Type of Transition					Cohesion	Flux
	Type 0	I	II	IIIA	IIIB		
1970 to 1980	0.8986	0.0443	0.0533	0.0038	0.0000	0.9024	0.1014
	UP	0.0139	0.0128	0.0003	-		
	DOWN	0.0303	0.0404	0.0036	-		
1980 to 1991	0.9322	0.0295	0.0358	0.0019	0.0005	0.9341	0.0678
	UP	0.0098	0.0123	0.0005	-		
	DOWN	0.0197	0.0235	0.0014	-		
1991 to 2000	0.9516	0.0191	0.0292	0.0000	0.0000	0.9516	0.0484
	UP	0.0096	0.0093	0.0000	-		
	DOWN	0.0096	0.0200	0.0000	-		
2000 to 2010	0.9691	0.0120	0.0189	0.0000	0.0000	0.9691	0.0309
	UP	0.0046	0.0063	0.0000	-		
	DOWN	0.0074	0.0126	0.0000	-		

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

Others interesting results can be obtained by Type II transitions, which involves a transition of only the neighbours in relative space, but the locality in question remains in the previous state. The probability of this transition type, as expected, decreases with each decade. As in the previous analysis, there is a greater probability of downward movement, i.e. higher probabilities of most populous neighbours become less populated.

We can investigate further highly populated communities that propelled the neighbours, i.e., MCAs that were populated above average and had less populous neighbours, while in the following period the neighbours became highly populated. In the interval between 1970 and 1980, we identified 14 MCAs (48 municipalities) that played this role. These minimum comparable areas are equivalent to the current territory of 48 municipalities, three of these are located in Pará state (North), 3 in Pernambuco and 3 in Bahia (Northeast), 6 in São Paulo state (Southeast), 3 in Paraná state (South), and finally, 27 municipalities in state of Mato Grosso and one in the Federal District (Midwest). Only 3 (Curitiba, Brasília and Cuiabá) of these 48 municipalities are capitals of their respective states. Between 1980 and 1991, 13 MCAs (25 municipalities) played this role, ten of these municipalities are located in Pará state (North); 1 in Rio Grande do Norte, 1 in Paraíba and 6 in Bahia (Northeast); 2 in Minas Gerais and 3 in São Paulo state (Southeast); and 2 in Santa Catarina state (South). Of these 25 municipalities, only 4 are capitals (Belém, Natal, João Pessoa and Salvador). Between 1991 and 2000, only 4 MCAs (5 municipalities) played this role, 1 in Sergipe (Northeast); 2 in Minas Gerais and 1 in São Paulo state (Southeast); and 1 in Santa Catarina (South). At this period, Sergipe's state capital (Aracaju) played this role of highly populated community that propelled the neighbours. Finally, between 2000 and 2010, also only 4 MCAs (8 municipalities) played this role, all of them in Northeast, 4 in Maranhão state (including the capital São Luís) and 4 in Ceara state[5]. Probably, these municipalities boosted population growth in theirs neighbourhoods. To investigate more deeply these municipalities is an interesting suggestion for future research. In relation to Type I transition, which occurs when only the locality moves, but its neighbours remain in the same category, this is less likely than Type II and is decreasing over the decades. Regarding the probability of Type IIIa, which occurs when both MCA and neighbors move in the same direction in the distribution, the results indicate that the probability of such transition became null between the decades of 1991 and 2000 and between 2000 and 2010. That is, in the last 20 years there were no transitions of minimum comparable areas together with their neighborhood within the urban hierarchy. Finally, the cohesion and flow measurements are decreasing over time. In other words, over the time the probability of the MCAs move in the same direction of its spatial lag between different classes decreases, and the flux measurement indicates that there is a decreasing instability in behaviour of MCAs relative to its neighbours in the population distribution, as already evidenced.

Spatial Markov Chains

From the traditional Markov matrix, Rey (2001) suggests an extending modification, so that the transition probabilities are conditioned on the initial relative population class of its spatial lag. The spatial Markov matrix, as called by Rey, speaks to the question of whether a locality's transition in the relative population distribution is related to the relative population of its neighbours. As explained by Rey, the spatial provides a great deal of information regarding the transitions of regions and the possible association between the direction and rate of the transitions and the regional context faced by each economy.

---

[5] The listings of municipalities for each decennial interval are in Table A1 in the Appendix.

A spatial Markov transition probability matrix was constructed to analyse the spatial-temporal dynamics of relative population distribution, i.e., considering the possible influence from neighbours on the transition of regions. The standard contiguity neighbours matrix (W) is used for estimation of following spatial Markov matrices, reported in Table 10.

Table 10 – Spatial Markov Transitions Probabilities Matrix for Brazilian MCAs population, 1970-2010

Spatial Lag	t_i	t_i+1			
		1	2	3	4
1	1	0.9546	0.0448	0.0006	0.0000
	2	0.1088	0.8391	0.0522	0.0000
	3	0.0000	0.1088	0.8407	0.0505
	4	0.0000	0.0000	0.0507	0.9493
2	1	0.9096	0.0894	0.0010	0.0000
	2	0.0850	0.8470	0.0670	0.0009
	3	0.0011	0.1105	0.8475	0.0409
	4	0.0000	0.0000	0.0700	0.9300
3	1	0.8866	0.1134	0.0000	0.0000
	2	0.0559	0.8410	0.1030	0.0000
	3	0.0009	0.0986	0.8489	0.0516
	4	0.0000	0.0000	0.0811	0.9189
4	1	0.8628	0.1326	0.0000	0.0047
	2	0.0631	0.7474	0.1836	0.0059
	3	0.0000	0.0761	0.8073	0.1166
	4	0.0000	0.0000	0.0657	0.9343

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

Initially, some evidences can be seen from Table 10. Firstly, spatial background appears to play an important role in the dynamics of relative population distribution. In other words, the neighbours of a MCA have an impact on its transition probabilities over time. If the spatial context did not work, then the four conditional matrices should be the same and are equal to the traditional Markov matrix (Table 7). But in fact it is the opposite. Specifically, a chi-squared test of the difference between each of the spatial conditioned transition submatrices against the overall (a-spatial) transition matrix rejects the null hypothesis that these matrices are equal at 1%[6]. Secondly, different spatial contexts have different effects on transition for regions. Specifically, the probability of upward transitions will increase for MCAs with neighbours in high classes. For example, for a MCA in the first quartile with neighbours in the same class, the probability of moving upward is 4.48%, while if it is adjacent to localities in fourth quartile this probability increases to 13.26%. A similar phenomenon to this occurs also for MCAs originally in classes 2 and 3.

In table 10, as in Table 7, the medium classes (2 and 3) have more probability of inter-class mobility than extreme classes (1 and 4). However, considering the spatial dimension, we can observe that the MCAs grouped into medium classes have a higher probability of a downward transition if your neighbours are in a less populated class (class 1). The opposite happens if the neighbours are the most populous class (class 4). This evidence that the MCAs in the third quartile is more likely to move within the distribution to a larger class when they are near the most populated places, certainly means that there is an overflow of the population of large cities to medium-sized cities. The latter evidence highlights again the major role of

[6] Table A2 in appendix.

medium-size localities in the processes of urban agglomeration that occurred in Brazil during the last 40 years, even considering the spatial context.

Furthermore, as suggested by Rey (2001), it is possible to know the influence of spatial dependence on the transition probability comparing the elements of a traditional transition matrix with the elements of the spatial Markov matrix. For example, ignoring spatial context (Table 7), the probability of a MCA in the third quartile to move down to the second quartile is 9.71%, this probability rises to 10.88% if the neighbours are in the first quartile (less populated class). We can also observe, by comparing the traditional and spatial matrix, that the less populous MCAs that have highly populated neighbours decreases the probability of persistence in the same class distribution. Specifically, ignoring the spatial context, the probabilities of MCAs to in the first and second quartiles are 92.08% and 82.49%, respectively. These probabilities are reduced to 86.28% and 74.74% when high-populated neighbours surround these locations. Additionally, we can explore steady state distribution implied by each estimated conditional transition probability matrix from Table 11, calculated as steady state distribution that was defined in section 2.3. The steady state distribution spatially conditioned is presented in Table 11.

Table 11 – Steady State Distribution for Brazilian MCAs population, 1970 – 2010

Spatial Lag	Population (%)			
	1	2	3	4
1	0.5472	0.2284	0.1125	0.1119
2	0.3225	0.3401	0.2101	0.1273
3	0.1570	0.3130	0.3239	0.2062
4	0.0537	0.1167	0.2939	0.5357

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

The long run distribution for MCAs with neighbouring relatively less populated (class 1) has 54.72% of localities in the first quartile and 11,19% in the fourth quartile, for example. On the other hand, the long run distribution for MCAs with neighbouring relatively high populated (class 4) has just 5.37% of localities in the first quartile and 53.57% in the fourth quartile. According to Gallo and Chasco (2007), concentration of the frequencies in some of the classes, that is, a multimodal limit distribution, may be interpreted as a tendency towards stratification into different convergence clubs. As can be noted in main diagonal, there would be a higher concentration of frequency on a particular class according to the spatial lag that can be an evidence of different convergence clubs according to spatial lag.

Finally, we can determine the speed with which the urban municipalities move within the relative population distribution, conditioned to spatial lag. Table 12 displays the expected time for a locality to move from class i to class j for the first time based on equation (8) for each submatrices in Table 10, i.e., conditional on quartile in population distribution of its neighbours.

As can be seen in Table 12, MCAs with relative population in the first quartile with neighbours in the first quartile return to the first quartile after 1.83 decades, after leaving the first quartile. This time is 18.62 decades for localities in the first quartile with neighbours high populated. Furthermore, MCAs in first class with neighbours also in the first quartile enter the third class 84.52 decades after leaving the first quartile, on average. On the other hand, this time frame falls to 23.31 decades if the spatial lag is in third quartile.

Table 12 – Spatial Markov First Mean Passage Time in Decade for Brazilian MCAs population, 1970 - 2010

Spatial Lag	t_i	t_i+1			
		1	2	3	4
1	1	1.83	22.27	84.52	240.90
	2	17.98	4.38	63.34	219.72
	3	36.31	18.33	8.89	156.38
	4	56.02	38.04	19.71	8.93
2	1	3.10	11.23	39.10	135.51
	2	23.10	2.94	28.35	124.74
	3	37.06	14.30	4.76	97.93
	4	51.35	28.59	14.29	7.86
3	1	6.37	8.82	23.31	70.82
	2	47.37	3.20	14.49	62.00
	3	63.37	16.52	3.09	47.51
	4	75.71	28.86	12.34	4.85
4	1	18.62	9.04	15.78	26.86
	2	126.64	8.57	8.26	20.26
	3	163.11	36.46	3.40	13.19
	4	178.33	51.68	15.22	1.87

Notes: Estimates by the Author. Database from Demographic Census for 1970, 1980, 1991, 2000 and 2010.

4. Conclusions

The objective of this paper was to explore the behavior of the population size distribution of Brazilian Minimum Comparable Areas (MCAs) covering all Brazilian territory between 1970 and 2010, revisiting the traditional rank-size rule and Markov chain approaches. In order to bring up more accurate information on the dynamics and evolution of the population distribution, the spatial dependence is introduced through the analysis of LISA Markov and Spatial Markov Chains, both developed by Rey (2001). The second and main objective of the present thesis was to model population growth dynamics of Brazilian Minimum Comparable Areas (MCAs) in order to assess the determinants of population growth of these units between 1970 and 2010 and examine the existence and magnitude of spatial interaction and spatial spillover effects associated with these determinants.

Initially, the non-parametric normal kernel density functions estimates brought evidence that behavior in the distribution shape may indicate that divergence in population size of MCAs is decreasing. The Zipf's law estimation indicates that the population distribution is, decade by decade, moving away from Pareto law. In other words, this result shows that in the Brazilian case, over time, less and less the ranking of cities is influenced by its size. In the estimation of quadratic rank-size equation, the curvature presents downward concavity; there is a negative correlation between ranking variation and size.

The traditional Markov chain approach brings as main evidence the high probabilities on the main diagonal indicating a low interclass mobility, high-persistence of MCAs to stay in their own class size from one decade to another over the whole period. As suggested by Rey (2001), a spatial Markov transition probability matrix was constructed to analyse the spatial-temporal dynamics of relative population distribution, i.e., considering the possible influence from neighbours on the transition of regions. The results bring evidence that different spatial contexts have different effects on transition for regions. Specifically, the probability of upward transitions will increase for MCAs with neighbours in high classes. Another interesting results is that the MCAs grouped into medium classes have a higher probability of a downward transition if their neighbours are in a less populated class (class 1). The opposite happens if the neighbours are the most populous class (class 4). This evidence highlights again the major role of medium-size localities in the processes of urban agglomeration that occurred in Brazil during the last 40 years, even considering the spatial context. In relation to the LISA Markov approach, we found evidence of stability in the population distribution behavior over time that corroborate with the normal non-parametric kernel density functions estimates.

Additionally, we investigated further highly populated communities that propelled the neighbours (MCAs that were populated above average and had less populous neighbours, while in the following period the neighbours became highly populated). It was identified that some municipalities, mainly in the north and northeast have played this role in the past 40 years, including some capitals of their respective states. Investigate more deeply the municipalities with this feature is an interesting suggestion for future research.

References

AMEMIYA, T. Advanced econometrics. Harvard University Press, Cambridge, 1985.

ANDRADE, T.A.; SERRA, R.V. O Desempenho das cidades medias no crescimento populacional brasileiro no period 1970/2000. Cidades Médias Brasileiras, IPEA, 2001.

ANSELIN, L. Local Indicators of Spatial Association – LISA. Geographical Analysis, 27, 1995.

ANSELIN, L. Spatial externalities, spatial multipliers, and spatial econometrics. International Regional Science Review, v.26, pp.153-166, 2003.

BLACK, D.; HENDERSON, V. Urban evolution in the USA. Journal of Economic Geography, v.3, pp. 343-372, 2003.

DOBKINS, L.H.; IOANNIDES, Y.M. "Dynamic Evolution of the U.S. City Size Distribution," Discussion Papers Series, Department of Economics, Tufts University 9916, Department of Economics, Tufts University, 1999.

DURANTON, G. Some foundations for Zipf's law: product proliferation and local spillovers. Reg Science and Urban Economics, v.36, pp. 542–563, 2006.

GABAIX, X. Zipf's Law for Cities: An Explanation. The Quarterly Journal of Economics, vol. 114, n. 3, 1999.

GABAIX, X.; IOANNIDES, Y. M. The evolution of city size distributions, Handbook of Regional and Urban Economics, in: J. V. Henderson & J. F. Thisse (ed.), Handbook of Regional and Urban Economics, edition 1, volume 4, chapter 53, pages 2341-2378, Elsevier, 2004.

GAN, L; LI, D.; SONG, S. Is the Zipf law spurious in explaining city-size distributions?. Economics Letters, v. 92, n. 2, p. 256-262, 2006.

GALLO, J.L.; CHASCO, C. Spatial Analysis of Urban in Spain, 1900-2001. Empirical Economics – Springer, 2007.

HAMILTON, J.D. Time series analysis. Princeton University Press, Princeton, 1994.

IBGE. Sinópse do Censo Demográfico 2010 – Brasil. Available at: http://www.censo2010.ibge.gov.br/sinopse/, 2011.

KEMENY, J.; SNELL, L. Finite Markov Chains. Springer, 1976.

MONASTERIO, L. M.. Dinâmica Espacial da População e dos Trabalhadores na Manufatura no Brasil (1872-1920). In: VII Encontro Nacional da Associação Brasileira de Estudos

Regionais, 2009, São Paulo. Anais do VII Encontro Nacional da Associação Brasileira de Estudos Regionais, 2009

MORO, S.; SANTOS, R.J. The Characteristics and Evolution of the Brazilian Spatial Urban System: Empirical Evidences for the Long-Run, 1970-2010. Texto Para Discussão 474, Cedeplar – UFMG, 2013.

OLIVEIRA, C.A. Evolução da Distribuição do Tamanho das Cidades Brasileiras: 1936-2000. Anais da ETENE, Banco do Nordeste, 2004a.

PARETO, V. Cours d'Economie Politique. Geneva, Switzerland, 1896.

REIS, E.; PIMENTEL, M.; ALVARENGA, A.I. Áreas Mínimas Comparáveis Para os Períodos Intercensitários de 1872 a 2000, IPEA, 2010.

REY, S.J. Spatial Empirics for Economic Growth and Convergence. Geographical Analisys, vol. 33, n. 3, 2001.

SIMON, C.J. Human Capital and Metropolitan Employment Growth. Journal of Urban Economics, v. 43, 1998.

ZIPF, G.K. Human behavior and the principle of least effort. Addison-Wesley, Cambridge, 1949.

Appendix

Table A1 – High-Populated Municipalities that Propelled the Neighbours

1970-1980		1980-1991		1991-2000		2000-2010	
Municipality	FU	Municipality	FU	Municipality	FU	Municipality	FU
Terra Santa	PA	Belém	PA	Aracaju	SE	Presidente Sarney	MA
Faro	PA	Redenção	PA	Marataizes	MG	Pinheiro	MA
Oriximiná	PA	São Geraldo do Araguaia	PA	Itapemirim	MG	Pedro do Rosário	MA
Santa Filomena	PE	Conceição do Araguaia	PA	Jundiaí	SP	São Luís	MA
Ouricuri	PE	Xinguara	PA	Itajaí	SC	Itapagé	CE
Santa Cruz	PE	Rio Maria	PA			Tejuçuoca	CE
Mata de São João	BA	Piçarra	PA			Catunda	CE
Camaçari	BA	Floresta do Araguaia	PA			Santa Quitéria	CE
Dias d'Ávila	BA	Sapucaia	PA				
Araguari	ES	Pau d'Arco	PA				
Coronel Fabriciano	ES	Natal	RN				
Campinas	SP	João Pessoa	PB				
Itapevi	SP	Conceição do Coité	BA				
Saltinho	SP	Monte Santo	BA				
Piracicaba	SP	Madre de Deus	BA				
São Roque	SP	Salvador	BA				
Araçariguama	SP	Senhor do Bonfim	BA				
Curitiba	PR	Andorinha	BA				
Foz do Iguaçu	PR	Uberaba	ES				
Santa Terezinha de Itaipu	PR	Delta	ES				
Poxoréo	MT	Atibaia	SP				
Cana Brava do Norte	MT	Itapira	SP				
Campo Verde	MT	Sorocaba	SP				
Campinápolis	MT	Balneário Arroio do Silva	SC				
Cuiabá	MT	Araranguá	SC				
Jaciara	MT						
Novo São Joaquim	MT						
Juscimeira	MT						
Ribeirão Cascalheira	MT						
Alto Boa Vista	MT						
São José do Xingu	MT						
Canarana	MT						

Confresa	MT
Santa Terezinha	MT
Querência	MT
Dom Aquino	MT
Primavera do Leste	MT
Luciara	MT
Barra do Garças	MT
Nova Xavantina	MT
São Félix do Araguaia	MT
São Pedro da Cipa	MT
Cocalinho	MT
Porto Alegre do Norte	MT
Araguaiana	MT
Água Boa	MT
Vila Rica	MT
Brasília	DF

Notes: Elaboration by the Author. FU = Federal Unit (State).

Table A2 - Test of difference between the conditional transition matrix against the overall transition matrix

Spatial Conditioned Transitions Submatrices	Chi2	P-value	Degrees of freedon
1	61.8955	0.0000	9
2	24.1559	0.0041	9
3	26.3733	0.0018	9
4	148.6507	0.0000	9

Notes: Estimates by the Author.