REGIONAL SOCIO-ECONOMIC TRANSFORMATION IN BRAZIL

A thesis
submitted in fulfilment
of the requirements for the degree
of
Doctor of Philosophy in Economics
at
The University of Waikato
by

VALENTE JOSÉ MATLABA

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

2012

Abstract

The regional income disparities in Brazil are well-known. Since the 1930s, such income disparities have declined only slightly. This thesis combines traditional economic theory with insights from regional science and economic geography to explain the development pattern in Brazil throughout the 20th century, using a wide range of data sets. It contributes to the consolidation of the field of New Economic Geography because some tools employed in this thesis have not yet had widespread use in the literature. The thesis also brings new insights for the understanding of Brazil's development process. The key finding of the thesis is that there has been almost time-invariant spatial autocorrelation in Brazil's growth process that impedes the lagging regions from catching up. The reason for this is that there is a clear cluster of contiguous rich regions (Southeast and South) – i.e. the core - characterised not only by high real income levels and high market potential, but also by the fact that they have the largest markets and are the platform of the global economy in Brazil. In contrast, there is another cluster of contiguous poor regions (North and Northeast) – the periphery - that has low real income levels, low market potential and low market access. The agglomeration of population and economic activity explains the observed concentration pattern. Although there were some efforts made through regional development policy to narrow the gaps amongst the regions, the agglomeration forces are very strong in Brazil. The creation of Brasília did not offset these agglomeration forces, partially because place-based policies matter. The creation of Brasilia obviously had major implications for Brasilia itself, but did not offset the agglomeration forces that led to the dominance of São Paulo. Similarly, investment in other lagging regions may not offset the advantages of São Paulo and Rio de Janeiro, but they may improve economic conditions in the lagging regions themselves if income transfers or subsidies are done for efficient industries.

Note on Publications

Working and conference papers have been published from this thesis as follows.

Working Paper:

Matlaba, V. J., Holmes, M., McCann, P., & Poot, J. (2011). *A century of the evolution of the urban system in Brazil*. Working Paper #12/11, Department of Economics, Waikato Management School, University of Waikato, Hamilton, New Zealand. Available at: http://econpapers.repec.org/paper/waieconwp/

Conference Papers:

Matlaba, V. J., Holmes, M., McCann, P., & Poot, J. (2011). *Classic and spatial shift-share analysis of state-level employment change in Brazil*. Paper presented at the Young Researchers in Economic Geography Workshop (28 October), Faculty of Spatial Sciences, University of Groningen, The Netherlands, 42pgs.

Matlaba, V. J., Holmes, M., McCann, P., & Poot, J. (2010). *A century of the evolution of the urban system in Brazil*. Paper presented at the 57th Annual North American Meetings of the Regional Science Association International (10-13, November), Denver, Colorado, the United States, 32pgs.

Matlaba, V. J., Holmes, M., McCann, P., & Poot, J. (2010). *Agglomeration externalities and 1981-2006 regional growth in Brazil.* Paper presented at the 57th Annual North American Meetings of the Regional Science Association International (10-13 November), Denver, Colorado, the United States, 32pgs.

Matlaba, V. J., Holmes, M., McCann, P., & Poot, J. (2010). *A century of the evolution of the urban system in Brazil*. Paper presented at the 14th Annual Waikato Management School Student Research Conference (18 October), University of Waikato, Hamilton, New Zealand, 25pgs.

Matlaba, V. J., Holmes, M., McCann, P., & Poot, J. (2010). *Employment growth in Brazil (1981-2006): A Glaeser et al.'s model approach.* Paper presented at the 6th Australasian Development Economics Workshop (10-11 June), University of Western Sydney, Parramatta, Australia, 36pgs.

Acknowledgements

To reach this stage I worked hard, but also I received crucial support from some people and institutions. I thank my parents, Lisete and José, for their care and encouragement throughout the process of my education. I also thank my sisters, Inês, Adélia, and Orpa; my brothers, Viriato and Vicente; my sister-in-law, Suzy; and all of my nieces and nephews, especially Dayana, Denise, and Rosma for their cheers.

Thanks a million to my PhD supervisors, Professor Mark Holmes, Professor Philip McCann, Professor Jacques Poot, and Professor Arthur Grimes, for their encouragement and outstanding supervision and for their timeless trust that I should reach this stage. I am sure that other students would like to have had the privilege that I had of being supervised by such a great supervisory panel. I will always be indebted to them for everything that they have done for me while living and studying in New Zealand, for their unconditional support, assistance and guidance that went beyond PhD supervision. They are amongst the greatest human beings I ever met.

I thank my examiners, Professor Viv B Hall and Professor Carlos Roberto Azzoni, for their very helpful comments on my thesis.

I thank Waikato Management School at the University of Waikato for financial support for part of my PhD studies and for all the facilities they provided. I was hugely fortunate to study at this institution given the excellent studying environment. To have my own office and computer available for me 24/7 to carry out my PhD research was crucial for my successful studies. The quality of life that Hamilton city offers is also acknowledged. I also thank WMS Department of Economics staff, especially Chairperson Dr Dan Marsh, and the department administrators, Ms Maria Fitzgerald and Ms Leonie Pope, for their quick responses to my student needs.

I thank my former professors at the University of São Paulo, especially Professor Maria Cristina Cacciamali, Professor Basília Aguirre, and Professor Gilberto Tadeu Lima for all their support.

I thank my friends and colleagues, especially Tinh Thanh Doan and Peck Leong Tan for the great time we had together in Sydney and, alongside Bentry Mkwara, for all their support and discussions at WMS and their invaluable comments on my work. I also thank Alessandra Faggian for her encouragement; Cláudio Dimande for our discussions on various subjects (I remember the great time we had together in São Paulo city and our decision at that time in undergraduate education that we had to study until PhD!); Steven Bond-Smith for our discussions and the great time we had together in Groningen; Rodrigo Alves Teixeira and Nadiejda Watanabe Deane Sá for our various discussions and their support since when we were undergraduate colleagues at University of São Paulo. I thank Bill Cochrane, Glen Stichbury, Yi An, Ernesto Mandlate, Wladimir Teixeira, Anderson Stanciole, Rodrigo Moreno Serra, and Pedro Pota for their support.

I thank all University of Waikato Library staff for their attention to library-related student needs.

I thank the staff and students at the Faculty of Spatial Sciences at the University of Groningen in the Netherlands for their hospitality while I was a visiting PhD student for two months in 2011. The opportunity to participate in two excellent excursions within the Netherlands was also appreciated.

I thank the staff of the School of Public Policy at the George Mason University in the United States, especially Professor Kingsley E. Haynes for their hospitality while I was a visiting PhD student in November 2010.

Finally, I am very grateful to Motu Economic and Social Policy Research for funding assistance (Foundation for Research, Science and Technology Grant MOTU-0601 Infrastructure).

Table of Contents

Appendices ……………………………………………………………….......... 201

List of Tables

List of Figures

CHAPTER 1 – INTRODUCTION

1.1. Problem statement

Brazil is a Federation of 27 States[1] located in South America. It shares a border with most of this region's countries, namely French Guiana, Suriname, Guyana, Venezuela, Colombia and Peru in the north; Bolivia and Paraguay to the West and Argentina and Uruguay in the South. At its East Coast lies the Atlantic Ocean. With a total area of about 8.5 million square kilometres, in 2010, its population was around 191 million inhabitants and its real GDP was approximately BR$ 3,678 billion (in BR$ of 2000)[2].

Figure 1.1: Brazil in South America's Map

Source: www.infoplease.com/atlas/southamerica.html

[1] When considering Distrito Federal as a state.
[2] The average commercial exchange rate BR$/US$ in 2010 was about R$ 1.76 per US$. This leads to a 2010 Real GDP of around US$ 2,090 billion (see IPEA). GDP data and population data are from IPEA (Institute of Applied Economic Research) and IBGE (Brazilian Institute of Geography and Statistics).

Brazil is well-known for having regional disparities among the highest in the world (Silveira-Neto & Azzoni, 2006, and the references therein). Regional inequality is important to study given that it limits growth potential and well-being across regions. This thesis explains the problem of uneven regional economic development in Brazil. The understanding of this problem is important, especially for Brazil, a country where regional concentration of income and economic activity is almost timeless. For example, the share of income and population (as a percentage of the nation's GDP and nation's population) for the two most developed regions were as follows. The southeast region and the south region accounted for around 63 percent and 56 percent, and 16 percent and 17 percent of the nation's GDP in 1939 and 2008, respectively[3]. With respect to the population, these two regions were quite stable as well, varying from around 44 percent to 42 percent, and from 12 percent to 15 percent from 1940 to 2008, respectively.

The changes for the other regions were also only slight[4]. The north and northeast regions accounted for 3 percent and 5 percent, and 17 percent and 13 percent of the nation's GDP in 1939 and 2008, respectively. Regarding population, the shares of these two regions doubled from 4 percent to 8 percent and declined from 35 percent to 28 percent from 1940 to 2008, respectively. When considering the metropolitan area data, the situation of a relatively stable pattern of concentration is quite similar: only two metropolitan areas, São Paulo and Rio de Janeiro, together, concentrated around 32 percent in 1939 and 26 percent in 2008 of the nation's GDP. However, there has been a twofold rampant concentration of population in these two metropolitan areas for which, when their populations are considered together as a percentage of the national population, there has been an increase from about 9 percent in 1940 to 19 percent in 2008.

Regarding differences in living standards, over time the ratios of the regions' per capita income to the nation's are also quite stable. These ratios varied for regions from 1939 to 2008 as follows: 1.39 to 1.32 for the southeast; 1.26 to 1.14 for the south; 0.53 to 0.64 for north; and 0.48 to 0.47 for the northeast. At the

[3] See Brazil's regions in Figure 2.1, section 2.2 below.
[4] The only region that appears to have had a relatively large change is Middle-west (also called Centre-west). This region's shares varied from 2 percent to 9 percent from 1939 to 2008 for GDP; from about 3 percent to 7 percent from 1940 to 2008 for population; and from around 0.80 to 1.27 from 1939 to 2008 for the ratio between this region's per capita GDP and the nation's. These changes appear to be correlated with the Brasilia creation in the late 1950s. The economic effects of the creation of Brasilia will be analysed in detail in Chapter 6.

metropolitan area level, these ratios varied from 3.88 to 1.78 for São Paulo and from 3.51 to 0.88 for Rio de Janeiro. Overall, the disparities in income, population and living standards in about 1939 resemble those observed in 2008 for these three indicators. Therefore, from 1939 to 2008, the changes in regional disparity for those three key socio-economic variables seem to be insignificant, especially for data at the regional and state levels.

Moreover, analysing the 20th century, Marcelo de Paiva Abreu states that, "(…) overall, despite its structural changes over the 20th century, Brazil's relative position [in the world] remains unchanged" (IBGE, 2006, p. 356). However, Brazil's position has improved hugely since 2000: in contrast to a real GDP growth rate of 17 percent only from 1991 to 2000, from 2000 to 2008 Brazil's real GDP growth rate was 35 percent (see Chapter 2 of this thesis, Table 2.1). These growth rates allowed Brazil's GDP to become among the highest in the world and Brazil is expected to overtake Britain in 2011 as the world's sixth largest economy[5]. Considering that better regional equity in Brazil would increase growth rates for the regions lagging behind, narrow the income gap between the regions as a consequence and also improve overall well-being, the outlined evidence is also consistent with findings of econometric models. Consequently, studies to explain regional disparities in Brazil, considering various sub-periods since 1939, have demonstrated a slow, rather than fast, income convergence nationwide or just within the regions (see, for instance, Ferreira, 2000; Magalhães, Hewings & Azzoni, 2005; Silveira-Neto & Azzoni, 2006). This pattern is consistent with that observed in 'many arenas of the global economy' (McCann, 2008, p. 356; Poot, 2004, p. 6).

1.2. Background and Research Objectives

The study of regional income disparity in Brazil has typically taken four approaches. These approaches mostly use data at the macro-regional and state levels. The first approach, which is essentially descriptive, finds that, for the whole period considered, the regional disparities are quite stable because changes

[5] See, for instance, http://www.economist.com/blogs/dailychart/2011/11/focus.

in the regions' shares in the nation's GDP are small (Rolim, 2008; Gomes, 2002)[6]. Concluding the analysis of development in Brazil, Gomes (2002, p. 25) states:

> In the pleasant side, there has been convergence of per capita GDPs among the Brazilian states, although this has not been true for every subperiod between 1947 and 1999. As to the official regions (North, Northeast, Center-West, Southeast, and South), the trend is less clear, the more so because of troublesome Northeast. All in all, in the last half-century, growth at a reasonable speed has been the rule for states and regions. Social conditions have also improved everywhere. The bad news is that disparities, be they economic or social, among regions, states, and municipalities remain great. The worn-out expression "the two Brazils" still holds, as anyone can testify looking at the map with p. c. GDPs of the country's municipalities: a divide between a rich South and a poor North is easily seen. And, worse than all this, in the last fifteen years or so, convergence among state p.c. GDPs has stopped, if not been reversed.

The second approach employs traditional econometric models, namely a neoclassical growth model (Barro & Sala-i-Martin, 1992, 2004). First, it finds that while inequality within the (macro) regions falls, between the regions it increases. Secondly, absolute convergence in income was in place until the 1980s'; since then, the convergence process is stable (Azzoni, 1997b; Ferreira, 2000). As Ferreira (2000, pp. 484-485) points out,

> (…) In the 1970s, a decade of generally high rates of per capita income growth, convergence was restricted mainly to the states located in the south-east, south and centre west regions (only in five of the 15 states located in the north and north-east, the poorest regions, the per capita income gap with respect to the national average was reduced in this period). After 1980, simultaneously to the dramatic reduction in growth rates, the speed of convergence among the rich states decelerated, while the poor states, in the north and north-east, started to catch up. As a consequence of these different influences, the estimates of the speed of absolute and conditional convergence moved in opposite directions between the two periods.

The third approach employs Exploratory Spatial Data Analysis (ESDA) without any regression modelling (Mossi, Aroca, Fernández & Azzoni, 2003; Silveira-Neto & Azzoni, 2006; Gondim, Barreto & Carvalho, 2007). Such studies

[6] For example, Rolim (2008, p. 7, Table 1) found that from 1985 to 2004 the highest change in the regional share was that for the southeast region, -4.7 percent (see also Gomes, 2002, p. 10, Figure 2).

found two income clusters in Brazil: a cluster of low income in the northeast and another of high income in the southeast; they also found that states with wealthier neighbours are more likely to grow faster. So, there is no evidence of national convergence. Finally, the fourth approach, which is scarce in the literature, acknowledges the role of space and estimates both spatial and conventional econometric models. This approach finds that after introducing spatial dependence in the analysis of the regional disparity, the convergence rate of income per capita is higher and also that there have been clubs of convergence due to spatial autocorrelation in Brazil's growth process (Magalhães, Hewings & Azzoni, 2005; Silveira-Neto & Azzoni, 2006; Resende, 2011)[7]. Indeed, Magalhães, Hewings & Azzoni (2005, p. 17) note:

> (…) it is possible to infer from the results in hand that, although some convergence among states is taking place, it seems to be more of a regional phenomenon or perhaps some type of club convergence than a global convergence process. States like São Paulo would be a dominant force in one club while the Northeast states would form a second group or club.

Moreover, by means of a multiple spatial scales study, Resende (2011, pp. 650-651) found for the period 1991-2000 that the dispersion of per capita income decreased 13.2 percent for clubs of rich regions and increased 9.7 percent for clubs of poor regions in Brazil. In other words, two conclusions can be drawn from the evidence of these four approaches in the literature. First, that national convergence is unclear and depends on the chosen sub-period of analysis. Secondly, irrespective of the approach used, there is a consensus that the found convergence rate is very small, which makes regional disparities almost timeless and generates a clear pattern of core-periphery (Brakman, Garretsen & van Marrewijk, 2001). The objective of this thesis is to demonstrate the role of second nature geography in the growth of regions in Brazil. In doing so, this study argues that the latter two literature approaches outlined above explain Brazil's regional disparities more than the former two.

[7] These studies also employ a Barro & Sala-i-Martin (1992) growth model, but have a merit of incorporating spatial interactions in their regression modelling.

1.3. Thesis Relevance

The development of regions can be explained either by first nature geography or second nature geography, or by both. Historically, traditional economic theory explains growth and development of regions by “first nature” geography only. The “first nature” geography factors include natural resources, climate, soil quality, natural harbours, and navigable rivers; to these factors are added exogenous factors (e.g. culture, type of management, etc) that lead to differences in technology and institutions across regions (Krugman, 1991a, 1991b, 1998; Brakman, Garretsen & van Marrewijk, 2001; Mellinger, Sachs & Gallup, 2003; Hanson, 2003; Poot, 2004; Bosker, 2008). Many regions lagging behind in Brazil are endowed with most of those “first nature” geography factors, which illustrates that the traditional economic theory fails in solely emphasising “first nature” geography to explain regional disparities.

However, the regional science and economic geography literature also consider the role of second nature geography. The second nature geography factors are a result of interaction between agglomeration forces and diffusion forces within and among the regions. Agglomeration forces include access to large markets, the presence of a large variety of goods and services, and an efficient labour market. Diffusion forces involve congestion, pollution, high competition between goods’ suppliers, and higher prices of immobile factors of regions such as land and buildings (Mellinger, Sachs & Gallup, 2003; Hanson, 2003; Poot, 2004; Bosker, 2008). It is the interplay between (some of) the elements of the two “natures” of geography that explain regions’ development.

This thesis is important in four respects. First, it considers elements of both “first nature” geography and “second nature” geography to understand regional economic disparities in Brazil. Secondly, although some of the tools of analysis employed are considered by the international literature, their use in this thesis for Brazil’s economic history is original. Thirdly, several previous studies clearly pointed out the regional disparities and concentration of economic activities in Brazil, but they do not explain Brazil’s peculiar nature of concentration, i. e. the fact that, irrespective of the performance of the national economy, the gaps between regions essentially do not narrow. Finally, some of the tools used in this thesis (for example, the application of the corrected power law method in the analysis of city size distribution using longitudinal metropolitan area data; spatial shift-share methodology, etc) have only received analytical attention in the international literature and lack empirical application. At most, their use is in an experimental phase and is not yet widespread for empirical investigation among scholars. Thus, it is hoped that this thesis will bring more consensus related to the validity of these techniques in the theoretical debate in the fields of regional science and economic geography.

1.4. Thesis Hypotheses and Research Questions

This study asks the following questions:

(1) Why is there still significant concentration of economic activities in Brazil?

(2) Why, in a span of 70 years, has there been no clear sign of a fall in the disparity of the three key socio-economic indicators as shown in 1.1 above?

The posited answers to these questions, which become this thesis' hypotheses, are:

(1) The essentially time-invariant pattern of concentration of economic activities in Brazil holds because the patterns of spatial interactions among regions have not been changed significantly since the Second World War and earlier;
(2) The economy needs simultaneous shocks (i.e. more place-based policies) to significantly reduce concentration in Brazil;
(3) The mechanisms of the growth process cannot be clearly detected, so, it is difficult to understand how they function in Brazil;

The rich and accessible databases from IPEA and IBGE allow for application of a set of techniques outlined below to test these hypotheses.

1.5. Research Methods and Data Sources

This thesis combines tools of the traditional economic theory with those from the regional science and economic geography literature. From these frameworks, the thesis is able to evaluate and test whether the stability of the spatial / regional concentration holds in Brazil. To test this regularity, a set of tools has been employed. The quantitative methods are as follows: (1) the neoclassical growth and convergence method (Barro & Sala-i-Martin, 1992, 2004; Ozgen, Nijkamp & Poot, 2010), (2) the power law for size of cities (Gabaix, 1999a, 1999b; Overman & Ioannides, 2001; Gabaix & Ioannides, 2003; Ioannides & Overman, 2003;

Anderson & Ge, 2005; Gabaix & Ibragimov, 2006; Soo, 2005, 2007; Córdoba, 2008; Bosker, 2008; The New School, 2010), (3) the classic shift-share methodology (Andrikopoulos, Brox & Carvalho, 1990; Selting & Loveridge, 1994; Ray & Harvey, 1995; Dinc, Haynes & Qiangsheng, 1998; Dinc & Haynes, 1999, 2005; Yasin, Alavi, Sobral & Lisboa, 2010), (4) exploratory spatial data analysis of the shift-share components, Moran's *I* and cluster analysis (Moran, 1950; Cochrane & Poot, 2008) and (5) spatial shift-share (Nazara & Hewings, 2004; Mitchell, Myers & Juniper, 2005), (6) the Seemingly Unrelated Regression (SUR) models (Zellner, 1962) for the levels of GDP, GDP per capita and population and for growth of GDP per capita, as well as spatial GDP per capita growth models (Barro & Sala-i-Martin, 1992, 2004; LeSage & Pace, 2009; Ozgen, Nijkamp & Poot, 2010). The thesis employs quantitative methods and uses secondary data obtained mainly from IPEA and IBGE[8].

1.6. Structure of the thesis

Chapter 2 outlines the regional science and economic geography theories, as opposed to the traditional economic theory. It uses a variety of indicators to illustrate why the former two theories explain the regional disparities in Brazil better than the latter (which has two main literature branches as pointed out in section 1.2). With the support of theoretical interpretation of the evidence, the key finding is that the lagging regions in Brazil are so because their disadvantages in "second nature" geography outweigh their advantages in "first nature" geography. One of the most important elements of economic geography theory related to "second nature" geography is increasing returns to scale that arise in a situation of a market of monopolistic competition due to agglomeration of economic activities (Krugman, 1991a, 1991b). The effects of such agglomeration on regional growth are analysed in the Chapter 3, which employs the Glaeser, Kallal, Scheinkman & Shleifer (1992) model and related seminal work (Henderson, Kuncoro & Turner, 1995; Combes, 2000) to test whether agglomeration externalities, such as those of Marshall-Arrow-Romer (MAR), Porter, and Jacobs (Marshall, 1890; Arrow, 1962; Romer, 1986; Jacobs, 1969; Porter, 1990; De Groot, Poot and Smit, 2009) play a role in regional growth in Brazil. From a range of models' specifications, in

[8] The software packages used in chapters 2 to 6 are Microsoft Excel and Stata, except in chapter 4, in which besides these softwares were also used Geographic Information System (GIS).

line with MAR and Porter theoretical frameworks, the evidence suggests that specialisation is important for regional growth in Brazil.

The concentration of economic activity is positively correlated with market potential (or market scale, or home market effect) (Krugman, 1995). Chapter 4 tests the regularity of market potential over the 20th century for the urban areas of Brazil by the power law (or the rank-size rule), which was originally introduced by Pareto (Gabaix, 1999a, 1999b; Gabaix & Ibragimov, 2006; Eeckhout, 2004, 2009; The New School, 2010). Despite strong political aspects (e.g. dictatorship, 1940-1945; 1964-1984; democracy, 1945-1964; 1989-2008), economic (closed economy, 1964-84; trade liberalisation, 1989-2008), waves of immigration (e.g. the late 1930s) and internal migration (e.g. the creation of Brasilia) that shaped the city size distribution in Brazil, the power law still holds[9]. This stochastic growth process in Brazil was also confirmed by application of panel unit root tests for the whole sample which provided evidence for Gibrat's law from 1984. Chapter 4 finds that although the absolute value of the power parameter is smaller than 1 (one is the level predicted by its special case, Zipf's law), there is convergence to 1, indicating an increasing concentration of population in the 100 largest urban areas in Brazil, which is consistent with the found positive correlation between growth of the 100 largest urban areas and their initial sizes since 1940. The findings from the power law provide new insights for the understanding of concentration of economic activity (or, interregional variation in the market potential). This indicates that, unless a significant exogeneous shock occurs, a large variation in size of cities for a certain size threshold is expected in Brazil.

Chapter 5 uses the traditional shift-share method to analyse growth of employment in regions (states) of Brazil. (For this method's details, see, for example, Dunn, 1960; Esteban-Marquillas, 1972; Arcelus, 1984; Berzeg, 1978, 1984; Haynes & Machunda, 1987; Selting & Loveridge, 1994; Dinc, Haynes & Qiangsheng, 1998; Dinc & Haynes, 1999; Fotopoulos, Kallioras & Petrakos, 2010.) The key finding is that north and centre-west states grew fastest due to better industry-mix [10] and competitive effects. As the regional science and economic geography literature pay particular attention to the effect of spatial interaction on growth of regions, this chapter also applies the standard techniques

[9] This includes the traditional power law and its corrected version.
[10] This is consistent with their lowest averages in Hirschman-Herfindahl Index calculated across sectors within the states from 1981 to 2006.

of the exploratory spatial data analysis (ESDA) such as Moran's *I* and Moran scatterplots (Getis, 1991; Anselin, 1995; Cochrane & Poot, 2008; Le Gallo & Kamarianakis, 2010) for two shift-share components, the industry-mix and competitive effects, both of which show a positive spatial autocorrelation. The analysis is complemented by the most recent development of the shift-share method (Nazara & Hewings, 2004), which directly incorporates the role of neighbour regions for regions' growth, from which it was found that the states that grew fastest usually performed well in the potential spatial spillover effect and in the spatial competitive effect. The latter findings and those from ESDA illustrate the role of geographical location for regional growth in Brazil.

Chapter 6 uses the urban areas defined in Chapter 4 to test for the economic effects of the creation of Brasilia city. Chapter 6 ties up the empirical law confirmed in Chapter 4 (the power law) with growth of regions. Two hypotheses are tested, i) the regional development policy, and ii) the agglomeration forces theory. This thesis test allows for evaluation of whether the effects of agglomeration forces (that cause concentration of income and population in the two largest cities of São Paulo and Rio de Janeiro) confirmed by the power law in Chapter 4 more than offset the effects of large scale regional development policy (the creation of Brasilia) on the pattern of economic activity established before Brasilia's creation (Krugman, 1994; Krugman & Venables, 1995; Skilling, 2001; Crawford, 2004; Poot, 2004; McCann, 2009a). Chapter 6 is a natural experiment as Brasilia's birth was an exogenous shock created by the policy makers to change the pattern of concentration of economic activity from the southeast and south regions as well as from coastal Brazil to the countryside. The creation of Brasilia and the underlining events mark an important change in Brazil's development process: before 1956 Brazil was essentially an agrarian country, but after that year, and along with the development polices (industrialisation, infrastructure development, and urbanisation) that took place under the term of President Juscelino Kubitschek (1956-1961), the country's industries were boosted.

Using seminal neoclassical works in growth and convergence (Barro & Sala-i-Martin, 1992, 2004; Ozgen, Nijkamp & Poot, 2010) and employing different models' specifications, Chapter 6 shows the extent to which the convergence model holds for Brazil. The key findings of this chapter are that even though the convergence model holds (however, again, the convergence rate is small and falls from about 3 percent to 1 percent from 1949 to 2008), Brasilia either had no effect or had a crowding-out effect on the spatial pattern of economic activities (measured by GDP per capita) in Brazil, impeding growth of its neighbour regions. After the 1960s, Rio de Janeiro lost importance, with a recovery in the late 2000s. In 2008, Rio and São Paulo had a similar to each other positive effect on growth of GDP per capita of their neighbour regions. The increased importance of the largest city of São Paulo is confirmed by an increase in the significance level (from 1949 to 1970, and stable since then) of its positive effect on growth of neighbouring regions. This finding, therefore, implies that the effects of agglomeration forces more than offset those of the policy behind Brasilia's creation, something that the policy makers did not expect.

Finally, Chapter 7 summarises the main findings of the thesis. It evaluates the extent to which the regional science and economic geography frameworks better explain the persistence of high concentration of income and of economic activity in Brazil, briefly considers policy implications of the findings and their caveats, and the avenues for future research.

CHAPTER 2 - ECONOMIC GEOGRAPHY AND UNEVEN REGIONAL SOCIO-ECONOMIC DEVELOPMENT IN BRAZIL

2.1. Introduction

This chapter describes the regional economy of Brazil. In doing so, it argues that (under)development of Brazilian regions obeys spatial patterns. This argument illustrates the limitation of the mainstream neoclassical economic theory which ignores that economic activities are not homogeneous in space (Krugman, 1991a, 1991b, 1998; Brakman, Garretsen & van Marrewijk, 2001 and the references therein; Mellinger, Sachs & Gallup, 2003; Fujita & Krugman, 2004[11]). However, in recent decades, many scholars have recognised the importance and implications of space for the distribution of economic activities and growth of regions (Anselin, 1995; Nazara & Hewings, 2004; Autant-Bernard, Mairesse & Massard, 2007; LeSage & Pace, 2009). The question asked here is: why are the north and northeast regions of Brazil, considering the selected socioeconomic variables below, the less developed in the country?

This chapter demonstrates that states' underdevelopment in Brazil is due to the fact that the advantages of "first nature" geography are more than offset by disadvantages of "second nature" geography for those backward states. For example, even though northeast states are coastal states, they still lag behind due to their problems with the second nature geography; that is, they cannot benefit from agglomeration or cluster economies that arise as a result of economies of scale and increasing returns at the firm level, due to lack of market access, given that they are located far from the biggest national markets of São Paulo and Rio de Janeiro[12]. Within the overall structure of this thesis, this chapter is important for two reasons. First, it uses the most extensive and updated regional data in Brazil; secondly, it provides theoretical explanations for regional disparities in Brazil. These aspects contrast previous descriptive studies which even though they do make an effort to analyse data, they do not relate these data to the pertinent economic theory. The theoretical approach used in this Chapter is the

[11] The ideas discussed in this Fujita & Krugman's paper are similar to those presented in the form of a chapter by Krugman in Clark, Feldman & Gertler (2003).

[12] This problem is known in the literature as a problem of Peripherality. For a discussion of its implications for growth both at regional and national levels, see McCann (2003) and the references therein, Poot (2004) and Hanson (2003). The latter author does not use the term 'Peripherality' but the explanations for agglomeration of economic activity that he provides in some sections of his work, focusing on the United States, imply the same idea of 'Peripherality' discussed by the former two authors.

new economic geography (Krugman, 1995; McCann, 2003; Bosker, 2008) which highlights that differences in performance of regional economies depend on location factors such as first nature geography and second nature geography.

The first nature geography factors include natural resources, climate, soil quality, natural harbours, and navigable rivers; to these factors are added exogenous factors (e.g. culture, type of management, etc) that lead to differences in technology and institutions across regions. The second nature geography factors are a result of interaction between agglomeration forces and diffusion forces within and among the regions. Agglomeration forces include proximity to large markets, the presence of a large variety of goods and services, and an efficient labour market. Diffusion forces involve congestion, pollution, high competition between goods' suppliers, and higher prices of immobile factors of cities such as land and buildings (Mellinger, Sachs & Gallup, 2003; Hanson, 2003, pp. 479-481; Poot, 2004, pp. 7-10; Bosker, 2008, pp. 1-7). This chapter is organised as follows. The next section briefly describes the data used and their sources. Section 2.3 discusses the indicators selected[13]. Lastly, section 2.4 provides the concluding remarks.

2.2. Data and Sources

To show and discuss the scale of economic concentration in Brazil, the regional variation of the following socio-economic indicators is considered: income as measured by GDP shares and by per capita GDP; total population and population density (or market scale); share of economically active population in total population (or labour force participation); human capital proxied by education levels (or knowledge intensity); potential unemployment rate (or resources underutilisation), average total employment, and labour productivity measured by total product (GDP or output) per worker (or efficiency and capital intensity); share of manufacturing employment in the total economically active population (or manufacturing concentration or industrialisation); size of manufacturing firms; electricity consumption per inhabitant as a proxy for economic activity (or, resource intensity); poverty rates as measures for a negative home market effect (or deprivation); and the Gini Index as a measure of inequality (or fairness). These indicators are considered at the national, regional, and state

[13] This selection was determined by data availability.

levels of Brazil (for states and regions of Brazil, see Figure 2.1). The data used were obtained from two official data sources in Brazil: IPEA – Instituto de Pesquisa Econômica Aplicada (or, Institute of Applied Economic Research, www.ipea.gov.br) and IBGE – Instituto Brasileiro de Geografia e Estatística (or, Brazilian Institute of Geography and Statistics, www.ibge.gov.br.). Both sources are rich in regional socio-economic data about Brazil.

Figure 2.1: States and Regions in Brazil[14]

Note: Center West stands for Centre-west.

Source: http://www.brazilmycountry.com/brazil-map.html#regions%20map

[14] In this thesis, middle-west and centre-west expressions are used interchangeably.

2.3. Descriptive Results

2.3.1. Gross Domestic Product, GDP

Total GDP data (see Table 2.1) indicate that there is great disparity in the proportion of national income across states and regions. However, there is evidence of convergence in the form of a decline in the share of São Paulo, Rio de Janeiro and Minas Gerais. In 1939, São Paulo alone accounted for approximately one-third of the national GDP. When considering the location of this state, southeast, it is found, in 1939, that just three states – São Paulo, Rio de Janeiro, and Minas Gerais – represented more than 60 percent of the nation's income. On the other hand, there were a limited number of states that had similar income shares. For example, Table 2.1 shows that only Rio Grande do Sul achieved an income share comparable to one of the states in the southeast region, namely Minas Gerais. Outside of the southeast and south regions, the two states that had a considerable income share are found in the northeast. These states are Bahia and Pernambuco, both of which had almost identical shares in national income in 1939. Overall, the southeast and south regions, together, accounted for approximately 80 percent of income in Brazil at that time (see Table 2.1).

Except for some convergence, with the share of the southeast and south declining to 73 percent by 2008, the described pattern did not significantly change from 1939 to 2008. Most changes observed were small changes within the regions. For example, in the southeast, São Paulo and Espírito Santo increased their national income shares at the expense of a loss of Rio de Janeiro's and Minas Gerais's national income shares. This situation was reversed in the south region as the leading state of Rio Grande do Sul, lost rather than gained its national income share at the expense of Paraná and Santa Catarina which had increases in their national income share, but the overall share of income for the southern region remained practically unchanged between 1939 and 2008.

Outside the southeast and south regions, Bahia managed to keep its share in national income, Pernambuco faced a half-decline, approximately, in its share from 1939 to 2008. Other north, northeast, and centre-west states kept very low income shares, except the young Distrito Federal. This state, from almost zero share in 1960, reached a share comparable to that of Bahia State in 2008. Bahia is an old and prominent state in the northeast. Overall, evidence shows that there has been a modest convergence over the period, but this convergence process seems to apply between the regions only, rather than within the regions and the nation simultaneously (see also Mossi, Aroca, Fernández & Azzoni, 2003).

Even though the southeast experienced a decline in its share, it still has the highest share (see Table 2.1). The increase in north and middle-west shares coincides with a fall in northeast share. The northeast region, however, is a backward region. This pattern is known in the literature as Kuznets hypothesis which argues that during the development process of a country, the regions (or groups of people) that have lower shares of income in the early stages of the country's development tend to increase their shares over time, which reduces inequality (Kuznets, 1955).

Table 2.1: Evolution of States and Regions' GDP Shares, 1939-2008

State	1939	1950	1960	1970	1980	1991	2000	2008
Acre	-	-	-	0.13%	0.12%	0.13%	0.15%	0.22%
Amazonas	1.09%	0.72%	0.85%	0.69%	1.11%	1.68%	1.71%	1.54%
Amapá	-	-	-	0.11%	0.08%	0.16%	0.18%	0.22%
Pará	1.57%	1.00%	1.38%	1.10%	1.55%	2.04%	1.72%	1.93%
Rondônia	-	-	-	0.10%	0.27%	0.42%	0.51%	0.59%
Roraima	-	-	-	0.03%	0.04%	0.09%	0.10%	0.16%
Tocantins	-	-	-	0.00%	0.17%	0.18%	0.22%	0.43%
North Region	**2.66%**	**1.71%**	**2.23%**	**2.16%**	**3.34%**	**4.71%**	**4.60%**	**5.10%**
Alagoas	0.91%	0.85%	0.81%	0.68%	0.66%	0.71%	0.64%	0.64%
Bahia	4.45%	3.78%	4.23%	3.80%	4.33%	4.40%	4.38%	4.01%
Ceará	2.09%	2.09%	1.96%	1.44%	1.54%	1.83%	1.89%	1.98%
Maranhão	1.25%	0.79%	1.10%	0.82%	0.84%	0.81%	0.84%	1.27%
Paraíba	1.33%	1.47%	1.42%	0.71%	0.65%	0.85%	0.84%	0.85%
Pernambuco	4.41%	3.86%	3.47%	2.91%	2.53%	2.89%	2.64%	2.32%
Piauí	0.87%	0.43%	0.41%	0.37%	0.38%	0.46%	0.48%	0.55%
Rio Grande do Norte	0.79%	0.90%	0.89%	0.54%	0.63%	0.78%	0.84%	0.84%
Sergipe	0.63%	0.48%	0.49%	0.43%	0.39%	0.64%	0.54%	0.64%
Northeast Region	**16.73%**	**14.65%**	**14.78%**	**11.71%**	**11.96%**	**13.37%**	**13.09%**	**13.11%**
Distrito Federal	-	-	0.04%	1.26%	1.99%	2.32%	2.69%	3.88%
Goiás	1.19%	1.17%	1.41%	1.52%	1.70%	1.88%	1.97%	2.48%
Mato Grosso do Sul	-	-	-	-	1.09%	0.99%	1.08%	1.09%
Mato Grosso	0.95%	0.62%	1.01%	1.09%	0.61%	0.91%	1.22%	1.75%
Middle-West Region	**2.14%**	**1.79%**	**2.46%**	**3.87%**	**5.39%**	**6.10%**	**6.95%**	**9.20%**
Espírito Santo	1.21%	1.31%	1.05%	1.18%	1.47%	1.66%	1.96%	2.30%
Minas Gerais	10.26%	10.53%	9.97%	8.28%	9.42%	9.52%	9.64%	9.32%
Rio de Janeiro	20.34%	18.96%	17.04%	16.67%	13.73%	12.28%	12.52%	11.32%
São Paulo	31.10%	34.76%	34.71%	39.43%	37.71%	35.25%	33.67%	33.08%
Southeast Region	**62.91%**	**65.55%**	**62.76%**	**65.55%**	**62.34%**	**58.71%**	**57.79%**	**56.02%**
Paraná	2.95%	4.90%	6.41%	5.43%	5.76%	5.87%	5.99%	5.91%
Rio Grande do Sul	10.33%	8.97%	8.78%	8.60%	7.93%	7.74%	7.73%	6.58%
Santa Catarina	2.27%	2.43%	2.59%	2.68%	3.29%	3.49%	3.85%	4.07%
South Region	**15.56%**	**16.29%**	**17.77%**	**16.71%**	**16.97%**	**17.11%**	**17.57%**	**16.56%**
Brazil's Total GDP* (billion BR$ of 2000)	46.7	77.9	136.2	285.3	760	931.9	1101.3	1569.4

Note: States' shares as well as Regions' shares sum 100% separately. GDP data are in real terms. Before 1970, GDP data for Rondônia and Roraima states and for the Acre State (the three then named territories) are included in the Amazonas State, so for this period the Amazonas' GDP is the overall GDP for the four states. Distrito Federal was inaugurated in 1960, therefore it does not have data as a separate state prior to that date. On the other hand, over the same period, GDP data for Amapá are included in Pará (IPEA). Until 1970, GDP data for Mato Grosso also includes Mato Grosso do Sul which implies that Mato Grosso's GDP is also the overall for these two states; the separation of the two states with a birth of Mato Grosso do Sul was on 1 January 1979, even though the official law for separation was approved on 11 October 1977 (www.citybrazil.com.br). Before 1980, GDP data for Goiás also includes Tocantins; IPEA provides Tocantins GDP data for 1980, but the official separation between these two states which culminated in birth of Tocantins State was in 1988 (www.citybrazil.com.br).

***Note:** The annual commercial average exchange rate Brazilian Real per United States Dollar (BR$ per US$), for example, was 1.8302 real per dollar for 2000 and 1.8346 real per dollar for 2008 (IPEA); thus, Brazil's real GDP grew approximately from US$ 601.7 billion in 2000 to US$ 855.4 billion in 2008.

2.3.2. Per capita Real Gross Domestic Product

Table 2.2 compares the evolution of states', regions' and nation's per capita GDP. Over the 1939-2008 period, real economic growth (in terms of growth in real GDP per capita) in Brazil was 3.1 percent per annum on average. In 1939, Rio de Janeiro, not São Paulo, was the state with the highest per capita GDP. In that year, only five (out of 27) states, namely Rio de Janeiro, São Paulo, Rio Grande do Sul, Paraná, and Mato Grosso (in this order), had a per capita GDP higher than the nation's. Over the period, there had been a considerable increase in states' per capita GDP for all states. However, even though north and northeast states also had a growth in their per capita GDP, they still fell behind the national average.

There were only eight states that performed better than the nation, with Distrito Federal, Santa Catarina, and Espírito Santo joining the 1939 short-list in 2008. In this group of eight states, there was a dramatic change: São Paulo overtook Rio de Janeiro from 1960 to 2008, but the former state was, in turn, surpassed by Distrito Federal from 1970. Distrito Federal benefited by in-migration of some of Brazil's highly skilled labour (see how this state performs in human capital indicators, Figures 2.2 and 2.3 below) due to transfer of the country's capital from Rio de Janeiro to Brasília in 1960. While Distrito Federal started with a per capita GDP that was approximately 0.20 relative to the nation's in 1960, with an annual average growth rate of 9.2 percent, it reached a per capita GDP level that was almost three times relative to the nation's in 2008, which makes Distrito Federal's income levels 'uncommonly' high by Brazilian standards.

Mato Grosso, Rio Grande do Sul, Paraná, São Paulo, and Rio de Janeiro performed better than the nation from 1939 to 2008 in terms of per capita GDP. Due to their real income levels, those states had either an identical or lower annual average growth rate relative to the nation (see Table 2.2). In contrast, some states with lower per capita GDP over the period compared to the nation's, had the highest average annual growth rates (higher than 3.3%). They were Tocantins, Sergipe, Goiás, Espírito Santo, Minas Gerais, and Santa Catarina (up to 1970). Additionally, two of the poorest regions, middle-west and north, but not northeast, had average annual growth rate higher than the nation's (see Table 2.2). This evidence also supports convergence hypothesis.

Table 2.2: Evolution of States and Regions' real per capita GDP versus Brazil, 1939-2008

State	1939	1950	1960	1970	1980	1991	2000	2008	Average Annual growth rate, earliest year to 2008[15]
Acre	-	-	-	1.70	2.93	2.99	3.14	5.12	2.9%
Amazonas	0.88	0.82	1.18	2.06	5.91	7.46	7.15	7.25	3.1%
Amapá	-	-	-	2.80	3.52	5.18	4.29	5.71	1.9%
Pará	0.44	0.67	1.16	1.45	3.46	3.83	3.15	4.14	3.3%
Rondônia	-	-	-	2.67	4.17	3.47	4.27	6.20	2.2%
Roraima	-	-	-	2.30	3.97	4.02	4.09	6.13	2.6%
Tocantins	-	-	-	-	1.74	1.84	2.11	5.29	4.1%
North Region	**0.55**	**0.65**	**1.03**	**1.50**	**3.84**	**4.38**	**4.08**	**5.29**	**3.3%**
Alagoas	0.33	0.61	0.86	1.22	2.54	2.64	2.56	3.22	3.3%
Bahia	0.47	0.61	0.96	1.45	3.48	3.45	3.67	4.34	3.3%
Ceará	0.56	0.61	0.80	0.94	2.21	2.68	2.89	3.68	2.8%
Maranhão	0.46	0.39	0.60	0.79	1.61	1.54	1.68	3.16	2.8%
Paraíba	0.42	0.67	0.96	0.85	1.79	2.46	2.72	3.55	3.2%
Pernambuco	0.64	0.89	1.14	1.61	3.13	3.78	3.81	4.17	2.7%
Piauí	0.45	0.32	0.43	0.62	1.33	1.64	1.94	2.78	2.7%
Rio Grande do Norte	0.44	0.73	1.03	0.99	2.54	3.03	3.46	4.25	3.3%
Sergipe	0.52	0.59	0.88	1.37	2.59	3.99	3.40	5.06	3.4%
Northeast Region	**0.50**	**0.64**	**0.90**	**1.19**	**2.61**	**2.93**	**3.08**	**3.88**	**3.0%**
Distrito Federal	-	-	0.36	6.71	12.88	13.48	14.67	23.80	9.2%
Goiás	0.69	0.76	0.98	1.47	4.14	4.37	4.37	6.67	3.4%
Mato Grosso do Sul	-	-	-	-	6.06	5.18	5.76	7.34	0.7%
Mato Grosso	1.10	0.92	1.51	1.95	4.04	4.20	5.55	9.28	3.1%
Middle-West Region	**0.82**	**0.87**	**1.26**	**2.43**	**6.02**	**6.03**	**6.69**	**10.55**	**3.8%**
Espírito Santo	0.73	1.20	1.20	2.10	5.51	5.94	7.22	10.47	3.9%
Minas Gerais	0.59	1.04	1.39	2.06	5.35	5.64	6.07	7.37	3.7%
Rio de Janeiro	2.33	3.16	3.46	5.29	9.24	8.93	9.90	11.19	2.3%
São Paulo	1.99	2.96	3.64	6.33	11.45	10.40	10.20	12.66	2.7%
Southeast Region	**1.45**	**2.27**	**2.79**	**4.69**	**9.16**	**8.72**	**8.99**	**10.96**	**3.0%**
Paraná	1.23	1.80	2.04	2.24	5.74	6.48	6.95	8.76	2.9%
Rio Grande do Sul	1.45	1.68	2.19	3.68	7.75	7.89	8.45	9.51	2.8%
Santa Catarina	0.97	1.21	1.64	2.63	6.89	7.17	8.21	10.54	3.5%
South Region	**1.31**	**1.62**	**2.04**	**2.89**	**6.78**	**7.20**	**7.82**	**9.45**	**2.9%**
Brazil	1.04	1.50	1.93	3.06	6.39	6.35	6.63	8.28	3.1%

Note: GDP data are in thousands of BR$ in 2000. The assumptions made for states' total GDP data in the note beneath Table 2.1 above also hold. Therefore, before 1970 Amazonas state per capita GDP is the overall per capita GDP for Rondônia, Roraima, Acre, and Amazonas states. On the other hand, over the same period, per capita GDP for Pará is the overall per capita GDP for both Amapá and Pará states. Until 1970, per capita GDP for Mato Grosso is the overall per capita GDP for both Mato Grosso and Mato Grosso do Sul. Before 1980, per capita GDP for Goiás is the overall per capita GDP for both Goiás and Tocantins states.

---

[15] The expression used for calculations of growth rate is $g = \left[e^{\left(\frac{1}{T}\right)\ln\left(\frac{y_t}{y_{t-T}}\right)} - 1\right] * 100$, where y_{t-T} and y_t are the earliest and the latest year's GDP per capita and T is the time period between both years.

2.3.3. Population

Population data show that the northern states of Brazil are relatively less inhabited. The low levels of income in those states are therefore not surprising. However, there are states with low income even though they are highly populated. For instance, Bahia was the third most populated (see Table 2.3) in 1940 but it was ranked fifth in 1939 income share (see Table 2.1). The population distribution across states did not change between 1920 and 2008 though there were some changes in ranking among the top six states of São Paulo, Minas Gerais, Bahia, Rio de Janeiro, Pernambuco, and Rio Grande do Sul[16]. These states are leading states in their regions in terms of income. This finding is consistent with the argument of a circular causality between population and the location of economic activities: economic activities tend to agglomerate in regions with high population levels (home market effect); and population levels, in turn, tend to be high in regions with concentration of economic activities (Krugman, 1991a, 1991b, 1995; Brakman, Garretsen & van Marrewijk, 2001; Fujita & Krugman, 2004; Capello & Nijkamp, 2009). Table 2.3 reinforces the finding of the home market effect as the regions with the highest shares are those that also have some of their states among the top ten at the national level in both population and income.

Besides the home market effect of population size, modern economic geography suggests that population density is important too. The data in Table 2.4 show that all states increased their population density levels from 1920 to 2008. Two findings can be drawn. First, the richest states of southeast and of northeast are among those with the highest density levels in all years. This is consistent with the previous findings of a circular causation for agglomeration (Krugman, 1991a, 1991b, 1995; Brakman, Garretsen & van Marrewijk, 2001; Fujita & Krugman, 2004; Capello & Nijkamp, 2009).

[16] This apparent regularity for population distribution will be tested in terms of the theory of the rank-size rule in Chapter 4.

Table 2.3: Evolution of the Shares of the States and Regions in Brazil's Population, 1920-2008

State	1920	1940	1950	1960	1970	1980	1991	2000	2008
Acre	0.30%	0.19%	0.19%	0.23%	0.23%	0.25%	0.28%	0.33%	0.36%
Amazonas	1.14%	1.01%	1.01%	1.02%	1.03%	1.20%	1.43%	1.59%	1.76%
Amapá	0.12%	0.07%	0.07%	0.10%	0.12%	0.15%	0.20%	0.28%	0.32%
Pará	3.09%	2.22%	2.22%	2.20%	2.33%	2.86%	3.37%	3.61%	3.86%
Rondônia	0.08%	0.03%	0.03%	0.10%	0.12%	0.41%	0.77%	0.79%	0.79%
Roraima	0.03%	0.03%	0.03%	0.04%	0.04%	0.07%	0.15%	0.16%	0.22%
Tocantins	0.36%	0.40%	0.40%	0.48%	0.56%	0.62%	0.63%	0.70%	0.68%
North Region	**5.12%**	**3.96%**	**3.94%**	**4.16%**	**4.43%**	**5.56%**	**6.83%**	**7.46%**	**7.99%**
Alagoas	3.19%	2.31%	2.31%	1.80%	1.71%	1.67%	1.71%	1.65%	1.65%
Bahia	10.87%	9.50%	9.50%	8.48%	8.05%	7.94%	8.08%	7.91%	7.65%
Ceará	4.30%	5.07%	5.07%	4.73%	4.68%	4.44%	4.34%	4.33%	4.46%
Maranhão	2.85%	3.00%	3.00%	3.53%	3.21%	3.36%	3.36%	3.30%	3.33%
Paraíba	3.13%	3.45%	3.45%	2.86%	2.56%	2.33%	2.18%	2.05%	1.97%
Pernambuco	7.03%	6.52%	6.52%	5.86%	5.54%	5.16%	4.85%	4.60%	4.61%
Piauí	1.99%	1.98%	1.98%	1.83%	1.80%	1.80%	1.76%	1.66%	1.65%
Rio Grande do Norte	1.75%	1.86%	1.86%	1.65%	1.66%	1.60%	1.65%	1.62%	1.64%
Sergipe	1.56%	1.32%	1.32%	1.08%	0.97%	0.96%	1.02%	1.05%	1.05%
Northeast Region	**36.67%**	**35.00%**	**34.55%**	**31.81%**	**30.18%**	**29.25%**	**28.94%**	**28.15%**	**28.00%**
Distrito Federal	0.13%	0.17%	0.17%	0.20%	0.58%	0.99%	1.09%	1.21%	1.35%
Goiás	1.30%	1.61%	1.61%	2.29%	2.60%	2.62%	2.74%	2.98%	3.08%
Mato Grosso do Sul	0.43%	0.58%	0.58%	0.82%	1.07%	1.15%	1.21%	1.24%	1.23%
Mato Grosso	0.32%	0.45%	0.45%	0.47%	0.64%	0.96%	1.38%	1.46%	1.56%
Middle-West Region	**2.18%**	**2.81%**	**3.07%**	**3.78%**	**4.89%**	**5.72%**	**6.42%**	**6.89%**	**7.22%**
Espírito Santo	1.49%	1.82%	1.82%	1.68%	1.72%	1.70%	1.77%	1.79%	1.82%
Minas Gerais	19.20%	16.33%	16.33%	13.88%	12.33%	11.24%	10.72%	10.53%	10.47%
Rio de Janeiro	8.83%	8.76%	8.76%	9.50%	9.66%	9.49%	8.72%	8.39%	8.37%
São Paulo	14.98%	17.41%	17.41%	18.37%	19.08%	21.04%	21.51%	21.88%	21.63%
Southeast Region	**44.50%**	**44.32%**	**43.37%**	**43.43%**	**42.79%**	**43.47%**	**42.73%**	**42.60%**	**42.29%**
Paraná	2.24%	3.00%	3.00%	6.06%	7.44%	6.41%	5.75%	5.71%	5.59%
Rio Grande do Sul	7.12%	8.05%	8.05%	7.72%	7.16%	6.53%	6.22%	6.07%	5.72%
Santa Catarina	2.18%	2.86%	2.86%	3.04%	3.12%	3.05%	3.09%	3.11%	3.19%
South Region	**11.53%**	**13.91%**	**15.07%**	**16.81%**	**17.71%**	**15.99%**	**15.07%**	**14.89%**	**14.50%**
Brazil's Total Population (millions of people)	30.7	41.2	52.0	70.6	93.1	119.0	146.8	166.1	198.6

Note: Before 1960, a backcast for Distrito Federal's population was done.
States' shares as well as Regions' shares sum 100% separately.

Secondly, there are some poor states with relatively high density levels in all years (e.g. Alagoas and Sergipe). These states are agricultural. This result is also in line with the literature (Krugman, 1991a, 1991b). Krugman (1991a, 1991b) develops a simple two-region model in which agglomeration may arise in a region specialising in agriculture activities due to economies of scale associated with non-mobile inputs, share of income spent in manufactured goods and transport costs. The transport costs parameter determines regional convergence or regional divergence in the long-run version of Krugman's model. The evolving distribution of population in Brazil is consistent with this model. The poor states in Brazil are agricultural, with less developed infrastructure, and they have small scale manufacturing[17]. In 2008, comparison between the nation and regions shows that while the Amazonian region (north region and a part of centre-west) was still relatively less inhabited with very low population density, the economically most important regions of southeast and south had a density around four and two times of Brazil's, respectively.

Table 2.4: Evolution of States and Regions' Population Density versus Brazil, 1920-2008

State	1920	1940	1950	1960	1970	1980	1991	2000	2008*
Acre	0.62	0.54	0.75	1.05	1.41	1.97	2.73	3.55	4.14
Amazonas	0.23	0.27	0.32	0.46	0.61	0.92	1.33	1.68	2.14
Amapá	0.24	0.60	0.28	0.50	0.82	1.26	2.02	3.21	4.29
Pará	0.78	0.74	0.93	1.26	1.77	2.77	3.95	4.81	5.87
Rondônia	0.07	0.14	0.15	0.29	0.46	2.02	4.75	5.55	6.29
Roraima	0.03	0.05	0.08	0.13	0.18	0.34	0.97	1.22	1.84
Tocantins	0.40	0.58	0.76	1.18	1.82	2.58	3.30	4.19	4.61
North Region	**0.40**	**0.45**	**0.54**	**0.77**	**1.07**	**1.72**	**2.59**	**3.22**	**3.93**
Alagoas	34.26	33.30	39.45	46.05	57.43	71.48	90.00	98.44	112.59
Bahia	6.32	7.40	8.60	10.70	13.38	16.89	20.92	23.28	25.68
Ceará	8.98	14.07	18.23	22.54	29.71	36.02	43.50	49.41	56.75
Maranhão	2.53	3.57	4.82	7.69	9.22	12.31	14.79	16.51	19.00
Paraíba	17.19	25.43	30.29	32.88	42.26	49.14	56.57	60.31	66.28
Pernambuco	21.71	27.08	34.61	42.09	52.51	62.49	72.04	77.52	88.99
Piauí	2.48	3.33	4.15	4.94	6.70	8.52	10.23	10.96	12.40
Rio Grande do Norte	10.25	14.65	18.24	22.02	29.24	35.82	45.31	50.58	58.82
Sergipe	22.14	25.16	29.25	34.57	40.95	51.85	67.66	79.26	91.22
Northeast Region	**7.38**	**9.45**	**11.61**	**14.44**	**18.26**	**22.61**	**27.27**	**30.15**	**34.15**
Distrito Federal	6.88	12.04	13.76	24.38	93.14	203.93	275.00	347.56	441.82
Goiás	1.05	1.75	2.83	4.55	6.81	8.79	11.78	14.56	17.19
Mato Grosso do Sul	0.31	0.68	0.86	1.65	2.85	3.91	4.97	5.76	6.54
Mato Grosso	0.13	0.18	0.24	0.38	0.68	1.29	2.24	2.68	3.27
Middle-West Region	**0.43**	**0.65**	**0.99**	**1.67**	**2.86**	**4.27**	**5.85**	**7.13**	**8.53**
Espírito Santo	11.95	20.26	25.30	30.19	35.08	44.37	56.31	64.73	74.92
Minas Gerais	9.92	11.50	11.53	16.81	19.71	22.97	26.76	29.82	33.84
Rio de Janeiro	62.64	82.90	109.13	154.94	207.71	260.75	291.68	318.14	362.55
São Paulo	18.62	29.04	37.03	52.34	71.85	101.25	126.96	146.47	165.24
Southeast Region	**14.80**	**20.01**	**22.40**	**33.57**	**43.37**	**56.31**	**67.66**	**76.53**	**86.73**
Paraná	3.49	6.18	10.56	21.56	34.81	38.33	42.31	47.63	53.13
Rio Grande do Sul	7.65	12.20	15.57	20.37	24.91	29.06	33.95	37.48	40.39
Santa Catarina	7.04	12.40	16.55	22.43	30.39	38.00	47.59	54.25	63.24
South Region	**6.13**	**10.11**	**13.95**	**21.14**	**29.35**	**33.86**	**39.22**	**43.91**	**48.77**
Brazil's Total Population Density	3.61	4.90	6.08	8.35	11.02	14.08	17.21	19.55	22.27

Note: Population Density (Average) for each state is the ratio between state's population and state's area (in square kilometers). The Distrito Federal misses value for the area for the first three years of the table. It was assumed the area of 1960 subtracting it from the area of Goiás state so the total nation's area remained unchanged.
***Note:** For 2008, the currently (April 2011) reported area on the IBGE website has been used; the reported total area of Brazil from IBGE is slightly higher by 12,148 square kilometres than the sum of its states areas. For 2008, it was used the reported total area to calculate the national density. For the other years, the area of Brazil is equal to the sum of its states' areas.

[17] For instance, individually, these poor states, which are mostly located in north and northeast regions, had from 1981 to 2007 an average share in manufacturing employment at the regional and national levels of up to 11 and 1 percent only, respectively. For Tocantins state, those average shares were calculated for the period 1992-2007 due to missing data.

2.3.4. Share of Economically Active Population[18] in Total Population

Table 2.5 shows the distribution of the share of economically active population in total population (i.e., the labour force participation rate). The pattern is similar to other indicators as some of the states and regions have the highest shares in this indicator as well as in others. Excluding 1920 and 1940, there is a positive correlation between labour force participation rate and real income per capita. For both these indicators, the southeast and south regions have the highest levels while the north and northeast present the lowest levels (compare rows of these regions in Tables 2.2 and 2.5).

Since 1940, the development process has created disparities in labour force participation rates across states (see Table 2.5). In that year, the states with the highest rates of labour force participation (and with at least 63 percent which was the nation's rate) were São Paulo, Rio de Janeiro, Pará, Rondônia, Rio Grande do Sul, Bahia, Paraná, Tocantins, Roraima, and Mato Grosso. Five of these states were among those with the highest real per capita income in 1939; and while the other five states had a very low per capita income in 1939, their annual average growth rate of income per capita from 1939 to 2000 was above the nation's.

In 2000, the states with the highest labour force participation rates (and even much higher than the nation's rate of 46.6 percent) are Santa Catarina, Rio Grande do Sul, Distrito Federal, Roraima, Espírito Santo, São Paulo, Paraná, Goiás, Mato Grosso, and Minas Gerais. Again, six of these states are among those with the highest real income per capita in 2000, with three of the other four having an equal or higher annual average growth rate of real income per capita from 1939 to 2000 compared to the nation's and despite the fact that their level for that indicator (real income per capita) was lower than the nation's. These findings are consistent with the new economic geography literature as high labour force participation reinforces the home market (or, high market potential for states and

[18] According to IPEA, "Economically active population involves people who during all the 12 months or part of them prior to the date of the Census had done paid work, paid by money and/or goods or commodities including under licence, with remuneration for disease, with scholarships, etc, and those without remuneration that usually have worked 15 hours or more per week in an economic activity helping the person with whom they resided, or in a care institution or cooperative or, yet, as learners, interns, etc. Also considered in this condition were people aged 10 years or more who did not work in the 12 months prior to the reference date of the Census but in the last two months have been actively seeking work" (IPEA).

regions) (Krugman, 1991a, 1991b, 1995; Brakman, Garretsen & van Marrewijk, 2001; Clark, Feldman & Gertler, 2003; Capello & Nijkamp, 2009).

Table 2.5: Evolution of States and Regions' Labour Force Participation Rate versus Brazil, 1920-2000

State	1920	1940	1950	1960	1970	1980	1991	2000
Acre	37.9%	60.9%	63.1%	30.0%	30.0%	30.9%	34.4%	39.3%
Amazonas	28.9%	60.7%	61.9%	29.5%	28.2%	31.1%	32.6%	41.0%
Amapá	25.0%	58.1%	65.4%	27.5%	25.5%	28.0%	31.9%	39.4%
Pará	27.0%	65.0%	64.6%	30.7%	28.6%	30.2%	32.9%	40.2%
Rondônia	44.1%	64.5%	66.6%	31.7%	30.5%	34.9%	38.2%	46.6%
Roraima	23.9%	63.0%	63.3%	26.6%	28.0%	33.8%	38.8%	50.9%
Tocantins	26.7%	63.1%	63.3%	30.0%	28.5%	28.2%	33.9%	42.0%
North Region	**28.2%**	**63.4%**	**63.8%**	**30.2%**	**28.5%**	**30.5%**	**33.7%**	**41.4%**
Alagoas	22.3%	57.7%	61.0%	33.4%	30.4%	30.9%	34.1%	39.8%
Bahia	24.2%	63.8%	62.4%	32.4%	30.7%	32.1%	35.1%	42.7%
Ceará	21.5%	58.6%	61.4%	31.2%	28.8%	32.4%	35.6%	41.5%
Maranhão	24.5%	62.2%	63.1%	31.7%	32.5%	32.7%	32.2%	39.6%
Paraíba	22.0%	61.0%	62.3%	29.5%	28.3%	30.4%	34.5%	41.2%
Pernambuco	23.2%	59.9%	63.6%	31.5%	29.2%	33.2%	35.5%	42.5%
Piauí	22.0%	60.0%	61.6%	29.4%	28.8%	30.8%	34.0%	41.7%
Rio Grande do Norte	22.3%	61.9%	62.6%	28.9%	26.5%	31.3%	34.9%	40.7%
Sergipe	24.5%	62.3%	63.6%	34.0%	29.5%	31.0%	35.6%	42.4%
Northeast Region	**23.1%**	**61.1%**	**62.4%**	**31.5%**	**29.7%**	**32.0%**	**34.7%**	**41.6%**
Distrito Federal	-	-	-	-	33.2%	40.2%	43.7%	51.0%
Goiás	22.2%	60.0%	62.3%	30.5%	29.7%	35.1%	41.2%	48.2%
Mato Grosso do Sul	25.4%	61.9%	62.1%	31.0%	30.8%	36.7%	40.5%	47.5%
Mato Grosso	23.2%	63.0%	61.9%	30.5%	31.2%	33.8%	39.8%	47.9%
Middle-West Region	**23.0%**	**60.9%**	**62.2%**	**29.0%**	**30.6%**	**36.1%**	**41.2%**	**48.5%**
Espírito Santo	25.2%	61.0%	64.4%	29.7%	28.6%	35.1%	40.4%	50.7%
Minas Gerais	21.4%	61.7%	62.2%	30.5%	30.1%	35.4%	40.4%	47.7%
Rio de Janeiro	31.6%	66.3%	69.6%	32.2%	32.4%	38.2%	42.0%	48.1%
São Paulo	23.9%	66.3%	68.1%	34.8%	35.9%	41.6%	44.0%	50.2%
Southeast Region	**24.4%**	**64.4%**	**66.2%**	**32.7%**	**33.1%**	**39.0%**	**42.6%**	**49.2%**
Paraná	22.9%	63.4%	63.4%	33.0%	32.9%	37.5%	42.8%	49.0%
Rio Grande do Sul	22.2%	63.9%	65.3%	32.8%	34.0%	41.2%	45.2%	51.2%
Santa Catarina	24.0%	62.2%	62.1%	29.9%	30.4%	37.4%	43.5%	51.9%
South Region	**22.7%**	**63.4%**	**64.2%**	**32.4%**	**32.9%**	**39.0%**	**43.9%**	**50.5%**
Brazil	23.9%	63.0%	64.4%	32.0%	31.7%	36.3%	39.8%	46.6%

Note: Distrito Federal has missing values for the first four years of the table.

2.3.5. Human Capital

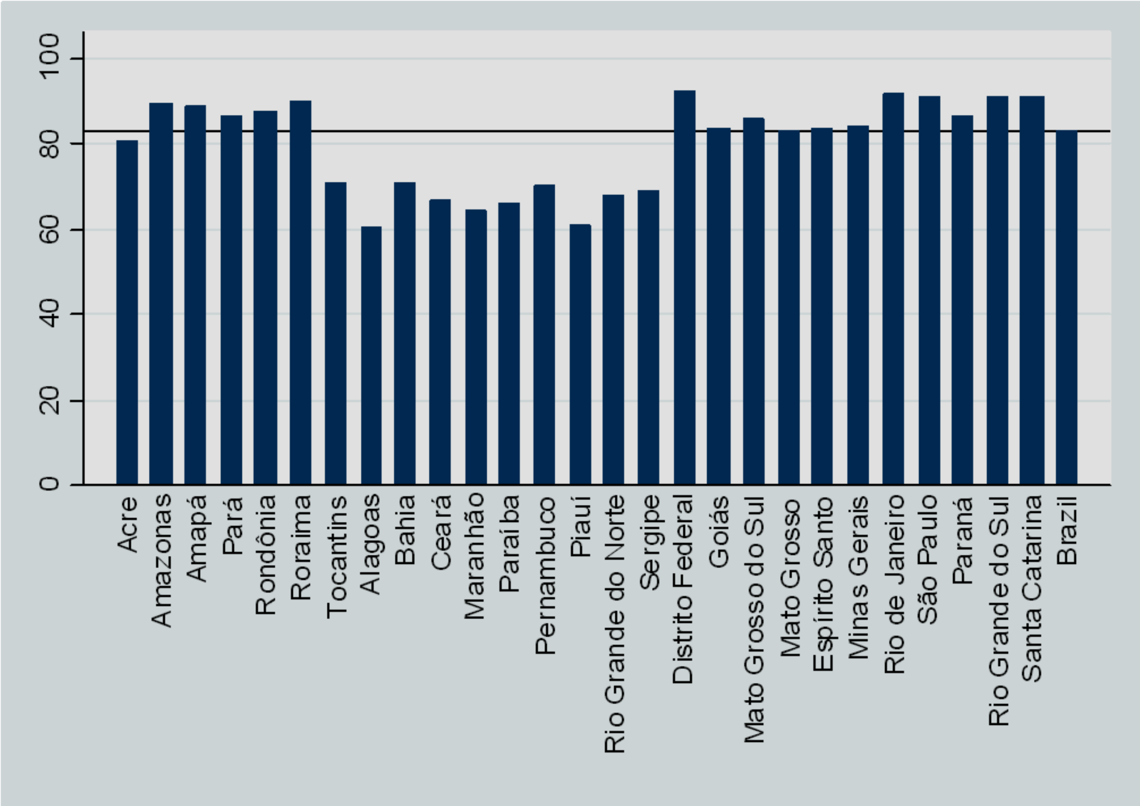
Two measures of education act as proxies for Human Capital[19]. The first is the literacy rate (100 minus the percentage of illiterate people aged 15 years and older)[20]. The average of this indicator's annual data from 1980 to 2006 has been taken (see Figure 2.2). This indicates that, given its development level (e.g. until the middle of the 1950s, Brazil was essentially an agrarian country), Brazil has a relatively high (approximately 83 percent; see the reference line in Figure 2.2) literacy rate. However, there is a strong variation across its states and regions. For instance, while the ratios between the following states' average and the nation's average is approximately 1.11 for Distrito Federal and Rio de Janeiro, 1.10 for São Paulo, Santa Catarina, Rio Grande do Sul, Roraima, Amapá, and Amazonas, it is only 0.80 for Ceará and Paraíba, 0.78 for Maranhão, 0.74 for Piauí, and 0.73 for Alagoas[21]. Therefore, there are huge gaps in literacy across Brazilian states.

[19] For these two human capital variables, the series is discontinuous starting from 1970. To calculate an average for a continuous series, the period from 1980 to 2006 was considered. The average calculated for these two variables for the 1970-2000 period using decadal data provided no significantly different result.

[20] According to IPEA, the rate of illerate people is "the percentage of people aged 15 years and older who know neither reading nor writing of a simple note".

[21] The latter five states are Brazil's northeast states. For a discussion about the negative effects of low human capital for the northeast's development, see also The Economist (2011), http://www.economist.com/node/18712379.

Figure 2.2: Average Literacy Rate: Brazil and its States, 1980-2006

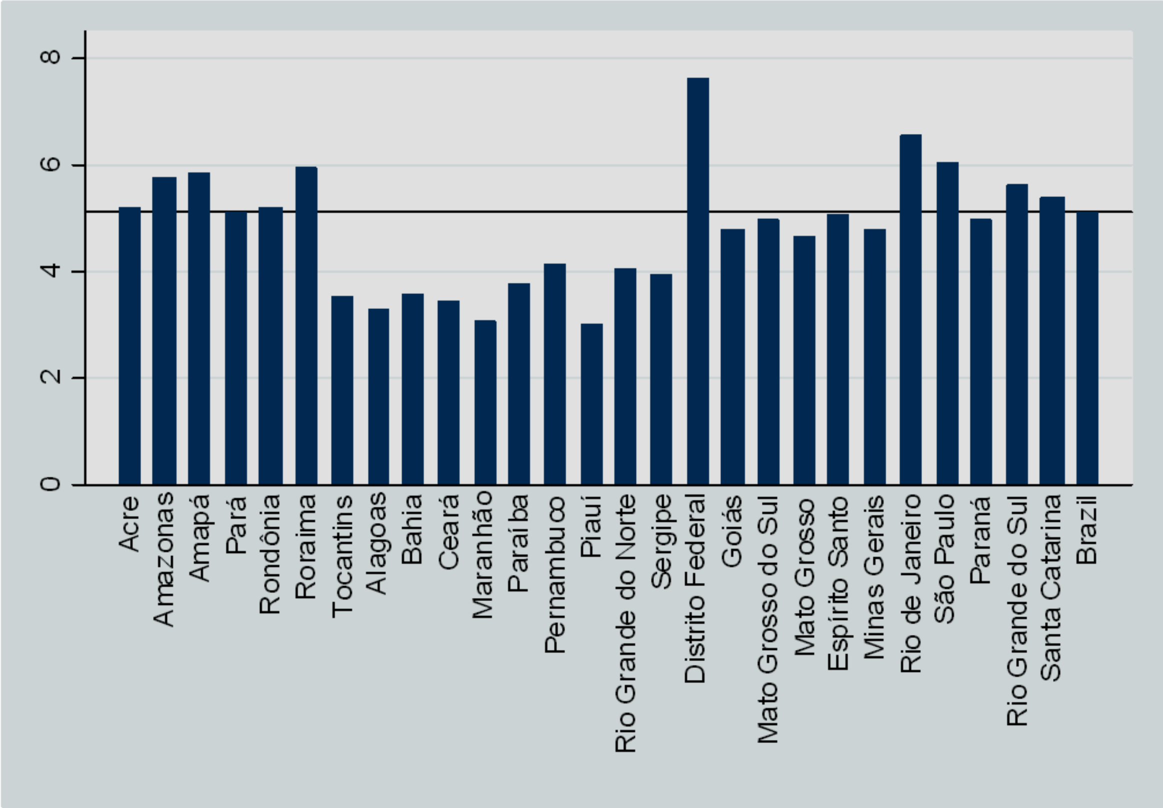


The second human capital measure is the average years of schooling for the population aged 25 years old or more. The average of this indicator's annual data from 1980 to 2006 has been taken (see Figure 2.3). The results are different from those of the literacy rate. For the schooling indicator, the average for Brazil from 1980 to 2006 is 5.11 schooling years only (see reference line in Figure 2.3), which is very low compared to that suggested by the information on literacy taking into account the fact that the age range between the two variables differs.

Additionally, there are great differences across states as well as between states' performance when compared to the nation's average. For example, the ratio between the following states and the nation for this indicator is approximately 1.49 for Distrito Federal, 1.28 for Rio de Janeiro, 1.18 for São Paulo, 1.17 for Roraima, 1.15 for Amapá, 1.13 for Amazonas, 1.10 for Rio Grande do Sul, and 1.06 for Santa Catarina. On the other hand, this ratio for some of the other states fell quite a long way behind; approximately 0.77 for Sergipe, 0.74 for Paraiba, 0.71 for Bahia, 0.69 for Tocantins, 0.68 for Ceará, 0.65 for Alagoas, 0.60 for Maranhão and Piauí. These latter eight states apart from

Tocantins [22] are located in the northeast region and are the lagging ones nationwide, which shows that both human capital indicators are consistent with the picture for population distribution, income, and economic activities in Brazil.

Figure 2.3: Average of schooling years: Brazil and its States, 1980-2006



2.3.6. Unemployment and Productivity

2.3.6.1. Potential Unemployment Rate

Table 2.6 shows that from 1981 to 2007 the nation's potential unemployment rate was quite stable. However, at regional level, all regions (except southeast and south) had a fall with northeast, middle-west and southeast having an "U" pattern. In 1981, three (out of 5) regions had national standard in their potential unemployment rates, but the other two regions were against the nation: for the north region, the rate was as high as 70 percent and for the south region was about 31 percent. This variation was also observed at state level. Overall, for instance, in 1981, excluding Distrito Federal, the highest potential unemployment rates, from

[22] This state is located in the north region but is among the poorest states in Brazil, and it shares a border with three northeast states, two of which are among the poorest in both the northeast region and the nation (see Figure 2.1, and Tables 2.1 and 2.2).

66 percent to 74.1 percent, were observed in northern states. On the other hand, in 1981, when considering a cut-off of 40 percent, the lowest potential unemployment rates (and smaller than the national average) were observed in three south states, in southeast states (except Rio de Janeiro) as well as in most northeast states (e. g. Sergipe, Piauí, Maranhão, Bahia, and Ceará) and middle-west state of Mato Grosso do Sul.

While in the south and southeast (except São Paulo and Rio de Janeiro) the potential unemployment rates were quite stable during the period, in the lagging regions' states of middle-west, northeast, and north, the rates declined significantly. In 2007, the potential unemployment rates were the lowest in lagging states (e. g. Tocantins, Rondônia) and all regions (except the leading region of southeast) had potential unemployment rates smaller than the national average. Overall, the states with a significant decrease in potential unemployment rate (or, with highest decrease in underutilization of resources, including human resources) were those with lower income per capita (or, lower market potential). This finding is consistent with Brazil's development and convergence hypothesis (Kuznets, 1955; Barro & Sala-i-Martin, 2004).

Table 2.6: Evolution of States and Regions' Potential Unemployment Rate, 1981-2007

State	1981	1991	1999	2007
Acre	73.2%			23.1%
Amazonas	63.6%			34.5%
Amapá	70.3%			32.1%
Pará	74.1%			39.6%
Rondônia	70.8%			21.4%
Roraima	71.2%			24.7%
Tocantins*	66.0%			17.2%
North Region	**70.0%**	**54.2%**	**63.4%**	**33.7%**
Alagoas	40.4%			34.3%
Bahia	37.9%			40.0%
Ceará	39.8%			40.4%
Maranhão	37.4%			24.0%
Paraíba	42.4%			34.7%
Pernambuco	40.2%			46.2%
Piauí	37.1%			22.2%
Rio Grande do Norte	46.0%			30.5%
Sergipe	29.2%			28.4%
Northeast Region	**39.2%**	**29.1%**	**26.6%**	**37.1%**
Distrito Federal**	73.2%			44.4%
Goiás	53.2%			36.4%
Mato Grosso do Sul	36.3%			22.5%
Mato Grosso	45.3%			28.1%
Middle-West Region	**40.4%**	**32.4%**	**29.1%**	**34.2%**
Espírito Santo	37.1%			40.4%
Minas Gerais	36.5%			35.7%
Rio de Janeiro	44.2%			55.3%
São Paulo	39.9%			44.6%
Southeast Region	**40.0%**	**32.8%**	**34.4%**	**44.7%**
Paraná	29.3%			33.2%
Rio Grande do Sul	31.4%			32.3%
Santa Catarina	32.8%			27.5%
South Region	**30.9%**	**23.6%**	**24.2%**	**31.6%**
Brazil	**40.4%**	**31.4%**	**47.1%**	**39.4%**

Note: The considered definition of potential unemployment rate is: (population minus children aged 14 and under minus persons aged 65 and over minus employment) divided by (population minus children aged 14 and under minus persons aged 65 and over). However, due to lack of data, there are the following restrictions: the population of both children aged 14 and under and of people aged 65 and over was not taken for Tocantins in 1981. 1991 and 1999 miss data by age group for the state level. The calculated potential unemployment rate is higher than the actual unemployment rate due to the fact that there were people employed or occupied in family businesses or farm who were considered unemployed.

Note: *From 1981 to 1991, employment data for Goiás is Goiás+Tocantins. To disaggregate these two states, the average percentage of Tocantins' employment share on Goiás' employment from 1992 to 2007 for the period 1981-1991 was assumed, and was subtracted the equivalent employment volume in Goiás employment. Employment was defined as the number of employed or occupied people in a paid professional occupation (IPEA). For details about the sectoral composition of employment across states, regions, and Brazil, see Chapters 3 and 5 below. In 1981, to split Tocantins' population from Goiás' population was assumed the proportion of the former state on the latter in 1980.

Note: **In 2007, due to lack of data, the proportion of children aged 14 or less and of people aged 65 or more in Goiás state was assumed for Distrito Federal in calculation of the latter states' potential unemployment rate.

2.3.6.2. Employment

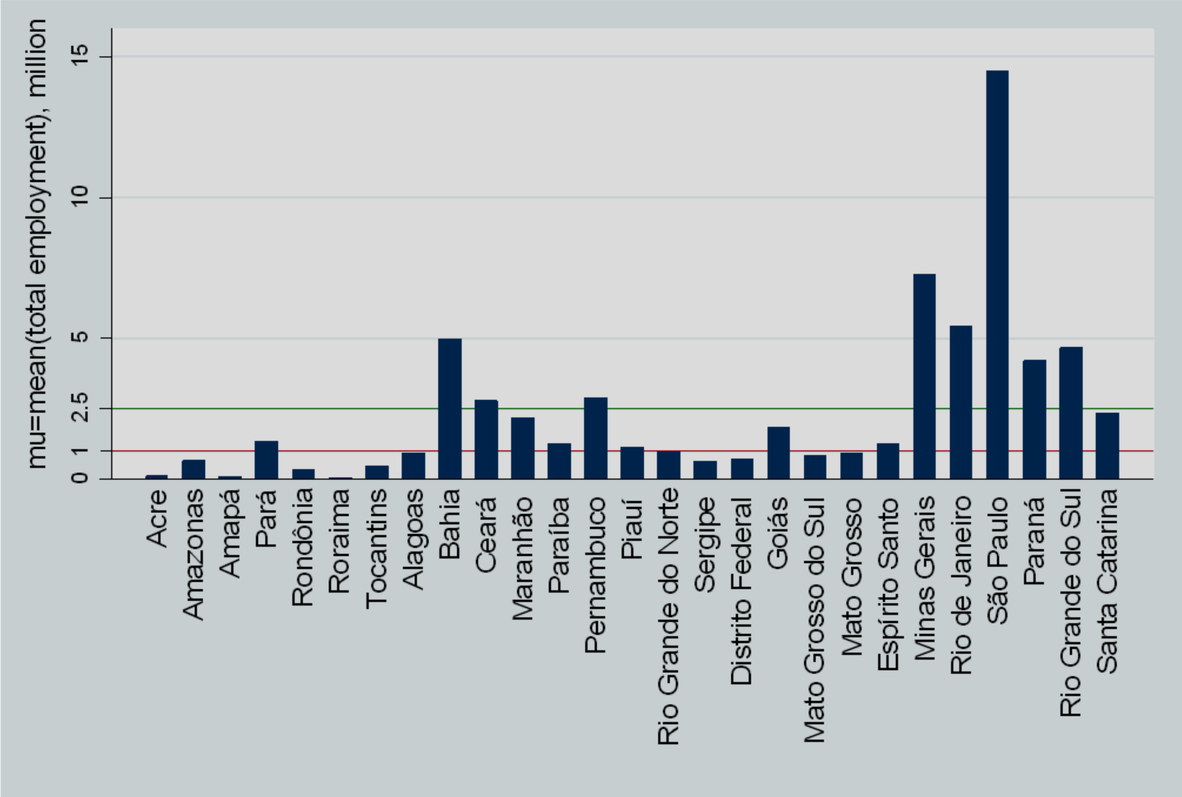
Figure 2.4 shows the states' average total employment[23] from 1981 to 2006. There are very few states that supplied a significant share of national employment. São Paulo was the state with the highest average employment (around 14.5 million out of the 65.2 million national average). The remaining 26 states can basically be slotted into three groups. The first group is a limited number of seven states that follow São Paulo state[24], namely (with their averages estimates, in millions, in brackets): Minas Gerais (7.3), Rio de Janeiro (5.5), Bahia (5.0), Rio Grande do Sul (4.7), Paraná (4.2), Pernambuco (3.0), and Ceará (2.8). This is a group of the usual southeast-south states and the three leading northeast states.

Two reference lines for 1 million and 2.5 million total employment averages were inserted in Figure 2.4. The second group of states, cannot reach a barrier of 2.5 million but are able to "break a barrier" of 1 million average total employment. These seven (with their averages estimates, in millions, in brackets) are: Santa Catarina (2.4), Maranhão (2.2), Goiás (1.9), Pará (1.4), Espírito Santo (1.3), Paraíba (1.3), and Piauí (1.2). For some of these states, these levels are a good achievement relative to their size. Finally, the third group is made up of the 12 states which cannot reach 1 million average of total employment; some of which "lack" economic activities, especially in the northern region of Brazil. These findings are consistent with the patterns shown by income and the other indicators.

[23] For employment, education, and Gini data for Tocantins state, there is no data from 1981 to 1990, a period for which a backast was done. To fill the missing years of 1991, 1994, and 2000 for all states as well as for Brazil, the average between the two neighbouring years of the missing years has been taken.

[24] However, the total employment average for these states is significantly smaller than that of São Paulo state. For example, the ratio between São Paulo's employment average and Minas Gerais's (which ranks second in Brazil in this indicator), is approximately 2.

Figure 2.4: Average Total Employment (million people): Brazil's States, 1981-2006



Note: Over the same period, the average total employed population in Brazil was approximately 65.2 million.

Following the discussion of employment, it is important to consider labour productivity. The results from Table 2.7 partially contradict those from Figure 2.4. Although northern states "lack" economic activity, most of them had productivity levels higher than those observed for the national level and also comparable with those of the states of the most developed regions. In 1991, for instance, Amapá and Roraima - two unimportant states in terms of their shares in national income - had the ratio between their labour productivity and the nation's equal to approximately 1.5 and 1.1, respectively. And also for 1991, this productivity ratio between Amapá and Rio de Janeiro as well as between Roraima and Rio Grande do Sul was equal to one. This performance, however, was not repeated in the years that followed for these two northern states because their employment growth rates were higher than their GDP growth rates.

The picture from the other regions is not surprising: overall, productivity levels are low in the northeast, and these levels are smaller than those observed at national level. In contrast, the southeast-south and centre-west[25] regions had high productivity levels. In particular, São Paulo and Rio de Janeiro had productivity levels much higher than those for the nation. These findings are consistent with the well-known regional disparity in Brazil (Ferreira, 2000; Azzoni, 2001; Mossi, Aroca, Fernández & Azzoni, 2003; Silveira-Neto & Azzoni, 2006).

Table 2.7: Evolution of Labour Productivity: States versus Brazil, 1980-2007

State	1980*	1991	2000	2007
Acre	20	12	11	11
Amazonas	29	28	25	18
Amapá	22	22	14	14
Pará	23	20	12	9
Rondônia	21	14	15	11
Roraima	22	16	11	12
Tocantins	4	4	4	9
North Region	**19**	**18**	**14**	**11**
Alagoas	8	7	7	8
Bahia	10	9	9	10
Ceará	6	7	6	7
Maranhão	5	4	3	6
Paraíba	6	6	7	8
Pernambuco	9	10	9	10
Piauí	4	4	4	5
Rio Grande do Norte	8	8	8	9
Sergipe	7	10	8	10
Northeast Region	**8**	**7**	**7**	**8**
Distrito Federal	68	32	34	51
Goiás	12	11	9	13
Mato Grosso do Sul	15	12	12	13
Mato Grosso	12	9	11	17
Middle-West Region	**18**	**15**	**14**	**20**
Espírito Santo	14	13	15	20
Minas Gerais	15	13	13	14
Rio de Janeiro	24	22	24	26
São Paulo	28	24	23	26
Southeast Region	**23**	**20**	**20**	**22**
Paraná	14	14	14	16
Rio Grande do Sul	17	16	16	17
Santa Catarina	16	15	16	18
South Region	**16**	**15**	**16**	**17**
Brazil	**17**	**15**	**15**	**17**

Note: Labour Productivity is GDP per unit of labour. For 1980, due to lack of employment data, the ratio between GDP of 1980 and Employment of 1981 was taken. This may be a problem for the results as states performed better than should be expected in the first column with data in this table. GDP data are in BR$ of 2000 (thousand).

[25] Distrito Federal stands with the growing and highest productivity levels. For example, the ratio between this state's productivity and the nation's grew from 2.13 in 1991 to 3 in 2007. This is not a surprise when considering the finding regarding the human capital for this state which plays a key role for states' income levels (for a quick understanding, see Figure 2.3 above).

2.3.7. Share of Manufacturing Employment in the Total Economically Active Population

Table 2.8 shows that from 1970 to 1995, the share of manufacturing employment in the total economically active population followed an inverted U-shaped curve in Brazil. This is also true for all regions and 20 states. However, the levels vary across states and regions. Given that the north, northeast and middle-west had the lowest shares of manufacturing in the total economically active population, this result indicates a skewed distribution of manufacturing employment, favouring southeast and south regions in Brazil. Indeed, in 1970, while southeast and south had 14 percent and 8 percent, respectively, in their shares of manufacturing employment in the total economically active population, the other regions had either 3 percent or 4 percent only. There was only a slight decrease in the gap of this indicator across regions because, in 1995, the southeast and south still had 11 percent and 12 percent, respectively, against either 4 percent or 5 percent only for each of the other regions.

These disparities are much clearer when considering the state-level data. In 1970, the states with the highest share of manufacturing employment in the total economically active population were (in this order): São Paulo, Santa Catarina, Rio de Janeiro, Rio Grande do Sul, Amapá, Minas Gerais, and Pernambuco. Of these states, four had much higher shares than the national average and there was a huge gap between these four states' indicator and those with the lowest. Eighteen states had the lowest shares (up to 4 percent only), in 1970, and were mostly north, northeast, and middle-west states.

In 1995, the states with the highest shares of manufacturing employment in the total economically active population were (in this order): Rio Grande do Sul, Santa Catarina, São Paulo, Paraná, Minas Gerais, Rio de Janeiro, Espírito Santo, Alagoas, and Amazonas. The first seven states are located in southeast and south regions while the latter two are northeast and north states, respectively. On the other hand, among the other 18 states, eight had either 1 percent, 2 percent or 3 percent (these are north and northeast states, excluding the middle-west state of Distrito Federal), and the remaining ten had either 4 percent or 5 percent (these are north and northeast states, excluding the middle-west states of Goiás, Mato Grosso do Sul, and Mato Grosso).

Therefore, overall, analysing all years in Table 2.8, the conclusion is that manufacturing (as percentage of total economically active population) is concentrated either in south and southeast states or in a limited number of seven states from the other regions, namely Amazonas, Amapá, Pará, Rondônia, Alagoas, Rio Grande do Norte, and Pernambuco. These results are consistent with the Marshal-Arrow-Romer (MAR) hypothesis which argues that agglomeration of economic activities in the regions arises due to knowledge externalities. These knowledge externalities, in turn, either boost or hurt growth of regions depending on whether the agglomerated activities are specialised or diverse (Glaeser, Kallal, Scheinkman & Shleifer, 1992; Henderson, Kuncoro & Turner, 1995; Combes, 2000; McCann, Mameli & Faggian, 2008; De Groot, Poot & Smit, 2009).

Table 2.8: Evolution of the Shares of States and Regions' Manufacturing Employment in the Total Economically Active Population, 1970-1995

State	1970	1975	1980	1985	1995
Acre	1%	1%	2%	2%	2%
Amazonas	4%	7%	13%	12%	7%
Amapá	7%	7%	7%	5%	1%
Pará	4%	5%	7%	5%	5%
Rondônia	4%	5%	5%	5%	5%
Roraima	3%	4%	3%	2%	2%
Tocantins	1%	1%	1%	1%	1%
North Region	**4%**	**5%**	**7%**	**6%**	**5%**
Alagoas	4%	6%	7%	6%	7%
Bahia	3%	3%	5%	4%	2%
Ceará	3%	4%	6%	6%	4%
Maranhão	1%	1%	2%	2%	1%
Paraíba	3%	4%	5%	4%	4%
Pernambuco	6%	7%	8%	7%	5%
Piauí	1%	2%	2%	2%	2%
Rio Grande do Norte	4%	5%	8%	6%	4%
Sergipe	4%	5%	7%	6%	3%
Northeast Region	**3%**	**4%**	**5%**	**5%**	**3%**
Distrito Federal	2%	4%	3%	3%	2%
Goiás	3%	4%	5%	5%	4%
Mato Grosso do Sul	3%	4%	5%	4%	4%
Mato Grosso	2%	2%	5%	4%	5%
Middle-West Region	**3%**	**4%**	**5%**	**4%**	**4%**
Espírito Santo	5%	7%	9%	8%	7%
Minas Gerais	6%	7%	10%	9%	8%
Rio de Janeiro	12%	13%	14%	11%	7%
São Paulo	20%	23%	26%	21%	14%
Southeast Region	**14%**	**16%**	**19%**	**16%**	**11%**
Paraná	5%	7%	9%	8%	8%
Rio Grande do Sul	10%	14%	16%	15%	15%
Santa Catarina	14%	17%	22%	19%	15%
South Region	**8%**	**12%**	**14%**	**13%**	**12%**
Brazil	9%	11%	13%	11%	8%

2.3.8. Size of Manufacturing Firms

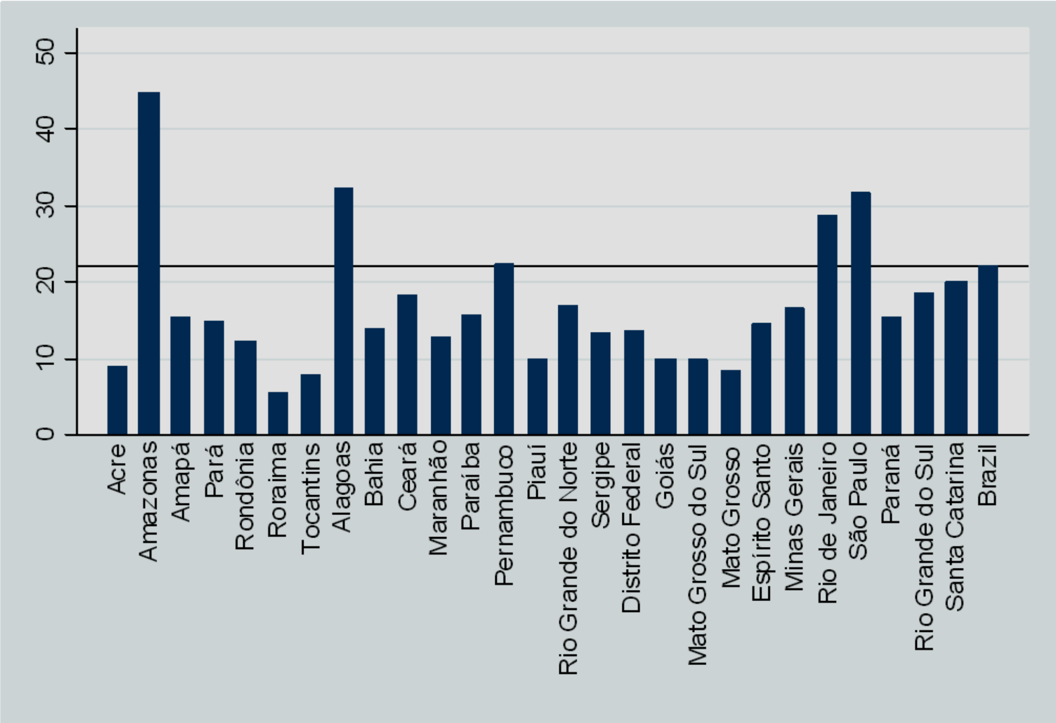
Figure 2.5 shows that the average size of manufacturing firms from 1970 to 1995 was 22 employees per manufacturing firm in Brazil (see reference line in Figure 2.5). However, there were some states that surpassed Brazil's average size. These were (with their employees per manufacturing firm in brackets and in this order): Amazonas[26] (44.83), Alagoas (32.33), São Paulo (31.67), Rio de Janeiro (28.67), and Pernambuco (22.5). Over the same period, these states, excluding Alagoas, were among those with the highest shares in national income and highest per capita income in their regions. The remaining 22 states are divided in two groups.

The first group involves 13 states where manufacturing firms average between 15.5 and 20.17 employees. The second group includes nine states that had 15 employees or fewer per manufacturing firm and were located in north, northeast, and middle-west regions. Overall, these results show that in Brazil firms' competition is only slight as for around only half the states (not for at least about 20 states), the average manufacturing firm is small (using a cut-off of fewer than 15.5 employees per manufacturing firm). As a result, Porter's hypothesis is only partially supported. It claims that within the regions, competition of specialised firms, rather than competition of diverse firms, promotes regional growth (Glaeser, Kallal, Scheinkman & Shleifer, 1992; Henderson, Kuncoro & Turner, 1995; Combes, 2000; McCann, Mameli & Faggian, 2008; De Groot, Poot & Smit, 2009)[27].

[26] The Amazonas performance in this indicator is due to a manufacturing belt in this state's capital, Manaus.

[27] This theoretical support for both Marshal-Arrow-Romer (MAR)'s and Porter's theories of knowledge externalities will be tested by econometric models in Chapter 3.

Figure 2.5: The Average Size of Manufacturing Firms: Brazil and its States, 1970-1995



Note: Manufacturing firm size is the ratio between manufacturing employment and manufacturing establishments. There are only six observation points for manufacturing establishments: 1907, 1970, 1975, 1980, 1985, and 1995. To hold the regularity of a 5-year observation, an interpolation for manufacturing establishments for 1990 was done before the calculations of the manufacturing firm sizes. Due to a lack of disaggregated data, the first year of the series was excluded. Tocantins misses manufacturing employment in 1990 for which a backcast was done using this state's manufacturing employment from 1992 to 2007.

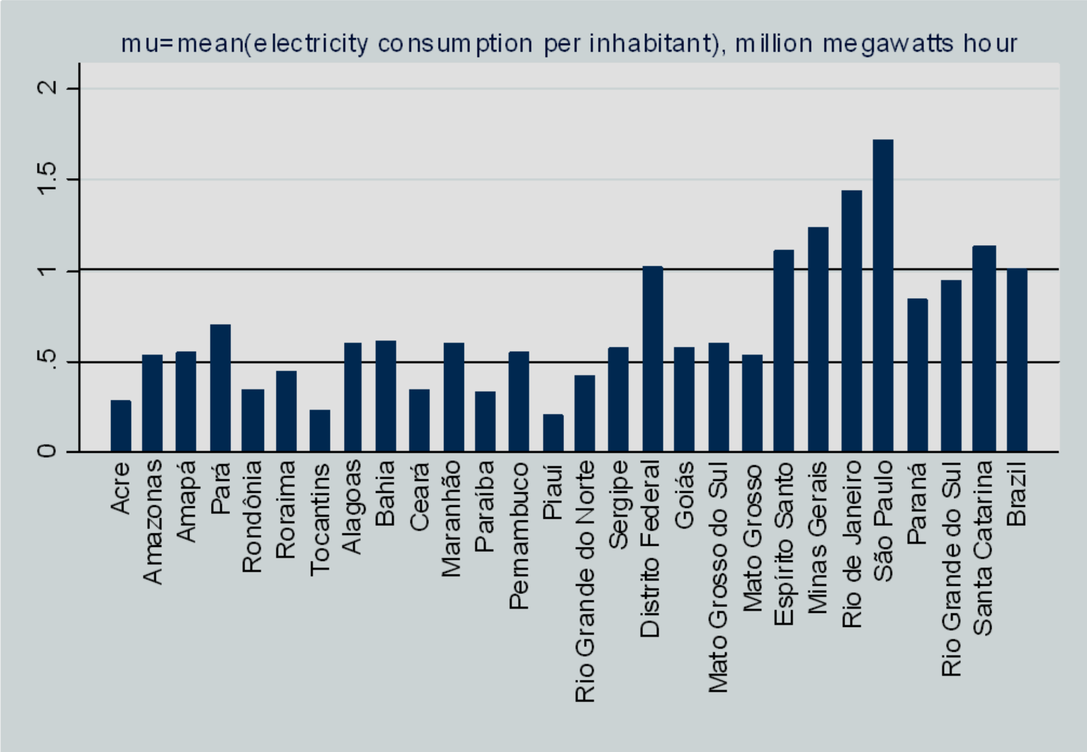
2.3.9. Electricity Consumption per inhabitant

A real measure of the level of economic activity that is free from possible "noises" is the average electricity consumption per inhabitant. As expected, the role of São Paulo in Brazil is clearly shown in Figure 2.6 below, where São Paulo has the highest average electricity consumption per inhabitant (approximately 1.72 million megawatts hour) from 1961 to 2004. This electricity consumption level is consistent with São Paulo's employment levels and its most developed manufacturing (sectoral data not shown here; for details, see Chapters 3 and 5 below) in Brazil.

As with the other indicators, there are few states that match São Paulo state. The second ranked state, Rio de Janeiro, fell quite far behind: the average electricity consumption per inhabitant for Sao Paulo state is approximately 1.19 times that of Rio de Janeiro. Minas Gerais, Santa Catarina, Espírito Santo, and Distrito Federal follow. These six states were higher than the national average, which was approximately 1.012 million megawatts hour (see the reference line for this value in Figure 2.6). All the other states had their averages below the national average.

However, there are 13 states that were able to break a barrier of 0.5 million megawatts hour per inhabitant, which are (approximated values are in brackets) Rio Grande do Sul (0.94), Paraná (0.84), Pará (0.70), Bahia (0.609), Alagoas (0.602), Maranhão (0.60), Mato Grosso do Sul (0.597), Sergipe (0.581), Goiás (0.575), Pernambuco (0.55), Amapá (0.542), Amazonas (0.538), and Mato Grosso (0.535). The latter three states (Amapá, Amazonas, and Mato Grosso) were closely followed by the other eight states which had much lower (fewer than 0.5 million megawatts hour) averages as a consequence of their very low levels of economic activity - associated with their low levels of income, even though those three latter states (Amapá, Amazonas, and Mato Grosso) have had high population density, especially since 1970 (see Table 2.4). The states with averages fewer than 0.5 million megawatts hour are essentially located in a contiguous north-northeast land area. Unless a huge shock occurs, firms will not choose these states as a location for their businesses. South America’s map of lights at night supports these results (see Figure 2.7). In sum, this indicator supports the argument of the role of initial conditions: firms tend to settle where others have already established; in turn, this attracts more companies to those locations (Krugman, 1991a, 1991b, 1995; Hanson, 2003).

Figure 2.6: Average Electricity Consumption per Inhabitant: Brazil and its States, 1961-2004



Note: Mato Grosso and Mato Grosso do Sul: the data for these two states are aggregated from 1961 to 1975. Since the latter state has much higher electricity consumption than the former, to fill the missing values over this period, three steps procedure was employed: first, it was imputed the electricity consumption of Mato Grosso into Mato Grosso do Sul's row. Second, it was filled the Mato Grosso cells by assuming that the average consumption share of this state over the period 1976 to 2004 was also observed from 1961 to 1975. Finally, it was replaced values for Mato Grosso do Sul by the difference between the imputed ones for this state and those estimated for Mato Grosso in the previous step. Regarding Tocantins, to fill the missing values from 1961 to 1988, it was assumed that the average share of this state's electricity consumption in Goiás consumption between 1989 and 2004 was also observed for the period from 1961 to 1988 to which a backast was done; then it was adjusted Goiás consumption over the period from 1961 to 1998 by taking from it the estimated electricity consumption for Tocantins. These adjustments kept the national electricity consumption unaltered. There is a missing value in 1997 for Distrito Federal which was filled by the average electricity consumption between 1996 and 1998. The data for electricity consumption is annual from 1961 to 2004; however, there are missing years for population data. Annual population data was obtained by interpolation for the missing years. The interpolation obeyed the pattern showed by population data from the sources.

Figure 2.7: South America at night

Note: The light in the south of Brazil is from Brazil's southern neighbour countries and that in the northwest is also from Brazil's neighbours in that region, not from Brazil's northern states (for details of countries' location in South America, see Figure 1.1). The image was taken on October 23, 2000 and re-published in June 2011 (http://earthobservatory.nasa.gov/IOTD/view.php?id=896; http://geology.com).
Source: NASA; http://geology.com.

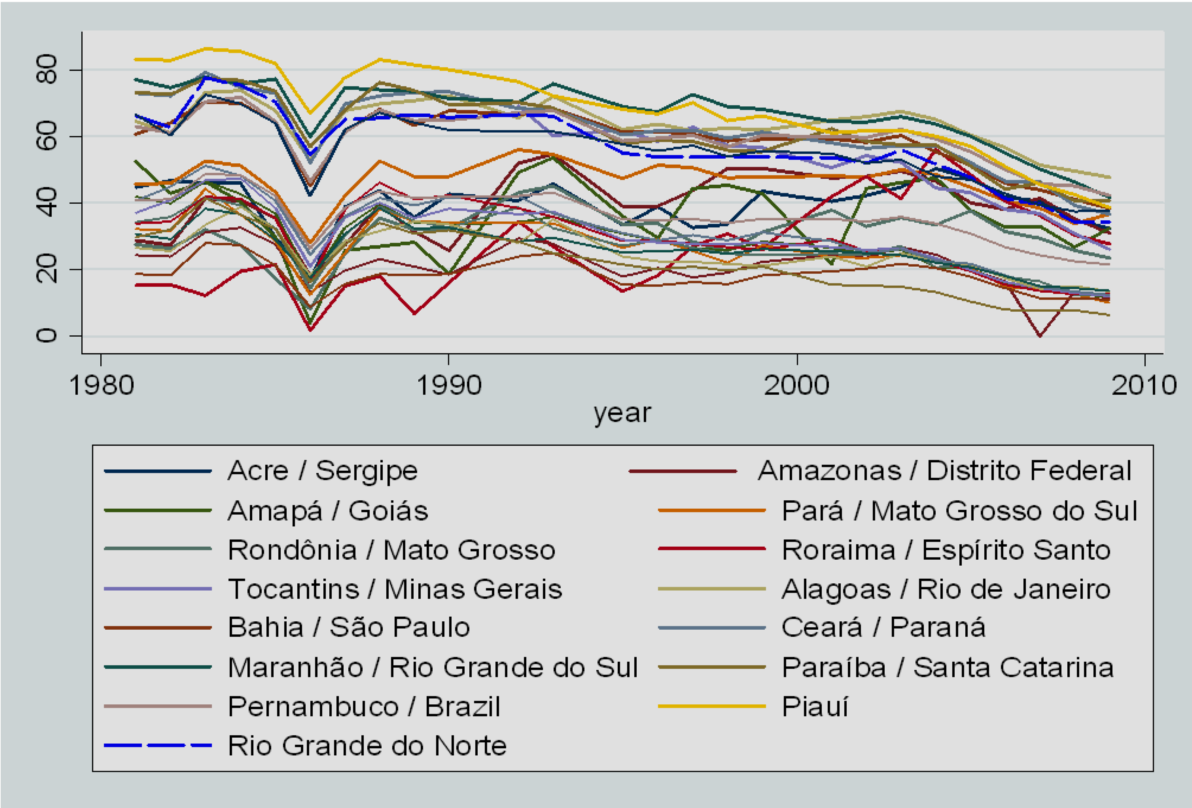
2.3.10. Poverty Rates[28]

Poverty rates are considered as measures for a negative home market effect (or deprivation). Figure 2.8 shows that all states in Brazil have reduced their poverty rates which indicates an improvement in local market potential. However, when considering the poverty level of Brazil as a reference, there are basically three groups of states. The first group involves northeast states: Alagoas, Piauí,

[28] According to IPEA, the poverty rate is measured as "the percentage of total population with domicile per capita income below the poverty line. The poverty line considered is the double of the extreme poverty line, an estimate of the value of a food basket with the minimum calories needed for an adequate individual living based on FAO and WHO recommendations" (IPEA). There are missing data for 1991, 1994, and 2000 for which interpolation was done. Until 1991, Tocantins has missing values.

Rio Grande do Norte, Sergipe, Bahia, Ceará, Maranhão, Paraíba, and Pernambuco. The second group includes northern states: Tocantins, Amapá, Pará, and also Acre since 1999. These two groups of states had poverty levels higher than the nation's average. The third group includes southeast-south states: Espírito Santo, Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Rio Grande do Sul, and Santa Catarina as well as the centre-west states: Distrito Federal, Goiás, Mato Grosso do Sul, and Mato Grosso. The states in the third group had much lower poverty rates than the national average. This finding shows a vicious circle: poverty is regionalised and its incidence is prominent in regions with the lowest levels in both income and economic activity; these regions are less urbanised (more agricultural), and have very low levels of human capital, which in turn reinforces poverty through a negative effect of low human capital on income.

Figure 2.8: Poverty Rates (Percentage): Brazil and its States, 1981-2009



2.3.11. Inequality

Figures 2.9 and 2.9a-2.9e show a variation in inequality across states and regions[29] versus Brazil, annually from 1981 to 2006. Looking at Brazil's figure, even though since 2000 the inequality indicator started to fall reaching its lowest levels in the time series, it is still high, at 0.56. The overall picture is that in all Brazilian regions there are states that presented inequality levels comparable with the nation's towards the end of the period. North and northeast regions had comparable levels of inequality and since the middle of the 1990s, states "clustered" in the same range of inequality, that is, inequality dispersion became smaller.

In contrast, in the south, all states showed inverted U-shaped inequality behaviour: inequality rose from 1981 to 1990 and then started to fall. Paraná and Rio Grande do Sul presented inequality levels higher than those of Santa Catarina. In the centre-west and southeast, when excluding the former region Distrito Federal, which had a rising inequality over the series, states within the regions had convergent inequality levels. The inequality levels to which states (particularly in the southeast region) and Brazil were converging were similar. In sum, inequality within Brazilian states remained high despite falling in the last seven years of the series, at above 0.50 for Brazil and for 26 states[30].

[29] The Gini Index is one of standard measures for inequality, ranging from zero to one. The higher the Gini Index, the higher the country or region's inequality.
[30] The only state with the smallest and more "acceptable" inequality is Santa Catarina.

Figure 2.9: Gini Indexes: Brazil, its States and Regions, 1981-2006

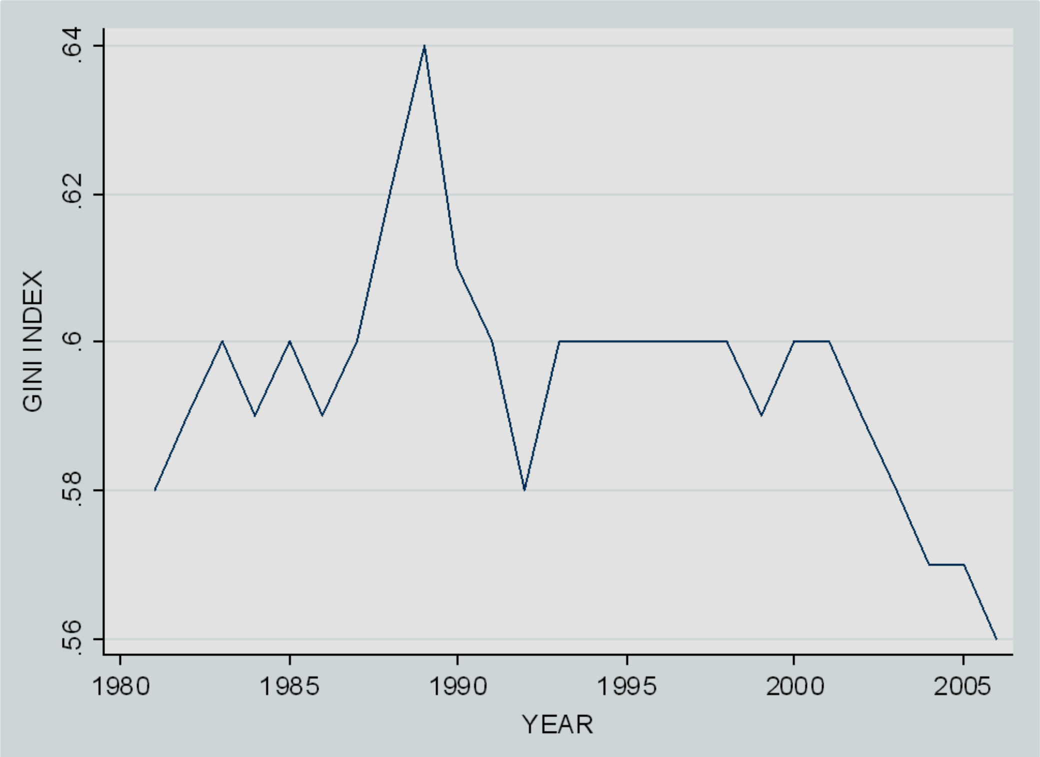


Figure 2.9a: North Region

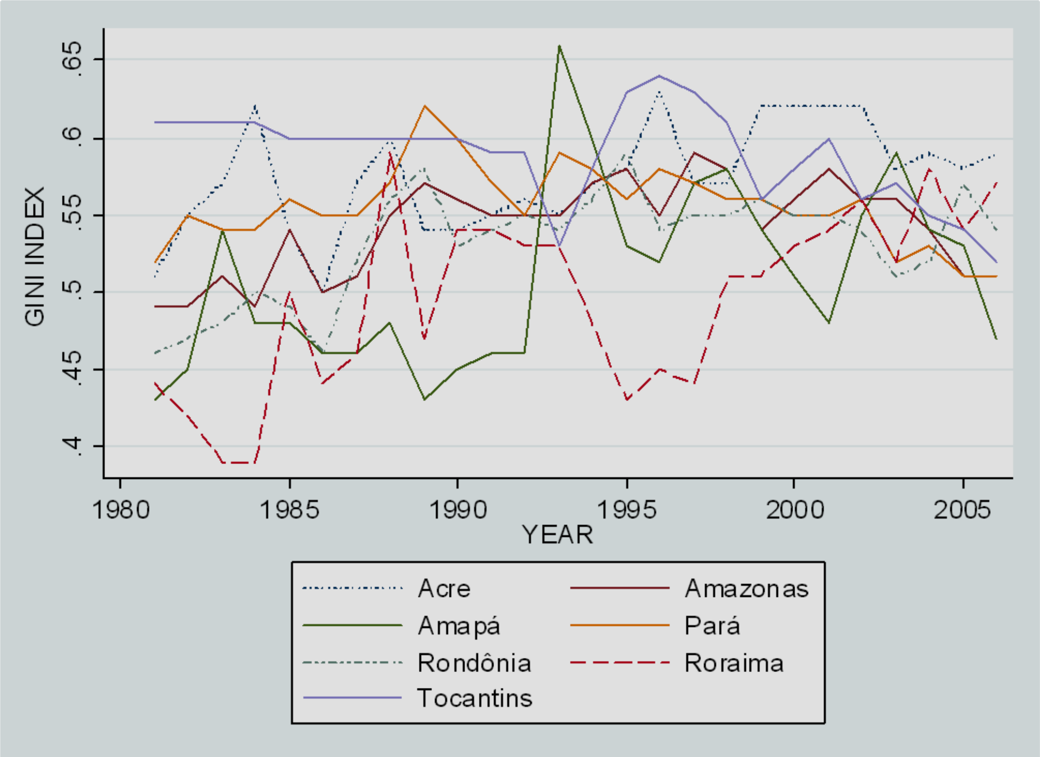


Figure 2.9b: Northeast Region

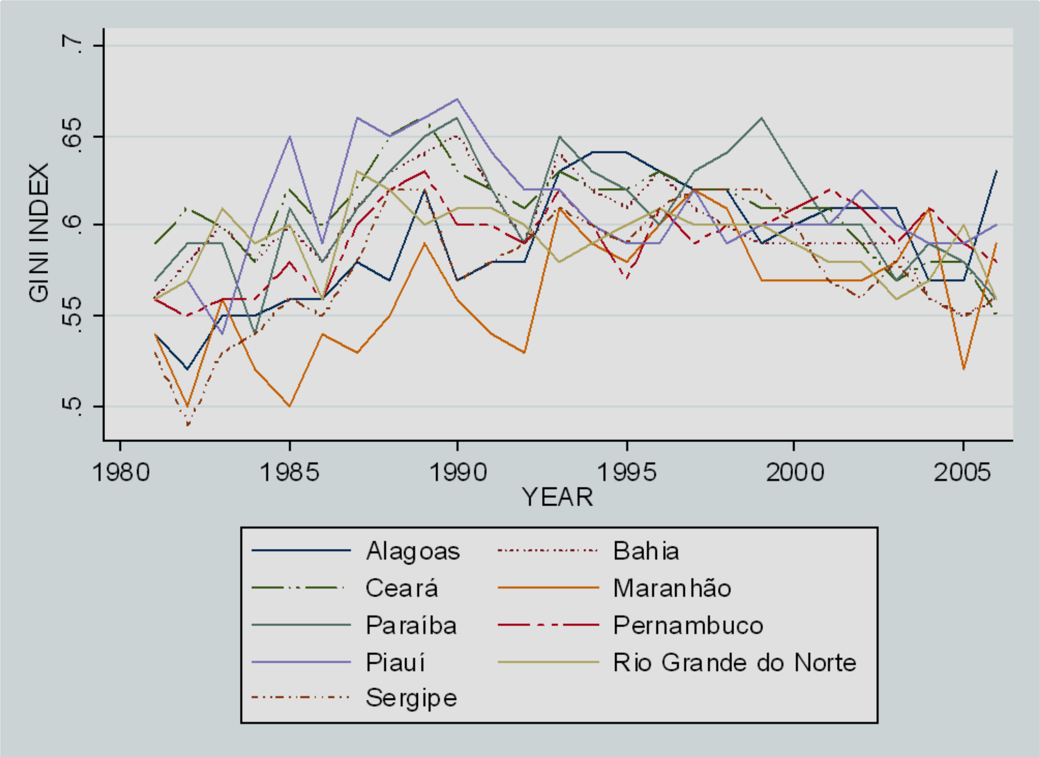


Figure 2.9c: Centre-west Region

Figure 2.9d: Southeast Region

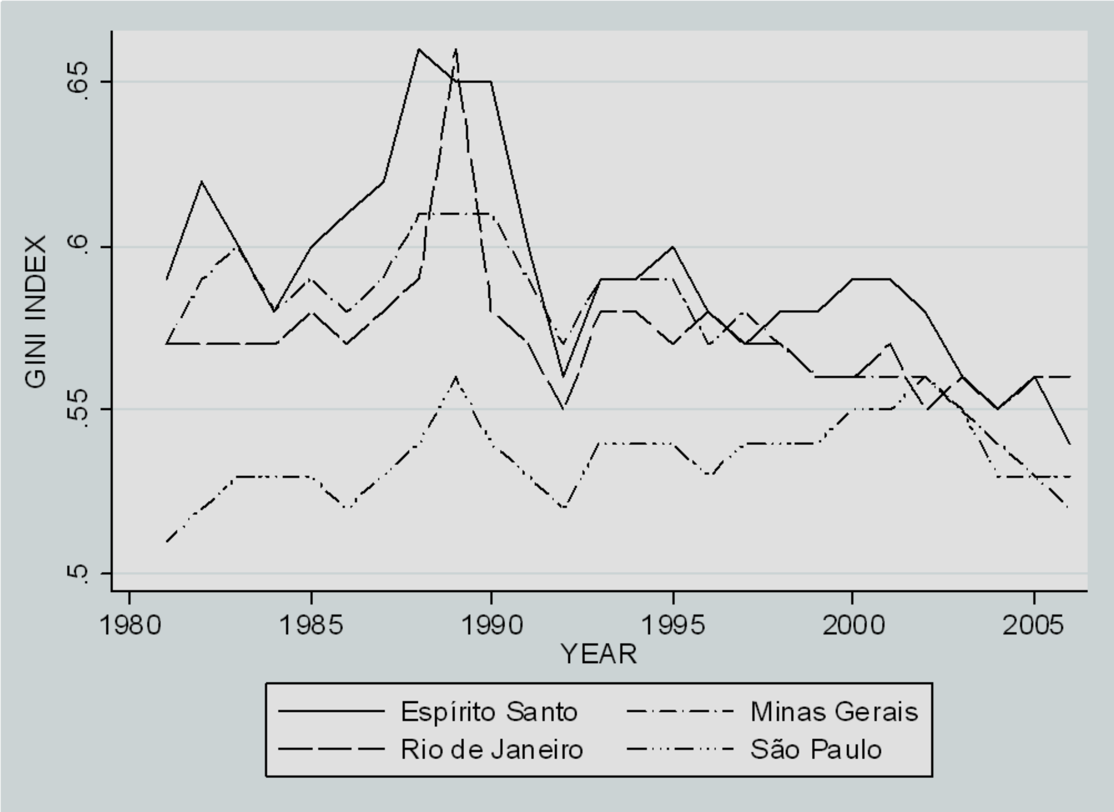
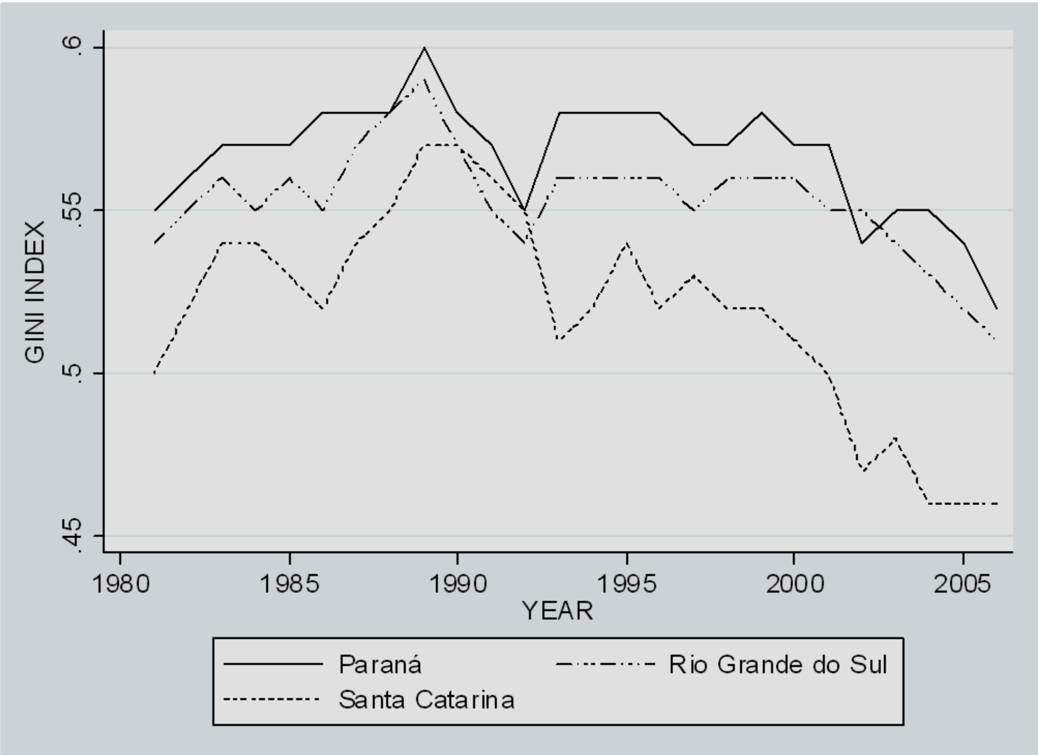


Figure 2.9e: South Region



2.4. Conclusion

This chapter assessed socio-economic regional data, mostly from 1920 (or 1939) to 2008 in Brazil. The patterns observed are consistent with the new economic geography and spatial economic theory (Anselin, 1995; Krugman, 1991a, 1991b; Brakman, Garretsen & van Marrewijk, 2001; Nazara & Hewings, 2004; LeSage & Pace, 2009), as opposed to the mainstream neoclassical a-spatial economic theory (Sandroni, 1994[31]; Varian, 1992). Although no structural model has been estimated in this chapter, three main conclusions can be drawn from the descriptive data analysis.

First, the path followed by Brazil has changed dramatically over the 20th century. Since the middle of the 20th century, the descriptive spatial data analyses indicate that there is no clear sign of a catch-up of the lagged regions as the improvement of these regions is only slight over the analysed period, and the finding seems to indicate a redistribution of income within the regions only.

Secondly, the concentration of economic activity and the resulting increase in regional income disparities in Brazil worsened from 1939 to 1950; then slightly improved from 1950 to 1960. From the latter year to the decades that followed, it worsened again. The decades 1960s and 1970s coincide with industrialisation in Brazil. The industrialisation process in Brazil was perverse because it concentrated investments in the southern regions, in part due to political power of these regions, which then stimulated in-migration of skilled workers from other regions searching for higher wages in the southern regions. This internal migration left the lagging regions with much less skilled labour. As a consequence, the spatial pattern of the Brazilian economy has changed, and the descriptive data show that there is no clear sign of a reversal because the change towards regional balance in Brazil is insignificant.

Thirdly, it was found that despite their coastal location, many states are still poor because the advantages of their location are more than offset by the disadvantages of smaller (lack of) home market as well as poor access to the largest national markets due to distance and infrastructure underdevelopment. The next four chapters (Chapters 3 to 6) empirically test the extent to which these findings are true using the new economic geography and spatial economic theory

[31] See, for example, entries for economics scholars David Ricardo (1772-1823), Carl Menger (1840-1921), and William Stanley Jevons (1835-1882).

as theoretical frameworks. These chapters contribute to the economic and regional science literature by discussing hitherto unexplored issues and methodologies regarding the evolution of regional economies in Brazil, an important emerging country in the world, therefore a relevant experiment in economic and regional sciences. In short, these chapters provide new insights for the understanding of some of the established laws (or theories) and those of which their robustness is still being tested.

CHAPTER 3 – AGGLOMERATION EXTERNALITIES AND 1981-2006 REGIONAL GROWTH IN BRAZIL

3.1. Introduction

Regional income convergence is a key area of debate among economists (Barro & Sala-i-Martin, 2004). If regions with a lower real income per capita grow faster than those with a higher real income per capita, economic growth can reduce interregional income inequality within a country. A standard model that explains the role of knowledge externalities in such convergence is the Glaeser *et al.* (1992) model (Glaeser *et al.*, 1992; Henderson *et al.*, 1995; Combes, 2000). Using a simple Cobb-Douglas production function, this model shows that the growth rate of employment is positively related to growth of technological knowhow and negatively related to the growth rate of wages. Assuming that wages are spatially equalised, employment growth differentials are then explained by the region-specific impact of knowledge externalities due to regional specialisation, competition and diversity, based on the theories of Marshall-Arrow-Romer (MAR), Porter, and Jacobs. The Glaeser *et al.* (1992) model aims to quantify each of these effects separately.

The theories advocated by Romer (1986) and Porter (1990) argue that the concentration of specialised industries in an area tends to benefit growth in that area because knowledge externalities enhance the productivity of clustered firms.[32] In contrast, Jacobs (1969) argues that a cluster of specialised industries in a specific area can actually reduce growth because specialisation tends to inhibit competition among firms, thereby limiting regional growth potential. When competition among firms is limited, there is no place for diversity and the effects of knowledge will be impeded insofar as not reaching out to broader sectors. Thus, the expected relationship between manufacturing industry employment growth and dynamic knowledge externalities and competition is ambiguous because it depends on whether the adopted perspective is that of MAR, Porter or Jacobs (De Groot *et al.*, 2009, p. 264).

Despite the international acceptance and utilisation of the Glaeser *et al.* (1992) empirical model of the impact of these knowledge externalities, the model has not yet been applied to the case of Brazil. The present chapter therefore estimates the

[32] These theories are summarised in Glaeser *et al.* (1992), Henderson *et al.* (1995) and De Groot et al. (2009).

model for Brazil. There has been significant income convergence among Brazilian states in recent decades. This chapter focuses on the role played by knowledge externalities on growth in manufacturing industry employment and, by implication, real incomes over the period 1981 to 2006. While there are many studies of economic growth, convergence and regional income inequality in Brazil, none to date has adopted the Glaeser *et al.* (1992) model. A number of common empirical estimation approaches are employed, based on cross-section data, panel data and pooled-periods cross-section data.

The chapter is structured as follows. The following section discusses the background literature on regional income convergence in Brazil. Section 3.3 provides a descriptive analysis of regional growth in Brazil. Section 3.4 describes the data used and their sources. Section 3.5 reports and discusses the regression results. Lastly, section 3.6 provides concluding comments.

3.2. Literature

Regional income convergence in Brazil has been widely studied. The literature has taken four approaches. The first approach estimates regression models using cross-section data. These regressions utilise the growth of per capita income over a given period in a region as the dependent variable and assume that this is negatively related to initial per capita income. This negative relationship is referred to as beta convergence (Azzoni, 1999, 2001, 2003; Silveira-Neto & Azzoni, 2006; Resende, 2011). Other predictors of growth are investment, education and the size of labour force (Ferreira, 2000). Other regressions explain per capita income of Minimum Comparable Areas (MCA) in terms of education, demographic indicators and the ratio of public expenditure over revenue of these areas (Rangel *et al.*, 2008). These studies find evidence in favour of absolute income convergence across Brazil's states.

Additionally, an analysis of economic growth between 1970 and 2000 explains growth in terms of a range of demographic and socioeconomic indicators (essentially education, per capita capital, and the crime rate). This study, by Coelho & Figueiredo (2007), focuses on municipalities' per capita income growth. Brauch & Monasterio (2007) focus on income convergence by means of Exploratory Spatial Data Analysis (ESDA) of the income shares of Minimum Comparable Areas. In an application of Quah's (1997) methodology to examine of distribution of income across Brazilian regions, Gondim *et al*. (2007) find convergence only within macro regions. Indeed, Magalhães *et al.* (2005, p. 17) point out that:

> (…) it is possible to infer from the results in hand that, although some convergence among states is taking place, it seems to be more of a regional phenomenon or perhaps some type of club convergence than a global convergence process. States like São Paulo would be a dominant force in one club while the Northeast states would form a second group or club.

In other words, this study and several others (e.g. Resende, 2011 and references therein) found clubs of convergence – a situation in which rich states and poor states converge within their macro regions, but in which interregional income disparity actually increases.

Silva & Silveira Neto (2007) apply Hanson's (1998) model and evaluate the role of knowledge externalities through the estimation of manufacturing industry employment growth across Brazilian states from 1994 to 2002. Manufacturing industry employment growth is regressed against the following variables: average wage per worker, relative size of establishments, forward/backward linkages, agglomerations, manufacturing diversity, and distance. Market linkages and manufacturing diversity are both found to be positively associated with growth of manufacturing industry employment, thus confirming Jacobs' and Porter's theories, while rejecting MAR theory.

The second approach to investigate regional income convergence is to use panel data analysis. The only study that uses this approach is Azzoni *et al.* (2000) who explain per capita income by geographic variables, labour force and human capital variables. They find conditional income convergence across Brazilian states.

The third approach applies time-series data analysis. For example, Barossi-Filho & Azzoni (2003) use a sample of 20 Brazilian states to study the convergence of state per capita GDP in terms of national time dummies, structural breaks and the lag of income. After performing unit root tests, they find state-level convergence within Brazilian macro-regions.

The second and third approaches described above are rather rare in the literature regarding regional income convergence in Brazil. Finally, a fourth approach is simply descriptive and measures the inequality pattern in terms of income inequality indicators or national income shares for selected time periods, rather than by means of estimating structural models. This approach finds either states' income convergence at a national level or states' income convergence only within macro-regions (Azzoni, 1997a; Andrade & Serra, 2001; Gomez, 2002; Mossi et al., 2003; and Rolim, 2008).

This chapter estimates Glaeser *et al.'s* model. The original Glaeser model is represented by the following equation:

$$\alpha \log\left(\frac{l_{t+1}}{l_t}\right) = -\log\left(\frac{w_{t+1}}{w_t}\right) + \log\left(\frac{A_{national,t+1}}{A_{national,t}}\right) + g(\text{specialization, competition, diversity, initial conditions}) + e_{t+1} \qquad (3.1)$$

On the left side of the equation (3.1), *l* is manufacturing industry employment. Taking two points of time, and assuming α=1, the dependent variable is manufacturing industry employment growth between period *t* and *t*+1. This is explained in terms of wage (*w*) growth (negatively), the national growth of technology (*A*), and by the *g* function which captures specialisation, competition, and diversity externalities, and initial conditions (e.g. human capital level or other variables considered relevant by the researcher); and lastly, e_{t+1} is a residual.

Glaeser *et al.*'s model was first used in an attempt to explain regional economic growth and regional employment convergence by dynamic externalities using manufacturing industry employment data in the US (Glaeser *et al.* 1992; Henderson *et al.* 1995) and manufacturing industry employment and services employment data in France (Combes, 2000) before it became more widespread. De Groot, Poot and Smit (2009) summarise the international literature covering the period 1997-2006. Based on meta-analysis – which consists of combining all empirical evidence in which at least one of the first three elements of the *g* function of equation (3.1) is tested for statistical significance – they found 322 articles that analyse dynamic externalities and that cited either Glaeser *et al.* (1992) or both Porter (1990) and Jacobs (1969).

Of these 322 articles, only 31 had a quantitative approach that was sufficiently similar to Glaeser *et al.* (1992) to permit a meta-analytic quantitative comparison and, together, these 31 articles yielded a total of 393 estimates. A summary of the significance tests conducted in the international literature is reproduced in Table 3.1. This table shows counts and percentages for each result in the following classification: “negative significant”, “negative insignificant”, “positive insignificant” and “positive significant” under the following three elements of agglomeration externalities: specialisation, competition and diversity.

Table 3.1: Counts of Statistical Significance of Knowledge Externalities in Empirical Estimates of the Glaeser et al. (1992) Model

	SPECIALISATION		COMPETITION		DIVERSITY	
	Count	%	Count	%	Count	%
Negative significant	60	37	16	20	17	11
Negative insignificant	33	20	13	16	40	26
Positive insignificant	16	10	19	24	37	24
Positive significant	53	33	31	39	58	38
Total	162	100	79	100	152	100

Source: De Groot, Poot & Smit (2009, p. 269)

On balance, Table 3.1 indicates that there is slightly more evidence to suggest the impact of specialisation is more negative than positive. In contrast, the evidence, on average, supports more of a positive effect for competition and diversity externalities. Nonetheless, there is considerable uncertainty in the empirical evidence given the percentage of statistically insignificant results. This varies between 32 percent in the case of competition and 77 percent in the case of diversity.

3.3. Growth in the regions of Brazil

This section first describes manufacturing employment growth and convergence among Brazilian regions. Secondly, it summarises some Brazilian economic trends since the 1980s and considers the association between manufacturing employment and per capita GDP in Brazil and in its macro-regions[33].

3.3.1. Convergence in Brazil: A reassessment

In line with the previous convergence literature summarised in Section 3.2, the following questions are asked: (1) Do Brazilian states grow at the same rate? (2) If the answer to the previous question is no, do they show manufacturing industry employment convergence from 1981 to 2006? These two questions are answered based on Table 3.2.

[33] It is considered per capita GDP rather than GDP itself, because the former is an indicator of labour productivity, which captures production externalities.

Table 3.2: Evolution of Manufacturing Industry Employment in Brazil, 1981-2006[34]

Macro-Region	State	SSR81	SSC81	SSR06	SSC06	AVAMEGS	AVAMEGR
North	Acre	3.22%	0.07%	2.97%	0.21%	7.13%	
	Amazonas	43.65%	0.89%	22.82%	1.59%	4.86%	
	Amapá	2.19%	0.04%	1.42%	0.10%	5.73%	7.46%
	Pará	43.15%	0.88%	62.19%	4.33%	8.92%	
	Rondônia	7.27%	0.15%	9.24%	0.64%	8.42%	
	Roraima	0.53%	0.01%	1.36%	0.09%	11.25%	
		100.00%		100.00%			
Northeast	Alagoas	5.63%	0.86%	3.77%	0.62%	1.27%	
	Bahia	22.07%	3.37%	21.73%	3.60%	2.82%	
	Ceará	25.54%	3.90%	26.84%	4.44%	3.08%	
	Maranhão	4.81%	0.74%	8.66%	1.43%	5.23%	2.88%
	Paraíba	6.15%	0.94%	8.86%	1.47%	4.34%	
	Pernambuco	23.73%	3.63%	15.18%	2.51%	1.09%	
	Piauí	2.78%	0.43%	4.55%	0.75%	4.84%	
	Rio Grande do Norte	5.26%	0.80%	6.26%	1.04%	3.58%	
	Sergipe	4.03%	0.62%	4.15%	0.69%	3.00%	
		100.00%		100.00%			
Centre-West	Distrito Federal	6.74%	0.18%	9.25%	0.50%	6.64%	
	Goiás	54.91%	1.46%	52.41%	2.81%	5.19%	5.38%
	Mato Grosso do Sul	26.30%	0.70%	17.27%	0.93%	3.69%	
	Mato Grosso	12.05%	0.32%	21.08%	1.13%	7.61%	
		100.00%		100.00%			
Southeast	Espírito Santo	1.85%	1.18%	3.37%	1.69%	3.99%	
	Minas Gerais	12.09%	7.71%	21.75%	10.89%	3.94%	1.59%
	Rio de Janeiro	18.11%	11.56%	12.01%	6.01%	-0.05%	
	São Paulo	67.94%	43.35%	62.87%	31.49%	1.28%	
		100.00%		100.00%			
South	Paraná	26.24%	4.23%	31.55%	6.47%	4.26%	
	Rio Grande do Sul	48.01%	7.74%	39.76%	8.15%	2.77%	3.52%
	Santa Catarina	25.76%	4.15%	28.70%	5.89%	3.96%	
		100.00%		100.00%			
	Brazil		99.91%		99.46%	2.56%	2.56%

Notes: SSR81 and SSC81 are States' Manufacturing Industry Employment Regional Shares and States' Manufacturing Industry Employment Country Shares in 1981, respectively, which were also calculated for 2006; RS1981 and RS2006 are Regional Manufacturing Industry Employment Country Shares; AVAMEGS=Annual Average of Manufacturing Industry Employment Growth at the State level=100*{[Natural Logarithm (Manufacturing Industry Employment in 2006/Manufacturing Industry Employment in 1981)]/25}; AVAMEGR=Annual Average of Manufacturing Industry Employment Growth at the Macro regional level=100*{[Natural Logarithm (Manufacturing Industry Employment in 2006/ Manufacturing Industry Employment in 1981)]/25}. The total values of the second and fourth columns are less than 100% due to data deficiencies discussed in the text.
Source: IPEADATA.

[34]This chapter focuses on manufacturing employment. It is worth mentioning that tertiary sector accounts for a significant share of employment. In the data used, the tertiary sector, which is aggregation of sectors of commerce, financial sector, services (not disaggregated further), and transportation & communications, accounted for about 60 percent of Brazil's employment in 2006. These components of tertiary sector are considered individually in the calculation of diversity externality, which is one of the key independent variables of the models estimated in this chapter. The components of tertiary sector are also considered in Chapter 5 in the decomposition of the right-hand side of identities of both classic and spatial shift-share methods.

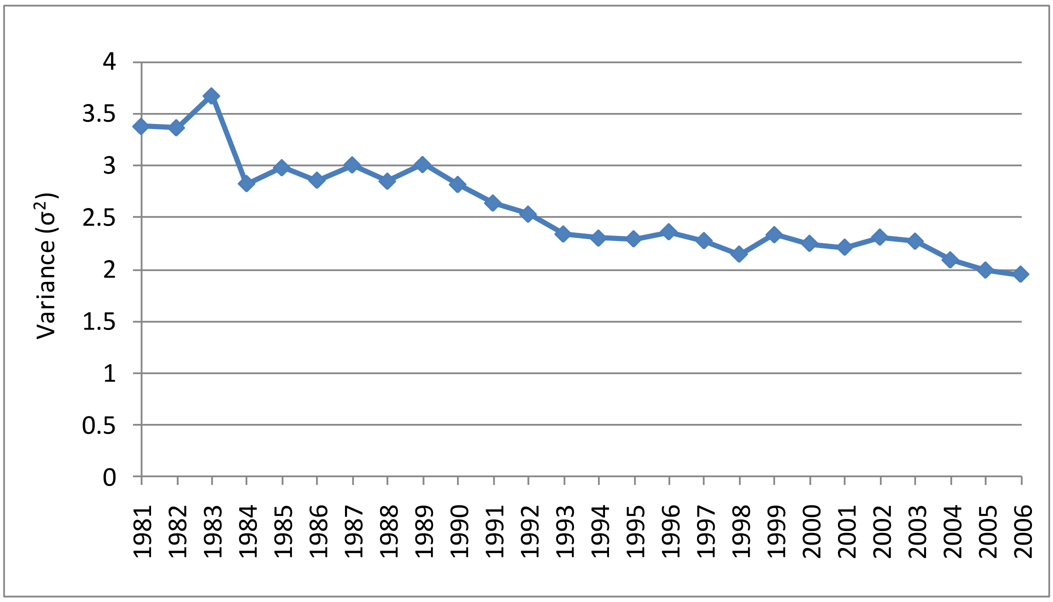
Table 3.2 shows the annual average growth rates of manufacturing industry employment from 1981 to 2006. At the national level, this growth rate was 2.56%. Growth rates vary across macro regions and, within macro regions, across states. At the macro regional level, the southeast, south and northeast grew at 1.59%, 3.52% and 2.88% respectively. These macro regions had relatively high manufacturing employment shares in 1981: 63.80%, 16.12% and 15.29% respectively.

In contrast, the centre-west and north, with national manufacturing employment shares in 1981 of only 2.65% and 2.05% grew, respectively, at 5.38% and 7.46%, on average, from 1981 to 2006. At the state level, both within the macro regions and the nation, states that grew fastest were those with lower manufacturing industry employment share in 1981. For instance, in the north macro region the smallest state of Roraima grew at 11.25%. This growth rate is above those for the two biggest states in the region, Amazonas and Pará, which grew at only 4.86% and 8.92%.

In the northeast, excluding Alagoas and Sergipe states, all other small states grew faster than the three that led manufacturing employment share in 1981. The growth rates for these leading states were Bahia 2.82%, Ceará 3.08%, and Pernambuco at 1.09%. In the centre-west, the leading states of Goiás and Mato Grosso do Sul were surpassed by the smaller ones, Mato Grosso and Distrito Federal. The growth rates for the former two states were 5.19% and 3.69% and for the latter two 7.61% and 6.64%. In the southeast, São Paulo and Rio de Janeiro grew less than Minas Gerais and Espirito Santo: the growth rates for the former two states were 1.28% and -0.05%; for the latter two 3.94% and 3.99%. Lastly, the south's biggest state, Rio Grande do Sul, grew 2.77%, which was less than 4.26% and 3.96% for Paraná and Santa Catarina, respectively.

Barro & Sala-i-Martin (2004, p. 464) show how absolute convergence can occur in relation to a given steady state. This is verified by calculating variance (σ^2) of income, product (or output) or another variable of the researcher's interest over time. If at the beginning of the period the variance across regions is higher (lower) than the variance at the steady-state level, the variance will decline (increase) over time. Assuming that there is a common steady-state for manufacturing industry employment across Brazilian states, Figure 3.1 shows absolute convergence across states with a downward trend variance of manufacturing industry employment from 3.68 in 1983 to 1.95 in 2006, towards its unknown steady-state. [35] So, in line with the previous outlined literature that demonstrated real income convergence in Brazil, it can be seen that manufacturing industry employment convergence has also been present.

Figure 3.1: Variance of Logarithm of Manufacturing Industry Employment Across Brazilian States



3.3.2. Recent Experience of the Brazilian Economy

The Brazilian economy experienced high inflation and stagnation during the 1980s and the beginning of the 1990s, and stabilisation plans failed (Macedo & Barbosa, 1997) until a successful stabilisation of the *Real Plan* in July 1994. However, the Cruzado Plan, implemented in February 1986, allowed the Brazilian economy to stabilise over six months, a short period in which a burst of GDP growth was observed. [36] Additionally, analysing the period 1980-1994, Abreu (2008b, p. 395) points out:

> In the years 1981-1983, during the administration of General João Figueiredo (1979-1985), the last of five successive military

[35] However, within macro-regions (see Figures 3.1a-3.1e in Appendix 3.1), the variance in the north and southeast is higher than that in the northeast, centre-west and south. Only the variance of the Southeast and South show a clear decline. This suggests that the macro-regions behave differently. This issue is reconsidered in the conclusion.

[36] The website of São Paulo's Regional Council of Economics (CORECON) presents a summary of measures of plans implemented in Brazil over the period 1986-1994 and their main results.

> presidents since the military coup of 1964, there was a sharp deterioration in the Brazilian GDP growth performance. Brazil suffered its most severe recession of the twentieth century. GDP fell 4.9 percent from its peak in 1980. After a brief recovery in 1984-1985 when GDP grew on average 7 percent per annum – years that also witnessed a transition from military to civilian rule (and ultimately a fully fledged democracy) – growth performance remained mediocre during the following decades. Between 1981 and 1994 GDP per capita increased on average less than 0.1 percent annually. And there was only limited improvement in the decade after 1994.

From 1994 to 2004, Abreu and Werneck (2008, p. 432) state that:

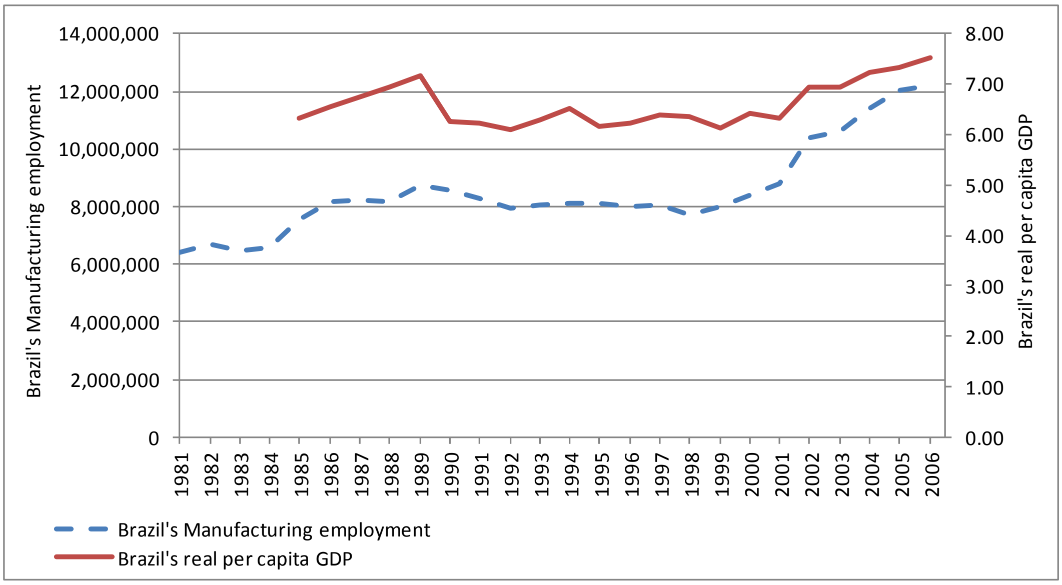
> In contrast with the previous fifteen years (1980-1994) there was some success in the period from 1995, in spite of many difficulties (…) But effective growth performance over the period continued to be mediocre: between 1994 and 2004 per capita GDP (gross domestic product) increased an average of only 0.9 percent per annum.

Therefore, with the exception of a few short time periods, the Brazilian economy performed poorly from 1981 to 2006. Despite this poor performance, a slight real income convergence was observed across the country's states (Ferreira, 2000; Rangel *et al.*, 2008). This sub-section investigates whether convergence in manufacturing industry employment is associated with the behaviour of per capita GDP at the state level in all macro-regions. To do this, the following questions are asked. First, is there a relationship between growth of manufacturing industry employment observed for Brazil and its per capita GDP? Secondly, if the answer to the first question is positive, is this correlation also present for states and macro-regions? Can states that follow the country's behaviour in terms of their manufacturing industry employment and per capita GDP growth be identified?

To answer the first question, manufacturing industry employment and real per capita GDP for Brazil are visually considered (see Figure 3.2). From 1985 to 1989 both graphs are upward sloping; from 1989 to 1992 both decline; from 1992 to 1997, while the manufacturing industry employment is basically constant, real per capita GDP increases from 1992 to 1994, followed by a decline between 1994 to 1995, and again positive growth from 1995 to 1997. From 1997 to 1998, there is a fall in manufacturing industry employment, followed by strong employment growth between 1998 and 2006. Per capita GDP declines between 1997 and 1999.

After 1999, per capita GDP also increases fairly strongly (but with oscillations) until 2006. Hence, while the fluctuations differ, manufacturing industry employment and per capita GDP in Brazil exhibit a positive relationship. Moreover, the sub-periods 1984-1988, 1989-1990 and 2000-2006 in which manufacturing industry employment of Brazil respectively grows, falls and grows again is remarkably consistent with the recent Brazilian economic cycles defined by Rolim (2008, p. 2, Figure 1). According to Rolim (2008), these aforementioned cycles refer, respectively, to the Brazilian presidencies of Sarney, Collor, and Lula.

Figure 3.2: Manufacturing Industry Employment *versus* real per capita GDP: Brazil



Note: *Per capita* GDP is in thousands of BR$ in 2000.

To answer the second question, the macro-regions individually are considered. To save space, employment and real Gross Regional Product (GRP) in the macro-regions has not been graphed. For the north the relationship between the two variables is ambiguous; it is either positive or negative depending on the sub-period within the 1985 to 2006 period. For the northeast, while the relationship is ambiguous for the three leading states of Caerá, Bahia and Pernambuco, for the small states of Alagoas, Maranhão, Paraíba, Piauí, Rio Grande do Norte and Sergipe there is a positive relationship between manufacturing employment growth and real GRP growth. Additionally, the trends in these latter states are similar to those of Brazil as a whole.

For the middle-west, the strength of the relationship depends on the chosen sub-period for all states. Goiás and Distrito Federal do not follow the pattern of the nation. In addition, Distrito Federal's per capita GRP for 2006 is almost three times the national per capita GDP.[37] In the southeast, the trends in São Paulo, Espírito Santo and Minas Gerais are similar and also correspond closely in most cases with the trends in Brazil as a whole. However, while per capita GRP of Minas Gerais follows that of Brazil, this is not the case for manufacturing industry employment in that state. Also, Rio de Janeiro presents "uncommon" behaviour: despite an increase of per capita GRP along the lines of Brazil's overall performance, the annual manufacturing industry employment growth rate is on average −0.05 percent. Finally, in the southern macro-region, the relationship between GRP per capita growth and manufacturing employment growth is ambiguous for all states. The trends in per capita real GRP are similar to that of Brazil's real per capita GDP. However, this is not the case for manufacturing employment, especially between 1989 and 1998.

In conclusion, there is a considerable variety and complexity in the patterns of regional growth in Brazil. The remainder of the chapter attempts to uncover the contributions of knowledge externalities, such as suggested by the theories of MAR, Porter, and Jacobs, and operationalised by the manufacturing employment growth model of Glaeser *et al*. (1992).

3.4. Data and Sources

This chapter makes use of two official data sources in Brazil. The first is IPEA - Instituto de Pesquisa Econômica Aplicada (or, Institute of Applied Economic Research).[38] This data source provides a variety of socio-economic data collected from public and private institutions. Regional data are available for municipalities, states and macro-regions.

This chapter employs three types of data. First, labour market data are used in which information on the number of employed people in each state from 1981 to 2006 was extracted. There are also state-level data on the percentage of employed people across nine sectors of economic activity. These are: (1) agriculture &

[37] In 1985, Distrito Federal's real per capita GRP was R$8,319.04 (in Brazilian R$ of 2000) while the national per capita GDP was R$6,336.53. By 2006, the gap increased to R$22,321.59 in Distrito Federal, compared with R$7,532.53 for Brazil as a whole. *Source*: www.ipea.gov.br.

[38] The IPEADATA can be found on the website www.ipea.gov.br.

fishing; (2) commerce; (3) construction; (4) electricity, water & gas; (5) finance; (6) manufacturing; (7) mining; (8) services; and (9) transportation & communications. For each state, annual employment by sector was calculated by multiplying the percentage of employed people in each sector by the total number employed people of that state. Aggregated across all states, the calculated sectoral employment accounts for between 92.22 percent and 99.21 percent of reported country-level employment between 1981 and 2006, after some values in years with missing data[39] have been imputed.

Secondly, the number of manufacturing sector establishments was considered. IPEADATA has one observation on manufacturing sector establishments for the following years: 1970, 1975, 1980, 1985 and 1995. The 1985 data were used to calculate the competition variable that was included in the cross-section models of manufacturing industry employment growth from 1985 to 1995 (see Appendix 3.2).

Thirdly, education data were used. The IPEADATA annually presents a variety of education indicators for each state from 1981 to 2007. The literacy rate (100 minus the percentage of illiterate people aged 15 years and older) was selected as the appropriate education indicator as it is a standard measure of the level of development.

However, there are two data caveats. First, there are missing data for 1991, 1994 and 2000. For these years, the literacy rate was interpolated by calculating the average between the neighbouring years for each of the years. In relation to employment, the following two-step procedure was applied: first, the average of the distribution of the employed population across sectors between the neighbouring years for the states with missing sectoral employment data was taken. Next, those percentages were applied to total employment in the state in each of the years listed above. This yielded for each state and each year manufacturing industry employment and employment in each of the other eight sectors. Ultimately, a complete time series of employment by sector and state for a period of 26 years was created.

The second data caveat is the level of data aggregation. The level of disaggregation of the data by region and by sector is an important issue for regional analysis because the more disaggregated the data are, the better will be

[39] A discussion of caveats about how the cells with missing data were filled follows further below.

the potential for understanding regional growth patterns. Comparing this chapter to Glaeser *et al.* (1992), there is a limitation in the manufacturing industry employment database. For example, this sector cannot be disaggregated further at the regional level in order to see the share of steel production, electronics manufacturing, etc in each region. This aggregation issue must be taken into account because with more disaggregated data, the results found here could still change in terms of sign and statistical significance (McCann *et al*., 2008). Consequently, while this chapter would ideally conduct this study at the level of disaggregated data by sub-sector within the manufacturing sector and also by municipality, it is in fact limited to focus on the whole manufacturing sector in the smallest regional level for which there are data for all of the variables of interest, namely the state level.[40]

The second data source is the Ministry of Transportation of Brazil. Its website provides a table of distances in kilometres between state capitals and other main Brazilian cities. Information of distance from the states' capitals to São Paulo's centre, which is the largest centre of economic activity in Brazil, was extracted. Data from the Ministry of Transportation of Brazil has been used because this institution has current distance data and also knowledge of the quality of infrastructure throughout the country, which is important for the estimation of distance between places (the actual distance between places is not necessarily the shortest distance).

Appendix 3.2 shows the nomenclature and definitions for the variables which are used in the econometric modelling that is discussed in the following section. Descriptive statistics for all variables are given in Appendix 3.3.

[40] The State of Tocantins has been excluded from the analysis because of missing data from 1981 to 1992 (see also Table 3.2 above).

3.5. Empirical Results and Discussion[41]

3.5.1. Cross-section Models[42]

Table 3.3 shows all cross-section models results. Two models in which the dependent variable is 1985-95 growth of manufacturing industry employment (columns equations 3.5.1.1 and 3.5.1.2) were estimated. Specialisation and competition are defined in the same manner as Glaeser *et al.* (1992). Diversity is defined as in Combes' (2000) study that applied Glaeser *et al.*'s model using data on manufacturing and services in the regions of France. Distance is defined as in Henderson *et al.*'s (1995) study of own industry employment in cities of the United States. Education is measured by the states' literacy rate. The second model uses all of the first model's variables except for the literacy rate which is replaced by the change in literacy between 1985 and 1995.

[41] This chapter has a caveat about the role of knowledge externalities on regional growth. In contrast to Glaeser et al, (1992), there is no estimation of wage equation due to lack of data.

[42] The test for multicollinearity in all estimated models in this chapter indicates that multicollinearity is not a concern. According to McCann *et al.* (2008, p. 10), multicollinearity is not present if the Variance Inflation Factor (VIF) is smaller than 10. The calculated VIF is relatively low. For cross-section models, the average VIF ranges from 2.11 to 4.09, and the highest VIF ranges from 2.72 to 6.52 associated to logarithm of competition and diversity variables, respectively. For the models that follow on subsections 3.5.2 and 3.5.3 below, the VIF values are much lower. For panel models with annual data, the average VIF ranges from 1.21 to 2.48, and the highest VIF ranges from 1.29 to 7.14 associated with distance and logarithm of specialisation variables, respectively, when excluding an-"outlier" case of VIF equals to 12.32 associated with the distance variable. For pooled-periods cross-section models, the average VIF ranges from 1.24 to 3.03, and the highest VIF ranges from 1.38 to 7.38 associated to specialisation and logarithm of specialisation variables, respectively, again, excluding an-"outlier" case of VIF equal to 9.61 associated with distance. Another way to test for multicollinearity is to correlate explanatory variables. Under this approach, multicollinearity is a concern if the correlation of a pair of variables is above the threshold absolute value of 0.80 (McCann *et al.*, 2008). The regressors' correlation across this chapter's cross-section models ranges from 0.0004 (between diversity and specialisation) to 0.6897 (between diversity and competition). On the other hand, correlation across variables of both panel models with annual data and pooled-periods cross-section models ranges from 0.39 (between distance and specialisation) to 0.55 (between logarithm of literacy and logarithm of diversity). In this latter range an-"outlier" case of absolute correlation of 0.90 (between São Paulo state's dummy and logarithm of distance) for the pooled-periods models has been excluded.

Table 3.3: Cross-section Models

Variables	***Eq. 3.5.1.1: OLS Model***	***Eq. 3.5.1.2: OLS Model***	***Eq. 3.5.1.3: OLS Model***	***Eq. 3.5.1.4: OLS Model***
Constant	-1.1089***	0.0370	-2.8412**	0.0163
	(-2.83)	(0.15)	(-2.13)	(0.18)
Specialisation	-0.1802	-0.1232	-0.0257	-0.0045
	(-1.11)	(-0.077)	(-0.15)	(-0.03)
Competition	0.2496**	0.2441**	0.3801***	0.3966***
	(2.22)	(2.30)	(2.60)	(2.95)
Diversity	-0.1730	0.0443	0.1227	0.2138
	(-0.79)	(0.23)	(0.85)	(1.29)
Distance	0.0001*	0.0139	0.0365**	0.0207
	(1.66)	(0.33)	(2.35)	(1.34)
Literacy	0.0150***		0.6166**	
	(2.85)		(2.06)	
1985-95 change in literacy		-1.3631		-0.8785
		(-1.69)		(-1.04)
Number of observations	26	26	26	26
R^2	0.7524	0.6846	0.7306	0.6930

Notes: Significance levels: *=10%, **=5%, and ***=1%; Dependent Variable: Growth = $\text{Log}\left(\frac{E_t}{E_{t-1}}\right)$, where E_t and E_{t-1} is Manufacturing industry employment in 1995 and 1985, respectively. Estimations are with robust standard errors. While in equations 3.5.1.1 and 3.5.1.2 the dependent variable is VDEP, in equations 3.5.1.3 and 3.5.1.4 it is VDEP2. Values of *t* statistics are in brackets. Equations 3.5.1.3 and 3.5.1.4 also differ from 3.5.1.1. and 3.5.1.2 in the following ways: in equations 3.5.1.3 and 3.5.1.4 the natural logarithm has been taken for all independent variables, as in Combes (2000), except for 1985-95 change in literacy in equation 3.5.1.4.

The models have been estimated by Ordinary Least Squares (OLS). In terms of the significant findings in equations 3.5.1.1 and 3.5.1.2 of Table 3.3, it can be noted that in both models, the coefficient of competition is positive and statistically significant. This result corroborates with both Porter and Jacobs' externalities. In both models the coefficient on distance is insignificantly different from zero showing independence between distance from (the largest market of) São Paulo and growth.[43] Finally, in equation 3.5.1.1, the coefficient of the literacy rate is positive as expected and also statistically significant. Hence the literacy rate is positively correlated with employment growth.

[43] In many of the models in this chapter, the coefficient of distance is either positive but not statistically significant or positive and statistically significant. This is against the expectation of a negative relationship with distance from states' to São Paulo's capital and economic growth. However, states that are far from São Paulo may have their growth more dependent on their neighbours' states, rather than on the São Paulo state market itself. This is consistent with the notion of 'clubs of convergence' (a group of states that converge within their macro-regions) stated in the conclusion section.

Equations 3.5.1.3 and 3.5.1.4 of Table 3.3 show models in which the model is specified as in Combes (2000). The dependent variable is manufacturing industry employment growth, VDEP2. The independent variables of the equations 3.5.1.3 and 3.5.1.4 are similar to those of 3.5.1.1 and 3.5.1.2 respectively; however, natural logarithms have been taken except for the change in literacy in equation 3.5.1.4.[44] The results from both models indicate that the coefficient of competition is again positive and statistically significant, supporting both Porter and Jacobs' externalities. Moreover, in model 3.5.1.3 the coefficients of distance from São Paulo and the literacy rate are also positive and statistically significant.

All cross-section models have high explanatory power: in equations 3.5.1.1, 3.5.1.2, 3.5.1.3, and 3.5.1.4 the adjusted R^2 is, respectively, 0.7524, 0.6846, 0.7306, and 0.6930. However, there are only 26 observations for each regression equation.

This chapter's cross-section models are comparable with previous studies. The findings that are simultaneously significant in this chapter and in the previous seminal papers by Glaeser *et al*. (1992) for the USA and by Combes (2000) for France were considered. Glaeser *et al.* (1992) estimate three regression models in which the dependent variable is 1956-1987 employment growth in each city-industry combination in the first model; 1956-1987 wage growth in each city-industry and 1956-1987 employment growth in the city after excluding the four biggest industries.

This chapter's results will be compared with those of Glaeser *et al.*'s first regression model. For this model these authors estimated four equations with specialisation in the 1st and 4th estimations and competition in the 2nd and 4th. They found a coefficient of specialisation that is, respectively, -0.0128 and -0.00799. This supports Jacobs' externalities. This result is rejected for Brazil because the specialisation coefficient in equations 3.5.1.1 and 3.5.1.2 is not statistically significant. For competition: Glaeser *et al*. (1992) found 0.587 and 0.561 which is consistent with Porter and Jacobs' hypothesis. Although the values are less than half, qualitatively identical results were found.

[44] Equations 3.5.1.2 and 3.5.1.4 were also estimated with the restriction that the coefficient of the change in the literacy rate is zero. The results displayed did not change much. For equation 3.5.1.4, while the coefficients of competition, specialisation and distance remained roughly the same (also in (in)significance), diversity remained positive but became statistically significant at 5% level.

Although Combes (2000) applied Tobit (for externalities indicators) and Probit (for regional dummies and density of employment) estimation methods rather than OLS to address the problem of truncated data, because his sample involved French plants of more than 20 workers[45], a comparison between Combes' results and this chapter's can still be made. This is because Combes also compared his results to those of Glaeser *et al.* who employed OLS instead of Probit and Tobit methods. Combes presents both Global Regressions Estimations (Table 1, p. 340) as well as Annual Global Regressions (Table 4, p. 352) both for manufacturing industry employment and services. Combes' findings for manufacturing industry employment are compared with estimations from this chapter.

Combes found -0.088 for specialisation on Global Regressions (p. 340) and -0.033 for specialisation on Annual Global Regressions (p. 352). In this chapter, negative coefficients were obtained in Brazil in equations 3.5.1.3 and 3.5.1.4 estimations, but they are statistically insignificant. For competition, Combes found for the aforementioned regressions -0.154 and -0.013, respectively whereas this chapter finds for equations 3.5.1.3 and 3.5.1.4 0.3801 and 0.3966 respectively, which supports both Porter and Jacobs' externalities. However, Combes used the inverse of this chapter's competition variable, which he named the size of plants. The expression is, "$\mathrm{size}_{\mathrm{z,s}} = \frac{\mathrm{emp}_{\mathrm{z,s}}/\mathrm{nbr}_{\mathrm{z,s}}}{\mathrm{emp}_{\mathrm{s}}/\mathrm{nbr}_{\mathrm{s}}}$, where $\mathrm{nbr}_{z,s}$ and nbr_s are the number of plants belonging to sector *s* in ZE *z* and France, respectively" (Combes, 2000, p. 337). Consequently, the findings from this chapter and those from Combes are consistent with respect to competition. For diversity, Combes found for the Global and Annual Global Regressions -0.051 and -0.026, respectively. This result supports both MAR and Porter's externalities. The coefficients of diversity in equations 3.5.1.3 and 3.5.1.4 are clearly inconsistent with this.

[45] The problem of truncation of employment data by firm size is not present in employment data used in this chapter. Therefore, the use of OLS method is appropriate,

3.5.2. Panel Models with Annual Data

This sub-section presents estimated coefficients of Annual Panel Models of six manufacturing industry employment growth equations shown in Table 3.4.[46] In equations 3.5.2.1 to 3.5.2.3 the dependent variable is VDEP, following Glaeser *et al*. (1992), whereas in equations 3.5.2.4 to 3.5.2.6 it is VDEP2, following Combes (2000). The role of education in growth was addressed by inclusion of the literacy rate in equations 3.5.2.1 and 3.5.2.4 as well as its annual change in equations 3.5.2.2 to 3.5.2.3 and 3.5.2.5 to 3.5.2.6. Equations 3.5.2.3 and 3.5.2.6 differ from 3.5.2.2 and 3.5.2.5 in that that the former include year and state dummies. The natural logarithm is taken for the specialisation, diversity and distance variables in equations 3.5.2.4 to 3.5.2.6 and for the literacy rate in the equation 3.5.2.4.

In all the six models, the coefficient of the lag of the dependent variable is negative as expected and also statistically significant. Because the dependent variable is in growth rates, the negative coefficient on the lagged growth rate is consistent with an autoregressive process in levels. The autocorrelation coefficient is one minus the reported coefficient and is therefore around 0.6 to 0.7. In equations 3.5.2.3 and 3.5.2.6, which are both models with year and state dummies, the coefficient of specialisation is positive and statistically significant, which supports both MAR and Porter's externalities. The coefficient of diversity is not statistically significant. The coefficients on the distance variable show that employment growth in Brazil is faster the further the state is from São Paulo.

In equation 3.5.2.1 the coefficient of the literacy rate is, as expected, positive and statistically significant. That is, the higher the literacy rate, the higher economic growth. Nevertheless, in equation 3.5.2.4 the coefficient of this variable is not statistically significant; the coefficient of the rate of annual change in literacy rate is also not statistically significant in equations 3.5.2.2, 3.5.2.3, 3.5.2.5 and 3.5.2.6. The explanatory power is relatively low for equations 3.5.2.1, 3.5.2.2, 3.5.2.4 and 3.5.2.5: 0.1029, 0.1121, 0.1396, and 0.1438, respectively. However, as expected, the R^2 is much higher for equations 3.5.2.3 and 3.5.2.6, which are both models with time and state dummies.

---

[46] Appendix 3B shows the descriptive statistics for the variables of the cross-section, panel and pooled-periods cross-section models. These latter two groups of models allow for some control of (time invariant) omitted variables and may yield therefore more accurate externalities parameter estimates.

Table 3.4: Panel Models with Annual Data

Variables	*Eq. 3.5.2.1: OLS Model*	*Eq. 3.5.2.2: OLS Model with Education Change*	*Eq. 3.5.2.3: OLS with year and states' dummies Model*	*Eq. 3.5.2.4: OLS Model*	*Eq. 3.5.2.5: OLS Model with Education Change*	*Eq. 3.5.2.6: OLS with year and states' dummies Model*
Constant	-0.1158** (-2.55)	-0.0262 (-0.83)	-0.5912** (-2.42)	-0.0351 (-0.15)	-0.0619* (-1.81)	-0.2695*** (-3.85)
Lag of the dependent variable	-0.2956*** (-4.46)	-0.2855*** (-4.59)	-0.3622*** (-8.02)	-0.3344*** (-5.21)	-0.3256*** (-5.43)	-0.3801*** (-7.66)
Specialisation	0.0298 (1.07)	0.0357 (1.26)	0.4676*** (3.49)	0.0616* (1.92)	0.0601* (1.89)	0.4370*** (4.73)
Diversity	0.0126 (0.32)	0.0336 (1.19)	-0.0381 (-0.68)	0.0639 (1.30)	0.0521 (1.30)	-0.0443 (-0.60)
Distance	0.0192** (2.36)	0.0159** (2.05)	0.0001*** (4.26)	0.0161** (2.41)	0.0166** (2.44)	0.0226*** (5.67)
Literacy rate	0.0012* (2.08)			-0.0060 (-0.11)		
Annual change in the literacy rate		-1.0081 (-1.34)	-0.5470 (-0.75)		-0.5596 (-0.77)	-0.4527 (-0.78)
Number of observations	624	624	624	624	624	624
R-Squared	0.1029	0.1121	0.3308	0.1396	0.1438	0.3584

Notes: Significance levels: *=10%, **=5%, and ***=1%; The dependent variable is manufacturing industry employment growth. Estimations are with robust standard errors. Values of *t* statistics are in brackets. While in the equations 3.5.2.1 to 3.5.2.3 the dependent variable is VDEP, in equations 3.5.2.4 to 3.5.2.6 it is VDEP2.

In equation 3.5.2.3 the dummy coefficient for 2002 was dropped. Dummies coefficients for 1983, 1984-1985, 1987-1988, 1990-2001, 2003, 2005 and 2006 are negative and statistically significant, while for the other years they are negative but not statistically significant. In relation to states' dummies coefficients Roraima and São Paulo were dropped. States' dummy coefficients for Acre, Amapa, Para, Rondonia, Bahia, Maranhão, Paraiba, Piaui, Sergipe, Distrito

Federal, Goiás, Mato Grosso do Sul, Mato Grosso, Espírito Santo, Minas Gerais, Rio de Janeiro, São Paulo, Paraná and Rio Grande do Sul are positive and statistically significant; those for Amazonas and Ceará are negative and statistically significant; for Alagoas, Pernambuco and Rio Grande do Norte are positive but not statistically significant. In equation 3.5.2.6 the dummy coefficient for 2002 was again dropped. All year dummies coefficients are positive, except for 1999, 2003 and 2006, which are negative – but in these three cases none of them is statistically significant. Regarding states' dummy coefficients, Amazonas and São Paulo have been dropped. All states' dummies coefficients are positive and statistically significant, except for Rio Grande do Sul, which had a negative and statistically significant coefficient as well as Santa Catarina, for which the dummy coefficient is also negative but not statistically significant.

3.5.3. Pooled-Periods Models

This sub-section presents models of average annual employment growth for pooled periods. The periods are 1981-1990, 1991-1998 and 1999-2006. For these models three steps were followed: first, time pools over the series were defined (see Appendix 3.2); secondly, the annual average growth within each period was calculated; thirdly, all of these average growth values taken at the state level were combined. This yields a total of 78 observations.

Two groups of three models were estimated. In all models, the change in literacy is used rather than literacy rate itself. Table 3.5 shows the results. In equations 3.5.3.1 to 3.5.3.3 the dependent variable is based on Glaeser *et al*. (1992), whereas in equations 3.5.3.4 to 3.5.3.6 this chapter follows Combes (2000). The models 3.5.3.1 and 3.5.3.4, estimated by OLS, involve the following explanatory variables: specialisation, diversity, distance, and change in the literacy rate. These models are expanded by the inclusion of year and state dummies in equations 3.5.3.2 and 3.5.3.5. Lastly, equations 3.5.3.3 and 3.5.3.6 are estimated as fixed effects models.[47] In equations 3.5.3.4-3.5.3.6 the natural logarithm was taken for all independent variables, except for the change in literacy.

[47] Fixed, random and between-effects models were estimated. Hausman and Wald tests involving the former two models accepted fixed effects model; however, estimations for which these tests were done did not use the robust standard errors. Distance and states' dummies have been excluded as they do not vary over time, but the coefficients displayed without these dummies do not significantly change compared to the robust standard error estimations. Therefore, a choice to maintain the fixed effects models estimated with robust standard errors was made. The other panel models not reported here do not affect the conclusions.

Table 3.5: Pooled-Periods Cross-Section Models

Variables	***Eq. 3.5.3.1: OLS Model***	***Eq. 3.5.3.2: OLS with year and states' dummies Model***	***Eq. 3.5.3.3: Fixed Effects Model***	***Eq. 3.5.3.4: OLS Model***	***Eq. 3.5.3.5: OLS with year and states' dummies Model***	***Eq. 3.5.3.6: Fixed Effects Model***
Constant	0.0466 (1.40)	-0.1365** (-2.01)	-0.0040 (-0.09)	0.0073 (0.39)	-0.0251 (-1.37)	0.0956*** (7.70)
Specialisation	-0.0079 (-0.59)	0.1256*** (3.94)	0.1049** (2.53)	-0.0033 (-0.25)	0.0988*** (4.32)	0.0741** (2.28)
Diversity	0.0044 (0.12)	-0.0472 (-0.66)	-0.0074 (-0.15)	-0.0039 (-0.11)	-0.0265 (-0.44)	0.0027 (0.07)
Distance	0.0115*** (3.42)	0.0001*** (7.08)	Dropped	0.0085*** (3.00)	0.0095*** (4.64)	Dropped
Change in literacy over successive periods	-0.3788*** (-4.21)	0.1810 (1.09)	-0.4012*** (-2.89)	-0.3999*** (-3.87)	0.1952 (1.22)	-0.3926*** (-2.88)
Number of observations	78	78	78	78	78	78
R-Squared	0.1990	0.7030	within=0.1750 between=0.011 overall=0.0064	0.1813	0.7086	within=0.1628 between=0.008 overall=0.0083

Notes: Significance levels: *=10%, **=5%, and ***=1%; The dependent variable is the rate of average annual manufacturing industry employment growth per period. The periods are 1981-1990, 1991-1998 and 1999-2006. Estimations are with robust standard errors. Values of *t* statistics are in brackets. In equation 3.5.3.2 the dummies for Roraima and São Paulo were dropped; Dummy coefficients for Amazonas and Bahia are negative and statistically significant; Dummy coefficients for Amapá, Pará, Rondonia, Distrito Federal, Goiás, Mato Grosso do Sul, Mato Grosso, Espírito Santo, Minas Gerais, Rio de Janeiro and Paraná are positive and statistically significant while those for Acre, Bahia, Maranhão, Piaui, Sergipe, Rio Grande do Sul and Santa Catarina are positive but not statistically significant. Other states' dummy coefficients are negative and not statistically significant.

In equations 3.5.3.4 to 3.5.3.6, the natural logarithm was taken for all independent variables, except for the change in literacy. In equation 3.5.3.5, the dummies for Amazonas and São Paulo were dropped. The dummy coefficients for Acre, Amapá, Pará, Rondonia, Roraima, Alagoas, Ceará, Distrito Federal, Goiás, Mato Grosso do Sul, Mato Grosso, Paraná, Alagoas, Paraíba, Piauí, Rio Grande do Norte, Sergipe and Minas Gerais are positive and statistically significant; those for Rio Grande do Sul and Santa Catarina are negative and statistically significant; for Ceará, Pernambuco and Rio de Janeiro they are negative but not statistically significant. The dummy for time period 1 (1981-1990) was dropped. The dummy for time period 2 (1991-1998) is negative and statistically significant. The dummy for time period 3 (1999-2006) is positive but not statistically significant.

The results obtained through OLS models (equations 3.5.3.1 and 3.5.3.4) indicate that the coefficient of distance is positive and statistically significant. Again, growth is stronger the further the region is from São Paulo. The coefficient of the change in literacy is negative and statistically significant. Both the distance effect and the change in literacy effect are consistent with the neoclassical convergence hypothesis. Since São Paulo has the highest level of income and the regions furthest away from São Paulo are the least developed, those regions will have the fastest growth rates. Similarly, when the rate of literacy increases fast, real incomes will increase and this also lowers the growth rate, in line with beta convergence.

OLS estimation with year and state dummy variables (equations 3.5.3.2 and 3.5.3.5) finds that the coefficient on specialisation is positive and statistically significant, which confirms both MAR and Porter's externalities. The coefficient of diversity is not significant in any of the models.

The fixed effects estimation (equations 3.5.3.3 and 3.5.3.6) finds again that the coefficient of specialisation is positive and statistically significant, supporting both MAR and Porter's externalities. Similar to McCann *et al.* (2008), it can be seen that there is an issue of stability of sign and statistical significance. By moving from the basic cross-section models to the panel models, the results with respect to the impact of the externalities appear unstable: while cross-section models confirm competition and diversity externalities, panel and pooled models "prove" the specialisation externality. [48] The results are summarised in Table 3.6.

With respect to competition externalities, these were confirmed in cross-section models and confirmed both Porter's and Jacobs' theories. However, because of a lack of establishments data (see the expression for competition in Appendix 3.2), the competition externality appears only in cross-section models. Therefore this is not comparable across models. With respect to specialisation, this externality was rejected in cross-section and simple pooled OLS models. However, it has been confirmed in annual panel and pooled models (both OLS with year and state dummy models), and in the fixed effects models. In these results it provided support for both MAR and Porter's theories. Finally, the presence of diversity externalities was only confirmed in the cross-section model without the education variable.

[48] The panel and pooled model results are unique and non-comparable with the previous seminal papers (Glaeser *et al*., 1992 and Combes, 2000) because both only employed cross-section data.

Table 3.6: A Comparison of Externality Impacts Across Models

	Cross-section			Annual Panel: OLS with Year and States' Dummies			Pooled OLS (a)			Pooled OLS with Year and States' Dummies (b)			Pooled (c): Fixed Effects		
Externality	MAR	Porter	Jacobs	MAR	Porter	Jacobs	MAR	Porter	Jacobs	MAR	Porter	Jacobs	MAR	Porter	Jacobs
Specialization				+	+					+	+		+	+	
Competition		+	+												
Diversity			+												

Notes: Sign confirms prediction of the externalities in a situation in which it is statistically significant. Empty areas mean that the externalities are statistically not significant. Because of lack of data, competition externalities have not been included in panel models, so areas are dark in these cases. For pooled models (a) to (c) refer to the estimated models as shown in the sub-section 3.5.3. Diversity externalities have been confirmed only in one cross-section model after imposing the restriction that the coefficient of education is zero.

3.6. Conclusion

Many studies have documented real income convergence among the Brazilian states. This convergence was also reinforced by this study through assessment of Brazilian states' manufacturing industry employment growth which shows that poor states grew faster than rich ones from 1981 to 2006. The main question of this study is why this happened. This chapter attempted to answer this question by analysing the importance of the theories of dynamic externalities proposed by Marshal-Arrow-Romer (Marshall, 1890; Arrow, 1962; Romer, 1986), Porter (1990) and Jacobs (1969) for the explanation of that convergence using Brazil's manufacturing industry employment data.

This chapter found that the aforementioned theories of dynamic externalities help to explain convergence among Brazilian states. First, the estimated cross-section models confirm competition along the lines of Porter and Jacobs, and also diversity externalities as suggested by Jacobs. Even though the findings partially contrast with previous studies, such as those by Glaeser *et al*. (1992) and Combes (2000), they are in accordance with other international studies summarised by De Groot *et al*. (2009). However, this partial confirmation must be judged with caution because the data used in the above studies are more detailed with respect to manufacturing industry employment than this study's data.

Secondly, using models not previously applied in this context by Glaeser *et al*. (1992) and by Combes (2000) in Tables 3.4 and 3.5, MAR and Porter's specialisation externalities were confirmed. Thirdly, the analysis of sign stability showed that while for cross-section models competition externalities under Porter and Jacobs' theories are confirmed, for the pooled OLS with year and state dummy models and fixed effects models, specialisation externalities become observed only under MAR's theory. This suggests that while MAR and Porter's (low competition or more specialisation) theories are valid in some macro-regions, Jacobs' (high competition or diversity) theory is important for the other regions. This result is consistent with the 'clubs of convergence' hypothesis of the convergence literature for Brazil (see, for example, Brauch & Monasterio, 2007; Gondim et al., 2007; Coelho & Figueiredo, 2007; Magalhães et al., 2005).

Finally, this chapter has three caveats. First, to understand Brazilian macro-regions' behaviour, it is important to employ disaggregated panel models that focus on individual regions because externalities operate differently in different parts of the country. Secondly, each type of externalities appears captured by specific models: on the one hand, cross-section models picture competition externalities under both Porter and Jacobs' theories and diversity also under Jacobs' theory; on the other hand, annual panel and pooled models capture specialisation externalities under MAR and Porter if those models are estimated by OLS with both year and state dummies or with fixed effects. Further research is needed on why changing the model specification from a cross-section to a panel approach leads to such apparently contradicting results. Thirdly, that infrastructure plays a role in growth and convergence of Brazilian states is acknowledged. Due to complexity of infrastructure (Daumal & Zignago, 2010), its effects on regional growth will be analysed. These three caveats will be investigated in future research.

CHAPTER 4 – A CENTURY OF THE EVOLUTION OF THE URBAN SYSTEM IN BRAZIL[49]

4.1. Introduction

Urbanisation has been a key area of debate among economists from the 1980s (Rosen & Resnick, 1980; Parr, 1985; Brakman, Garretsen, van Marrewijk & van den Berg, 1999; Black & Henderson, 1999, 2003; Duranton, 2007). The urban area plays an important role in the regional economy as the spatial unit where most economic activities occur. A standard method to test whether the distribution of cities is consistent with various theories of urbanisation is to check if the power law holds (Brakman, Garretsen & van Marrewijk, 2009). The power law (or Pareto distribution) holds when there is a negative loglinear relationship between the size and rank of cities, at least above a certain city size threshold.[50] This law, and specifically the slope of the loglinear relationship, is an important tool for understanding urban growth. If urban growth is a stochastic process where every city shares the same expected growth rate and the variance of the growth rate is also the same for each city (referred to in the literature as Gibrat's law), the distribution of city sizes is lognormal. The upper tail of the lognormal distribution closely resembles a Pareto distribution. Sometimes this Pareto distribution has a unitary slope which implies that the product of rank and size among the larger cities is constant, which is referred to as Zipf's law (Gabaix, 1999a, 1999b; Eeckhout, 2004, 2009).

Previous studies of the power law in various countries have four limitations. First, these studies often just use cross-sectional data on cities (Rosen & Resnick, 1980; Soo, 2005; Giesen, Zimmerman & Suedekum, 2009). Other works use panel data on cities but only for short continuous time periods (Song & Zhang, 2002; Xu & Zhu, 2009).[51] Those who exploit longer time series tend to use only one observation per decade (Parr, 1985; Overman & Ioannides, 2001; Delgado & Godinho, 2006; Moura & Ribeiro, 2006; Garmestani, Allen, Gallagher &

[49] Glen Stichbury of Waikato University provided helpful advice on Geographic Information System.

[50] Pareto pioneered the power law in his *Cours d'Economie Politique* (1896, 1897) (The New School, 2010). However, this law was first applied in economics to the distribution of income (Simon, 1955) rather than the distribution of cities.

[51] The only exceptions are Bosker (2008, chapter 5), who analyses the distribution of the 62 largest west-German cities from 1925 to 1999; and Giesen & Suedekum (2009) who test for Gibrat's and Zipf's laws for the 71 largest west-German cities at the national level and the mostly 20 largest cities at the state level from 1975 to 1997. Both studies employ annual city data.

Mittelstaedt, 2007). Secondly, although the urban area (or metropolitan area) is the most appropriate geographical unit of analysis on the grounds that agglomeration externalities are better captured in this spatial unit (McCann, 2001; Brakman, Garretsen & van Marrewijk, 2009),[52] many previous studies use data on the smaller, administratively defined, cities. This biases the power law parameter downward. Thirdly, even studies that employ urban agglomeration data can be deficient if they use only cross-sectional data from which it is impossible to test whether agglomeration takes place which would be reflected in an increase in the power law parameter over time (Rosen & Resnick, 1980; Pumain & Moriconi-Ebrard, 1997; Gabaix & Ioannides, 2003; Brakman, Garretsen & van Marrewijk, 2009). Finally, in the city size distribution literature, there are only a few exceptions (Black & Henderson, 2003; Gabaix & Ibragimov, 2006; Nishiyama, Osada & Sato, 2008; Bosker, 2008). These studies are more concerned with the robustness of the power law as, to investigate this law, they go beyond the use of the standard techniques.

This chapter overcomes these four limitations by using data on urban agglomerations at frequent intervals over a long time span to test the traditional power law and one of its modern specifications (Gabaix & Ibragimov, 2006). The chapter studies the size distribution of 185 urban areas in Brazil observed annually in 102 years from 1907 to 2008. While there are other power law estimates for Brazil (for example, Brakman, Garretsen & van Marrewijk 2009 suggest an estimate of 0.7815 based on 193 cities), this is the first application with a long and continuous time series of urban area populations. This study aims to test whether the three aforementioned laws hold concerning the size distribution of urban areas in Brazil: namely the power law among the large cities, Zipf's law as a special case of the power law, and Gibrat's law with respect to the entire distribution.

The dataset used is unique as its construction is based on a wide range of geographical and historical information on urban activity rather than on administrative definitions of cities. The power law suggests that there is a concentration of economic activity in large agglomerations. However, such

[52] Rosen & Resnick (1980, p. 170) note that, "For size distribution studies, the entire metropolitan area is the most desirable choice for an urban unit as it represents an integrated economic unit. Since many workers and consumers in a city often reside in the surrounding suburbs, it seems reasonable to include these areas in the definition of the city". Soo (2005, p. 242) adds: "Data for agglomerations might more closely approximate a functional definition, as they typically include surrounding suburbs where the workers of a city reside".

agglomerations are usually a combination of a core city together with surrounding smaller cities or towns. The spatial unit of measurement is therefore the urban area that is consistent with urban economic theory: a single or multiple core metropolitan area that has its boundary defined by a transition from predominantly urban to predominantly rural activity. Using this definition of urban areas, the power law is confirmed for the 100 largest urban areas of Brazil. It also confirms the lognormal distribution by Kolmogorov-Smirnov and Shapiro-Francia tests (Stephens, 1974; Stata software, 1996-2012) for all urban areas. Gibrat's law (Gibrat, 1931; Eeeckhout, 2004, 2009) was also confirmed after 1984 by panel unit root tests for the whole sample of urban areas and (to a smaller extent) by a panel of individual urban areas for both Levin-Lin-Chu and Im-Pesaran-Shin tests (see Levin, Lin & Chu, 2002; Im, Pesaran & Shin, 2003; Baum, 2003; Bosker, 2008; Stata software, 1996-2012), but the law was rejected prior to 1984 in favour of the mean reversion hypothesis.

Conversely, panel unit root tests for the 100 largest urban areas rejected Gibrat's law for all sub-periods: while for the sub-period 1907 to 1939 the rejection of this law favoured the mean reversion hypothesis only at 10 percent significance level; for the sub-periods 1940 to 2008 a highly significant positive correlation between growth of urban areas and their initial sizes was found. This chapter rejects Zipf's law, but finds support for the increasing economic importance of urban agglomeration in the process of economic development in Brazil. The traditional power law parameter for the size distribution of urban areas increases from 0.63 in 1907 to 0.89 in 2008. The robustness of the results is checked by employment of the Gabaix & Ibragimov (2006) correction method from which was found a power parameter that ranges from 0.60 in 1907 to 0.84 in 2008. From Gabaix & Ibragimov's (2006) method, a panel model pooling the same range of urban size distributions provides a power law parameter equal to 0.68, which is within the range of cross-sectional estimation using both traditional and corrected power law equations.

The chapter is organised as follows. The next section discusses the background literature on the power law. Section 4.3 describes the data used and their sources. Section 4.4 briefly outlines the characteristics of the recent structural transformation in Brazil to provide the context. Section 4.5 discusses the empirical results. Lastly, section 4.6 provides concluding remarks.

4.2. Literature

The power law of the distribution of cities' sizes is a property that applies to many distributions with fat tails. Income distribution is another socioeconomic example of a fat (right) tail distribution and in fact it was to this distribution that the power law was first applied to by Pareto at the end of the 19th century. The New School (2010, paragraph 10) states:

> [Pareto] argued that in all countries and times, the distribution of income and wealth follows a regular logarithmic pattern that can be captured by the formula: Log N = log A + m log x where N is the number of income earners who receive incomes higher than x, and A and m are constants.

This law was subsequently applied to the distribution of German cities as early as 1913 by Auerbach (see Gabaix, 1999a, 1999b; Overman & Ioannides, 2001; Gabaix & Ioannides, 2003; Ioannides & Overman, 2003; Anderson & Ge, 2005; Soo, 2005, 2007; Córdoba, 2008; Bosker, 2008). Auerbach denoted the variables of the power law equation as follows: N is population size of the city with rank x, with the largest city ranked 1, the second largest city ranked 2, and so on; A and m are parameters: the former is the intercept that equals the expected value of the logarithm of the largest city and the latter is the slope which equals the power law parameter. These two parameters are usually estimated by OLS (the alternative is the Hill estimator, see Hill 1975).[53] The power law parameter m is a negative number of which the absolute value is known as α (or q) in the city size distribution literature. Zipf (1949) emphasised the special case in which $\alpha = 1$; consequently, this particular case is known as Zipf's law (or the rank-size rule).

The estimate of α indicates the degree of city size distribution skewness. If Zipf's law does not hold there are two possibilities: i) if $\alpha > 1$, the city size distribution is more uneven and the biggest city is larger than Zipf's law predicts; ii) if $\alpha < 1$, the city size distribution is more even and the biggest city is smaller than Zipf's law expects.[54]

Of particular interest for the study of development of the urban system in a country is the change in the power law parameters over time. The change in the

[53] In contrast to the original work by Pareto and by Zipf (1949), some studies put city rank on the left side of the equation and city size on the right. See Eeckhout (2004, 2009) and Bosker (2008) and the references therein.

[54] See Brakman, Garretsen & van Marrewijk (2009), Chapter 7.

intercept shows the expected growth in the largest city. The change in the slope parameter suggests whether the distribution of city sizes is becoming more uneven or even. When agglomeration is becoming more important, the slope parameter increases over time and in fact, this is what was found for the US (Black and Henderson, 2003). As noted in the introduction, the power law for larger cities is consistent with Gibrat's law describing the process of urban growth. Gibrat's law assumes independence between city growth rate and city size. When this law holds, the rank size rule is stable over time. In other words, the ratio of the largest city size to each of the other city sizes does not change over time. This urban system stability has economic implications for the distribution of employment, market areas, city innovation potential as a result of the volume of research in that city, variety of goods and services in the city, housing markets, etc. The dependence of economic aggregates of the region or country on the urban system is exactly in the spirit of Christaller's and Lösch's urban hierarchy theories that connect the complexity of economy of the urban area to the area size (Krugman, 1996; McCann, 2001; Mori, Nishikimi & Smith, 2005, 2008; Duranton, 2002, 2007).

City size distribution studies differ in sample size, degree of development of the studied country and in either rejecting or confirming Zipf's law. The literature has taken three approaches. The first approach uses cross-section data on cities to test Zipf's law and finds a power parameter either greater than 1 (Rosen & Resnick, 1980; Soo, 2005) or less than 1 (Eeckhout, 2004; Garmestani, Allen & Gallagher, 2008). The second approach makes use of a range of urban (or metropolitan) area cross-sections to comparatively reject Zipf's law (Rosen & Resnick, 1980; Brakman, Garretsen & van Marrewijk, 2009) or confirm Zipf's law (Gabaix, 1999a, 1999b; Gabaix & Ioannides, 2003; Ioannides & Overman, 2003). The third approach tests for both Zipf's law and Gibrat's law simultaneously. This approach employs panels of cities or urban (metropolitan) areas. Among studies using this approach, most reject both laws (Pumain & Moriconi-Ebrard, 1997; Song & Zhang, 2002; Black & Henderson, 2003; Moura & Ribeiro, 2006; Delgado & Godinho, 2006; Soo, 2007; Bosker, 2008; Xu & Zhu, 2009), but there is some support (Giesen & Suedekum, 2009).

4.3. Data and Sources

This chapter uses two official data sources from Brazil: IBGE - Brazilian Institute of Geography and Statistics and IPEA - Institute of Applied Economic Research. The websites of these two institutes are sources of rich socio-economic data at national and regional levels. The administrative geographical unit used is the municipality. Municipality population data from all censuses from which municipality population data are available at this level were obtained (see Table 4.1); that is, covering the period from 1907 to 2008. The sample of 185 urban areas has been built up from 1,409 municipalities in Brazil (for details about municipalities included in every urban area, see Appendix 4.1; a note at the bottom of this appendix explains how merges and splits of municipalities were addressed for the defined urban areas).

The construction of urban areas took four steps. First, a sum of the population of contiguous municipalities in 2008 was taken. The contiguity was checked by means of 2009 IBGE Brazilian States Political maps.[55] These maps were complemented with Google maps.[56]

The definition of urban areas used implied that some crossed state boundaries.[57] Therefore in some cases, an urban area is a collection of contiguous municipalities that belong to neighbouring states. The reason is essentially historical. Information on municipality history and of splits and merges of municipalities over time comes from the IBGE population data files themselves.[58] Other sources were consulted (Tenenbaum, 1996; Fausto, 1999) regarding the history of regions and settlements in Brazil

Secondly, the urban area definition for 2008 was applied back to 1907. The urban area growth was observed both in terms of an increase in population of the municipalities and birth of new contiguous municipalities (a detailed appendix is available upon request). Thirdly, a smoothing was applied to these population data under the assumption that some observed changes are inconsistent with the underlying demographic processes. This smoothing took account of neighbouring

[55] http://www.ibge.gov.br/
[56] http://maps.google.co.nz/
[57] In fact, to avoid compromising the originality of the tests for the power, Zipf and Gibrat laws for Brazil, the areas definitions created by bureaucrats and politicians have not been used. For a discussion of the importance of using functional rather than administrative urban areas, see Holmes & Lee (2010).
[58] www.citybrazil.com.br has also been used. However, the material on this website is essentially based on IBGE information.

municipalities as well as temporal changes. Fourthly, urban area populations were estimated for years without data from the official sources by interpolation.[59] Comparison between the calculated and original data for years in which both types of data are available (see Table 4.1) suggest that the smoothing and interpolation do not distort the analysis: the correlation between the original and the adjusted data is around 0.98.

In a discussion about sample quality for the power law test, Resende (2004, p. 1547) notes the importance of using heterogeneous samples of cities. Due to data limitations the sample of this chapter does not include all urban areas in Brazil, unlike Eeckhout's (2004) USA study. Yet, the sample is heterogeneous in that it involves urban areas of all sizes, in contrast with other studies that only use the largest metropolitan areas (such as: Gabaix & Ibragimov, 2006; Black & Henderson, 1999). This urban area size heterogeneity is achieved even though some urban areas that appeared to have strongly oscillating populations over time have been dropped. These may be considered outliers.

Table 4.1: Original Municipality Population Data

YEAR	SOURCE	THE SOURCE OBTAINED DATA BY
1907 to 1912	IBGE	Estimate
1920	IPEA	Census
1936 and 1937	IBGE	Estimate
1939	IBGE	Estimate
1940	IPEA	Census
1950	IPEA	Census
1960	IPEA	Census
1970	IPEA	Census
1980	IPEA	Census
1985	IBGE	Estimate
1991	IPEA	Census
1996	IPEA	Estimate
1999 to 2008	IBGE	Estimate*

*2000 is from Census.

The data on urbanisation in Brazil have caveats that originated in 1938 during the Getúlio Vargas presidency when the government elevated all municipalities to city status despite the economic structure of some municipalities not fulfilling the requirements of an urban economy. As a consequence, this overestimated urbanisation in Brazil (Veiga, 2003). After an analysis of law amendments that

[59] The population of Brasilia, the capital of Brazil that was started to be built in 1956, was extrapolated back from 1960.

established new municipalities in Brazil, Resende (2004, p. 1544) points to "non-rigorous criteria for the creation of municipalities (...)".

In recent times, this urban population data problem has been solved by the use of satellite data on urban activity from EMBRAPA,[60] but these data refer only to the cross-section of Brazilian urban areas that correspond to the 2000 population census. Although some authors advocate their use for the analysis of city size distributions (Kinoshita et al., 2008), these satellite mapping data also have limitations and are subject to criticism (Doll & Muller, 2000).

4.4. Structural Transformation in Brazil

4.4.1. Brazil's Recent Economic History

Since the arrival of the Portuguese in April 1500 and subsequent colonisation, Brazil has undergone many phases of strong social, political, economic and cultural change. This sub-section briefly describes the main events that influenced the city size distribution from 1907 to 2008. In order to do so, six periods are defined. The first period is 1907-1930, referred to as "Development of the Republic" (Lobo, 1996, p. 426). This period is characterised by labour immigration that was needed to facilitate growth of manufacturing. Although manufacturing grew as a result of Foreign Direct Investment and exports, the economy was essentially dependent on exports of coffee. The fall in coffee prices during the 1929 depression reduced state revenue that was necessary for import of machinery which the industrialisation policy depended on.

The second period is 1930-1945 (The Vargas Era). This period is characterised by: i) national integration policies, combination of authoritarian, totalitarian and fascist elements and the beginning of the imports substitution process (Lobo, 1996, p. 428); ii) the increase in internal migration (Fausto, 1999, p. 234); and iii) the immigration restriction policy which reduced population growth in the 1930s (Lobo, 1996; Bethell, 2008; Silva, 2008).

The third period is 1945-1964 (Democracy or "Developmental State"). This period is marked by: i) the Kubitschek Government (1956-1961) that adopted an economic policy inspired by Rostow's theory of take-off.[61] It concentrated

[60]EMBRAPA stands for Empresa Brasileira de Pesquisa Agropecuária (Brazilian Enterprise of Farming Research). These satellite data are available on http://www.urbanizacao.cnpm.embrapa.br/conteudo/base.html

[61] This theory argues that development has mainly two stages: at the first stage the government should focus on developing regions that have the 'preconditions of self-sustained growth' (Lobo,

investment in certain areas (Minas Gerais, São Paulo and Rio de Janeiro) where the preconditions for self-sustained growth existed (Lobo, 1996, p. 428); ii) the investment for construction of Brasilia city, inaugurated in 1960, and another migration wave from the northeast to São Paulo (Lobo, 1996, p. 429); iii) incentives for national manufacturing intensified the imports substitution process (Fausto, 1999; Abreu, 2008).

The fourth period is 1964-1984 (Dictatorship or "Authoritarian State"). The main characteristics are (Lobo, 1996; Fausto, 1999; Abreu, 2008a): i) the combination of economic stagnation and inflation ('stagflation'); ii) the annual average real income growth is 11.2% over the 'economic miracle' (1969-1974); iii) income concentration; iv) little political rights and freedom; strong regulation of the economy and creation of public institutions (1967-1974); v) oil shocks (1974-1980) causing macroeconomic instability; vi) redistribution of product that harmed the northeast and benefited the middle-west, north and south regions; vii) protectionism, contractionist policies, and falling output (1981-1983).

The fifth period is the short period 1985-1989 (Democratic Transition). This period is characterised by hyperinflation and stagnation. Lastly, the sixth and most recent period is 1989-2008 (trade liberalisation and the return to democracy (Lobo, 1996)). The main events are (Lobo, 1996; Abreu, 2008b; Abreu & Werneck, 2008): i) the structural reforms under Collor de Mello (1990-1992) and Itamar Franco (1992-1994) presidencies; ii) the policies that aimed to balance inflation and unemployment were more successful after mid-1994; iii) however, as Abreu & Werneck (2008, p. 432) point out, "(…) between 1994 and 2004 per capita GDP (gross domestic product) increased [at] an average of only 0.9 percent per annum"; and (iv) despite trade liberalisation, the Brazilian economy remains relatively closed over this period. Prideaux (2009, p. 16) notes that "Brazil's imports and exports taken together were equivalent to 22% of its GDP in 2007, compared with 23% for America".

Finally, it is worth noting that the Brazilian economy has been marked by strong state intervention throughout the 1907 to 2008 period. Politics played an active role that shaped the socio-economic structure and the city size distribution.

1996, p. 428) in order to 'take-off' the development; then, at the second stage, that development is expanded to the less developed regions. The problem with this "selective support" of regions dependent on their development stage is that it creates regional inequality from the outset.

4.4.2. Urbanisation in Brazil

Figure 4.1 shows that the population of Brazil grew from 20.3 million in 1907 to 191.9 million inhabitants in 2008, which implies an annual average growth rate of 2.2 percent. Table 4.2 presents the evolution of the urban population in the sample of urban areas. The urban population defined by the sample of this chapter increased from 53 percent of the total population in 1907 to 70 percent in 2008. For comparison, the urban population share was estimated by the UN Secretariat to have been 36.2 percent in 1950, increasing to 86.5 percent in 2008. The smallest urban area in 1907 was Goianésia do Pará with a population of 200. It remained the smallest urban area until 1945, after which Caracaraí took that place. The latter's population was 18,789 in 2008. The largest city in 1907 was Rio de Janeiro, with a population of just over 1 million. Its population had increased to 4.8 million in 1960. From 1961, São Paulo became the largest city, with a population of close to 20 million in 2008. The average urban area population increased from 58,401 in 1907 to 730,383 in 2008.

Figures 4.2 and 4.3 visually display the evolution of the urban system in Brazil between 1907 and 2008 (see Appendix 4.2 for changes in urban areas' population across decades). Clearly, the average population of urban areas increased 12-fold over the century. However, population growth was not at a steady rate over the century. Over time, population growth has changed in an M-shaped pattern. Population growth first peaked in the 1910s and then dropped to a low in the 1930s. After that, growth increased again until the 1950s, followed by a drop and subsequent stabilisation of the growth rate by the 1990s. The second part of this "M" pattern is consistent with the law of diminishing returns of land use. In other words, the increase of the urban area population is limited by the contiguous land area.

Figure 4.1: Population of Brazil, 1907 to 2008

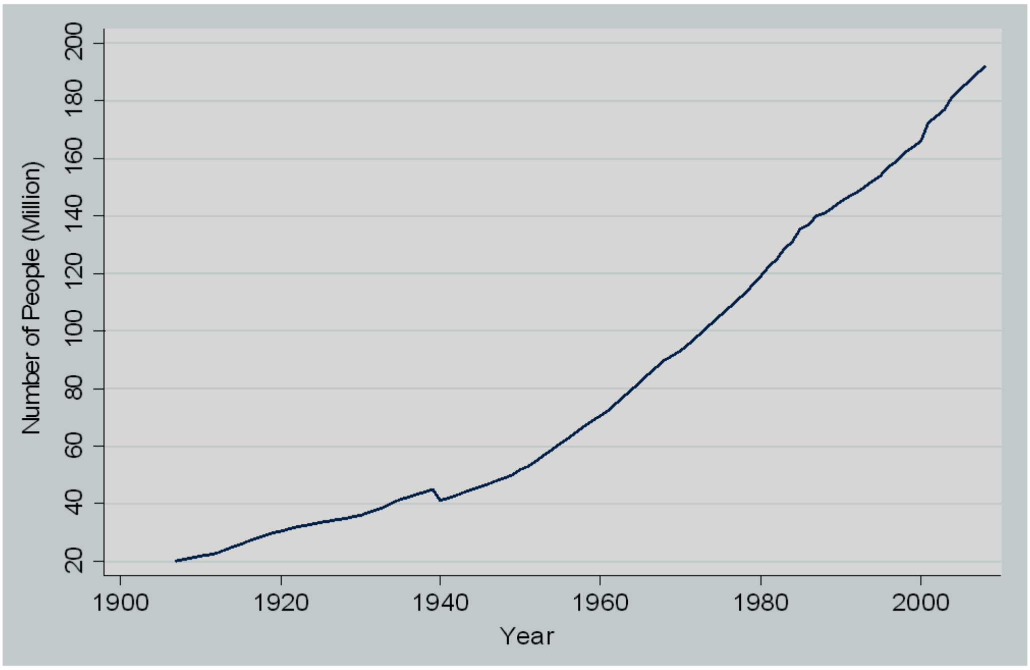


Table 4.2: National Population, Urban Population, and Urban Areas Sample (N=185)

Year	Total Population (1)	Total Sample Urban Population (2)	Total Sample as a Percentage of Total Population (3) = [(2)/(1)]*100	Minimum Urban Area Size	Maximum Urban Area Size	Average Urban Area Size	Percentage of Urban Population*
1907	20,253,609	10,804,332	53.35	200	1,039,082	58,401	
1910	21,819,738	11,670,719	53.49	300	1,103,057	63,084	
1920	30,559,034	14,675,734	48.02	600	1,378,865	79,328	
1930	36,000,000	18,098,944	50.27	787	1,814,562	97,832	
1940	41,169,321	20,431,303	49.63	1,200	2,203,345	110,439	
1950	51,941,078	26,507,511	51.03	869	3,137,977	143,283	36.2
1960	70,624,622	37,592,468	53.23	3,321	4,811,937	203,202	44.9
1970	93,134,846	52,516,454	56.39	4,421	8,063,414	283,872	55.8
1980	119,011,052	73,585,193	61.83	6,000	12,465,119	397,757	67.4
1990	145,000,000	93,571,199	64.53	8,577	14,800,000	505,790	74.8
2000	166,112,518	112,609,413	67.79	10,457	17,296,131	608,699	81.2
2008	191,943,158	135,120,951	70.40	18,789	19,859,740	730,383	86.5**

Notes: For the years without data on the sources, total population was assumed based on the smoothness of the population curve. The minimum area is Goianésia do Pará from 1907 to 1945 and Caracaraí from 1946 to 2008. The maximum area is Rio de Janeiro from 1907 to 1960 and São Paulo from 1961 to 2008. ** It refers to 2010.

*Source: The United Nations Secretariat, World Urbanization Prospects: The 2007 Revision, http://esa.un.org/unup.

Figure 4.2: The Urban Population of Brazil, 1907

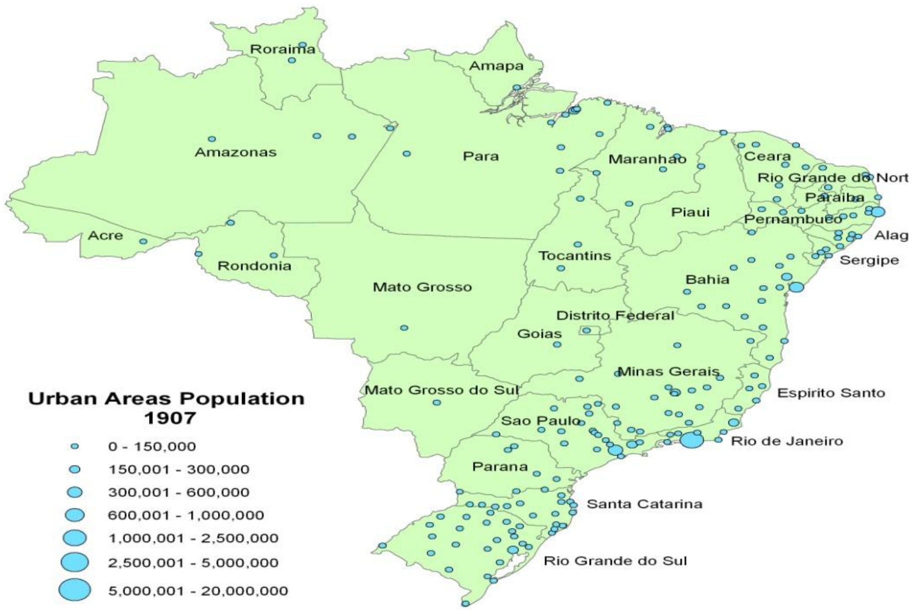
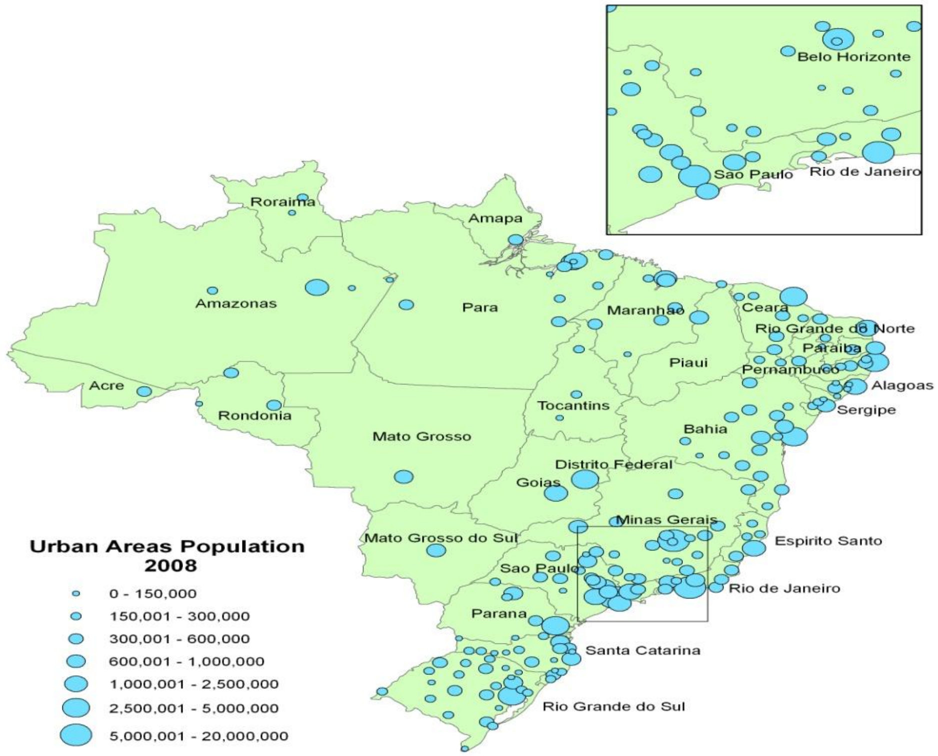


Figure 4.3: The Urban Population of Brazil, 2008



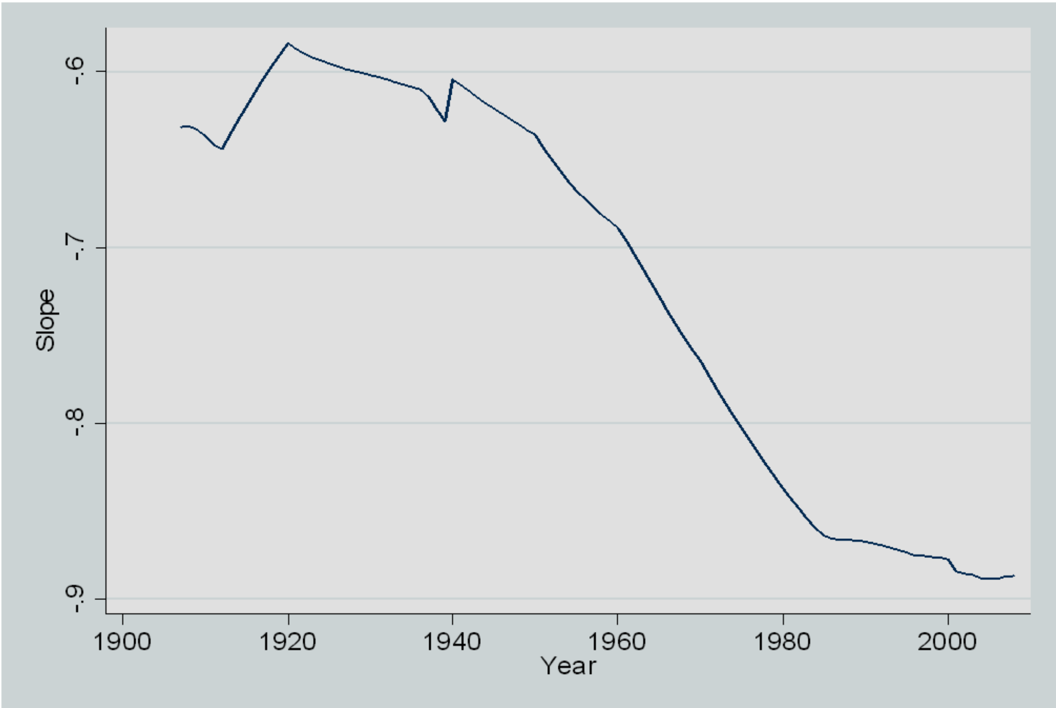
4.5. Empirical Results

4.5.1. Traditional Power Law Equation

This subsection first presents OLS estimates of the power law parameters for the 100 largest urban areas. The intercept shows a steep growth of the largest urban area: first, Rio de Janeiro from 1907 to 1960 (log size increases from 13.51 to 14.76), then São Paulo from 1961 to 2008 (log size grows from 14.82 to 16.64).

Figure 4.4 shows that the slope coefficient of the loglinear equation decreases from -0.63 in 1907 to -0.89 in 2008. Looking at the whole period, the figure clearly shows Gibrat's law of proportional city growth does not hold, because the slope is not constant. Moreover, Zipf's law does not hold either because the power law parameter is less than one. Except for 1939-1940, the city size distribution has become increasingly uneven. In other words, there is a city size divergence (or a convergence to the Zipf's law).

Figure 4.4: The Slope of the Power Law Equation: Brazil, 1907 to 2008



The slope behaviour can be divided into three phases. Phase I: the slope is fairly static around -0.6 from 1907 to 1939. This period included manufacturing industry development in Brazil but the economy was essentially agricultural, producing and exporting coffee - the main source of state revenue. Imports of machines to develop manufacturing were partially limited by a fall in state revenue as a result of the Great Depression (1929-1933). Labour immigration, to be employed in manufacturing, was only beginning. Phase II: the slope steeply falls from -0.60 to -0.86 from 1940 to 1983. In this period industrialisation occurs either by import substitution or by industrialisation policies that had sectoral targets. Policies favoured urban areas, immigration restriction and internal migration. Politically, this period alternates between two extremes: democracy (1945-1964) and dictatorships (1940-1945; 1964-1983). Phase III: the slope is relatively stable, changing from -0.86 to -0.89 from 1984 to 2008. Trade liberalisation and weak economic growth characterise this period. Politically, this period represents a return to democracy after 20 years of dictatorship.

Comparing Brazil with the USA provides an interesting contrast. The absolute value of the slope parameter based on the distribution of US metropolitan areas is about one for the entire 20th century. This confirms both Gibrat's law (Krugman, 1996, p. 400) and Zipf's law (Anderson & Ge, 2005, p. 758, footnote 1; Nitsch, 2005, p. 92; and Rossi-Hansberg & Wright, 2007, p. 598) for the US. Another interesting comparison is China. While for Brazil both laws are rejected, for China, Zipf's law is accepted before the 1979 Reforms and Gibrat's law is rejected after the reforms (Anderson & Ge, 2005, p. 758). While simple inter-country comparisons tend to trivialise complex differences between countries, it is nonetheless clear that stages of industrialisation and development of the market economy have a major impact on city size distributions.

The fit of the power law / Pareto distribution is very good in Brazil: the Adjusted R-Squared is a V-shaped curve that starts close to 0.98 in 1907 and ends close to 0.99, but with a global minimum at 0.93 in 1960. The power law fits least between the 1940s and the 1970s. This period coincides with Phase II of the development of the power law slope (1940-1983) during which industrialisation occurred. These findings contrast with Black & Henderson (2003) using US metropolitan areas decadal data from 1900 to 1990. These authors found that power law fits better during the industrialisation phase (e.g. R^2 is: 1900: 0.981; 1910: 0.979), and is worse for the recent decades as the US has become a services-oriented economy (e.g. R^2 is: 1980: 0.957; 1990: 0.952).[62]

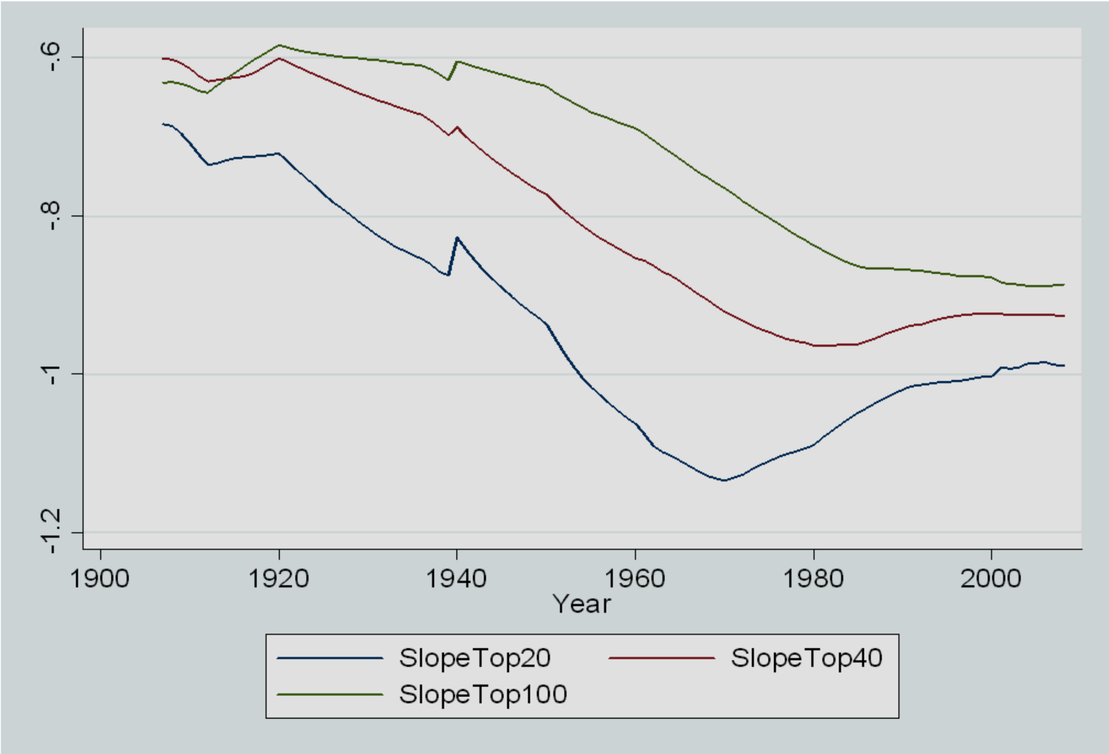
As a robustness check, several cut-offs were tried with respect to defining the large urban areas. When restricting the sample to only the 10 or 40 largest urban areas, the rank-size rule holds approximately for certain sub-periods, but the slope

[62] Industrialisation of the United States economy occurred predominantly between 1880 and 1900 (http://www.britannica.com).

parameter varies over time. For example, with 10 urban areas, the slope decreases from -0.96 to -1.04 from 1926 to 1939 and then increases from -1.04 to -0.95 from 1982 to 1999. Considering the top 40 urban areas, the slope decreases from -0.95 to -0.96 from 1975 to 1980, then increases from -0.96 to -0.95 from 1981 to 1989. In both cases, outside of these time intervals the rank-size rule does not hold.

Some studies (Song & Zhang, 2002; Black & Henderson, 2003; Soo, 2007; Xu & Zhu, 2009) find that the slope increases with the movement to the upper tail of the distribution. This is also largely the case with the Brazilian data. Considering the top 10, 20, 40, and 100 urban areas and ignoring the sub-period 1907-1914 in which the absolute value of the slope of the top 100 urban areas is greater than that of the top 40, there is a monotonic increase in the slope with the movement to the upper tail from 1914 to 1953. The slope increases with the movement to the upper tail for the top 20, 40, and 100 urban areas' cut-offs for 94 (out of 102) years of the series (see Figure 4.5).

Figure 4.5: The Slope for the 20, 40, 100 largest Urban Areas in Brazil: 1907 to 2008



The estimated α in this chapter is consistent with the meta analysis of Nitsch (2005) who found that the power parameter is less than 1 if estimation is based on metropolitan areas (rather than city proper data) and uses post-1900 data. It is also consistent with Soo (2005) who found that the average α using urban agglomeration data is less than 1. Soo (2005) criticises Rosen & Resnick’s earlier (1980) study that suggested a value of 1 with urban agglomeration data. This chapter found that the absolute estimated power parameter from 1907 to 2008 is less than 1. Rosen & Resnick’s (1980, p. 171) estimation for metro areas is either below 1 for Brazil, Italy and Mexico or above 1 for France and India. However, for the US it is 1, which confirms the rank size rule. Soo (2005, p. 253) used urban agglomeration data of the 1990s for 26 countries to find a power parameter less than 1 for 22 cases, approximately 1 for three cases and higher than 1 for one country. Soo’s (2007) estimation using urban areas in Malaysia varies from 1.08 in 1957 to 0.86 in 2000. Gangopadhyay & Basu (2009) used urban agglomerations data for India and China. For both countries, the parameter is very large and even close to 2. For example, for India, estimation suggests around 1.9 (between 1980 and 2000). For China, it is around 1.8-2 for the same period. Brakman, Garretsen & van Marrewijk (2009, pp. 318-319) used urban agglomeration data to find power parameter greater than 1 for nine out of 22 countries.

The result indicates that the power parameter has increased in absolute value from 0.63 in 1907 to 0.89 in 2008. This result rejects Zipf’s and Gibrat’s laws, implying that urban areas’ distribution in Brazil is more even than Zipf’s law predicts.[63] As a consequence, the ratio of the largest urban area (São Paulo) to the second largest (Rio de Janeiro) is 1.23 for 2008. However, also for 2008, the ratio of the largest urban area to the third and the fourth largest is respectively, 3.89 and 5.35 rather than 3 and 4.

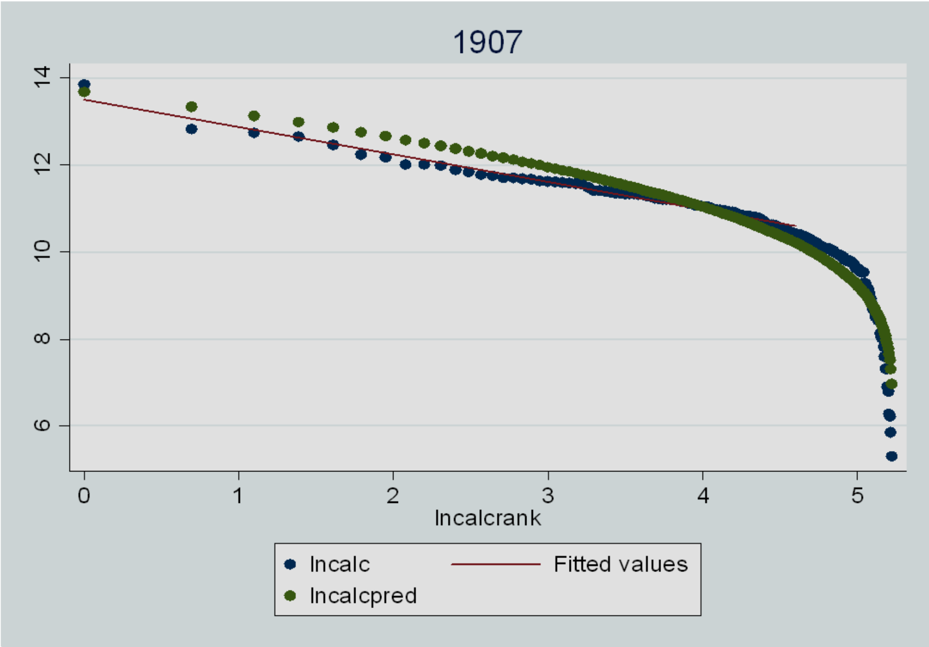
One of debates in the power law literature is whether the Pareto distribution fits the data better than the lognormal distribution. Using 2000 Census for all US cities, Eeckhout (2004, 2009) argues that both curves can fit the data equally well. On the other hand, using the same database, Levy (2009) pointed out that the lognormal curve fits better to the middle and bottom cities of the distribution, whereas the Pareto line fits the upper tail better.

[63] In this subsection, the conclusion about Gibrat’s law is based on the theoretical fundamentals that relate this law with Zipf’s law (Gabaix, 1999b; Eeckhout, 2004). However, an econometric test for Gibrat’s law is applied in the subsection below.

The power law fits well for the 100 largest urban areas in Brazil, which is illustrated by an adjusted R^2 of at least 0.93 over the series. To test whether Eeckhout or Levy's conclusion applies to the Brazilian data, a cut-off of the top 100, 40, 20 and 10 urban areas has been considered. Broadly speaking, a support for Eeckhout's claim of the equivalence of the Pareto and lognormal distributions for the larger cities for the period up to the 1950s was found. For more recent years, Levy's conclusion, that the lognormal of the entire distribution is inappropriate for the tail of the largest cities, is also correct for Brazil. This is illustrated by comparing the fit of both distributions in both 1907 and 2008, see Figures 4.6 and 4.7. (For a change across decades in the quality of the fit of both the lognormal and Pareto curves in the city size distribution, see Appendix 4.3.) Therefore, considering the entire series, the conclusion is that both the Pareto and lognormal distributions fit well to the cities' data from 1907 to 1943 (Eeckhout, 2004, 2009; Giesen, Zimmerman & Suedekum, 2009). From 1944 to 2008 this chapter's estimates support for the lognormal distribution fitting better to the middle and bottom cities (Levy, 2009), whereas Pareto distribution describes better the very upper tail cities (Levy, 2009; Giesen, Zimmerman & Suedekum, 2009).

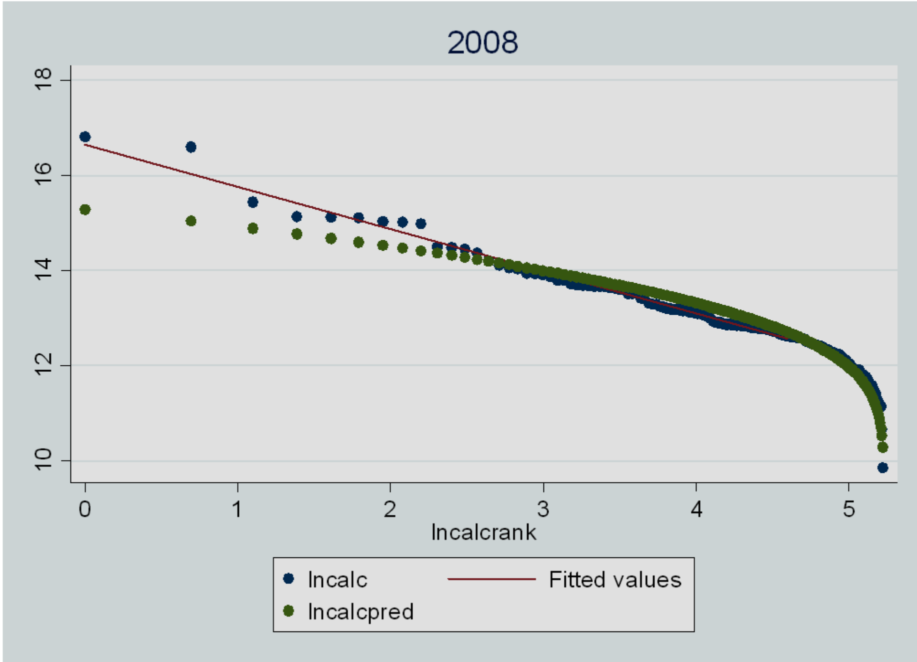
Giesen, Zimmerman & Suedekum (2009), using 2,075 German settlements in 2006 and 25,359 USA cities in 2000, supported Eeckhout (2004) by showing that the lognormal fits these countries' city data. In Brazil's case the fit of a lognormal is not perfect because, depending on the period, it either does not match some observations in the upper tail or on "swelling" segment between the middle and the bottom of the distribution (for instance, compare Figures 4.6 and 4.7). Giesen, Zimmerman & Suedekum (2009) find that data from Germany and the USA are better described by a "Double Pareto Lognormal" (DPLN), which is a lognormal with a Pareto fit for both the upper and lower tails of the distribution. Since the sample is of urban areas in Brazil rather than all Brazilian urban areas in the database, a comparison between these authors' findings and this study is only done with the upper tail. The finding of this chapter differs from Giesen, Zimmerman & Suedekum (2009) in the sense that the Pareto fitting improvement in the upper tail is conditioned by the urban areas' cut-off and by the chosen period as shown in the comparison between this chapter's result and Levy's (2009).

Figure 4.6: Pareto and Lognormal Distributions: Brazil, 1907



Note: lncalc is the natural logarithm of population; lncalcrank is the natural logarithm of population rank; lncalcpred is the predicted logarithm of population from fitting a lognormal distribution; Fitted values line refers to the power law estimated on the 100 largest cities.

Figure 4.7: Pareto and Lognormal Distributions: Brazil, 2008



Note: lncalc is the natural logarithm of population; lncalcrank is the natural logarithm of population rank; lncalcpred is the predicted logarithm of population from fitting a lognormal distribution; Fitted values line refers to the power law estimated on the 100 largest cities.

4.5.2. Corrected Power Law Equation

This chapter focuses on the traditional power law equation for which there is a wide range of studies to compare internationally[64]. Considering developments of this method over the last decade, however, this sub-section tests the Gabaix and Ibragimov (2006) correction[65] for the power law to check the robustness of the results from the traditional power law equation. The parameters of the latter equation, according to part of the literature, are biased (Gabaix and Ibragimov, 2006; Nishiyama, Osada & Sato, 2008; Bosker, 2008). By applying Gabaix & Ibragimov's (2006) correction method, the criticism by Nishiyama, Osada & Sato (2008) that the parameters of the traditional power law equation are biased is addressed, since estimations using approaches from each of these two studies are consistent with each other (Nishiyama, Osada & Sato, 2008, p. 708)[66].

The results of the corrected power law equation are as follows. The intercept shows a steep growth of the largest urban area: first, Rio de Janeiro from 1907 to 1960 (log size increases from 13.37 to 14.62), then São Paulo from 1961 to 2008 (log size grows from 14.67 to 16.44). The behaviour of the slope (which is the full line in Figure 4.8) resembles that for the traditional power law; however, in this case, the power law fits slightly better. Over the series, the Adjusted R-Squared is at least 0.95[67]. Both of the bounds for the 95 percent confidence interval, i.e the lower bound (long-dash line) and upper bound (short-dash line), are consistent with the slope (Nishiyama, Osada & Sato, 2008; Bosker, 2008). The power

[64] Most studies test the traditional power law using city data; others use metropolitan area or urban area data. Both types of studies have been outlined and discussed in the section 4.2 and in the subsection above. The studies that employ the corrected power law equation (Gabaix and Ibragimov, 2006) using metropolitan area or urban area data are few (Gabaix and Ibragimov, 2006; Nishiyama, Osada & Sato, 2008).

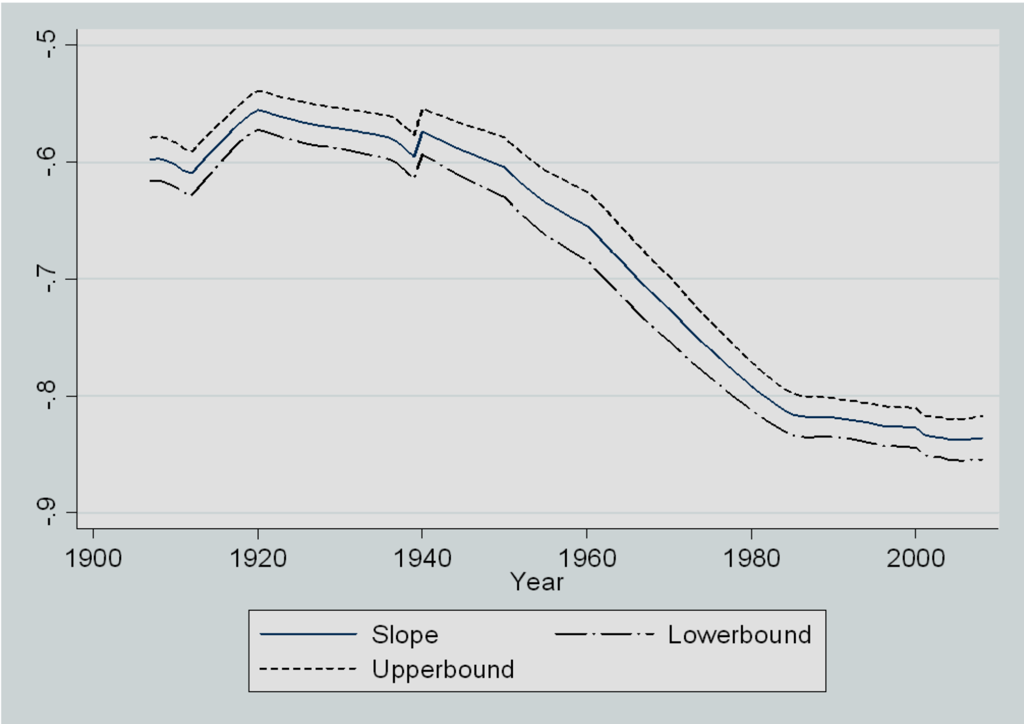
[65] According to Gabaix & Ibragimov (2006, p. 3, equation 1.4), after the shift on the rank, the Pareto equation for city data becomes: $\log(\text{Size}) = c - d\log(\text{Rank} - 1/2)$, where c is the intercept and d is the power parameter. Both parameters are estimated by OLS.

[66] Other branch of slightly different literature analyses the similarities between the lognormal and q-exponential city size distributions (González-Val, Ramos-Gutiérrez & Sanz-Gracia, 2011). This study uses all country city data while this chapter considers urban area data as truncated, because, due to data limitations, a large sample of urban areas rather than all urban areas in Brazil has been used. According to the authors, there is a problem in comparison of city data across countries due to variations in city definitions (González-Val, Ramos-Gutiérrez & Sanz-Gracia, 2011, p. 6, footnote 3). However, the interesting result of their study is that both q-exponential and lognormal are suitable to describe city data and complement each other. Therefore, the employment of q-exponential for this chapter's dataset should yield fitting results that resemble those from lognormal distribution.

[67] This is evidence that the power law parameters found using Gabaix & Ibragimov's (2006) method are more robust. As a result, all tests applied in this subsection focus on this method.

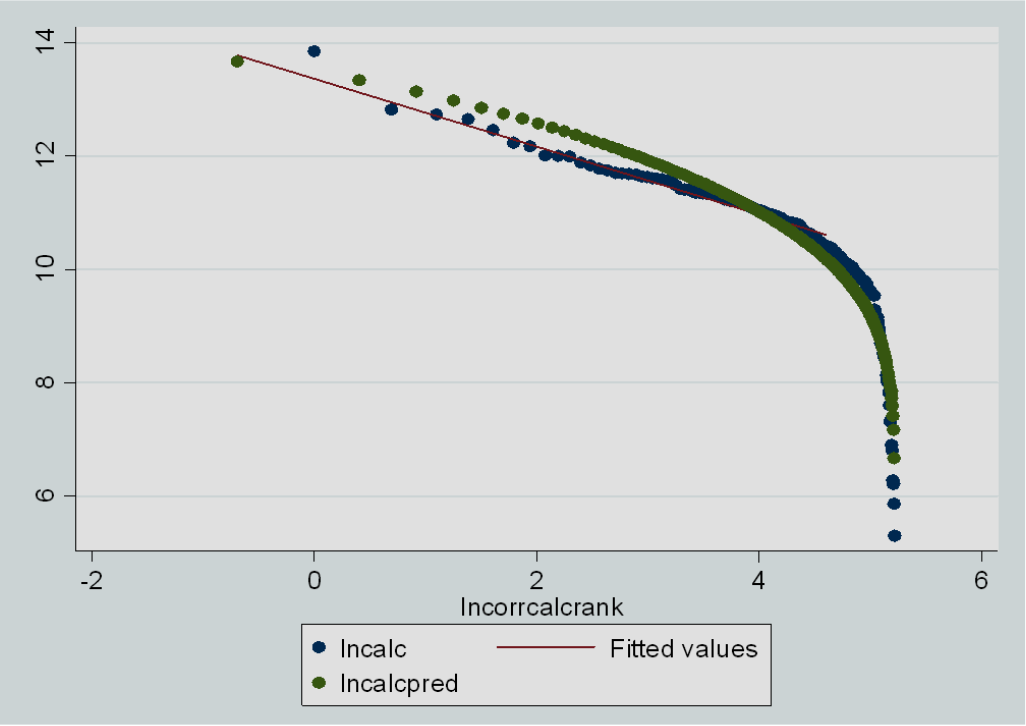
parameter has increased in absolute value from 0.60 in 1907 to 0.84 in 2008[68]. This result rejects both Zipf's and Gibrat's laws, implying that urban areas' distribution in Brazil is more even than Zipf's law predicts. Thus, the theoretical implications of the results from both methods are identical. The fitting of both Pareto and lognormal distributions (Eeckhout, 2004, 2009; Skouras & Ioannides, 2010; Levy, 2009) observed and discussed in the previous subsection is supported by the plotting of both distributions after the ranks have been corrected. Despite a relative steepness of curves for the corrected power law method, the time series shows results consistent with those shown for 1907 and 2008 using the traditional power law equation. Figures 4.9 and 4.10, and Figures 4.6 and 4.7 resemble each other, respectively.

Figure 4.8: The Slope of the Corrected Power Law Equation: Brazil, 1907 to 2008



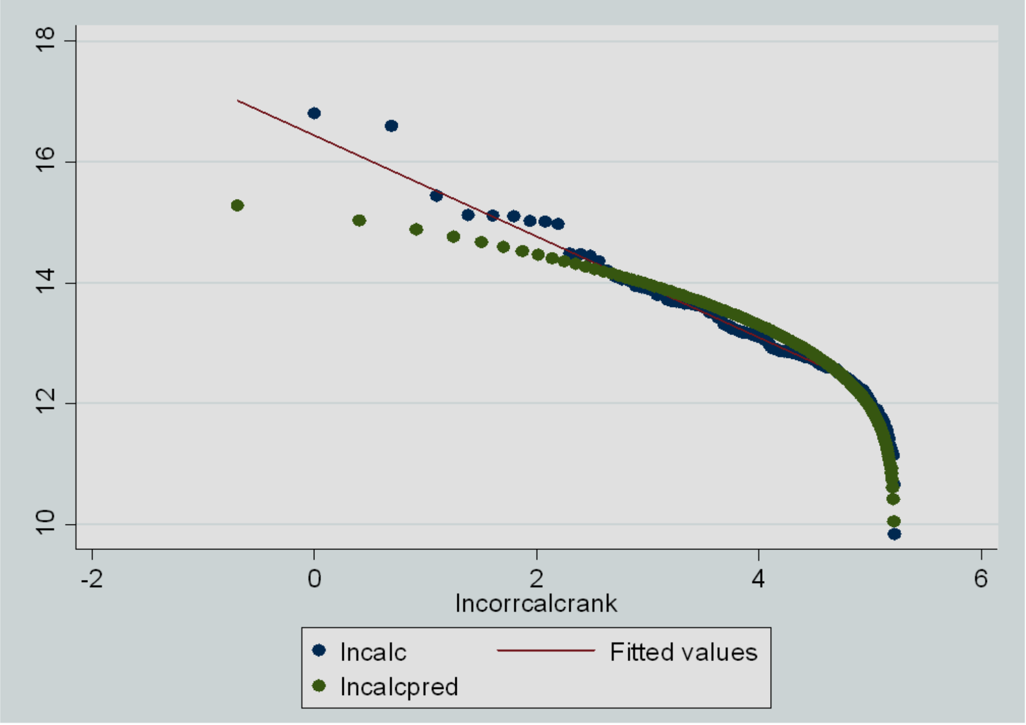
[68] Therefore, the corrected method shows that although there is a convergence to Zipf's law, the speed is smaller.

Figure 4.9: Corrected Pareto and Lognormal Distributions: Brazil, 1907



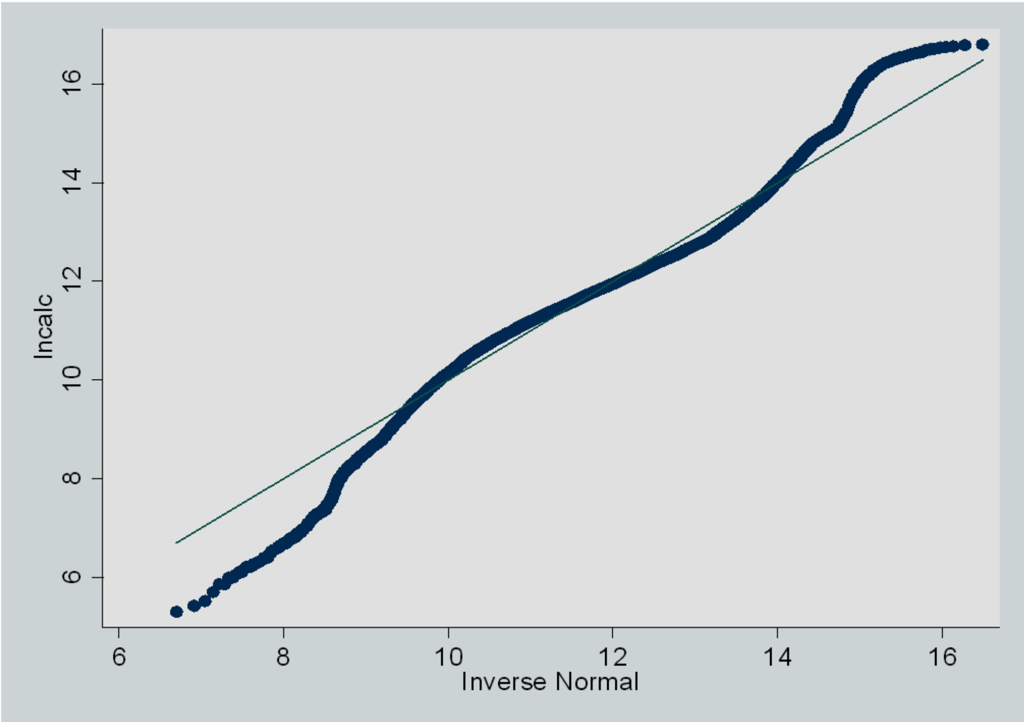
Note for Figures 4.9 and 4.10: lncalc is the natural logarithm of population; lncorrcalcrank is the natural logarithm of population corrected rank; lncalcpred is the predicted logarithm of population from fitting a lognormal distribution; Fitted values line refers to the power law estimated on the 100 largest cities.

Figure 4.10: Corrected Pareto and Lognormal Distributions: Brazil, 2008



From 1907 to 2008, while the sample mean (μ=lnSize) increases from 10.33 to 12.78, the standard deviation (sigma=sd(lnSize)) decreases from 1.31 to 0.98. The robustness of the lognormal fit is tested by quintile plot of the size of urban areas in the sample against quintiles of normal distribution, over the 1907-2008 period. The quintile plot of the size of urban areas (see Figure 4.11) deviates from the normal curve, essentially for the largest and the smallest areas. This is consistent with the rejection of normality hypothesis by the Shapiro-Wilk test (W=0.9673) at 1 percent significance level. However, both the Kolmogorov-Smirnov (KS)[69] test of goodness of fit (with D=0.0510; *p*-value=0.0000) and the Shapiro-Francia (with W'=0.967; *p*-value=0.3853) tests[70] accepted the null hypothesis, i. e. that the size of urban areas is normally distributed. Since Gibrat's law is econometrically tested by a panel unit root, there are no concerns with the implications of these conflicting results from the three normality tests.

Figure 4.11: Quintile Plot for the Size of urban areas: Brazil, 1907 to 2008



Note: lncalc is logarithm of urban area's population. This figure represents quintiles of size against quintiles of a normal distribution (thin line).

[69] Eeckhout (2004, p. 1434) also performs this test finding a much smaller D statistic with a *p*-value of 1 percent.

[70] Both Shapiro-Francia and Shapiro-Wilk tests for normality were applied, and the latter test rejected the normality hypothesis. These tests were chosen because the literature demonstrated that the Shapiro-Wilk test is more powerful than the Kolmogorov-Smirnov test (see Stephens, 1974). A break by the sub-periods 1907-1939, 1940-1983, and 1984-2008 does not significantly alter the results of these tests and of the quintile plots.

Additionally, the power law parameters were estimated from the corrected method for the size distribution of the 100 largest urban areas using a panel model with pooled data. Table 4.3 shows the results. The estimated intercept (in logarithm) is 14.73 and the power parameter is -0.68 (which is equal to the average slope across the whole series). Clearly, the estimates with pooled data also rejected Zipf's law. This model also shows that the power law is a property of the cross-sectional distribution of city sizes. The model explains the variation of urban areas (i.e. population of each urban area) relatively well. This is also a temporary theoretical confirmation that Gibrat's law did not apply to the Brazilian data.

Table 4.3: Panel Model with Pooled Data

Variables	**Pooled Model Estimates**
Constant	14.73 (577.60)
Logarithm of City Rank-1/2	-0.68 (-100.00)
Number of Observations	10202
Adjusted R-Squared	0.4950

Notes: Dependent variable is logarithm of city population. Values of *t* statistics are in brackets.

4.5.2.1. Tests of Gibrat's Law

To confirm the theoretical rejection of Gibrat's law (Gibrat, 1931; Eeeckhout, 2004, 2009), panel unit root tests are performed. In particular, the following specification is tested (Black & Henderson, 2003, p. 351; Bosker, 2008, p. 134)[71]:

$$\Delta lnS_{it} = \alpha + \delta_t + \gamma lnS_{it-1} + \epsilon_{it} \qquad (4.1)$$

On the left side of equation 4.1 is the population growth of urban area *i* between the periods *t*-1 and t. On the right side, α is the model's intercept, δ_t is a deterministic trend, S_{it-1} is the size of population in the urban area *i* in the period (t-1), and ϵ_{it} is the model's residual in time *t*. The null hypothesis to be tested is whether $\gamma = 0$ against the alternative that $\gamma < 0$. If the null hypothesis cannot be rejected, then the unit root is accepted, i. e., that urban areas' growth and their

[71] The period and the sample are large enough to employ panel unit root tests for Gibrat's law.

initial sizes are independent. Otherwise, the alternative hypothesis that there is a mean reversion in urban areas growth process will be confirmed.

Figure 4.12: Urban areas' growth, 1907 to 1908

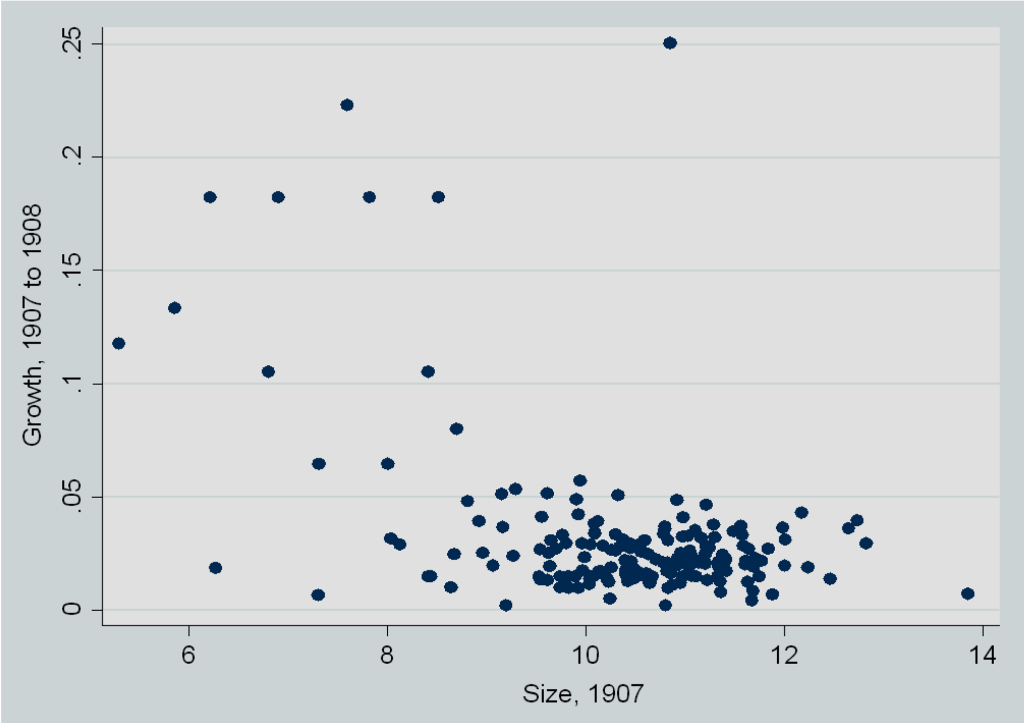
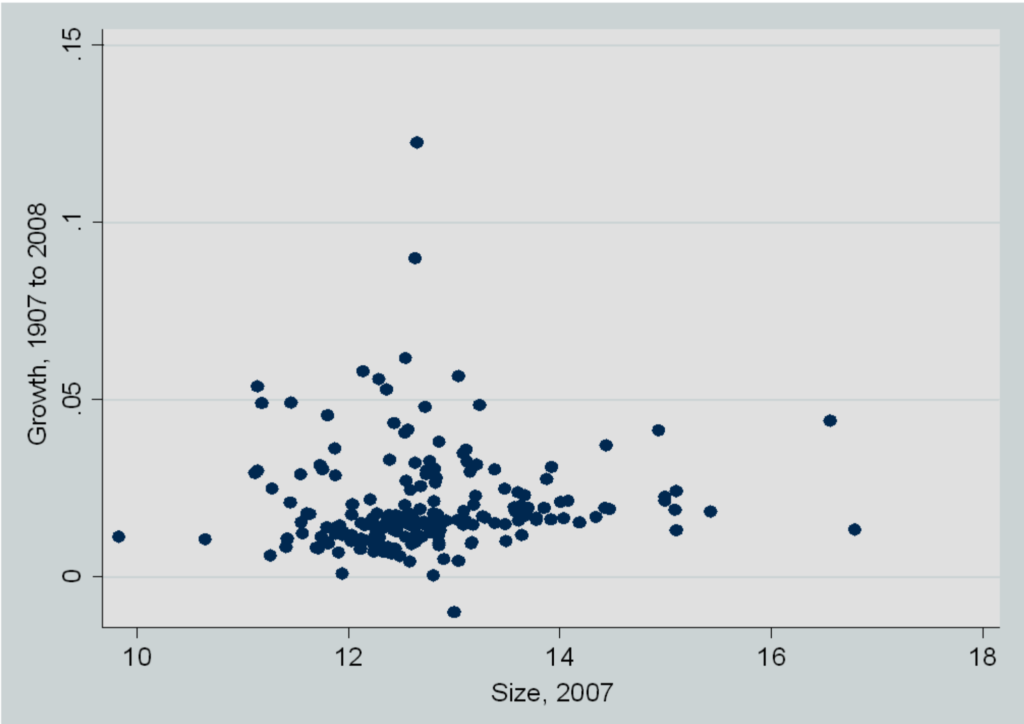


Figure 4.13: Urban areas' growth, 2007 to 2008



From the inspection, Figures 4.12 and 4.13 show unclear functional relationships between urban areas' growth and their sizes. Both figures indicate that urban areas' growth from 1907 to 1908 and from 2007 to 2008 and their urban areas' sizes in 1907 and 2007, respectively, are independent. This is consistent with Gibrat's law. The equation (4.1) is estimated for the overall sample and for the largest 100 urban areas (Black & Henderson, 2003, p. 353) and also for individual urban areas (Bosker, 2008, p. 135).

The subsection 4.4.1 notes that the Brazilian economy faced structural socio-economic transformation over the 1907-2008 period. This is reflected on the behaviour of the power parameter (see Figures 4.4 and 4.8). As pointed out by some previous studies, Gibrat's law breaks over time (see, for instance, Anderson & Ge, 2005; Glaeser, Ponzetto & Tobio, 2011, pp. 7-8). Consequently, this chapter tests Gibrat's law for each of the three phases (or "shocks") in which the power parameter has a unique behaviour.

Table 4.4: Tests of Gibrat's Law

Period	Variables	Overall Sample	100 largest urban areas[1]
1907-1939	$\ln S_{it-1}$	-0.0068***	-0.0010*
		(-18.15)	(-1.72)
	Time fixed effects	Yes	Yes
	Number of Observations	5920	3168
	Adjusted R-Squared	0.0861	0.0327
1940-1983	$\ln S_{it-1}$	-0.0036***	0.0047***
		(-7.55)	(6.26)
	Time fixed effects	Yes	Yes
	Number of Observations	8140	4356
	Adjusted R-Squared	0.0414	0.0389
1984-2008	$\ln S_{it-1}$	-0.0001	0.0010***
		(-0.25)	(2.72)
	Time fixed effects	Yes	Yes
	Number of Observations	4625	2475
	Adjusted R-Squared	0.0445	0.0855

	Panel unit root tests on urban areas sizes t-statistic (p-value)		
Test (Overall Sample)[2]	**1907-1939**	**1940-1983**	**1984-2008**
Levin-Lin-Chu	**-12.4184**(0.000)	-0.0934(0.463)	**-10.1184**(0.000)
Im-Pesaran-Shin	-1.1549 (1.000)	**-1.7243** (0.000)	-1.7752(0.519)

[1]The differences between these results and those found considering the corrected ranks are very small (e. g. the Adjusted R-Squared in this table is less than 2 percent smaller in periods I and III and less than 10 percent smaller in period II); for the upper part of the table, the values in brackets are *t*-statistics. *,**,***=Significant at 10%, 5%, and 1%, respectively. For Levin-Lin-Chu test, the t statistic for panel unit root test is adjusted t*. For Levin-Lin-Chu test, the null hypothesis is: H_0: Panels contain unit roots, against the alternative hypothesis that Ha: Panels are stationary. For the Im-Pesaran-Shin, the t-statistic reported is t-bar, and the hypotheses are: H_0: All panels contain unit roots against Ha: Some panels are stationary. Also, for this test, in contrast to Levin-Lin-Chu test, the higher the p-value, the higher the probability of acceptance of the null hypothesis.
[2]Bold results are a significant rejection of null hypothesis.

The upper part of Table 4.4 shows that, for the first two sub-periods, for a panel of the whole sample, the unit root hypothesis is rejected at 1 percent significance level and the hypothesis of a mean reversion is confirmed, which shows a rejection of Gibrat's law and evidence of a negative relationship between growth of urban areas and their initial sizes (Black & Henderson, 2003, p. 353)[72]. The robustness of the unit root results is confirmed by Levin-Lin-Chu and Im-Pesaram-Shin[73] unit root tests (see the bottom of Table 4.4), which also rejected the null hypothesis[74] for panels of the overall sample. However, for the sub-period 1984-2008, the null hypothesis cannot be rejected because even though the coefficient of $\ln S_{it-1}$ is -0.0001, it is not significant even at 10 percent significance level (t-statistic is -0.25 only), which supports Gibrat's law, i.e., for this sub-period city growth and city size were independent. The results for the 100 largest urban areas rejected the unit root (Gibrat's law) in favour of mean reversion hypothesis at 10 percent significance level for the sub-period 1907-1939. Nevertheless, since 1940, the correlation between initial city size and city growth is positive and statistically significant at 1 percent level, which indicates that larger urban areas had higher growth, implying increase of agglomeration of population in the 100 largest urban areas.

In order to evaluate whether urban areas' growth is correlated to urban areas' own initial sizes, the Levin-Lin-Chu and Im-Pesaran-Shin unit root tests[75] (see Levin, Lin & Chu, 2002; Im, Pesaran & Shin, 2003; Baum, 2003; Bosker, 2008) were applied for panels of individual urban areas. Table 4.5 summarises the results by sub-period and by significance levels, namely 1, 5, and 10 percent. Again, the null hypothesis tested is whether Gibrat's law holds. Table 4.5 shows

[72] The structural change process that leads to this mean reversion in population growth is consistent with changes since the second-half of 20th century in the other socio economic indicators. For instance, the average years of schooling in Brazil for people aged 25 years old or more grew from 3.4 to 6.9 years of schooling from 1960 to 2007 (IPEA). Education is negatively correlated with birth rate.

[73] The spirit of these tests is identical to that of the specification (4.1), and both include panel means and time trends to test γ. However, both differ given that while Levin-Lin-Chu assumes that all cities' panels have the same autoregressive parameter, and includes one lag ADF regressions, the Im-Pesaran-Shin has no lags included in ADF regressions and allows each city panel to have its autoregressive parameter, γi(www.stata.com; Baum, 2003). For a brief literature review on these tests for Gibrat's law, see Bosker (2008, chapter 5).

[74] For Im-Pesaram-Shin unit root test, the null hypothesis is confirmed for the first sub-period. This result is not inconsistent with that of Black & Henderson's model because the coefficient of $\ln S_{it-1}$ from equation 4.1 is almost zero.

[75] In the panel unit root tests for the overall sample and individual urban areas, common autoregressive parameters, panel means, and time trends were included.

there is a variation in the number of urban areas in which Gibrat's law is rejected in terms of significance levels and whether either the Levin-Lin-Chu (LLC) or Im-Pesaran-Shin (IPS) test is applied. However, one key finding can be pointed out. Gibrat's law holds for the third sub-period in which, compared to the previous sub-periods, the highest number of urban areas for which the null hypothesis is rejected is the smallest when a comparison between LLC and IPS tests is done. This indicates, for the third subperiod, a stochastic growth process in Brazil. These findings are consistent with previous ones that indicated that Gibrat's law held in recent decades, but for the largest urban areas size has actually been even more important for growth, which is reflected by a positive correlation between growth and size for the 100 largest areas since 1940.

Table 4.5: Panel unit root tests for individual urban areas

Period	Significance level for rejection of null hypothesis	Levin-Lin-Chu (Number of urban areas)	Im-Pesaran-Shin (Number of urban areas)
1907 to 1939	1%	16	11
	5%	27	3
	10%	27	9
	Total	**70**	**23**
1940 to 1983	1%	17	74
	5%	16	8
	10%	6	10
	Total	**39**	**92**
1984 to 2008	1%	12	5
	5%	25	10
	10%	21	10
	Total	**58**	**25**

4.6. Conclusion

This chapter used a unique dataset to analyse the evolution of the size distribution of urban areas in Brazil by means of a fixed sample of 185 urban areas observed annually from 1907 to 2008. Four conclusions concerning Brazil's urban system growth process can be drawn from the estimations. First, the absolute value of the traditional power law parameter (α) of the size distribution of the 100 largest urban areas increases from 0.63 to 0.89 from 1907 to 2008. Although the power law holds, Zipf's law is rejected. To verify the extent to which the power law is stable irrespective of the method used, this chapter analysed cross-sections and time series dimensions simultaneously through employment of a panel model with pooled-periods data using the corrected power law equation (Gabaix & Ibragimov, 2006)[76]. In this case, the absolute value of the power law parameter is equal to 0.68, which also rejects Zipf's law. Gibrat's law was tested by the panel unit root tests (which test the log-linear relationship between growth of urban areas and their initial sizes) for the whole sample of urban areas for each of the three sub-periods from 1907 to 2008 in which the power parameter has a regular behaviour. The key finding was that while Gibrat's law is rejected in favour of the mean reversion hypothesis for the 1907-1983 sub-period, after 1984 it holds. The robustness of this finding is confirmed by the Levin-Lin-Chu and Im-Pesaran-Shin panel unit root tests for the whole sample and (to a smaller extent) for individual urban areas.

Conversely, panel unit root tests for the 100 largest urban areas rejected Gibrat's law at 10 percent significance level only, providing evidence for a weak mean reversion for the relationship between urban areas' size and their growth rate for the sub-period 1907-1939. Since 1940, a positive and significant (at 1 percent level) relationship was found between both variables. The latter evidence, combined with the convergence to Zipf's law, indicates an increase in agglomeration of population in the 100 largest urban areas over the last seven decades of the analysed period. Secondly, α parameter increases with the movement to the upper tail of the distribution, as shown in the literature. This illustrates that the regularity that Zipf's law states is stronger for the largest areas in Brazil. For example, the inequality among the 40 largest urban areas is higher than that observed when considering the largest 100 urban areas, although for the

[76] This method's cross-sectional absolute estimates of the power parameter vary from 0.60 in 1907 to 0.84 in 2008.

former there is evidence of a decline in absolute value of the slope parameter since the late 1970s (see Figure 4.5 above). Thirdly, two remarkable regularities were found in Brazil. While the industrialisation period is associated with the power parameter fall, the pre- and post-industrialisation periods are related to a relatively stable parameter. The fit of the power law OLS model is least for the industrialisation phase, which is the intermediary stage of development of Brazil in which the power parameter steeply falls. However, this model performs better during the pre- and post-industrialisation period of Brazil. Finally, the tests indicated that both the Pareto and lognormal distributions describe to some extent the urban areas' size distribution during the 20th century in Brazil. These four conclusions are consistent with theories that argue that increasing returns to scale arise as a result of agglomeration of economic activities (Brakman, Garretsen, and van Marrewijk, 2009).

This chapter has a caveat that will be addressed in future research. Given that a fixed sample of urban areas has been used, there is no evaluation of the impact of birth and death of urban areas on the city size distribution.

CHAPTER 5 – CLASSIC AND SPATIAL SHIFT-SHARE ANALYSIS OF STATE-LEVEL EMPLOYMENT CHANGE IN BRAZIL[77]

5.1. Introduction

The Brazilian economy has gone through a remarkable transformation since the difficult times of the last quarter of the 20th century. Brazil is now seen as one of the engines of global economic growth and together with Russia, India and China making up the often cited BRIC acronym. During the next decade, Brazil is expected to overtake Britain and France and become the world's fifth largest economy, with São Paulo possibly the world's fifth wealthiest city (The Economist, 2009).

Such rapid national development begs the question of whether the benefits are being reaped in all regions, with poorer ones catching up, or whether the gap between the rich and poor regions is widening. At present, Gross State Product (GSP) per capita in Rio de Janeiro and São Paulo is 50 percent higher than Brazil's Gross Domestic Product (GDP) per capita, but in the northeastern states of Piaui and Maranhão, GSP per capita is less than 30 percent of Brazil's GDP per capita.

To address such a question one would ideally carry out a formal econometric analysis along the lines of neoclassical or endogenous growth models (e.g. Barro and Sala-i-Martin, 2004). Alternatively, one might consider the dynamic adjustments suggested by models of the New Economic Geography (e.g. Brakman *et al*., 2001). In either case, a first requirement is the availability of reliable regional accounts data at sectoral and aggregate levels, plus a range of socio-economic indicators. In Brazil such subnational accounts data have been, until recently, rather incomplete or difficult to compare over time.

However, sub-national demographic and employment data are available on a consistent basis for several decades. Chapter 3 exploited such data to identify the impact of Marshall-Arrow-Romer, Porter and Jacobs' externalities in manufacturing by means of the Glaeser *et al*. (1992) approach. Here a broader approach to analyse state growth in Brazil involves considering all production sectors simultaneously. For this purpose, this chapter starts with the conventional shift-share accounting framework, which decomposes total growth in a region in

[77] Bill Cochrane of Waikato University provided helpful advice on computational aspects.

terms of national, industry-mix, and competitive shift effects (Dunn, 1960; Esteban-Marquillas, 1972; Arcelus, 1984; Berzeg, 1978, 1984; Haynes & Machunda, 1987; Dinc, Haynes & Qiangsheng, 1998; Dinc & Haynes, 1999). Despite criticisms and various alternative formulations, the classic shift-share approach remains popular after half a century of application (Knudsen & Barff, 1991; Hoppes, 1991; McDonough & Sihag, 1991; Loveridge, 1995; Knudsen, 2000; Shields, 2003).

However, the approach is extended in this chapter in five ways. First, the classic shift-share components are tracked over five consecutive quinquennia, starting in 1981. This provides a dynamic perspective on the shift-share decomposition. Secondly, a new structural change effect is defined and calculated to show that most states have been creating jobs in industries that nationally became more prominent and shed jobs in industries that contracted nationally, i.e. states generally did not go against the trends. Thirdly, a wide range of alternative shift-share decompositions proposed in the literature are calculated to show that these refinements lead to interpretations that remain very similar to those of classic shift-share analysis. Fourthly, the spatial patterns in the shift-share decomposition are identified by means of exploratory spatial data analysis (ESDA). Fifthly, Nazara and Hewings's (2004) spatial shift-share taxonomy is used to add a spatial component for each state into the shift-share decomposition, namely a measure of spatially weighted employment growth in surrounding states. Nazara and Hewings also introduce additional industry-specific spatial components into shift-share, but because the focus of this chapter is on regional aggregates rather than individual industries, the latter spatial shift-share taxonomy can be simplified to four-component decomposition. In this decomposition, the spatial component has an intuitively attractive interpretation, namely the regional rate of growth one might expect in the presence of full spatial spillover of growth in surrounding regions, after controlling for national industry-specific growth. Nazara and Hewings refer to this effect as (minus) the neighbour-nation regional shift effect.

Together with the classic shift-share decomposition, the spatial analysis provides evidence of a catching up of peripheral regions in Brazil, although agglomeration effects ensure that the dominance of the states of the south east remains. The results of this dynamic and spatial shift-share analysis are therefore consistent with those of the econometric literature on regional development in Brazil (see e.g. Rolim, 2008; Daumal & Zignago, 2010).

The period under consideration is 1981 to 2006. The reasons for the choice of this period are twofold. First, the available sub-national data are only complete and consistent for this period only. Secondly, this period covers a wide range of socio-economic and political conditions in Brazilian economic history: economically, it includes sub-periods of depression (1981-1983; 1986-1993) and prosperity (1984-1985; 1994-2006); politically, it includes dictatorship (1981-1984), democracy (1990-2006), and a combination of both regimes (1985-1989); institutionally, in addition to political changes themselves, it presents a sub-period of a relatively closed economy from 1981 to 1989 and another of a gradual trade liberalisation since 1989 (Lobo, 1996; Abreu, 2008a, 2008b; Abreu & Werneck, 2008). Interestingly, it will be seen later in this chapter that the fundamental driving forces of growth (or decline) as measured by dynamic spatial shift-share analysis remain robust under such dramatically changing circumstances.

Previous shift-share studies of growth in various countries often only consider non-spatial effects (e.g. Haynes & Dinc, 1997; Keil, 1992; Hoppes, 1991; Barff & Knight, 1988; Blien & Wolf, 2002).[78] This is also the case for Brazil (Chahad *et al*., 2002; Rolim, 2008; Nogueira & Lopes, 2008; Amorim & Da Mata, 2008). Some studies incorporate in the shift-share method the implications of international trade for the regional economy (Markusen *et al*., 1991; Noponen *et al*., 1998; Gazel & Schwer, 1998; Dinc & Haynes, 2005; Fotopoulos *et al*., 2010), but the analysis is not developed in that direction in this chapter, given that Brazil's international trade accounted over the period considered for no more than 24 percent of GDP (and was in fact 60 years earlier higher than during the 1981-2006 period). Instead, this chapter integrates the non-spatial classic shift-share methodology (Andrikopoulos *et al*, 1990; Selting & Loveridge, 1994; Ray & Harvey, 1995; Yasin *et al.*, 2010) with exploratory spatial data analysis (ESDA) of the shift-share components (Cochrane & Poot, 2008) and the methodology developed in Nazara & Hewings (2004), which

[78] However, Le Gallo & Kamarianakis (2010) combine exploratory spatial data analysis (ESDA) tools with spatial seemingly unrelated regressions (SUR) modelling to explain productivity disparities across European regions from 1975 to 2002.

explicitly incorporates spatial effects in the shift-share taxonomy to explain growth of regions.[79]

The chapter is organised as follows. Section 5.2 briefly presents the classic shift-share methodology. Section 5.3 describes the employment data used and their sources. Section 5.4 summarises the main facts of the Brazilian economy since 1981 and presents the results of the classic shift-share. Section 5.5 analyses structural change in Brazil in terms of the shift-share components. Section 5.6 presents alternative shift-share formulations. Section 5.7 performs exploratory spatial data analysis (ESDA) for industry-mix and competitive effect components of the classic shift-share method. Section 5.8 provides simple spatial shift-share decomposition, building on Nazara & Hewings (2004). Finally, section 5.9 provides concluding remarks.

5.2. Classic Multi-Period Shift-Share Analysis

This section briefly presents the classic shift-share method. This method decomposes the change in employment as follows (e.g., Cochrane & Poot, 2008, p. 55):

$$\Delta E_{ir}^{t} \equiv E_{ir}^{t} - \mathrm{E}_{ir}^{t-1} \equiv NE_{ir}^{t} + IM_{ir}^{t} + CE_{ir}^{t} \tag{5.1}$$

where:

$$NE_{ir}^{t} = g_{00}^{t} \times E_{ir}^{t-1} \tag{5.2}$$

$$IM_{ir}^{t} = (g_{i0}^{t} - g_{00}^{t}) \times E_{ir}^{t-1} \tag{5.3}$$

$$CE_{ir}^{t} = (g_{ir}^{t} - g_{i0}^{t}) \times E_{ir}^{t-1} \tag{5.4}$$

The terms in the above equations are defined as:

E_{ir}^{t-1} = Employment in the i^{th} industry in the r^{th} region at time *t*-1.

E_{ir}^{t} = Employment in the i^{th} industry in the r^{th} region at time *t*.

NE_{ir}^{t} = National Growth Effect on industry *i* in the r^{th} region between (*t*-1) and *t*.

IM_{ir}^{t} = Industry-Mix Effect on industry *i* in the r^{th} region between (*t*-1) and *t*.

[79] Mitchell *et al*. (2005) apply Nazara & Hewings's (2004) spatial shift-share decomposition to data on Australian regions. Mayor & López (2008) combined a variety of spatial analysis tools with the shift-share method.

CE_{ir}^{t} = Competitive Effect on industry *i* in the r^{th} region between (*t*-1) and *t*.
g_{ir}^{t} = Growth rate of employment in industry *i* and region *r* between (*t*-1) and *t*.
g_{i0}^{t} = Growth rate of nationwide employment in industry *i* between (*t*-1) and *t*.
g_{00}^{t} = Growth rate in nationwide total employment between (*t*-1) and *t*.

Using (5.1) to (5.4) if the researcher aggregates employment in each region *r* over industries i and define g_{0r}^{t} as the growth rate of total employment in region *r* between times (*t*-1) and *t*, this growth rate can be decomposed into a national growth rate; a growth rate due to the industry-mix and a residual that is referred to as the competitive growth rate c_{0r}^{t}. Hence,

$$g_{0r}^{t} \equiv g_{00}^{t} + m_{0r}^{t} + c_{0r}^{t} \qquad (5.5)$$

in which the growth component due to industry-mix is defined by

$$m_{0r}^{t} \equiv \sum_i w_{ir}^{t-1} (g_{i0}^{t} - g_{00}^{t}) \qquad (5.6)$$

with w_{ir}^{t-1} the fraction of employment in region *r* that is in industry *i* at time (*t*-1). Equation (5.6) shows that the industry-mix growth rate is a weighted average of national sectoral growth rates, minus national aggregate growth, with the weights being the shares of the various sectors in regional employment at the beginning of the period under consideration.

5.3. Data and Sources

This chapter uses employment data from IPEA - Institute of Applied Economic Research (www.ipea.gov.br). IPEA is one of the official public data sources in Brazil and it presents a variety of socio-economic data collected from public and private institutions, mostly at the state level.

Data have been collected for all 27 states (including Distrito Federal; for states' boundaries, see Figure 2.1 in Chapter 2) from which information on the number of employed people in each state by sector was extracted from 1981 to 2006. These sectors are: (1) agriculture & fishing; (2) commerce; (3) construction; (4) electricity, water & gas; (5) finance; (6) manufacturing; (7) mining; (8) services and (9) transportation & communications. State sectoral employment was calculated by multiplying the reported percentage of employed people in each sector by reported total employment in each state.

The five selected periods to analyse employment growth are: 1981-1986, 1986-1991, 1991-1996, 1996-2001, and 2001-2006. Although there are data to calculate annual changes, the use of five-year periods provides some control for cyclical employment fluctuations, as suggested by Thirlwall (1967) (see Barff & Knight, 1988, pp. 3-4). Using periods of equal duration addresses Brown's (1969) criticism that varying the periods may lead to the risk of an undue influence of sudden employment (or income) changes in atypical years (Barff & Knight, 1988, p. 6; Knudsen & Barff, 1991, pp. 427-428; Knudsen, 2000, pp. 179-180).

There are missing employment data for all states in 1991. To address this problem, the distribution of employment across sectors was simply interpolated between 1990 and 1992 and the interpolated shares subsequently applied to the known state total employment. Additionally, there were missing employment data for Tocantins from 1981 to 1991. Here it was assumed that total employment growth was identical to known state population growth over the sub-periods 1981-86, 1986-91, and 1991-96. Sectoral shares in Tocantins were assumed to have been the same in 1981, 1986, and 1991 as observed in 1996.[80]

[80] An alternative assumption would have been to backcast the 1981-1991 Tocantins sectoral shares from 1996 by means of the observed trends in national sectoral shares. This has very little impact on the results reported in the tables in this chapter.

5.4. Results of Classic Shift-Share Analysis

This section outlines the main characteristics of the events that shaped the performance of the Brazilian economy from 1981 to 2006; then, using employment data, it presents the results of the non-spatial shift-share analysis (Tables 5.1 to 5.3). In terms of the economic history of Brazil, 1981-2006 can be subdivided into three periods as follows.

Period I: 1981-1984 (the final part of the dictatorship or "Authoritarian State" period, which started in 1964). The main characteristics are (Lobo, 1996; Fausto, 1999; Abreu, 2008a): i) the combination of economic stagnation and inflation ('stagflation'); ii) little political rights and freedom; iii) oil shocks (1974-1980) causing macroeconomic instability; iv) economic redistribution that harmed the northeast and benefited the middle-west, north and south regions; v) protectionism, contractionist policies, and falling output (1981-1983).

Period II: 1985-1989 (democratic transition). This period is characterised by poor economic performance as a result of hyperinflation and stagnation.

Period III: 1989-2006 (Trade liberalisation and the return to democracy). The main facts are (Lobo, 1996; Abreu, 2008*b*; Abreu & Werneck, 2008): i) the structural reforms under Collor de Mello (1990-1992) and Itamar Franco (1992-1994) presidencies; ii) the policies that aimed to reduce and stabilise inflation and unemployment were more successful after mid-1994; however, iii) as Abreu & Werneck (2008, p. 432) point out, "(…) between 1994 and 2004 per capita GDP (gross domestic product) increased [at] an average of only 0.9 percent per annum".

Summarising economic growth over the 1981-2006 period, Abreu (2008b, p. 395) states:

> In the years **1981-1983**, during the administration of General João Figueiredo (1979-1985), the last of five successive military presidents since the military coup of 1964, **there was a sharp deterioration in the Brazilian GDP growth performance.** Brazil suffered its most severe recession of the twentieth century. GDP fell 4.9 percent from its pick in 1980. **After a brief recovery in 1984-1985 when GDP grew on average 7 percent per annum** – years that also witnessed a transition from military to civilian rule (and ultimately a fully fledged democracy) – growth performance remained mediocre during the following decades. **Between 1981 and 1994 GDP per capita increased on average less than 0.1 percent annually.** And there was only **limited improvement** in **the decade after 1994**. (Bolds added)

Table 5.1: National Employment by Sector, 1981-2006

	% shares and total 1981	1981-86 change (%)	1986-91 change (%)	1991-96 change (%)	1996-01 change (%)	2001-06 change (%)
Agriculture and Fishing	32.2	5.7	13.1	2.0	-5.0	10.6
Commerce	11.2	29.8	28.1	15.8	19.6	44.4
Construction	8.7	-4.9	9.9	13.4	14.4	17.6
Electricity, Water and Gas	0.8	9.5	4.5	-8.7	-2.7	18.3
Financial Sector	2.1	20.1	-3.2	-20.1	-1.9	127.5
Manufacturing	15.5	27.5	1.0	-3.0	9.5	38.6
Mining	0.6	40.5	-5.0	-40.8	22.8	46.3
Services	24.6	30.8	26.5	20.4	19.2	23.2
Transportation and Communications	4.3	9.7	20.5	11.1	24.6	22.8
National Employment	**41,397,661**	**17.7**	**16.2**	**9.1**	**11.1**	**26.5**
GDP growth rate per capita	**---**	**-2.4**	**7.4**	**12.9**	**-9.8**	**12.2**

Notes: National employment in 1981 may be compared with an estimated national population of 122,500,000 inhabitants (estimate based on interpolation from IBGE – Brazilian Institute of Geography and Statistics population census data). The average annual GDP per capita growth rate from 1981 to 2006 is 4.1percent.

A comparison of national and sectoral employment growth between the five-year sub-periods show that the periods 1981-1986 and 1996-2001 stand out, and productivity declined even though employment increased. Commerce, financial sector, services, and transportation and communications are industries that had significant employment growth throughout the 1981-2006 period. Employment change was rather volatile in the other sectors.

Table 5.2 provides the classic shift-share decomposition of total employment growth in Brazil’s states in terms of the national, industry-mix and competitive components for the five sub-periods. The states have been ranked according to the five-period average total employment growth rate (from high to low). Roraima had the highest average five-year growth rate (75.1 percent) and Rio de Janeiro the lowest (10.9 percent).

The ranking is consistent with the finding of long-run regional convergence already elaborated in chapters 3 and 4. In Chapter 3 it was concluded that in particular, specialisation has been responsible for regional convergence. The observed convergence is a result of a reduction in concentration of economic activities that essentially benefited middle-west and north Brazil, rather than the traditional large markets of São Paulo and Rio de Janeiro.[81] This can be seen from two features in Table 5.2. First, the top ten states in terms of total employment growth are in all five-year sub-periods either from north or middle-west regions. The only other states that "infiltrated" the top ten in the ranking are Rio Grande do Norte in the 1981-1986 and 1986-1991 periods; Maranhão, Piauí, Sergipe, Paraíba, and Ceará in 1991-1996; Alagoas in 1996-2001; and Sergipe and Paraíba in 2001-2006. All these states are from the northeast region, which is contiguous to the north and middle-west regions. In the five-period averages from 1981 to 2006, Brazil's richest states of São Paulo and Rio de Janeiro occupied the lower end of the employment growth ranking. These two states were 23rd and 27th respectively.

Table 5.3 shows state Location Quotients (hereafter LQ). LQ<1 indicates that the area is less specialised than the nation in a particular sector; LQ>1 means the area is highly specialised in a specific sector. Based on tables 5.2 and 5.3, three main questions are addressed.

The first question is to identify the states that have a high competitive growth rate (Table 5.2) and to check how this is linked to the LQ (Table 5.3). The north and middle-west states occupied the top nine positions in competitiveness.[82] Their positive competitive effect suggests that their sectoral employment grew generally above national sectoral employment.[83] However, this effect decreased for most of these states, and eight of the other states, from the first to the second sub-period and this amplified in the third and the fourth sub-periods. However, the competitive effect recuperated over the last sub-period.[84] The behaviour of the competitive effect in the last sub-period is consistent with the successful policies

[81] Chahad, Comune, and Haddad (2002) found a similar result when analysing employment change from 1985 to 1997 in Brazil. However, such findings contradict previous studies for the period 1960-1970 in which centripetal forces were apparently stronger than centrifugal forces, with high growth of the number of firms, the number of people employed, and gross value of production in the main metropolitan centres (São Paulo and Rio de Janeiro or former Guanabara) (Enders, 1980). By comparison, Fotopoulos & Spence (1999, p. 1737), analysing a change of manufacturing industry establishments from 1984 to 1988 in Greece, found an outward movement of manufacturing plants from Central Greece and Athens (the largest markets in Greece). Another comparison is proved by Hanham & Banasick (2000, pp. 110-111) who cite Edgington (1994) who found that the peripheral regions of the north and south of Japan have had faster manufacturing employment growth during the 1980s than the metropolitan regions of central Japan.
[82] While the middle-west states of Goiás and Mato Grosso do Sul are not in this list, the competitiveness component in the latter state is closer to that of the group.
[83] Except: Tocantins in the 1986-1991 period; Rondônia and Distrito Federal in the 1991-1996 period; and Mato Grosso in the 2001-2006.
[84] The average competitive effect across all states is as follows: 1981-1986: 18.4 percent; 1986-1991: 8.7 percent; 1991-1996: 8.6 percent; 1996-2001: 6 percent; and 2001-2006: 16.3 percent.

to control inflation and reduce unemployment in the trade liberalisation period. The impact of policies was properly captured by the observed national effect in Table 5.2 which has a minimum in the third sub-period, even though there is a caveat that the stabilisation policies have benefited the northern and middle-west states rather more than the whole nation (see states with the lowest levels in competitive effect and those with a negative competitive effect in Table 5.2 for the last two sub-periods). This result may be interpreted as positive as it shows a process of employment deconcentration across states.

The analysis of the LQs in Table 5.3 indicates that the most competitive states, due to specialisation (with an LQ higher than 1.5) in faster growth industries nationwide, are basically located in the north and middle-west regions[85]. One of the reasons for the dominance of these two regions is that, being historically lagged regions, the development of infrastructure helped boost employment in all sectors over the study period.[86] Considering the first sub-period, there are five sectors (out of nine) that grew fast with a growth rate of at least 20 percent nationally. Among these five sectors, northern and middle-west states have a LQ greater than 1.5 in three sectors, such as commerce, mining and services. On the other hand, São Paulo and Rio de Janeiro accounted for most national employment in the financial sector and manufacturing. These two sectors, however, experienced weak growth between the second and fourth sub-periods (in the third sub-period both even had negative growth rates) before they recovered in the last sub-period due to successful policies implemented in this sub-period.

Based on the 1981-1986 sub-period, Table 5.2 suggests that there are some other competitive states, such as Espirito Santo, Santa Catarina, Ceará, Minas Gerais, Alagoas, Maranhão, Paraíba and Bahia. However, all but one of these states had a negative competitive effect for at least two of the sub-periods. According to LQs in Table 5.3, these states are relatively specialised in the following sectors, with an LQ of at least 1.40: agriculture and fishing (Espírito Santo, Santa Catarina, Minas Gerais, Alagoas, Maranhão, and Bahia),

[85] Although, overall, the northern states were the less specialised and Table 5.2 shows that they had the lowest Hirschman-Herfindahl Index in 1981.

[86] It is worth noting that in 1981 Roraima had an LQ of 1.54 for commerce, 2.33 for construction, 3.7 for electricity, water and gas, 1.38 for financial sector, 0.37 for manufacturing, 8.42 for mining, 1.28 for services, and 1.58 for transportation and communication. This state had the highest annual average growth rate (11.25 percent) for manufacturing industry employment from 1981 to 2006.

construction (Ceará and Paraíba), electricity, water and gas, and manufacturing (Santa Catarina), and mining (Espírito Santo and Bahia).

The second question is whether the observed total regional employment growth rates in Table 5.2 are consistent with the earlier described economic history of Brazil. Table 5.2 indicates that, as expected, the core regions of southeast and south had generally lower employment growth rates than the lagging north and middle-west and some northeast states. This trend is compatible with high specialisation of the lagging regions for three (out of five) of the faster growing sectors of economic activity in Brazil.[87]

The third issue is whether the differences in total state employment growth rates are due to differences in industry-mix at the state level relative to the national economy or whether these differences are due to the competitive advantage that a specific state has relative to the national economy. Table 5.2 shows that the top six (out of 27) states – in terms of the five-period average employment growth rates – are the only states that have had a positive industry-mix effect in five-year periods.[88] Again, these states are either from the north or the middle-west regions of Brazil and appear to have had an industry structure that has been more beneficial than that of the other states, even during periods in which, for some sectors, the nation's sectoral growth rate was less than average growth. Additionally, these top six states had the highest competitive effect over time as a result of a high LQ in six (out of nine) sectors in 1981 (see Table 5.3) and specialisation in sectors with a growth rate larger than that observed for the nation. These sectors are: commerce, services, and transportation and communication in the first period; construction and services in the second period; services in the third and the fourth periods; and, finally, commerce and services in the fifth period.

[87] Among the other four sectors, the three sectors in which north and middle-west had a comparative disadvantage are agriculture & fishing, manufacturing, and financial sector. These latter two sectors had some of the highest growth rates in the sub-period 2001-2006, 127.5 percent and 38.6 percent, respectively (see Table 5.1).

[88] Excluding Rio de Janeiro, which had a positive industry-mix effect in all of the five-year periods, but had the lowest five-period average total employment growth due to a consistently high negative competitive effect.

Table 5.2: Classic Shift-Share Decomposition of Total Employment Growth Rate in Brazil's States

State*	1981-86				1986-91				1991-96				1996-2001				2001-06				Five-Period Averages			
	g_{0r}^{86}	g_{00}^{86}	m_{0r}^{86}	c_{0r}^{86}	g_{0r}^{91}	g_{00}^{91}	m_{0r}^{91}	c_{0r}^{91}	g_{0r}^{96}	g_{00}^{96}	m_{0r}^{96}	c_{0r}^{96}	g_{0r}^{01}	g_{00}^{01}	m_{0r}^{01}	c_{0r}^{01}	g_{0r}^{06}	g_{00}^{06}	m_{0r}^{06}	c_{0r}^{06}	g_{0r}	g_{00}	m_{0r}	c_{0r}
Roraima	121.8	17.7	1.7	102.4	62.7	16.2	3.2	43.4	54.9	9.1	2.7	43.1	37.0	11.1	5.0	20.9	99.4	26.5	0.0	72.9	75.1	16.1	2.5	56.5
Rondônia	114.7	17.7	3.3	93.8	40.1	16.2	2.9	21.0	8.1	9.1	2.4	-3.3	28.2	11.1	3.9	13.2	121.4	26.5	1.0	94.0	62.5	16.1	2.7	43.7
Amapá	48.2	17.7	4.2	26.2	56.7	16.2	5.1	35.5	98.2	9.1	4.9	84.2	20.0	11.1	3.7	5.2	87.1	26.5	2.4	58.2	62.0	16.1	4.0	41.9
Acre	96.3	17.7	5.5	73.1	30.5	16.2	4.4	10.0	16.4	9.1	3.8	3.6	34.0	11.1	5.0	18.0	123.5	26.5	0.0	97.1	60.2	16.1	3.7	40.3
Distrito Federal	150.6	17.7	5.5	127.4	24.5	16.2	4.1	4.3	8.3	9.1	4.5	-5.3	24.8	11.1	5.1	8.6	43.8	26.5	4.1	13.1	50.4	16.1	4.7	29.6
Pará	39.3	17.7	4.0	17.5	36.3	16.2	2.9	17.2	16.2	9.1	2.6	4.5	53.3	11.1	3.0	39.2	95.2	26.5	2.1	66.7	48.1	16.1	2.9	29.0
Amazonas	42.1	17.7	4.8	19.6	36.7	16.2	-0.1	20.6	13.0	9.1	1.5	2.4	26.1	11.1	3.4	11.5	93.0	26.5	3.0	63.5	42.2	16.1	2.5	23.5
Tocantins	12.7	17.7	-7.2	2.1	12.0	16.2	0.3	-4.4	98.6	9.1	-4.0	93.4	22.0	11.1	-3.2	14.1	26.2	26.5	-4.7	4.4	34.3	16.1	-3.8	21.9
Mato Grosso	52.0	17.7	-1.8	36.1	53.0	16.2	-0.4	37.2	10.7	9.1	-1.7	3.3	24.1	11.1	-1.7	14.7	16.4	26.5	-2.0	-8.0	31.2	16.1	-1.5	16.7
Rio Grande do Norte	23.7	17.7	-1.0	7.0	25.2	16.2	0.6	8.5	11.8	9.1	0.2	2.5	10.4	11.1	-0.6	-0.1	26.7	26.5	-1.4	1.6	19.6	16.1	-0.4	3.9
Mato Grosso do Sul	24.1	17.7	-0.6	7.0	20.7	16.2	1.9	2.7	11.9	9.1	1.5	1.4	8.8	11.1	-0.6	-1.7	28.4	26.5	-0.4	2.3	18.8	16.1	0.4	2.3
Espírito Santo	16.6	17.7	-3.0	1.8	24.3	16.2	-0.3	8.4	9.2	9.1	-0.8	0.8	15.0	11.1	-1.1	5.0	27.3	26.5	-1.5	2.3	18.5	16.1	-1.3	3.7
Sergipe	12.2	17.7	-2.8	-2.8	16.4	16.2	-0.7	1.0	13.2	9.1	-0.9	5.0	7.2	11.1	-1.2	-2.7	36.4	26.5	-2.2	12.1	17.1	16.1	-1.5	2.5
Santa Catarina	20.1	17.7	-1.8	4.2	16.4	16.2	-2.2	2.4	7.8	9.1	-2.5	1.2	17.8	11.1	-2.0	8.7	22.2	26.5	1.4	-5.6	16.9	16.1	-1.4	2.2
Piauí	8.8	17.7	-7.8	-1.1	24.1	16.2	-0.7	8.6	15.1	9.1	-1.7	7.7	7.4	11.1	-4.7	1.0	28.4	26.5	-5.7	7.5	16.8	16.1	-4.1	4.8
Ceará	20.6	17.7	-3.0	5.9	16.7	16.2	-0.5	1.0	12.3	9.1	-0.6	3.7	12.7	11.1	-3.3	5.0	21.3	26.5	-1.9	-3.3	16.7	16.1	-1.9	2.5
Goiás	0.1	17.7	-0.7	-16.9	40.7	16.2	1.4	23.2	2.1	9.1	0.9	-7.9	15.4	11.1	0.5	3.8	23.8	26.5	0.0	-2.8	16.4	16.1	0.4	-0.1
Minas Gerais	19.0	17.7	-1.1	2.4	17.4	16.2	0.5	0.8	8.2	9.1	-0.2	-0.7	9.8	11.1	-0.5	-0.8	25.2	26.5	-1.3	0.1	15.9	16.1	-0.5	0.3
Alagoas	15.7	17.7	-4.6	2.6	20.5	16.2	-1.0	5.3	-7.1	9.1	-1.4	-14.7	32.1	11.1	-3.6	24.6	16.6	26.5	-5.5	-4.4	15.6	16.1	-3.2	2.7
Maranhão	13.2	17.7	-6.4	1.9	19.8	16.2	-0.8	4.5	29.8	9.1	-2.5	23.2	3.9	11.1	-6.7	-0.5	10.9	26.5	-6.0	-9.5	15.5	16.1	-4.5	3.9
Paraíba	19.9	17.7	-4.2	6.4	16.2	16.2	1.3	-1.2	13.2	9.1	-0.3	4.3	-3.4	11.1	-3.0	-11.5	31.4	26.5	-3.0	7.9	15.4	16.1	-1.9	1.2
Bahia	20.5	17.7	-4.2	6.9	18.8	16.2	-0.1	2.7	3.5	9.1	-1.1	-4.4	10.8	11.1	-3.8	3.4	22.1	26.5	-4.5	0.1	15.1	16.1	-2.7	1.7
São Paulo	20.0	17.7	4.6	-2.3	8.2	16.2	-1.0	-7.0	10.0	9.1	0.4	0.4	9.2	11.1	3.2	-5.1	26.0	26.5	4.4	-4.8	14.7	16.1	2.3	-3.7
Pernambuco	14.4	17.7	-1.4	-1.9	16.6	16.2	0.4	0.1	3.0	9.1	0.0	-6.1	12.5	11.1	-1.3	2.7	18.0	26.5	-3.0	-5.5	12.9	16.1	-1.1	-2.1
Rio Grande do Sul	10.7	17.7	-0.8	-6.3	17.3	16.2	-0.7	1.8	6.2	9.1	-0.9	-2.0	8.4	11.1	-1.0	-1.7	17.6	26.5	-0.8	-8.1	12.0	16.1	-0.8	-3.2
Paraná	8.0	17.7	-3.1	-6.7	12.9	16.2	0.2	-3.4	8.9	9.1	-0.5	0.3	7.7	11.1	-0.9	-2.5	21.8	26.5	-0.5	-4.2	11.8	16.1	-1.0	-3.3
Rio de Janeiro	13.6	17.7	5.0	-9.1	8.1	16.2	2.1	-10.2	5.2	9.1	3.2	-7.1	5.1	11.1	5.2	-11.2	22.5	26.5	3.3	-7.3	10.9	16.1	3.8	-9.0

*Ranked from highest to lowest in terms of 5-period average percentage of total employment growth rate.

Table 5.3: Location Quotients and Hirschman-Herfindahl Index of Brazil's States, 1981; Ordered from north to south

State	Agriculture and Fishing	Commerce	Construction	Electricity, Water and Gas	Financial Sector	Manufacturing	Mining	Services	Transportation and Communication	Hirschman-Herfindahl Index (HHI)*
Acre	*0.400*	**1.571**	*0.468*	**2.156**	**1.619**	0.809	1.170	**1.599**	**1.819**	0.227
Amazonas	*0.296*	**1.797**	1.020	**2.021**	1.019	1.453	1.255	1.161	1.381	0.194
Amapá	*0.191*	**1.598**	1.481	1.308	0.741	0.830	**2.674**	**1.516**	**2.042**	0.216
Pará	*0.409*	**1.803**	1.049	1.301	0.867	0.792	**2.774**	1.345	**1.789**	0.198
Rondônia	*0.445*	**1.911**	1.149	**1.800**	0.603	0.754	**3.184**	1.202	**1.969**	0.186
Roraima	*0.239*	**1.544**	**2.328**	**3.701**	1.384	*0.373*	**8.416**	1.282	**1.576**	0.188
Tocantins	**2.351**	**1.756**	0.194	*0.179*	*0.003*	*0.010*	*0.040*	*0.023*	*0.444*	0.614
Alagoas	**1.712**	0.625	0.873	*0.486*	*0.390*	0.569	0.843	0.683	0.670	0.353
Bahia	**1.710**	0.841	0.757	0.811	0.513	*0.446*	1.426	0.674	0.656	0.351
Ceará	1.013	0.889	**2.104**	0.763	*0.399*	0.934	0.537	0.821	0.619	0.213
Maranhão	**2.086**	0.628	0.594	0.740	*0.229*	*0.227*	1.006	0.535	0.511	0.479
Paraíba	1.043	0.964	**2.477**	0.847	*0.438*	*0.477*	1.294	0.861	0.725	0.223
Pernambuco	1.230	1.157	0.936	1.079	0.619	0.764	*0.334*	0.858	0.919	0.241
Piauí	**1.629**	0.585	**2.209**	*0.486*	*0.340*	*0.247*	*0.494*	0.575	0.542	0.339
Rio Grande do Norte	1.077	0.908	1.320	1.033	0.629	0.613	**5.410**	1.009	0.945	0.218
Sergipe	1.358	0.651	1.140	0.929	*0.426*	0.606	**6.714**	0.776	1.166	0.256
Distrito Federal	*0.113*	**1.583**	1.219	**1.983**	**2.658**	*0.432*	*0.053*	**1.899**	**1.740**	0.276
Goiás	1.179	1.055	0.902	0.833	0.864	*0.453*	**1.681**	1.132	0.937	0.249
Mato Grosso do Sul	1.012	1.113	1.173	1.146	0.790	0.597	0.689	1.110	1.205	0.219
Mato Grosso	1.327	1.087	0.885	1.337	0.859	*0.362*	**5.546**	0.853	1.088	0.255
Espírito Santo	1.417	0.798	0.982	0.669	0.732	0.656	**1.511**	0.793	0.985	0.275
Minas Gerais	1.190	0.833	0.964	0.809	0.745	0.689	1.110	1.075	0.925	0.246
Rio de Janeiro	*0.178*	1.227	1.181	**1.633**	**1.714**	1.229	0.747	**1.569**	**1.704**	0.225
São Paulo	*0.371*	1.167	0.917	1.006	**1.679**	**1.864**	*0.365*	1.168	1.130	0.207
Paraná	**1.551**	0.917	0.686	0.803	0.802	0.573	*0.491*	0.761	0.825	0.309
Santa Catarina	1.402	0.730	0.546	1.485	0.645	1.174	1.165	0.676	0.897	0.276
Rio Grande do Sul	1.190	0.894	0.797	0.994	0.872	0.983	1.128	0.909	0.889	0.237

Note: values higher than 1.5 are in bold, and those lower than 0.5 are in italics; * $HHI_i = \sum_j s_{ij}^2$ where s_{ij} is the share in state *i* of the industry *j* in local (State) total employment.

Conversely, 16 of the other states had a negative five-period industry-mix average as a result of a highly negative industry-mix component at least in one of the five-year sub-periods. This finding indicates that, with respect to the five-period average, these 16 states were more harmed by the national poor performance through following the nation's trend in the sub-periods in which the nation had a negative (or low positive) sectoral employment growth rate, because they were endowed with industries that were growing less than average. It is not a surprise that those 16 states also had the smallest (even negative) competitive average effect over the study period.

5.5. Structural Change

This section and the next two replicate for Brazil Cochrane and Poot's (2008) shift-share analysis of employment change in New Zealand. This section investigates whether states sectoral growth rates followed the national trend. The approach to answer this question is to decompose the industry-mix effect from equation (5.6) in section 5.2 as follows:

$$\sum_i w_{ir}^{t-1}\,(g_{i0}^t - g_{00}^t) \equiv \sum_i w_{ir}^{t}\,(g_{i0}^t - g_{00}^t) + \sum_i (w_{ir}^{t-1} - w_{ir}^{t})(g_{i0}^t - g_{00}^t) \quad (5.7)$$

The second term of the right-hand side of the equation above measures the effect of changing industry composition on the regional employment growth rate. This will be referred to as the structural change effect. The industry-mix effect calculated by means of end-of-the period weights will be referred to as the modified industry-mix effect.

The states among the top ten in terms of the total employment growth rate (see Table 5.2) also have the highest (positive) all five-year periods average for the modified industry-mix effect (Table 5.4). These states are: Acre, Amazonas, Amapá, Pará, Rondônia, and Roraima from the north region and Distrito Federal, Goiás, and Mato Grosso do Sul from the middle-west region. Rio de Janeiro and São Paulo now have a modified industry-mix effect at levels comparable to those of northern and middle-west states. As shown earlier, northern and middle-west states benefited from growth in sectors in which they were highly specialised (three out of nine) in 1981, even though that growth was larger in some sectors than others depending on the period (Table 5.1). In terms of magnitude, the average modified industry-mix effect in Brazil as a whole was largest for 1986-1991 and smallest for the 2001-2006 period.

Table 5.4: The Modified Industry-Mix and Structural Change Effects on Employment Growth in Brazil

State**	Modified Industry Mix Effect					Structural Change Effect				
	1981-86	1986-91	1991-96	1996-01	2001-06	1981-86	1986-91	1991-96	1996-01	2001-06
Acre	6.91	4.48	4.80	3.05	-2.82	-1.43	-0.09	-1.01	1.90	2.81
Amazonas	5.13	1.67	2.63	4.25	0.83	-0.31	-1.74	-1.14	-0.84	2.15
Amapá	6.06	4.88	4.02	6.27	1.58	-1.82	0.20	0.85	-2.62	0.80
Pará	5.05	3.77	3.30	3.68	1.93	-1.03	-0.82	-0.68	-0.71	0.15
Rondônia	4.88	3.55	2.64	3.27	-2.40	-1.63	-0.64	-0.26	0.65	3.40
Roraima	5.65	3.43	4.98	4.03	-1.45	-4.00	-0.25	-2.28	0.93	1.42
Tocantins	-6.93	-0.20	0.67	-2.01	-2.02	-0.27	0.46	-4.63	-1.17	-2.72
Alagoas	-4.00	0.30	-0.65	-4.12	-3.97	-0.60	-1.26	-0.79	0.55	-1.55
Bahia	-3.44	0.75	-0.46	-2.78	-2.89	-0.78	-0.81	-0.67	-1.01	-1.57
Ceará	-0.34	0.73	-0.41	-1.22	-0.43	-2.65	-1.18	-0.16	-2.12	-1.46
Maranhão	-5.23	0.05	-1.75	-4.71	-3.83	-1.20	-0.86	-0.72	-2.00	-2.18
Paraíba	-1.00	1.45	0.30	-1.02	-1.44	-3.24	-0.20	-0.56	-2.02	-1.55
Pernambuco	0.27	1.32	0.90	-1.20	-0.45	-1.72	-0.97	-0.89	-0.14	-2.57
Piauí	-4.57	0.28	-0.86	-4.53	-4.39	-3.23	-0.97	-0.87	-0.14	-1.26
Rio Grande do Norte	0.47	1.43	0.87	0.46	0.51	-1.51	-0.86	-0.66	-1.03	-1.91
Sergipe	-1.74	0.77	0.82	-0.32	0.37	-1.05	-1.48	-1.68	-0.89	-2.55
Distrito Federal	6.66	4.74	5.06	6.21	5.99	-1.18	-0.67	-0.52	-1.14	-1.85
Goiás	0.89	2.08	2.09	2.52	2.09	-1.61	-0.72	-1.18	-2.05	-2.05
Mato Grosso do Sul	0.62	2.53	2.13	1.71	1.12	-1.24	-0.63	-0.65	-2.29	-1.48
Mato Grosso	0.16	0.86	0.23	-1.08	-0.89	-1.99	-1.23	-1.95	-0.62	-1.16
Espírito Santo	-1.23	0.69	0.37	-0.32	1.61	-1.73	-0.96	-1.14	-0.80	-3.15
Minas Gerais	0.26	0.84	0.94	0.13	0.62	-1.35	-0.38	-1.12	-0.62	-1.97
Rio de Janeiro	6.07	3.09	4.35	5.98	5.85	-1.04	-0.95	-1.18	-0.81	-2.57
São Paulo	5.49	0.17	1.78	4.15	6.71	-0.91	-1.20	-1.36	-0.98	-2.34
Paraná	-1.17	0.76	0.56	0.23	2.81	-1.88	-0.59	-1.08	-1.18	-3.31
Santa Catarina	-0.63	-1.49	-1.29	-0.06	2.50	-1.17	-0.69	-1.26	-1.94	-1.12
Rio Grande do Sul	1.07	0.10	-0.16	-0.45	1.32	-1.83	-0.82	-0.79	-0.60	-2.09

** In ‘from North to South’ geographical order. There are states with zero sectoral employment as follows: in the mining sector: Acre in 1981, 1986, 1996, 2001, and 2006; Alagoas in 1996, Roraima in 1981, and for Amapá in 1986 and 1996; and in the financial sector: Amapá in 1996. In all these cases the population growth rate has been used to act as a proxy for the employment growth rate over the sub-periods to estimate sectoral employment in each of those years. This assumption yielded results that are consistent with the overall pattern of employment data in Brazil. However, a lack of regional population data in 1981 and 1986 was overcome by interpolation of the corresponding population time series.

The structural effect is negative in all but eight cases (Acre, 1996-2001 and 2001-2006; Amazonas, 2001-2006; Amapá, 1986-1991, 1991-1996 and 2001-2006; Pará, 2001-2006; Rondônia and Roraima, 1996-2001 and 2001-2006; Tocantins, 1986-1991; and Alagoas, 1996-2001). The former seven states are in the north, while the latter is in the northeast region. This indicates that the sectoral trends in those eight states were different from the nation in the specified periods. However, given that the number of these cases is small, the overall conclusion is that most states have generally not gone against the national trend in terms of structural change. Hence, when a sector grows faster (slower) than average, its share in employment increases (decreases) in almost all regions. The positive structural change effects occur predominantly in the fourth and fifth five-year periods. These represent a period of a relative prosperity in terms of Brazil's GDP per capita increase, although the fourth sub-period is worse than the fifth. In the last period only the northern states (six out of seven) have gone against the national trend (i.e. had a positive sign in structural change effect).

Table 5.5: Persistence in Regional Employment Change and its Components

Compare Ranking	Regional Growth Rate	Industry-Mix Growth Rate	Competitive Growth Rate
81/86 with 86/91	0.5788***	0.5788***	0.5940***
81/86 with 91/96	0.2509	0.8675***	0.2851
81/86 with 96/01	0.6368***	0.9267***	0.6197***
81/86 with 01/06	0.5372***	0.8932***	0.6020***
86/91 with 91/96	0.2039	0.8187***	0.2070
91/96 with 96/01	0.0812	0.8968***	0.1453
96/01 with 86/91	0.6410***	0.6917***	0.6654***
86/91 with 01/06	0.4243**	0.4316**	0.4359**
91/96 with 01/06	0.4908***	0.7454***	0.3748*
96/01 with 01/06	0.3944**	0.8639***	0.3962**

Notes: The table reports Spearman's Rank Correlation Coefficients. Coefficients are significant at the *10%; **5%; and ***1% levels.

Table 5.5 compares Spearman rank correlation coefficients for states' growth rates, the industry-mix growth rates, and the competitive growth rates across five-year periods. As in Cochrane & Poot (2008), the highest rank correlation coefficients are found for the industry-mix growth rates. This indicates that a change in regional industry mix is a relatively long-term process and there is therefore relative high temporal persistence in the state ranking based on this shift-share decomposition component.

The ranking based on the competitive growth rate is clearly more volatile over time. In several cases, the rank correlation is statistically insignificant. The same is true with respect to regional growth rates. Clearly, the relative performance of the states over the 1991/96 period was very different from previous periods (1981/86 and 1986/91) but also subsequently (1996/01).

5.6. Alternative Formulations

One of the criticisms of classic shift-share analysis is that the industry-mix effect interacts with competitive effect. In other words, it is difficult from the shift-share identity to isolate regional performance that truly depends on a region's strengths because a region can grow faster either as a result of an 'appropriate' mix of industries that are also going well elsewhere, or as a consequence of being specialised (i.e. a high LQ) in a buoyant industry which is not found elsewhere. This section reviews and applies some of the shift-share extensions that were done to isolate the interaction between the industry-mix and competitive effects in a region's growth (Loveridge & Selting, 1998, pp. 43-49; Cochrane & Poot, 2008).

The first extension considered is the calculation of Esteban-Marquillas homothetic employment, which is the employment that a region *r* would have had in industry *i* if the share of industry *i* in regional employment was the same as the share of industry *i* in national employment:

$$EH_{ir}^{t-1} = \frac{E_{i0}^{t-1} E_{0r}^{t-1}}{E_{00}^{t-1}} \tag{5.8}$$

Hence, homothetic employment would be the same as actual employment if, and only if, LQ = 1. The decomposition of competitive effect using equation (5.8) is:

$$CE_{ir}^{t} \equiv CEH_{ir}^{t} + AE_{ir}^{t} \equiv \left(g_{ir}^{t} - g_{i0}^{t}\right) \times EH_{ir}^{t-1} + \left(g_{ir}^{t} - g_{i0}^{t}\right) \times (E_{ir}^{t-1} - EH_{ir}^{t-1}) \tag{5.9}$$

The CEH_{ij}^{t} measures the comparative advantage of region's sector *i* compared to the nation ($g_{ir}^{t} > g_{i0}^{t}$) and *AE* is the Esteban-Marquillas' allocative effect which depends on the extent to which the region *r* is specialised in the industry *i* (i.e. homothetic employment differs from actual employment).

The Esteban-Marquillas' extension can also be applied to the industry-mix effect. This is referred to as Esteban-Marquillas' second decomposition,

$$E_{ir}^{t} - E_{ir}^{t-1} \equiv \Delta E_{ir}^{t} \equiv NEEM2_{ir}^{t} + IMEM2_{ir}^{t} + CEH_{ir}^{t} + AE_{ir}^{t} \tag{5.10}$$

$$NEEM2_{ir}^{t} = g_{i0}^{t} \times EH_{ir}^{t-1} \tag{5.11}$$

$$IMEM2_{ir}^{t} = g_{i0}^{t} \times (E_{ir}^{t-1} - EH_{ir}^{t-1}) \tag{5.12}$$

CEH_{ir}^{t} and AE_{ir}^{t} are defined as in (5.9); $NEEM2_{ir}^{t}$ is the Esteban-Marquillas modified national growth effect on industry i in the r^{th} between times (t-1) and t, $IMEM2_{ir}^{t}$ is the Esteban-Marquillas modified industry-mix effect on industry i in the r^{th} region between times (t-1) and t.

Cochrane & Poot (2008, p. 67) cite Kheil (1992) to show that:

$$\textstyle\sum_i NEEM_{ir}^{t} = \sum_i NE_{ir}^{t} \quad \text{and} \quad \sum_i IMEM2_{ir}^{t} = \sum_i IM_{ir}^{t} \tag{5.13}$$

It can be seen that CEH_{ir}^{t} and CE_{ir}^{t} are closely linked via the location quotient LQ_{ir}^{t} as follows:

$$CEH_{ir}^{t} = \frac{CE_{ir}^{t}}{LQ_{ir}^{t}} \tag{5.14}$$

in which the location quotient LQ_{ir}^{t} is again defined as the ratio of the share of industry i in region r over the share of industry i in the nation (as reported for 1981 in Table 5.3).

Other authors also used homothetic employment in their extensions. Based on equations (5.2) and (5.3), Bishop and Simpson (1972) created alternative expressions for national growth and industry-mix effects:

$$\Delta E_{ir}^{t} = E_{ir}^{t} - E_{ir}^{t-1} \equiv NEBIS_{ir}^{t} + IMBIS_{ir}^{t} + CE_{ir}^{t} \tag{5.15}$$

$$NEBIS_{ir}^{t} \equiv g_{00}^{t} \times E_{ir}^{t-1} + (g_{i0}^{t} - g_{00}^{t}) \times EH_{ir}^{t-1} \tag{5.16}$$

$$IMBIS_{ir}^{t} = (g_{i0}^{t} - g_{00}^{t}) \times (E_{ir}^{t-1} - EH_{ir}^{t-1}) \tag{5.17}$$

The new components of the three equations above are:

$NEBIS_{ir}^{t}$ = the Bishop-Simpson modified national growth effect on industry i in the r^{th} region between (t-1) and t.

$IMBIS_{ir}^{t}$ = the Bishop-Simpson modified industry-mix effect on industry i in the r^{th} region between (t-1) and t;

The relationship between different measures introduced above is tested by Pearson correlation coefficients for each period and each measure for the 27 States with nine industries, i.e. 243 observations per period. The results are given in Table 5.6. *IM* is highly correlated with *IMBIS* and *IMEM2*, except for the 2001-2006 period in which the correlation between *IM* and *IMEM2* is low; *CE* is highly correlated with *CEH*; *NEBIS* is highly correlated with *NEEM2*. *IM* and *CE* are uncorrelated in three of the sub-periods.[89] These results are qualitatively similar to those found by Cochrane & Poot (2008) for New Zealand. However, even more of the 28 correlation coefficients per period are statistically significant (positive or negative) in the Brazilian case than in the New Zealand case.

It is also useful to consider a comparison between the findings in this chapter and those of Loveridge & Selting (1998, p. 52). However, Loveridge & Selting calculated the shift-share component extensions for Minnesota from 1979 to 1988 by using income rather than employment. They also calculated correlations for the entire study period, rather than for divided small sub-periods. Therefore, Loveridge & Selting's (1998) results are verified here for each sub-period by considering only significant correlations in both studies. Identical results in both studies are: *AE* and *CEH*: the correlation is approximately -1; *IMBIS* and *IM*: positive correlation; *NEEM2* and *IM*: positive correlation; *IMEM2* and *IM*: 0.50 for both studies when considering the first sub-period, but other sub-periods still exhibit a statistically significant positive correction between these measures, except the last one; *NEBIS* and *IM*: positive correlation; *NEBIS* and *CE*: positive correlation, however, in the Brazilian case, the correlation is much higher; *NEBIS* and *NEEM2*: identical positive correlation of 0.9; *IMBIS* and *NEEM2*: positive, even though the correlation in Brazil is much higher; *IMBIS* and *IMEM2*: identical positive correlation of 0.8.

In general, it can be concluded that while the extensions are theoretically attractive, in practice the information contained in the alternative measures can often be proxied by the basic, and easily interpretable, measures. The cross-study comparison shows that this is the case for the Brazilian, US and New Zealand data. However, as will be shown in the next two sections, extensions that introduce a spatial dimension add an important and informative component to shift-share analysis.

[89] These two effects only have a correlation for the 1981-1986 period (0.510) and for 1986-1991 (0.418) that is significant at 1 percent and 5 percent levels, respectively.

Table 5.6: Simple correlations between shift-share components for the 27 States of Brazil

1981-1986

	IM	*CE*	*CEH*	*AE*	*NEBIS*	*IMBIS*	*NEEM2*	*IMEM2*
IM	1							
CE	**0.510**	1						
CEH	**0.436**	**0.965**	1					
AE	-0.287	**-0.806**	**-0.934**	1				
NEBIS	**0.583**	**0.990**	**0.943**	**-0.770**	1			
IMBIS	**0.998**	*0.455*	*0.381*	**-0.237**	**0.532**	1		
NEEM2	**0.587**	**0.996**	**0.955**	**-0.790**	**0.994**	**0.535**	1	
IMEM2	**0.519**	*-0.471*	**-0.512**	**0.508**	*-0.386*	**0.571**	*-0.388*	1

1986-1991

	IM	*CE*	*CEH*	*AE*	*NEBIS*	*IMBIS*	*NEEM2*	*IMEM2*
IM	1							
CE	*0.418*	1						
CEH	**0.566**	**0.660**	1					
AE	**-0.549**	**-0.525**	**-0.986**	1				
NEBIS	*0.468*	**0.929**	**0.714**	**-0.603**	1			
IMBIS	**0.999**	*0.387*	**0.554**	**-0.541**	*0.439*	1		
NEEM2	**0.519**	**0.993**	**0.687**	**-0.558**	**0.930**	**0.490**	1	
IMEM2	**0.762**	-0.270	0.131	-0.208	-0.168	**0.783**	-0.158	1

1991-1996

	IM	*CE*	*CEH*	*AE*	*NEBIS*	*IMBIS*	*NEEM2*	*IMEM2*
IM	1							
CE	-0.012	1						
CEH	-0.378	**0.682**	1					
AE	*0.381*	**-0.676**	**-1.000**	1				
NEBIS	0.054	**0.974**	**0.615**	**-0.609**	1			
IMBIS	**0.998**	-0.067	*-0.416*	*0.419*	0.002	1		
NEEM2	0.070	**0.996**	**0.646**	**-0.640**	**0.975**	0.015	1	
IMEM2	**0.857**	**-0.525**	**-0.673**	**0.672**	*-0.454*	**0.884**	*-0.453*	1

1996-2001

	IM	*CE*	*CEH*	*AE*	*NEBIS*	*IMBIS*	*NEEM2*	*IMEM2*
IM	1							
CE	0.231	1						
CEH	0.330	**0.833**	1					
AE	-0.322	*-0.421*	**-0.853**	1				
NEBIS	**0.613**	**0.816**	**0.687**	-0.356	1			
IMBIS	**0.999**	0.219	0.321	-0.320	**0.605**	1		
NEEM2	*0.467*	**0.968**	**0.845**	*-0.471*	**0.899**	*0.455*	1	
IMEM2	**0.977**	0.019	0.157	**-0.239**	*0.454*	**0.980**	0.268	1

2001-2006

	IM	*CE*	*CEH*	*AE*	*NEBIS*	*IMBIS*	*NEEM2*	*IMEM2*
IM	1							
CE	0.365	1						
CEH	0.348	**0.985**	1					
AE	-0.324	**-0.947**	**-0.989**	1				
NEBIS	*0.420*	**0.995**	**0.979**	**-0.942**	1			
IMBIS	**0.994**	0.259	0.241	-0.221	0.316	1		
NEEM2	*0.437*	**0.997**	**0.980**	**-0.942**	**0.996**	0.334	1	
IMEM2	0.178	**-0.851**	**-0.845**	**0.819**	**-0.814**	0.288	**-0.807**	1

Notes: **Bold – Significant at 1% level (2 tailed);** *Italics – Significant at 5% level* (2 tailed). Correlation coefficients of absolute value of 0.8 or above are underlined.

5.7. Exploratory Spatial Analysis of Shift-Share Components

This section examines the spatial distribution of the industry-mix and competitive effects of the traditional shift-share decomposition. The tools of spatial autocorrelation analysis that are used include Moran's *I* and cluster maps (Moran, 1950; Getis, 1991; Anselin, 1995; Cochrane & Poot, 2008; Le Gallo & Kamarianakis, 2010).[90] Spatial autocorrelation is increasingly recognised as a major issue in econometric analysis, because the levels of many socio-economic variables are not random in space. In other words, those levels depend on the geographical location of any given region *r*. It often matters whether region *r* has many neighbours or is relatively isolated (Moran, 1950; Anselin, 1989; Biles, 2003; Nazara & Hewings, 2004; Autant-Bernard, Mairesse & Massard, 2007). Researchers who ignore the problem of spatial autocorrelation are more likely to estimate misguided models.

One global (i.e. summary across space) measure of spatial autocorrelation is Moran's *I*, which is defined as follows:

$$I = \frac{\frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{n} W_{ir}(z_i-\bar{z})(z_r-\bar{z})}{\sigma^2(z)} \tag{5.18}$$

In this equation z_i is a variable observed at location *i* with *i*=1,…, *n* (*n*=27 in the application to Brazilian states below), W_{ij} is the spatial weights matrix that portrays interaction between all pairs of regions *i* and *r* (*i*; *r*=1,…, 27); $\bar{z}$ is the sample average of z and $\sigma^2(z)$ is the sample variance of z. The spatial weights matrix can be created by means of software or manually. Moran's *I* autocorrelation measure ranges from -1 to +1. Positive values of Moran's *I* indicate positive spatial correlation, negative values suggest that all regions are surrounded by regions that are "opposites" (in practice this is rarely observed), and a zero Moran's *I* the absence of spatial correlation. The interpretation of Moran's *I* is based on the four quadrants in which the plot of the measure of the regions' interactions against the variable of interest belongs:

> The four different quadrants of the scatterplot correspond to the four types of local spatial association between a region and its neighbours: HH denotes a region with a high value surrounded by regions with high values; LH a region with a low value surrounded

[90] Although there are a number of other spatial autocorrelation indicators such as Geary's ci, Getis and Ord's G1i, and Getis and Ord's G2i (see Pisati, 2010; Newton, 2001).

by regions with high values, and so on. Quadrants HH and LL (respectively LH and HL) refer to positive (respectively negative) spatial autocorrelation indicating spatial clustering of similar (respectively dissimilar) values. (Le Gallo & Kamarianakis, 2010, p. 6)[91].

The simplest spatial interaction matrix is one in which interaction is determined by contiguity, with "1" in the original matrix indicating contiguity and "0" indicating non-contiguity. To create weights, the matrix is row-standardised (each row element is divided by the row sum).

A geographic evaluation of spatial autocorrelation is achieved by LISA (Local Indicators of Spatial Association) because these indicators allow the researcher to identify "outlier regions". This is illustrated by significance and cluster maps in which values of the variable of interest are geo-coded, and the levels are indicated by colour or shading on a map.

A LISA is a statistic that satisfies two criteria (Moran, 1950; Cochrane & Poot, 2008, p. 71; Le Gallo & Kamarianakis, 2010, p. 6):

(i) the LISA for each observation gives an indication of significant spatial clustering of similar values around that observation;
(ii) the sum of the LISA for all observations is proportional to a global indicator of spatial association.

The local version of Moran's I statistic is a LISA and expressed as follows:

$$I_i = (z_i - \bar{z}) \sum_{j=1}^{n} W_{ij} \, (z_j - \bar{z})] \qquad (5.19)$$

and hence

$$I \equiv \frac{1}{n\sigma^2(z)} \sum_{i=1}^{n} I_i \qquad (5.20)$$

There are two important issues in the analysis of local Moran's I_i, as defined in equation (5.19) above:

Firstly, the local Moran's I_i is not approximately normally distributed. This difficulty has been overcome in practice in a relatively straightforward manner by using a conditional

[91] These quadrants are also known in the spatial econometrics literature as: High-High=hot spots; Low-High=spatial outliers; High-Low=spatial outliers, Low-Low=cold spots, and locations with no significant spatial autocorrelation.

> randomisation or permutation approach to yield empirical pseudo significance levels.
> A second complicating factor arises from the fact that the LISA statistics for individual locations will tend to be correlated which, along with the related problem of multiple comparisons, will lead to a flawed interpretation of the level of significance. Anselin suggests employing either the Bonferroni or Sidak correction to account for the multiple comparisons. However, the assumption of multivariate normality in the case of the Sidak correction is unlikely to be met by spatial data, while a Bonferroni correction may be too conservative. (Cochrane & Poot, 2008, p. 71).

In what follows, Moran's *I* scatterplots and cluster maps are presented for the Industry-Mix (IM) and Competitive Effect (CE) components of the classic shift share analysis of section 5.4. The chosen values for IM and CE for each of the cluster maps are the averages across the five sub-periods. The spatial weights matrix for Moran's *I* is a simple first order row-standardised "queen's contiguity" matrix of Brazil that was created in Microsoft Excel. Queen's contiguity means that regions are considered contiguous if they have either a common border or a common edge.[92].

Moran's *I* scatterplots for both IM and CE were estimated and the Moran's *I* significance levels were calculated by an OLS regression of the spatially weighted value for all regions outside any particular region against the value of the variable in that particular region.[93] This OLS regression is precisely what is represented by equation (5.18). The Moran scatter plot for the IM effect is displayed in Figure 5.1. Moran's *I* (i.e. the slope of the regression line) is positive (0.4563) and statistically significant at the 1 percent level. This indicates that there is a clear pattern of a positive spatial association for the IM effect. This is also supported by the cluster map (see Figure 5.2), which shows a clear pattern of a contiguous area with high levels in industry-mix (i.e. hot spots). This area involves the following states (middle-west-northwest space): Distrito Federal, Goiás, Mato Grosso, Rondônia, Acre, Amazonas, Roraima; and two pairs of an "island" neighbouring states, which are (southeast) São Paulo-Rio de Janeiro and (northeast) Rio Grande do Norte-Paraiba, and an isolated "island" Amapá.

[92] On the other hand, rook contiguity and bishop contiguity consider regions as contiguous if and only if they share a common border and a common edge, respectively. These definitions can directly be downloaded from www.s4.brown.edu/s4/Training/Modul2/GeoDa2.pdf.
[93] To address the problem of small sample size, the data were pooled to obtain 135 observations (or, 27 states times 5 sub-periods).

Figure 5.1: Moran's I Scatterplot, Industry-Mix (pooled 5 sub-periods of 5 years)

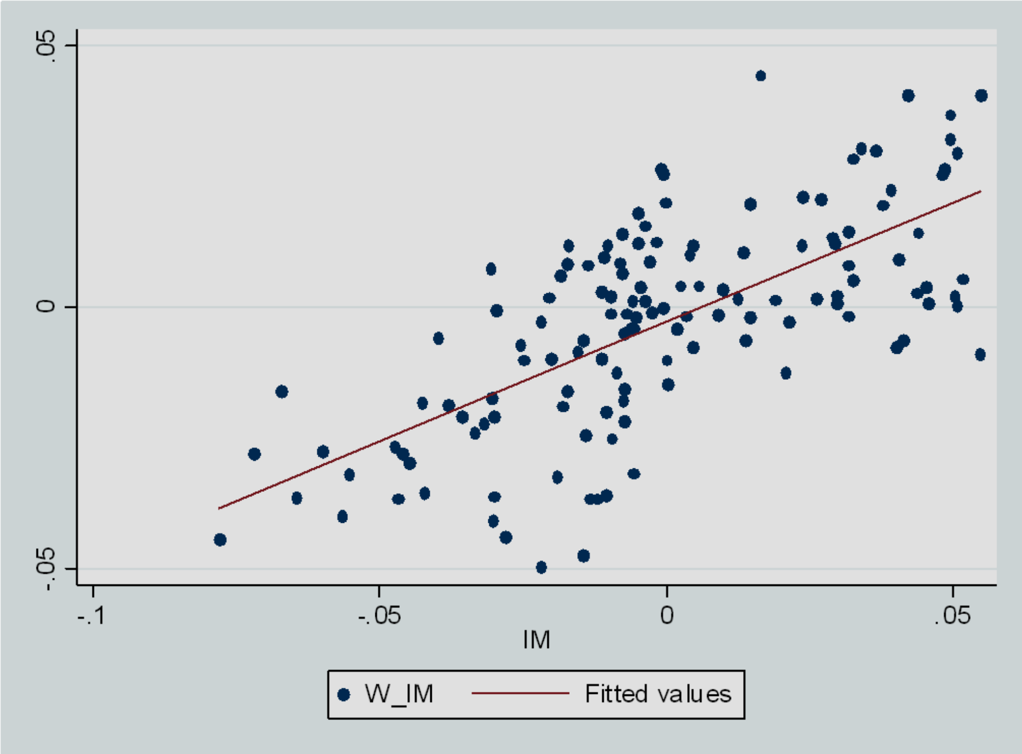
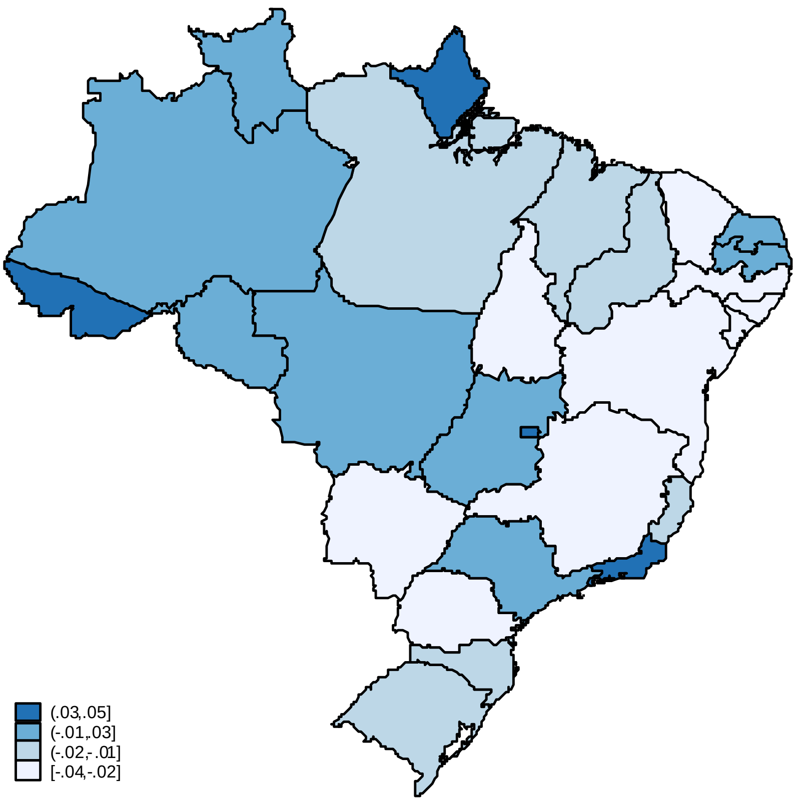


Figure 5.2: Industry-Mix Cluster Map (average of 5 sub-periods of 5 years)



On the other hand, there is another cluster of contiguous states with low industry-mix levels (cold spots), which are (in south-east and north-northeast land areas): Minas Gerais, Bahia, Tocantins, Sergipe, Alagoas, Pernambuco and Ceará, and (in centre-west and south land area): Mato Grosso do Sul-Paraná.

For the CE effect, the Moran's *I* scatter plot (Figure 5.3) also shows a positive (0.3375) and statistically significant Moran's *I*. However, comparing Figures 5.1 and 5.3 it is clear that there is greater spatial correlation in the industry-mix effect than in the competitive effect. A similar result was observed by Cochrane and Poot (2008). The cluster map Figure 5.4 shows hot spots located in (contiguous) northern states of Rondônia, Acre, Amazonas, and Roraima; northeast states of Pernambuco-Paraíba-Rio Grande do Norte; southeast-centre-west states of Espírito Santo, Minas Gerais and Mato Grosso do Sul, and in "islands" Amapá, Tocantins, and Distrito Federal. On the other hand, the cold spots are found along the east coast of Brazil, namely in the southeastern states of São Paulo and Rio de Janeiro and in the northeast-north land area (Piauí, Maranhão, and Pará), which overlaps an area with a relatively poor industry-mix, and "islands" of Goiás and Rio Grande do Sul.

Figure 5.3: Moran's I Scatterplot, Competitive-Effect (pooled 5 sub-periods of 5 years)

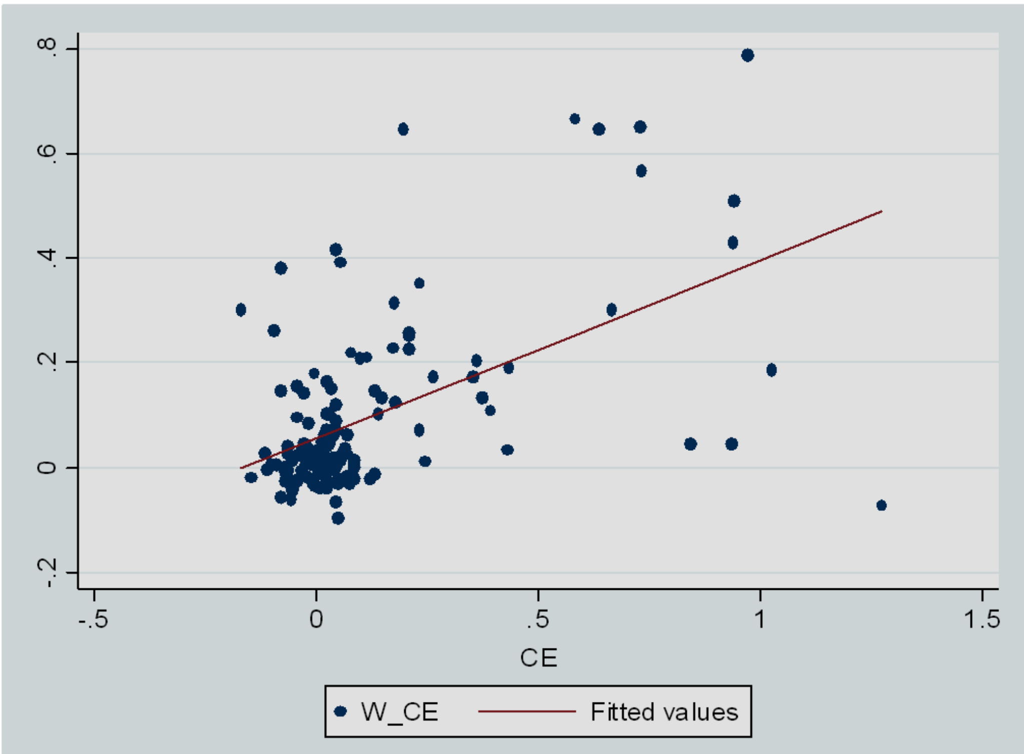
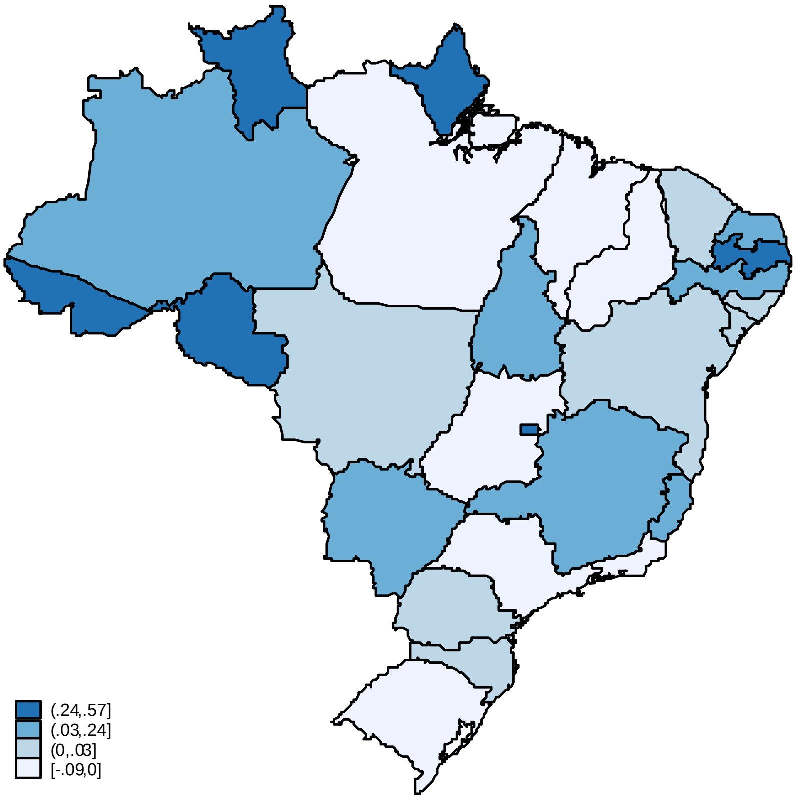


Figure 5.4: Competitive Effect Cluster Map (average of 5 sub-periods of 5 years)



The economic interpretation of the results above is that two clusters of extremes (High-High *versus* Low-Low)[94] can be observed, which is consistent with the positive spatial autocorrelation across states in Brazil generally and the argument that scale economies may arise as a consequence of local agglomeration of economic activities (Krugman, 1991b). The evidence, based on the two shift-share components, in favour of economic agglomeration theory is as follows: the industry-mix result indicates low specialisation for many states.[95] This finding reconfirms many previous studies for Brazil (Rolim, 2008; Daumal & Zignago, 2010, pp. 747-748, and footnote 22, p. 747) that found convergence across states. However, this convergence is due to the improvement of the industry-mix (i.e. greater diversity) for the less developed middle-west and northern states rather than specialisation. [96]

[94] This result suggests interstate mobility among businesses may be low.

[95] On the map for industry-mix average, these states are: Amapá, Roraima, Amazonas, Acre, Rondônia, Mato Grosso, Goiás, Distrito Federal, São Paulo, Rio de Janeiro, Rio Grande do Norte, and Paraiba. These states are essentially from north and middle-west which are the regions benefited from convergence from 1981 to 2006.

[96] A conclusion based on cross-sectional data may be misguided. Although northern states were more specialised in the 1981 cross-section (see Table 5.3), the analysis of the long-run trend of the so-called Hirschman-Herfindahl Index (see the main text) indicates that those lagging states

In fact, a long-run analysis reveals that, when considering annual data for sectoral employment of all states except Tocantins[97], the average Hirschman-Herfindahl Index (HHI) is quite stable in Brazil. This index is defined as follows:

$$HHI_i = \sum_j s_{ij}^2 \qquad (5.21)$$

Where: s_{ij} is the employment share of the sector *j* in state *i*'s total employment.

Figure 5.5: Hirschman-Herfindahl Index, Brazil

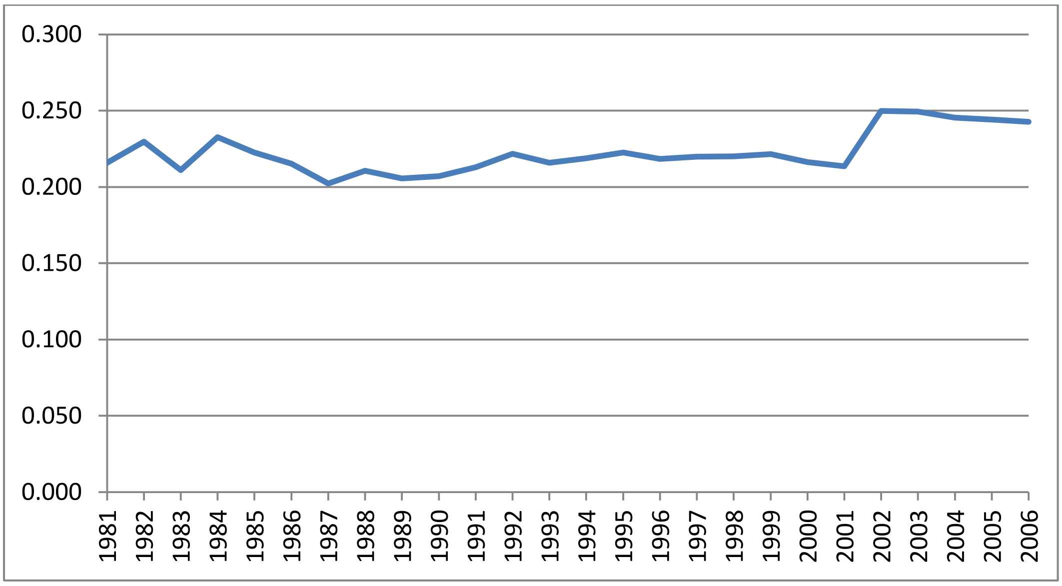


Figure 5.5 shows that HHI oscillates between 0.20 and 0.23 from 1981 to 2001. From 2001, it increases to 0.25 in 2002, then falls to 0.24 in 2006. Nevertheless, the average of HHI for the period 1981-2006 varies considerably across states. Overall, the northern states, which had the highest growth rates, were less specialised (higher industry-mix effect) with the lowest averages of HHI (lesser than 0.20) and the northeastern states were more specialised (lower industry-mix effect) with the highest averages of HHI (higher than 0.22). Generally, a high average industry-mix indicates that the industrial structure of the fastest growth states has been diversified. On the other hand, the result from Figure 5.4 clearly shows a higher performance of the northern states which are some of the lagging ones, while the most developed south-eastern states of São

increased their competitiveness over time due to lesser specialisation relative to both the nation and the other states.

[97] This state has been excluded due to a lack of state sectoral data from 1981 to 1991.

Paulo and Rio de Janeiro had relatively lower competitiveness. The explanation for higher growth for the lagging regions is as follows. Due to their low income level and their early stage of development, small increases in capital, average education and infrastructure improvement have a large effect on their growth rates. This result supports the neoclassical beta convergence hypothesis (see also Resende, 2011 and the references therein).

5.8. Spatial Shift-Share Analysis

While section 5.7 investigated the spatial properties of the classic shift-share components, this section adds a new spatial component to the shift-share accounting framework in order to investigate regional growth of the 27 states in Brazil from 1981 to 2006. The regional growth rate is decomposed according to the taxonomy of spatial shift-share developed by Nazara & Hewings (2004). The growth rate for sector *i* from time (*t*-1) to *t* in region *r* is linked to the interaction between regions as defined by spatial contiguity. The incorporation of a spatial effect on the growth rate of sector *i* in region *r* is done by means of a four step procedure. First, the spatial contiguity matrix (27x27) for the 27 states in Brazil that was used in the previous section is used again here. Spatial contiguity is indicated by "1" if states share a border or an edge, or zero otherwise[98]. Secondly, this spatial contiguity matrix is row-standardised by taking the ratio between each cell and the sum of its matrix row. Thirdly, values of each cell of the row-standardised spatial weights matrix are multiplied by values of the corresponding sector employment in the states. Fourthly, the percentage changes of the spatially weighted sectoral employment from time (*t*-1) to *t* is defined as the spatial growth rate of the sector *i* in the region *r*.

Nazara & Hewings (2004, p. 480) express these steps for both sector *i* and region *r* on the right-hand side of their equation (4) that is reproduced here [equation (5.22)]. This equation defines the spatially-weighted growth rate of a region's *r* neighbours. To account for the neighbour effect in a region's *r* growth rate, Nazara & Hewings replace the nation's sector *i* growth rate, G_i, by region *r's* neighbour sector *i* growth rate, which is labelled gS_{ir}^t in equation (5.22) [Nazara & Hewings, 2004, p. 480, equation (5)]

$$gS_{ir}^t = \frac{\sum_{k=1}^R \tilde{w}_{rk} E_{ik}^t - \sum_{k=1}^R \tilde{w}_{rk} E_{ik}^{t-1}}{\sum_{k=1}^R \tilde{w}_{rk} E_{ik}^{t-1}} \tag{5.22}$$

where: gS_{ir}^t = Growth rate of employment in industry *i* and region's *r* neighbour between (*t*-1) and *t*; $\tilde{w}_{rk}$ is the element of row-standardised spatial weights matrix $\widetilde{W}$ that captures interactions between regions *r* and *k*; E_{ik}^{t-1} and E_{ik}^t are, respectively, employment in the i^{th} industry in the k^{th} region at time (*t*-1) and *t*.

[98] Hence queen contiguity is again adopted. The Distrito Federal is a region within Goiás state. They are assumed to share a border.

The decomposition of employment growth rate for sector *i* from the period *t*-1 to *t* in the region *r* after the spatial effects have been incorporated in the classic shift-share method is as follows: substituting Nazara & Hewings (2004, pp. 480-481) equation (6) in their equation (5), the following four shift-share components are obtained:[99]

$$\Delta E_{ir}^{t} = (E_{ir}^{t} - E_{ir}^{t-1}) = NE_{ir}^{t} + IM_{ir}^{t} + PSE_{ir}^{t} + SCE_{ir}^{t} \qquad (5.23)$$

The first two terms of the right-hand side of equation (5.23) are from the classic shift-share method, as defined in section 5.2 (equations 5.2 and 5.3). The new terms that refer to spatial effects for growth of regions are:

$$PSE_{ir}^{t} = (gS_{ir}^{t} - g_{i0}^{t})E_{ir}^{t-1} \qquad (5.24)$$

$$SCE_{ir}^{t} = (g_{ir}^{t} - gS_{ir}^{t})E_{ir}^{t-1} \qquad (5.25)$$

The terms in the above equations are defined as:[100]

PSE_{ir}^{t} = Potential Spatial Spillover Effect. It is the regional growth a region would have if spatial autocorrelation is 1, i.e. the corresponding regional growth rate is identical to the spatially weighted regional growth rates of surrounding regions. But surrounding regions could have the same industry-mix effect if there is spatial correlation in industry mix (which there is in the Brazilian case), so the potential spillover effect must be adjusted by subtracting the industry mix growth component, g_{i0}^{t}. PSE_{ir}^{t} is referred to as a potential spatial spillover effect because it is the maximum possible effect. The real spatial autocorrelation is likely to be much less than 1 so the actual spatial spillover will be far less than the potential spatial spillover.

SCE_{ir}^{t} = Spatial Competitive Effect (or, the negative of neighbour-nation regional shift effect defined by Nazara & Hewings, 2004, p. 481).

[99] From Nazara & Hewings (2004, pp. 480-481) seven components can be identified. However, at the regional level, two components individually add to zero. These are: neighbour industry-mix effect and regional industry-mix effect (or, the negative own-region industry-mix effect). And there is a double counting for the other two: the neighbour-nation regional shift effect is equal to minus the neighbour-region regional shift effect. Thus, these components are excluded and a simplified version of the spatial shift-share identity with only four components is used.

[100] The definition of g_{i0}^{t} is in section 5.2, and that of gS_{ir}^{t} is in the above equation (5.22).

5.8.1. Spatial Shift-Share: Results for a Simple Contiguity Spatial Weights Matrix

This sub-section presents the results of the spatial shift-share method. The average regional growth rates for the five sub-periods are compared with the average for each of the four components in all states. Figure 5.6 shows state growth rates and the national growth rate.[101] It can be seen that there are three groups of states. The first group grew faster than the nation and had the highest average growth rates. This group includes: Roraima, Rondônia, Amapá, Acre, Distrito Federal, Pará, Amazonas, Tocantins, Mato Grosso, Rio Grande do Norte, Mato Grosso do Sul, and Espírito Santo. The second group had very similar growth rates to the nation. This group consists of Santa Catarina, Alagoas, Minas Gerais, Maranhão, Goiás, Ceará, Sergipe, and Piauí. Finally, the third group includes the remaining seven states which had a growth rate smaller than the national rate.

Figure 5.6: State growth rates versus national growth rate (average of 5 sub-periods of 5 years)

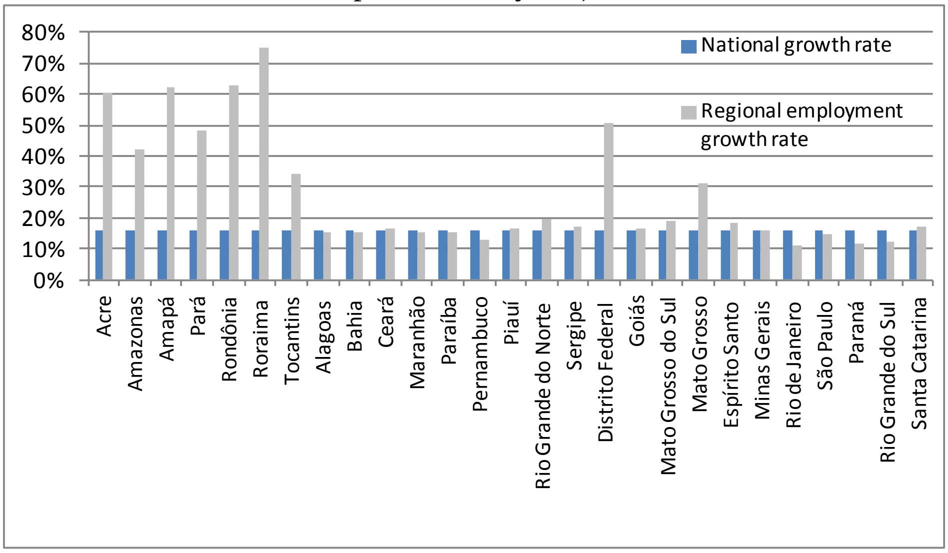


Figure 5.7 shows the regional growth rate and the industry-mix effect. The states with the highest growth rates also had a positive national industry-mix effect, i. e. those endowed with industries that were growing faster than average. These are seven states, namely Roraima, Rondônia, Amapá, Acre, Distrito Federal, Pará, and Amazonas. Rio de Janeiro and São Paulo are the only non-fast growing states that join this group with a positive national industry-mix effect. On

[101] In all figures that follow, the light shading refers to regional growth rate and the dark one to the defined components of the spatial shift-share.

the other hand, other states lacked important industries in terms of growth which yielded a zero (for Rio Grande do Norte, Goiás, and Mato Grosso do Sul), or a negative effect (for all the other 15 states) in this component. However, the industry mix effect is small relative to regional growth performance in all states.

Figure 5.7: State growth rates versus national industry-mix effect (average of 5 sub-periods of 5 years)

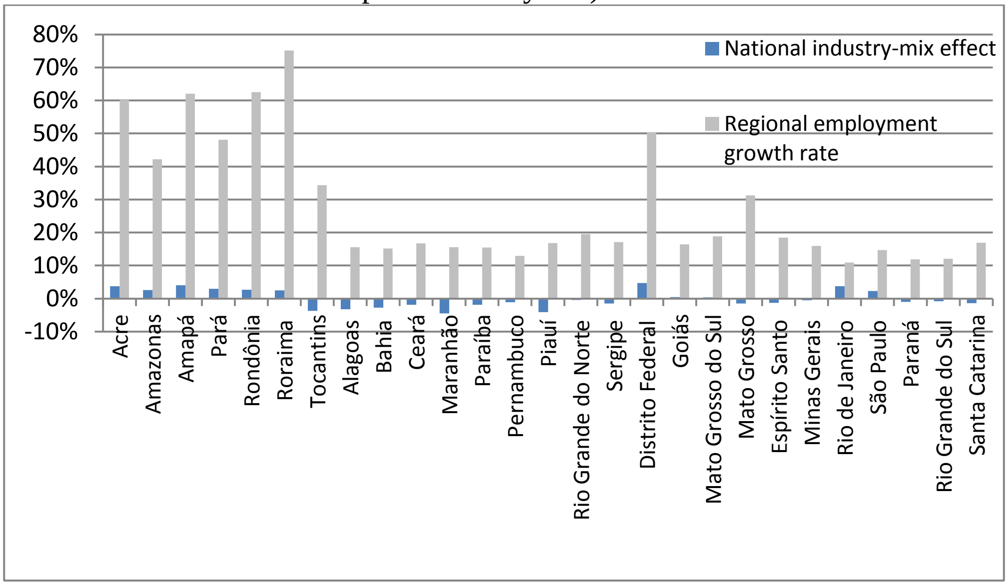
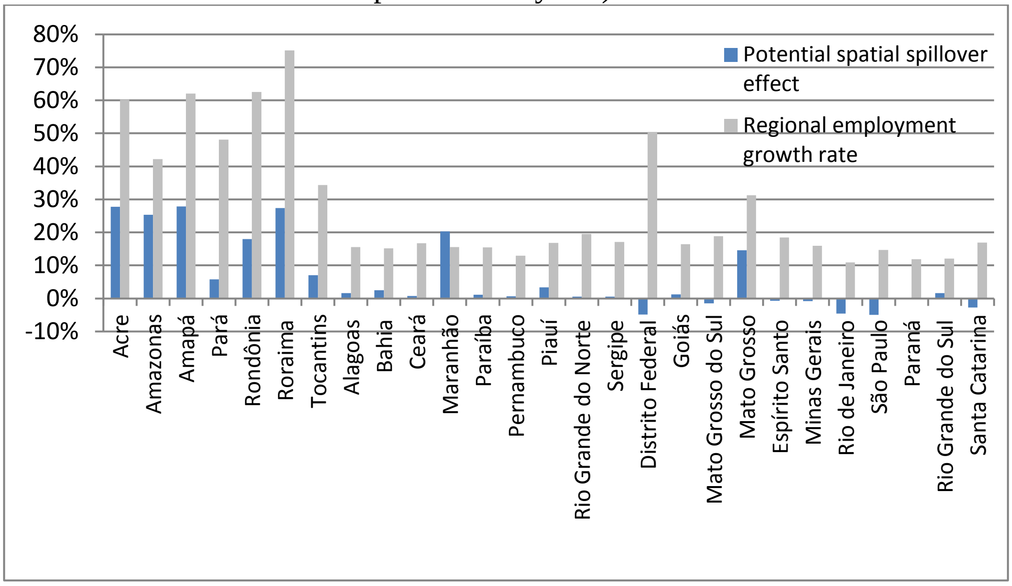


Figure 5.8 compares regional growth rates with the potential spatial spillover effect. Among the states that grew fastest are those that had the highest (positive) potential spatial spillover effect, i. e. their neighbouring states grew faster than the expected growth based on industry composition.[102] These are (ordered according to the size of spatial spillover effect, at least 15 percent): Amapá, Acre, Roraima, Amazonas, Maranhão, Rondônia, and Mato Grosso. Some other states still had a positive potential spatial spillover effect (but only up to 7 percent), such as Tocantins, Pará, Piauí, Rio Grande do Sul, Bahia, Alagoas, Goiás, Paraiba, Ceará, Pernambuco, Rio Grande do Norte, and Sergipe. The remaining seven[103] states had a negative potential spatial spillover effect[104], indicating that they were surrounded by states with weak growth relative to the expected growth according to the industry composition.

[102] Excluding Maranhão.
[103] Excluding Paraná for which this effect was equal to zero.
[104] It is worth noting that both São Paulo and Rio de Janeiro had the lowest level of this effect, -5 percent.

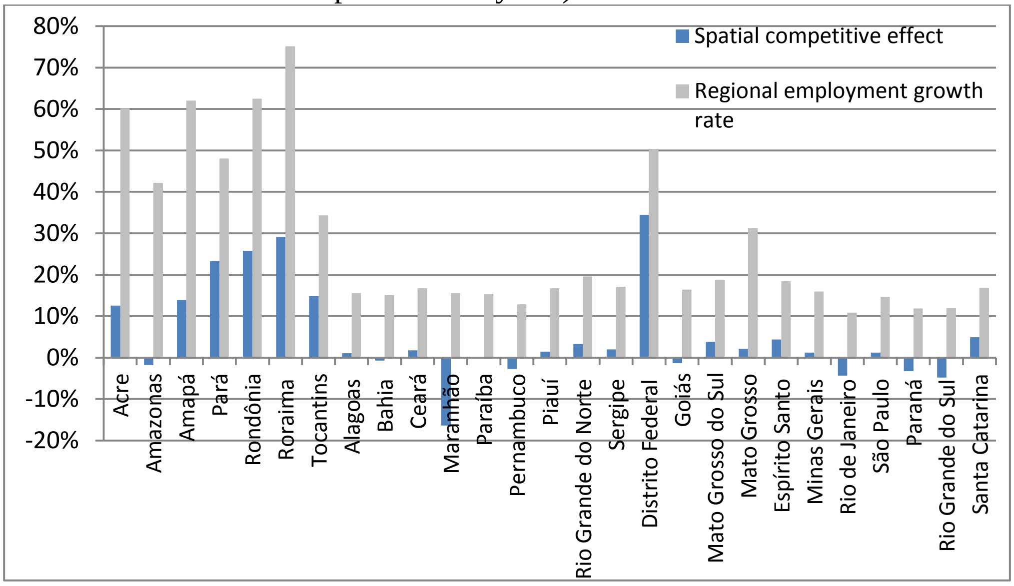
Figure 5.8: State growth rates versus potential spatial spillover effect (average of 5 sub-periods of 5 years)



Finally, Figure 5.9 shows the regional growth rate and the spatial competitive effect. Most of the states that grew fastest also had the highest (positive) spatial competitive effect, i. e. they grew faster than the surrounding regions. These are (ordered according to the size of spatial competitive effect, at least 13 percent): Distrito Federal, Roraima, Rondônia, Pará, Tocantins, Amapá, and Acre. Some other states still had a positive spatial competitive effect (but only up to 5 percent), such as Santa Catarina, Espírito Santo, Mato Grosso do Sul, Rio Grande do Norte, Mato Grosso, Ceará, Sergipe, Alagoas, Piauí, São Paulo, and Minas Gerais. For the remaining nine states, the spatial competitive effect was zero for Paraíba and negative for the other eight states[105] due to smaller growth rates relative to growth rates of their neighbouring states, which indicates that their poor performance is particularly due to their own weaknesses.

[105] It is worth noting that for one fast grower, the northern state of Amazonas, this effect is also negative. Maranhão, Rio Grande do Sul and Rio de Janeiro had the lowest levels for this component, -16, -5, and -4 percent, respectively.

Figure 5.9: Growth rates versus spatial competitive effect (average of 5 sub-periods of 5 years)



Inspection of the sub-periods indicated that the sub-period 1991-96 was atypical. Only four states grew fastest, namely Tocantins, Amapá, Roraima, and Maranhão, and among the other states, most had a moderate growth rate, between 8 and 16 percent. The characteristics for this period are that it had the lowest national growth rate and very low levels for the other three components, national industry-mix effect, potential spatial spillover effect, and spatial competitive effect for almost all states. On the other hand, the sub-period 1996-2001 stands out as the one with very negative spatial competitive effect for nine states, mostly located in the north and northeast regions.

5.8.2. Spatial Shift Share: Results for an alternative Spatial Weights Matrix

A valid question in spatial shift-share analysis is whether the results are sensitive to the definition of the spatial weights matrix. In order to investigate this issue, an alternative spatial weights matrix is considered in this sub-section. This alternative row-standardised spatial weights matrix takes into account population data and distance data for the beginning year of each of the five sub-periods. Population data used refer to the 27 urban areas that constitute Brazilian state capitals. These urban areas were defined through observation of contiguous municipalities taken together in 2008 (see Appendix 4.1 of Chapter 4). The urban areas population data were calculated by interpolation for the years in which there are no data from the official sources, that is, for the first three beginning years of the five sub-periods.[106] For the other two beginning years, 1996 and 2001, the correlation between the original and calculated data is 0.99, which shows the resemblance between both types of data. For consistency with the other previous three years, the calculated data have been used in this study.

The Municipality Population Data used for construction of the urban areas were obtained from the Institute of Applied Economic Research (IPEA) and the Brazilian Institute of Geography and Statistics (IBGE).[107] The matrix of distances between the 27 Brazilian state capitals was obtained from Brazil's Ministry of Transportation. The spatial weights matrix used to measure the states' interactions is based on the gravity model, which relates distance between regions and population size of those regions (see Getis, 1991, pp. 29-30; Bavaud, 1998, pp. 157-158; McCann, 2001, p. 202; Brakman, Garretsen & van Marrewijk, 2001, pp. 265-270), and is defined as: [108]

(i) $$w_{rk}^{*(t-1)} = D_{rk} / \sqrt{P_r^{t-1} P_k^{t-1}} \qquad (5.26)$$

(ii) $$w_{rk}^{t-1} = w_{rk}^{*(t-1)} / (sum\ of\ w_{rk}^{*(t-1)}\ over\ k) \qquad (5.27)$$

[106] These years are 1981, 1986, and 1991.

[107] The IPEA data source has been briefly introduced in section 5.3. IBGE is one of the most important public socio-economic data sources and is the institution that conducts censuses in Brazil.

[108] In fact, the gravity equation suggests that the spatial interaction between regions is inversely related to distance between pairs of regions and positively related with the product of economic size of the two respective regions. Here population is used as an indicator of the scale of regional economy.

Where D_{rk} is distance between regions r and k; P_r^{t-1} and P_k^{t-1} is population sizes of the capitals r and k at time (t-1), which is the initial year of sub-period under consideration.

Comparing the obtained results using this alternative spatial weights matrix with those above that used the queen contiguity matrix of spatial weights, it turns out that the results are very similar for all components in all states for each of the five-year sub-periods from 1981 to 2006 as well as for the averages for whole period.[109] Therefore, in the Brazilian context, the first-order spatial weights queen contiguity matrix and the spatial matrix based on the gravity model can substitute for each other because both yield the same results.

5.9. Conclusion

This chapter applied different techniques to analyse employment growth across 27 states in Brazil from 1981 to 2006. Three key conclusions can be drawn from the analysis. First, from the classic shift-share method it can be concluded that higher employment growth rates of the less developed regions are due to these regions' comparative advantage associated with high performance of the industry-mix and competitive effect components irrespective of the national structural change. This evidence confirms previous studies that found regional convergence in Brazil (Rolim, 2008; Daumal and Zignago, 2010). The reason for this convergence appears to be an improvement of diversity of the economies of the less developed regions (i.e. northern states) given that they had the smallest Hirschman-Herfindahl Indexes as well as higher performance in the industry-mix component, rather than specialisation as previous studies have pointed out. Secondly, examination of the industry-mix and competitive effect components employing exploratory spatial data analysis (ESDA) provided evidence of a positive spatial association for both components. This result supports agglomeration economies and beta convergence theories, as previously found (Resende, 2011, and the references therein), because, compared with the 1960s, nowadays economic activities are slightly less concentrated in the southern and more developed regions of Brazil.

Thirdly, the chapter provided a simplified version of Nazara & Hewings's (2004) spatial shift-share taxonomy from which the role of spatial autocorrelation

[109] Graphs are not shown here but are available upon request.

in regional growth in Brazil could be quantified in a straightforward way. Growth differentials in favour of northern and middle-west states are basically associated with their strengths in two regional components of the spatial shift-share, namely potential spatial spillover effect and spatial competitive effect that, together, outweigh the poor performance on national industry-mix effect for those lagging states. On the other hand, most states in Brazil had lower growth rates that were associated with their low rates of both spatial components[110]. These results confirm the core-periphery framework which is associated with the importance of agglomeration forces in Brazil (Brakman, Garretsen & van Marrewijk, 2001). This association implies that, due to the large regional disparities and large scale of concentration in favour to the southeast-south regions (the core), a fast growth for the lagging regions (the periphery) is still less relevant to change the spatial pattern of economic activities, because the initial conditions that strongly favoured the core seem to have essentially permanent effects in Brazil. As a consequence, for instance, the observed modest growth rate for São Paulo (the core) still counts, given the scale of this state's economy, for much of the concentration of economic activities, population and income in Brazil.

This chapter has two caveats. First, the available data have high level of aggregation, i.e. the state level. Had employment data been available at the municipality level, rather than at the state level, this would have allowed a spatial regression approach to quantify the various components of regional growth. Given improving data availability in recent years, this could be an avenue for future research. Secondly, the employment data used were retrieved from IPEA website but originally come from Pesquisa Nacional por Amostra de Domicílios (or, 'National Survey by Sample of Households', PNAD/IBGE). These Surveys include both formal and informal employment. With the country's development, the found higher employment growth rates for lagging regions may partially be due to more formalisation of employment in these regions than in the core regions. This implies that the true convergence is probably smaller than the found, which reinforces the observed core-periphery pattern.

---

[110] It is well-known that economic variables considered in a study may be correlated with the unknown variables not included in the analysis. The unknown variables are assumed to be in the residual which its expected value is zero. This assumption also holds in this chapter. The effort of decomposing further the competitive effect of the classic shift-share method into potential spatial spillover effect and spatial competitive effect still leave a residual on the latter component of the spatial shift-share method.

CHAPTER 6 – A NATURAL EXPERIMENT IN BRAZIL: ECONOMIC IMPACTS OF THE CREATION OF BRASILIA CITY

6.1. Introduction

Economic growth and economic development are key areas of debate and concern among economists (Barro & Sala-i-Martin, 2004; Krugman, 1995). Many factors, such as the average skill level of the population, capital accumulation and population density are correlated with economic development, but one of the fundamental challenges in empirical economics is that the causality usually runs both ways. Hence an effective way to understand the economic development of a country is to evaluate the socio-economic effects of an exogenous shock, such as a natural disaster or a large public infrastructure project. It has also become clear in recent years that the impacts of such a shock are unlikely to be localised but instead are spatially distributed and that this needs to be taken into account in econometric estimation (Anselin, 1995; Nazara & Hewings, 2004; Autant-Bernard, Mairesse & Massard, 2007; LeSage & Pace, 2009). The exogenous shock studied in this chapter is the creation of Brasilia City in the Central Plateau of Brazil and the associated highways built from Brasilia to other regions. Unlike most other cities in Brazil and worldwide, Brasilia City was created by the vision and direction of a politician, President Juscelino Kubitschek. The city was inaugurated in 1960, and did not arise from the natural development of a national urban system such as described by Loschian or Christallian settlement patterns (McCann, 2001; Brakman, Garretsen & van Marrewijk, 2001; Portnov & Benguigui, 2010; Zipser, Mlek & Zipser, 2010; Wang, Liu & Chen, 2010).

In studying impacts of the creation of Brasilia, this chapter tests whether such a significant shock affects not just the area where the city was established but also regional growth elsewhere in the country. The question whether the spatial pattern of socio-economic indicators produced by the shock changes underlying developments in income (GDP), population and income per capita is addressed. Specifically, three hypotheses are tested (all stated relative to the counterfactual of what would have occurred without the creation of Brasilia), with most emphasis placed on testing the third hypothesis:

(1) The creation of Brasilia led to spatial changes in GDP across regions, affected by proximity to Brasilia and the highway network;
(2) The creation of Brasilia led to spatial changes in population across regions, affected by proximity to Brasilia and the highway network;
(3) The creation of Brasilia led to spatial changes in income per capita across regions, affected by proximity to Brasilia and the highway network.

Previous studies about Brasilia have typically taken three approaches. The first approach is historical / anthropological and discusses the social relations in Brasilia compared with those found nationwide (Epstein, 1973; Madaleno, 1996; Galanternick, 2003; Story, 2006; Kohlsdorf, Kohlsdorf & Holanda, 2009; Issitt, 2010; Skidmore, 2010). This approach finds that over time Brasilia frustrated the utopian expectations of its planners as it reproduced similar patterns of social classes by spatial segregation related to income that characterised other Brazilian cities; something the planners wished to avoid.

The second approach is architectural. This approach shows that the characteristics of Brasilia's buildings and green spaces (showing the influence of French and British modernist architecture) reflect its original Master Plan prepared by Lúcio Costa in 1957 (Mello, Wilheim & Costa, 1960; Sanders, 1973; Cornish, 1991; The Economist, 1998, 1999; Kubitschek, 2000; El-Dahdah, 2003; William, 2007, p. 342; Issitt, 2010). This approach finds that, "(…) Brasília is a clearly defined point of rupture, after which everything changes (…)" (Williams, 2007, p. 309). However, aspects of the city's built environment have been criticised: Brasília's strict zoning and large open spaces with their representation of power facilitated the operation of military rule (Madaleno, 1996, p. 273; del Rio & Siembieda, 2009). In addition, the creation of the city caused overspending by government that resulted in subsequent fiscal difficulties.

The third approach is economic and examined the effect of Brasilia and its associated highways[111] on development of the centre-west region, and on the

[111] Many highways were built to connect Brasilia to all corners of the country (radial highways), and others were upgraded over the same period (mostly longitudinal, transverse, diagonal, and connecting highways; DNIT, 2011). These affected the development of municipalities, particularly those located adjacent to these highways. For instance, the following neighbouring cities emerged during the construction of Brasília and its related highways: Cidade Livre (or Núcleo Bandeirante), 19/12/1956 and Taguatinga, 05/06/1958 (Williams, 2007, p. 344); Candangolândia (03/11/1956), Cruzeiro (30/11/1959), Brazilândia (created on 05/06/1933 and integrated into Brasília in 1960), Sobradinho (13/05/1960), Gama (12/10/1960) (Gonçalves, 2002) while of those that were already

country's political, socio-economic, and institutional transformation (Snyder, 1964; Katzman, 1975; Becker, 1977; Smith, 1987; Madaleno, 1996; IPEA, 2010b; Farret, 2001; Serra, Dowall, Motta & Donovan, 2005; Dowall & Monkkonen, 2007; Holanda, Ribeiro & Medeiros, 2008). The first five studies, considering the period between the 1950s and the 1990s, adopt a descriptive approach to either evaluate changes in population and migration patterns for municipalities located near the new highways (especially those that connect Brasília with northern regions, such as Belém-Brasilia highway) or to examine changes in the centre-west's share of national GDP since Brasília's creation. They find that, as a consequence of the creation of Brasilia, the centre-west region increased its population and GDP shares in the nation from around 4 and 2.5 percent in 1960 to 7 and 10 percent in 2007, respectively (and consequently transforming the city from one with below average labour productivity to one with average labour productivity).

Becker (1977) and Katzman (1975) study the effect of the Belém-Brasilia highway on regional development of the surrounding municipalities. Katzman's (1975) study, however, is the only study with a regression modelling approach. Kaztman measures the impact of the Belém-Brasilia highway on rural settlement. After regressing population growth for the period 1960-1970 against distance from Goiânia to county *i*, distance from county *i* to Belém, and a dummy variable for counties located on the highway, he finds that, "(…) Counties located on the highway grew significantly faster than those off the highway, but given location, access to Belém had no additional impact. Since counties off the highway grew at about the same rate as the rest of the state, there are no grounds for concluding that the highway merely drained population from established areas in the north" (Katzman, 1975, p. 103). This study also notes that among counties (municipalities) in the northern states, the municipalities on the Belém-Brasilia highway grew faster (e. g. Imperatriz has grown by 126 percent) than those off this highway (e.g. Chapada Sul Maranhense has grown by 11 percent). However

established towns, Planaltina (1810) and Brasilândia (1930s), became cities resulting from Brasília's sprawl (Kohlsdorf, Kohlsdorf & Holanda, 2009). This frustrated the planned target of 600,000 inhabitants for Brasília city (the current CDB) in the original Brasília Pilot Plan (Kohlsdorf, Kohlsdorf & Holanda, 2009). Other cities surrounding Brasilia were born over the decades that followed its inauguration in 1960. A description of the country's integration due to transportation infrastructure development and the number of cities born in the countryside as a consequence of Brasilia was provided by Gomes (2009) (see, for details, *Diário de Cuiabá* Magazine on http://www.diariodecuiaba.com.br/detalhe.php?cod=58833).

there are regions from south Goiás state (off the highway) that grew even faster (e.g. Araguaia Paraense grew by 246 percent). While this latter result may have been due to fiscal incentives and a boom in pepper exports (Katzman, 1975, p. 103)[112] the data do not show a clear delineation in population growth related to proximity to the highway.

Other studies by economists (Holanda, Ribeiro & Medeiros, 2008; Farret, 2001; Serra, Dowall, Motta & Donovan, 2005; Dowall & Monkkonen, 2007) focus on Brasília itself rather than its impact on the centre-west region or the Goiás state economy. They analyse the economic implications of the pilot plan and government policy on land use in Brasília and find that Brasília has a settlement pattern that leads to urban sprawl. This leads to worse spatial segregation than observed in other cities of a similar size (in terms of population and area) due to the high cost of living associated with the large costs of commuting, land, and infrastructure in Brasília (Serra, Dowall, Motta & Donovan, 2005; Dowall & Monkkonen, 2007).

This chapter fills a major gap in the analysis of Brasilia's economic impacts in three respects. First, previous empirical studies analyse the effect of Brasilia for the centre-west and north regions, and also for Brasilia itself, but not for the national economy. Secondly, this study provides new insights into the importance of agglomeration forces versus regional development forces in driving economic systems. This chapter's unique evidence shows that the effects of agglomeration forces outweigh the effects of regional development policy, which implies that even a huge localised shock (the creation of Brasilia) can be insufficient to reverse the path of the established agglomeration forces. Thirdly, at a methodological level, there has been to date no previous empirical study about Brasilia's effect that explicitly takes spatial autocorrelation into account. Spatial econometric tests and techniques are used to measure spillover effects in accounting for the effect of the creation of Brasilia on growth of Brazilian regions from 1939 to 2008.[113]

[112] Katzman (1975, p. 102; 104) estimates that the total population growth from 1960 to 1970 due to Belém-Brasilia highway for Goiás state and other states directly affected by this highway, such as Pará and Maranhão, is between 160,000 and 320,000 people. This suggests that only 8 percent of population growth was due to the highway.

[113] As Kubitschek (2000, p. 7) puts it, "(…) The population nucleus created in that far away region would spill over as a stain of oil allowing the interior to open eyes for a great country's future". Referring to Brasília's construction and its location, Kubitschek (2000, pp. 13-14) adds: "Built on strategic point, the highways that serve it – realise, with perfection, a true sewing of Brazil by its interior, approximating states that while geographically share a border, they were living so distant

This chapter is structured as follows. Section 6.2 presents historical background to Brasilia's creation. Section 6.3 outlines the chapter's theoretical framework. Section 6.4 describes the data used and their sources. Section 6.5 presents and discusses the empirical results, and section 6.6 concludes.

from each other, as if they did belong to different countries" (see also Figure 6.4; Galanternick, 2003; El-Dahdah, 2003, p. 48).

6.2. Brasilia's Historical Background

Brasilia is the capital of Brazil. With a total area of 5,788 square kilometres, its 2010 urban area population was approximately 2.5 million inhabitants (IBGE, 2011). Unlike other Brazilian cities and most other cities in the world, the birth of Brasilia was artificial, i.e. it was not the result of a gradual development process. Instead, the city was created over a period of three years and ten months from the beginning of its construction in 1956 until its inauguration on 21st April 1960 (Kubitschek, 2000, p. 34). Furthermore, it was built within a forest area that lacked economic activities at that time.[114]

The idea of transferring the capital from Rio de Janeiro was as old as 1716, when the Marquis of Pombal, Portugal, suggested the change of the capital to the countryside due to security concerns arising from a possible invasion from other imperialist powers, particularly France. In 1821, José Bonifácio de Andrada e Silva, a Brazilian politician serving the Portuguese Empire, again raised the idea of moving the capital to the countryside. This idea was ever-present in Brazilian Constitutions from 1891 onwards (Farret, 2001; Santos, 2008) and was reinforced in the 1934 Constitution. There had been some earlier attempts to move the capital to Brazil's interior, but due to political instability and lack of leadership prior to Kubitschek's presidency, the initiative lacked progress.[115]

The first two capital cities in the history of Brazil, Salvador in Bahia and Rio de Janeiro city are located on the coast (see Figures 6.1 and 6.2). The main arguments to transfer the capital from Rio de Janeiro to Brasília included: avoiding Rio's colonial connotations; to have a more united country; and to achieve better planning with a more geographically centralised Federal Government. The site chosen for the capital's location (latitude; -15.8^{o} south; longitude -47.9^{o} west, approximately), which is in the country's Central Plateau[116] was considered appropriate to spread development from the country's coast to the

[114] There are other international examples of planned cities, including: Alexandria, St. Petersburg, Canberra, Washington DC, Ottawa, Pretoria, Ankara, and Islamabad. Brasília is unique in the sense that its location was in an empty area far away from the country's largest cities (see Kubitschek, 2000, p. 365; Skidmore, 2010, p. 143; Figure 6.4; http://www.infobrasilia.com.br/bsb_h2p.htm#Outras%20cidades%20planejadas).

[115] See the following websites: http://educacao.uol.com.br/biografias/ult1789u484.jhtm; http://citybrazil.uol.com.br/df/brasilia/historia-da-cidade; Brazil's Ministry of Foreign Affairs, www.mre.gov.br; http://citybrazil.uol.com.br/df/historia-do-estado; Distrito Federal (The Federal District), www.distritofederal.df.gov.br.

[116] See Brasilia and Distrito Federal in Figures 6.1, 6.2 and 6.3.

country's interior[117], and to take advantage of then unexploited economic potential and natural resources in the interior. It was considered this would lead to an increase in the domestic market due to integration of the centre-west and northern regions into the national economy (Epstein, 1973; Kubitschek, 1978; Couto, 2001; Santos, 2008; Oliveira, 2008).

Figure 6.1: Brazilian states and their capital cities

Source: http://www.dholmes.com/master-list/brasil.gif

[117] This is known as the marcha para o Oeste under Vargas Era, or 'the march to the West'. In a speech of August 7, 1940, in Goiânia, President Vargas states: "Your plateau is a viewpoint of Brazil. It is imperative to locate in geographical centre of the country powerful forces able to irradiate and guarantee country's future expansion" (Couto, 2001, p. 47).

Figure 6.2: States and borders of Brazil

Source: www.portalsaofrancisco.com.br/alfa/brasil-mapas/

Figure 6.3: Brasília and the surrounding regions

Source: www.portalsaofrancisco.com.br/alfa/brasil-mapas/

Although there are some examples of the state's role for development in Brazil before 1956, for example, during the Vargas (1930-1945; 1951-1954) and Dutra (1946-1951) presidencies, it was in Kubitschek's term (1956-1961) that the state took a prominent role. Over the latter period, in general, there was considerable nationalism amongst Brazilians. This nationalism helped Brazilians to find a common ground that reflected a belief that industrialisation (especially in base industries) was a means for national development, reflecting an argument that the lack of industrialisation was the main explanation for the country's underdevelopment. This ideology was named nacional-desenvolvimentismo (or, 'the national-developmentism'). Kubitschek's ability to reconcile opposing political views and his urgency in completing the construction of the new capital yielded industrial production growth of 80 percent (in constant prices) from 1956 to 1961 and a real annual growth rate of 7 percent from 1957 to 1961 (see Skidmore, 1967, p. 164; Rabelo, 2003, p. 54; Almeida, 2004, p. 10; Story, 2006, and references therein; Oliveira, 2006; Oliveira, 2008; Bethell, 2008). This compared with approximately 6.5 percent (also in constant prices in BR$ in 1980) from 1950 to 1955 (IPEA), 3.6 percent for 1962-1963, and 5.3 percent from 1964 to 1968 (Almeida, 2004)[118].

In the 1950s, two-thirds of Brazil's territory was still "empty". For Kubitschek, Brasília was a strategy for development policies and represented advancement from the agrarian past (Kubitschek, 2000; Story, 2006; http://cpdoc.fgv.br). As a result of his concern about reliance on the agrarian past, he advocated the building of Brasília during an electoral campaign in Jataí, Goiás, in April 1955. On April 18, 1956, he signed, in Anápolis, Goiás, a law for the capital's change to be sent to the Congress.[119] In the next year, he appointed the engineer Israel Pinheiro as President of Novacap (Portuguese acronym for 'new capital'; Novacap - Commissão Planejadora da Nova Capital do Brasil, or, 'the Planning Commission of the New Capital of Brazil'), a company created with the responsibility of building the new capital (Story, 2006; www.citybrazil.com.br).

There was debate about the underdevelopment of Brazil over the 1950s. Two features received special attention: industrialisation and stabilisation. Due to the country's inflationary background, there were concerns that a poorly prepared development plan would create inflationary problems. President Kubitschek used a political argument that, to unify and transform the country, the inflation problem should not be a concern. He focused on economic growth based on funding

[118] Until 2010, the real GDP growth rates observed under the Kubitschek period were surpassed only by the miracle period (1968-1973) with an annual growth rate of approximately 11 percent (or, the 1970s decade taken together, i. e 1971-80, with a real growth rate of 8.5 percent per annum) (Almeida, 2004, p. 25). The other three decades since 1980 have had a real annual growth rate less than 4 percent (calculation using IPEA database).

[119] The Congress approved initial funding on 19 September, 1956 (Story, 2006, p. xviii).

industrialisation at all cost without taking into account the consequences. This action left his successor two main problems: inflation and external debt. The former problem was a result of an increase in money supply to fund development projects while the latter was due to the fact that Kubitschek's Target Plan was funded by borrowed external capital (Rabelo, 2003; Almeida, 2004; Story, 2006).

The change of the capital has to be understood in a context in which the state is a major player in planning development. There is also prestige embedded in the change of capital as it symbolises national progress (National-developmentism) (Oliveira, 2006; Rabelo, 2003; Story, 2006). This attitude's change under Kubitschek's presidency is stated by Madaleno (1996, p. 273):

> So, from the beginning of the republic in 1889 up to the popular dictatorship of Vargas (1930-45), Brazilian constitutions and politicians endorsed bills to study, plan and build a new interior capital. Yet not a single president, and the dictator even less so, took the responsibility, nor the heavy burden, of making roads through the cerrado (savanna) of the remote highlands. Finally, in 1955, the newly elected president Juscelino Kubitschek faced up to the issue of the new capital and firmly supported the idea that Brazil could never be a modern and progressive country while centred only on the overpopulated littoral [coast]. Brazilians should seek interior resources, establish new settlements in the central and northern territories, and the southeastern polarized industrial capital should re-direct its emphasis in order to conquer a wider market and hence promote the integration of the whole country.

Although Kubitschek's term was successful in terms of meeting his goals, the criticism that his presidency lacked long-run sustainability holds when considering inflation and deficits on balance-of-payments that took place from the middle of his term. The crisis that emerged obligated him to arrange a stabilisation agreement with the International Monetary Fund (Skidmore, 2010). However, it is argued that his optimism, confidence, and willingness to modernise Brazil helped Brazilians understand that their country could compete with more developed nations in the world (Snyder, 1964, p. 34; Skidmore, 2000, p. 144; Kubitschek, 2000; Almeida, 2004).

The standard view in the 1950s was that growth implies development (Story, 2006; Madaleno, 1996). This is clear in one of Kubitschek's statements (Dedina-Wagner, 1961, p. 7): "With the new capital it will come, (…), a time of abundance and a genuine fraternity that will permit indistinctly to all Brazilians the enjoyment of goods of culture and progress". Story adds (2006, p. 35): "What united the planners and supporters of Brasília—politically diverse as they were—was not just a faith in the potential of Brazil, but a belief in progress. They shared a worldview of linear development, the notion that nations and peoples passed through various necessary stages of evolution". Brazil's old colonial problems of great interpersonal and interregional disparities, with their social consequences[120], are still present (IBGE, 2006, pp. 11-24)[121].

To modernise the country, Kubitschek's administration elaborated the Plano de Metas (or, 'the Target Plan', or 'the Goals Plan') with 30 goals. To this plan, the construction of Brasília was added as the 31st goal (also called 'synthesis-goal') (Kubitschek, 2000, p. 447; Oliveira, 2009). This plan focused on two key sectors; energy and highways' infrastructure[122]. The radial highways starting in Brasilia were constructed and inaugurated at the same time as the construction of Brasilia (essentially in 1958 and 1959). Most of the other highways, such as longitudinal, transverse, diagonal and connecting highways, were also inaugurated between 1957 and 1960 (see the main highways network related to Brasilia in Figure 6.4).

[120] Analysing Brazil's economy over the twentieth-century, Marcelo de Paiva Abreu (IBGE, 2006, p. 356) concludes that, "(…) overall, despite its structural transformation over the last century, Brazil's relative position at the global level has not been changed".
[121] Some institutions were created aiming to reduce regional disparities. For example, in 1959, SUDENE – the development agency for the northeast region – was created.
[122] Some criticisms in the literature are that the Target Plan was less concerned with other important sectors such as health and education (Story, 2006; Rabelo, 2003). For instance, Rabelo (2003, p. 49, Table 1), shows that the investment target for energy, transportation and basic industry was of 42.4, 28.9, and 22.3 percent of the total investment target, respectively, while for education, the target was just 2.84 percent of the total investment target.

Figure 6.4: Infrastructure network related to Brasilia[123]

Source: Story (2006, p. 37).

This development created conditions for further infrastructure development that took place during the military rule (especially for connecting highways, from 1964 to 1974), under the following presidencies: Castelo Branco (1964-1967); Costa e Silva (1967-1969); Garrastazu Médici (1969-1974); and Ernesto Geisel (1974)[124]. As Story (2006, p. xxxix) notes:

> (…) the military dictatorship decided to follow the path to development and vision of modernity embodied in Brasília. The

[123] Distances from Brasilia to the other cities are in kilometres.
[124] It is worth noting that some highways, especially the connecting ones, were inaugurated in 1951, after a return of Getúlio Vargas to power (1951-1954).

new regime did not significantly alter Kubitschek's recipe for progress, but rather implemented policies consistent with the precedent he had so firmly implanted. Brasília provided the map and lexicon for Brazil's subsequent attempts to achieve development and modernity. It put forth a vision of modernity that was carefully planned and coordinated (…)".

Discussing development of Brazil's urbanism in the years after Brasília's creation, del Rio (2009a, p. 18) adds:

> Ambitious plans, large projects, and technocratic modernism were the trademarks of Brazilian urbanism and its large supporting bureaucratic and financial apparatus created by the military. Particularly during the late 1960s and the 1970s, as Segawa (1988) notes, the euphoria of Brazilian 'economic miracle', together with the 'planning syndrome' of the military regime, encouraged extensive public works and the relocation of investment and the population into the interior of the country. Until the early 1980s, public works endeavours throughout Brazil typically reflected this approach to urbanism: urban renewal plans; construction of highways and viaducts; eviction of *favelados*; and construction of low-income residential projects, new city and state administrative centers, university campuses, airports, bus and train terminals, hydroelectric power stations, and new independent and company towns.

The creation of Brasília has had implications for São Paulo and Rio de Janeiro metropolitan areas. It led to a fall of public sector jobs, more in the latter metropolitan area than in the former. Manufacturing has increased in both São Paulo and Rio de Janeiro metropolitan areas due to the attraction of international automobile companies such as Ford, Volkswagen, Mercedes-Benz, Vemag-DKW, Willys-Overland, and General Motors, in line with the industrialisation policy (Kubitschek, 2000). There has also been a change in population distribution and migration patterns in Brazil. People from the north and northeast migrated to these two traditional cities and to the growing and emerging cities in the interior looking for employment opportunities [125] (Skidmore, 1967; Matos, 2000; Skidmore, 2010; www.macalester.edu). As del Rio (2009*b*, p. 8) states:

> Vargas's, and later Kubitschek's, drive for industrialisation was pivotal in population explosion of Brazil's major cities in the 1940s and 1950s. Abreu (1987) noted that Rio de Janeiro's

[125] See also http://www.suapesquisa.com/historiadobrasil/governo_jk.htm

> metropolitan area grew by 103 percent between 1950 and 1960 and that in 1960, 53 percent of its population consisted of migrants – half of whom lived in areas just outside Rio, most in illegal land subdivisions – those without land titles, built without official approval, or without the infrastructure the developer was supposed to provide. The situation was similar in the São Paulo metropolitan area, where Meyer, Grostein, and Biderman (2004) registered a growth rate of 6.17 percent between 1950 and 1960, compared to the growth rate of 3.04 percent for Brazil as a whole in the same period.

This finding is also pointed out by del Rio (2009b, pp. 7-8):

> The ambitious economic, industrial, and national development programs set forth in 1956 by President Juscelino Kubitschek – who as a mayor of Belo Horizonte had built the Pampulha modernist complex – included the construction of a new national capital not only as a progressive national symbol but also as means to bring development into the country's interior. The territorial logic was to develop and modernize Brazil's heartland and populate the middle of the country. Kubitschek's agenda was pivotal in consolidating Brazilian modernism. His economic strategy was successful in promoting overall optimism and encouraging industrialization, but it also sped up migration urban areas and urbanisation (Skidmore 1999; MacLachlan 2003).

In short, the creation of Brasília, industrialisation, urbanisation, infrastructure and manufacturing development are hallmarks of Kubitschek's government. It all helped to modernise the country, differentiating itself from its agrarian past and putting it in a new developmental stage. The developments changed Brazil's landscape and population distribution during the period 1956-1961 and in the decades that followed. [126]

What is less clear-cut, however, is whether the creation of Brasilia fundamentally affected the spatial distribution of economic activity and population beyond its immediate environs. This chapter tests whether the creation of the new capital supplemented or offset the standard economic forces of regional income convergence and agglomeration that were already at work and which the creation of the new capital was designed to modify.

[126] However, there are remained drawbacks associated with unaffected patterns of income concentration. Kubitschek's government was less concerned with the problem of income concentration in Brazil. Consequently, Brazil's Gini index of inequality was very high in 1960 at 0.57 (Bértola, 2002, p. 10).

6.3. Theoretical Framework

This section has three parts. The first part briefly outlines the modern economic geography literature employed as the theoretical framework to explain the effect of the creation of Brasilia in Brazil. The second part briefly reviews the effect of very large scale regional development policies in other countries. Finally, the third part presents the methods applied and briefly discusses why they are adequate for this study.

6.3.1. Regional development policy *versus* agglomeration forces

This chapter tests two competing hypotheses. The first is that the major regional development policy characterised by the creation of Brasilia (and its accompanying highways) had a significant impact on the spatial development pattern of Brazil. The competing hypothesis is that the spatial development pattern is predominantly driven by the increasing importance of agglomeration forces centred on the two major cities, São Paulo and Rio de Janeiro, with at most a small and temporary impact of Brasilia. The question is therefore whether the creation of Brasilia reduced (or even offset) the agglomeration forces that made São Paulo and Rio de Janeiro the leading urban economies of Brazil. Before testing these hypotheses empirically, some of the key studies and evidence about the agglomeration forces in economic geography literature are briefly reviewed in this section.

The economic geography literature is very broad but its main aim is to understand how and why most economic activities ultimately end up only in certain locations. In other words, the economic geography literature asks why firms and people choose to agglomerate in particular locations. In order to answer this question, this subsection focuses on six characteristics of the regional economy that underpin firms' and people's location decisions.[127]

The first characteristic is firm productivity. Firms locate where their productivity is high. This often means location in regions with high population density (Ciccone & Hall, 1996). Such regions are also preferred by workers due to wage premia. High levels of productivity are associated with high technology, innovation, entrepreneurship, and skilled labour. Together, these four elements

[127] The choice of these characteristics is based on their appearance in empirical studies of the economic geography literature about the effects of those characteristics on agglomeration of economic activities and growth of regions.

make firms produce goods and services with high value added (McCann, 2009a, 2009b). The production of most high value-added goods and services takes place in cities (or urban areas). Some of the characteristics of leading cities' economic activities are that they are knowledge-intensive, attract university graduates (or otherwise have high human capital) (McCann, 2009a, 2009b), and provide a high level of amenities for their residents that outweigh the disutility associated with congestion and pollution (Glaeser, Kolko & Saiz, 2001).

The second characteristic is population size. With high transport costs, specialised firms prefer to produce in larger locations or where other firms have already chosen to produce in order to take advantage of a specialised labour pool (Glaeser *et al.*, 1992; Henderson, Kuncoro & Turner, 1995; Combes, 2000; Crawford, 2006; De Groot *et al*., 2009). These chosen locations tend to be near the market for the final products and suppliers (the so-called backward and forward linkages) (Krugman, 1991a, 1991b, 1994). In these places workers are more likely to find new jobs and a greater range of jobs. Regions with low transport costs (often rural regions or smaller towns or cities) tend to specialise in production of low value-added goods and services that are exported to larger and diverse markets (see Skilling, 2001; McCann, 2008, p. 362; McCann, 2009b, slide 45).

The third characteristic is population density. Densely populated areas are associated with high market potential (or, high home market effect). This leads to high firm sales and profits for certain classes of firm (particularly knowledge-intensive firms). Such firms, aiming to maximise profits, therefore tend to locate in regions with high population density, which for that reason also offer more jobs for local workers (Ciccone & Hall, 1996; Mori, Nishikimi & Smith, 2005, 2008).

The fourth characteristic is geographical location. High value-added firms prefer to produce in the core regions rather than on the periphery. This not only facilitates transactions with local markets but also helps trade because the effective distance for shipments of goods to other regions is relatively small (Disdier & Head, 2008) resulting in relatively lower transport costs. Additionally, the core is usually more closely tied to the global economy than are peripheral regions, and so peripheral regions may not fully converge to the core regions in terms of productivity and incomes per head. As McCann (2008, p. 363) states:

> The differences between the major centres and the relatively smaller centres will tend to increase. This is not to say that all

> major centres will increase relative to smaller centres, as it will also depend on the range of technologies and industries evident in particular cities, as different industries and technologies rise and fall over time. However, the arguments outlined earlier still do imply that globalization will lead to increasing differences between the fortunes of regions and cities even within the same country. **This is because particular major urban centres will benefit from the increasing scale advantages associated with being nodes in global trade networks** (bolds added).

Peripherality is one of the most important locational problems. The problem of peripherality has well been explored in the economic geography literature using the case of New Zealand (Poot, 2004; Crawford, 2006; Disdier & Head, 2008; McCann, 2003, 2008, 2009a, 2009b) and also of Finland (Poot, 2004). This problem is worsened in the current era of globalisation in which important costs of trade, the spatial transaction costs, have increased (McCann, 2008, 2009a, 2009b). The creation of Brasilia can be considered as a case of creating a major city in a peripheral region of Brazil, contrasting with the existence of the two major cities in Brazil's core, São Paulo and Rio de Janeiro.

The fifth characteristic is economic diversity (or the industry-mix effect). A competitive and diverse regional economy is more likely to prosper because there is a much higher speed of knowledge diffusion across firms and workers (Jacobs, 1969). If that particular regional economy is a core region, it is more likely to attract "footloose" manufacturing (Disdier & Head, 2008).

Finally, the sixth characteristic is the growth rate itself. The outcome from the first five characteristics will be a self-reinforcing growth of major agglomerations. The regions with high growth rates are more attractive for new firms and residents (immigrants). The long-run rate of growth, however, will depend on whether the regions meet key requirements for growth such as attraction and maintenance of highly skilled workers (through R&D activities and large-scale knowledge production and diffusion, for instance).

Together, the six characteristics of economic activity at the core, be it within a country or globally, result in agglomeration of economic activities in that region at the expense of other regions. Thus the agglomeration forces themselves lead to circular causation among the outlined characteristics and concentration of economic activities and income in some regions (Krugman, 1991a, 1991b, 1995; Brakman, Garretsen & van Marrewijk, 2001; Fujita & Krugman, 2004; Capello &

Nijkamp, 2009). Broadly defined economies of scale favour firms' profitability and workers' wages, and people enjoy higher living standards because of access to a high variety of goods and services. However, these benefits are counter-balanced by high rents and higher cost of living relative to other regions as land prices are bit up by firms and residents. Availability of land for city expansion and the efficiency of transport networks will therefore affect the patterns of agglomeration. Given both the agglomeration forces and these offsetting effects, only a few regions will achieve major scale (Krugman & Venables, 1995; McCann, 2008). The empirical task in this chapter is therefore to test whether the creation of Brasilia modified these forces of agglomeration in terms of the development of cities located different distances from Brasilia, São Paulo and Rio de Janeiro.

6.3.2. Regional development policy in other countries

Two types of regional development policies in other countries are relevant benchmarks for the Brasilia case study. The first type is the explicit creation of a new capital city. Broadly speaking, the arguments for the creation of a new capital city tend to be related to a country's military safety, national unity, national identity, and national integration[128] (Robinson, 1973; Tranter, 1990; Bowling, 1991; Kironde, 1993). Studies that deal with the creation of a new national capital city can have different focii. One is the architectural or environmental issues of the new capital city[129] (Robinson, 1973; Tranter, 1990; Moser, 2010). Such studies find that it is inadequate during the building of the Master Plan to focus on architectural and environmental issues only, thereby ignoring economic issues. In practice, over time, economic aspects dominate and lead to spatial segregation in the new city.

The second focus is the development of the new capital itself, irrespective of whether the new capital already existed as an urban area before it gained the status of national capital (Lovejoy, 1985; Kironde, 1993). Such studies discuss the requirements for a successful transfer of capital city, such as adequate resources,

[128] National integration is especially related to the infrastructure network (highways) as the location chosen for the new capital city is usually towards the central area of the country. This is true, for instance, for the new capital cities of Canberra (Australia), Lilongue (Malawi), Yamoussoukro (Ivory Coast), Abuja (Nigeria), Dodoma (Tanzania), Islamabad (Pakistan), Putrajaya (Malaysia), and Washington DC in the USA (compared to previous national capital cities) (Lovejoy, 1985; Bowling, 1991; Kironde, 1993; Moser, 2010). The first eight of these cities were created during the 20th century while Washington DC was built in the late 18th century.

[129] Also known as garden city studies, such as those for Canberra and Putrajaya cities.

master planning, availability of international skilled labour to train locals (Lovejoy, 1985), a strong personal drive for the new capital from key national political leaders, and support more broadly from government, a willingness of government officials to move, and a reduction in economic constraints (Kironde, 1993). Given these preconditions, studies find an increase in the rank of the new capital in the city size distribution. Using the case of Dodoma (Tanzania), Kironde (1993, pp. 443-444) points out:

> Dodoma soon became a major junction of the north-south and east-west road network. In 1948, it had just over 9000 people. By 1967, this had grown to over 23,000 but Dodoma's most dramatic growth was between 1978 and 1988 when the population grew nearly fivefold, and when the town jumped from the twelfth to the third rank in the national urban hierarchy… This is *prima facie* evidence that the capital transfer project has had a strong impact on Dodoma although the government has not moved, for, while it is admitted that Dodoma would have grown without being the earmarked capital, it is unlikely that it would have grown at such a high rate.

The second type of regional development policy that is relevant for considering the impact of Brasilia is a major investment project (in agriculture, manufacturing, and commerce) that focuses on a specific region, such as the Tennessee Valley Authority (TVA) scheme. The TVA project was launched in May 1933 by President Franklin Roosevelt as one of his measures to ameliorate effects of the Great Depression (Boyce, 2004). Garrison (1974) employs comprehensive shift share and entropy (a measure of industrial concentration) methods to analyse employment and earning changes by industry at the county level from 1959 to 1968 in the TVA Region. From the shift share analysis, Garrison (1974, pp. 52-53) finds that:

> most employment growth in the rural and small-town counties consisted of competitive gains (74.2 thousand of total growth of 90.8 thousand jobs for the two groups combined), while much of the growth in the urban and metropolitan areas may be attributed to national growth and industry mix forces affecting those industries already located there at the beginning of the study period. Most growth in the small-city group, 42.1 of 56.3 thousand jobs, also is attributed to competitive gains.

On the other hand, from the entropy method, Garrison (1974, p. 55) finds that,

> the increase in *f* [concentration indicator] values *within* the rural and small-town sets indicates dispersal of employment over time

> to counties which had relatively little employment at the beginning of the period, while the virtually constant numbers-equivalents within the small-city and urban sets indicates no such dispersal.

It should be noted that both types of studies described above (i.e. those about the creation of a new capital city and those relating to a major regional investment project) tend to be merely descriptive. They do not apply structural regression models to evaluate a general impact of a particular new capital city or investment project on the patterns of development across the national economy. It is the latter approach that will be adopted below to measure the spatial development impact of Brasilia.

6.3.3. Methods

The literature review of subsections 6.3.1 and 6.3.2 above leads to the hypothesis that agglomeration forces centred on São Paulo and Rio de Janeiro are likely to be important in determining the spatial pattern of development across Brazil, *but* may be moderated by the effect of a large scale regional development project, the creation of Brasilia. This will be tested below. In order to test these competing hypotheses (i.e. regional development policy *versus* agglomeration forces theory), three sets of models are proposed to help understand the extent to which Brasilia city affects the pattern of regional growth in Brazil. In particular, this chapter aims to test whether the creation of Brasilia affected development patterns of cities other than São Paulo, Rio de Janeiro and Brasilia itself.[130]

6.3.3.1. Cross-section Levels Models

The first equation employed is equation (6.1). The aim is to test whether the level of city GDP, after controlling for its pre-war (1939) GDP level, is affected by its proximity to each of the three cities. This cross-sectional relationship is tested over multiple years to examine whether there are changes over time in the effects of proximity to the three cities on the GDP of other urban areas. The method used for estimation of equations across the cross-sections is Seemingly Unrelated Regression (SUR). This method allows for estimation of parameters of a system

[130] Brasilia, Rio de Janeiro and São Paulo were therefore excluded from the sample of cities when estimating all models in this chapter. Definitions of variables are in Appendix 6.1.

of equations when the residuals are assumed to be correlated across equations (Zellner, 1962).

$$\mathrm{gdp}_{r,t} = \beta_0 + \beta_1 \mathrm{gdp1939}_r + \beta_2 \mathrm{dist\_b} + \beta_3 \mathrm{dist\_r} + \beta_4 \mathrm{dist\_s} + \text{error term}_{r,t} \quad (6.1)$$

where $\mathrm{gdp}_{r,t}$ is the real GDP of urban area (region) r in period t; $\mathrm{gdp1939}_r$ is region r's GDP in 1939 and dist_b, dist_r and dist_s are the distances (in kilometres) of the urban area to Brasilia, Rio de Janeiro and São Paulo respectively.[131]

In addition to estimating equation (6.1) for the level of GDP, the same approach is applied to estimating the impact on GDP per capita and on population, with identical functional forms. Specification (6.1) provides a descriptive approach with no restrictions on the parameters. The models from equation (6.1) are a precursor to the growth models that test a particular relationship derived from neoclassical growth theory. The reason for adopting both approaches is to check robustness of the results. From a methodological perspective, this chapter puts most emphasis on the growth models which are derived from a standard theoretical model of convergence (Barro & Sala-i-Martin, 1992, 2004; Ozgen, Nijkamp & Poot, 2010). This is specified below.

6.3.3.2. Cross-section Growth Models

The second approach tests the effect of Brasilia on per capita income growth of regions. The following econometric specification is commonly used in the literature (Ozgen, Nijkamp & Poot, 2010) to measure the impact of an exogenous shock on economic growth and convergence:

$$(1/T).\log(y_{r,t}/y_{r,t-T}) = \alpha - [(1 - e^{-\beta T})/T].[\log(y_{r,t-T})] + \gamma S_{r,t} + \text{other variables+error term} \quad (6.2)$$

where the dependent variable is the average annual growth rate of per capita income between t-T and t; $y_{r,t}$ is the per capita income in region r in the 12 month period ending at date t; T is the number of years spanned by the data; β is the

[131] Non-linear (quadratic and cubic) distance terms were also tested, but these were not significant in preliminary regressions and so are excluded henceforth. The recorded distances are those at present. Due to data limitations, the effect of *changes* in distance from regions to Brasilia, Rio or São Paulo as a result of changes to the highway network could not be evaluated.

annual rate at which a region's economy converges to a long-run steady state, and γ is the coefficient of the region-specific shock $S_{r,t}$. Virtually all studies of beta income convergence (so-named, because these studies aim to estimate β in equation (6.2)) adopt specification (6.2) or its linearised equivalent.

The emergence of Brasilia matters because of the potential for spatial interaction. Let us assume that the population of Brasilia is driven by exogenous political factors. The spatial interaction effect between Brasilia and region r at time t ($M_{rb,t}$) can be represented as:

$$M_{rb,t} = \chi(P_{r,t})^{\delta}(P_{b,t})^{\varepsilon}/(dist\_b)^{\varphi} \tag{6.3}$$

where the subscript b refers to Brasilia, $P_{r,t}$ is the population of region r at time t; and $dist\_b$ is the distance between the capital of region r and Brasilia, as before. The population of Brasilia is assumed to be at least 1, even in the earliest period (so that logs can be taken). The population of region r is endogenous at time t, but $P_{r,t}$ can be instrumented with $P_{r,t-T}$. The shock variable ($S_{r,t}$) in equation (6.2) is replaced by log $M_{rb,t}$ and, for simplicity, the beta convergence relationship is linearised to get:

$$(1/T).\log(y_{r,t}/y_{r,t-T}) = \text{constant} - \beta\log(y_{r,t-T}) + \delta\log(IP_{r,t}) + \varepsilon\log(P_{b,t}) - \varphi\log(dist\_b) + \text{other variables} + \text{error term} \tag{6.4}$$

where $IP_{r,t}$ is the predicted value of $P_{r,t}$ from the first stage regression of $P_{r,t}$ on $P_{r,t-T}$. Among the "other variables", the spatial interactions with Rio de Janeiro and São Paulo are assumed to matter too, so the full specification becomes:

$$(1/T).\log(y_{r,t}/y_{r,t-T}) = \text{constant} - \beta\log(y_{r,t-T}) + \delta\log(IP_{r,t}) + \varepsilon\log(P_{b,t}) - \varphi\log(dist\_b) - \xi\log(dist\_r) - \zeta\log(dist\_s) + \text{error term} \tag{6.5}$$

This model can be estimated period-by-period jointly for every $t > t$-T by means of the SUR estimator (and again the cross-sectional r refers to all regions excluding Brasilia, Rio de Janeiro and São Paulo). As with the first approach, parameter stability over time can be investigated. The impact of Brasilia on regional growth is measured by the (possibly time-varying) parameters ε and φ. The results from specification (6.5) are this chapter's main focus, because they allow for evaluation of the impact of Brasilia's creation relative to the power of agglomeration forces associated with São Paulo and Rio de Janeiro within a standard neoclassical growth model of income convergence (Barro & Sala-i-Martin, 1992, 2004; Ozgen, Nijkamp & Poot, 2010). If Brasilia has a positive impact on per capita incomes of cities that are closer to it, the coefficient on log(*dist_b*) is expected to be negative (i.e. a short distance to Brasilia results in higher income growth for a city). Similarly, if agglomeration forces associated with Rio de Janeiro or São Paulo are operative, the coefficients on log(*dist_r*) and log(*dist_s*) are expected to be negative.

6.3.3.3. Cross-section Spatial Growth Models

In testing the validity of the competing hypotheses (regional development policy *versus* agglomeration forces theory), it is important to test for spatial autocorrelation or spatial lags in the estimation strategy. With the introduction of spatial autocorrelation, specification (6.5) becomes:

$$(1/T).\log(y_{r,t}/y_{r,t-T}) = \rho \textstyle\sum_s w_{rs}\ [(1/T).\log(y_{s,t}/y_{s,t-T})] + \mathbf{X}\boldsymbol{\beta} + \text{error term} \qquad (6.6.1)$$

Equation (6.6.1) represents a spatial lag model. $\sum_s w_{rs}\ [(1/T).\log(y_{r,t}/y_{r,t-T})]$ is the spatially lagged dependent variable, ρ is the spatial autoregressive parameter, while **X** is a vector of the observations in region r on other explanatory variables from equation (6.5) which is multiplied by the correspondent vector of parameters, $\boldsymbol{\beta}$. Equation (6.5) can be extended alternatively to become:

$$(1/T).\log(y_{r,t}/y_{r,t-T}) = \mathbf{X}\boldsymbol{\beta} + \varepsilon \qquad (6.6.2).$$

where: $\varepsilon = \lambda \mathbf{W}\varepsilon + \mu$

Equation (6.6.2) represents a spatial error model. λ is the spatial autocorrelation parameter, μ is a vector of errors that are assumed to be independently and identically distributed, W is the matrix of spatial weights w_{rs} and the other variables and parameters are defined as in equation (6.6.1).

The spatial weights w_{rs} that reflect the interactions between regions are based on the inverse distance ($1/d_{rs}$), in kilometres, between a pair of regions r and s. In these models, the impact of the regions' interactions is captured by the ρ and λ parameters, respectively. The spatial weights matrix is row standardised and is symmetric; its main diagonal has zeroes as intra-regional interaction is ignored (LeSage & Pace, 2009).

6.4. Data and Sources

This chapter uses socio-economic data from two official sources in Brazil: IPEA - Institute of Applied Economic Research (www.ipea.gov.br) and IBGE – Brazilian Institute of Geography and Statistics (www.ibge.gov.br). These sources present a variety of socio-economic data, mostly at the state level, collected from public and private institutions about the Brazilian economy.

The geographical unit of analysis (region r) is the urban area. These were already defined in section 4.3 of Chapter 4. The initial sample size is 185 urban areas. Municipality data were used to aggregate up to the variables of interest. From IPEA and IBGE, data were obtained for the following two socio-economic variables. The first variable is total GDP in thousands of 2000 BR$. These data are essentially available on a decadal basis from 1939 to 1996, and annually from 1999 to 2008 (see Appendix 6.2).

The second variable is population or number of inhabitants. Ideally, the data for this variable would be compiled for the same years for which GDP data are available at municipality level; however, there are some years of missing population data. In those cases, population data were interpolated between neighbouring years (see, in Chapter 4 above, Appendix 4.1). Per capita GDP was formed by taking the ratio between the two variables.

Distance data are obtained from each urban area to the following three cities: Brasília, Rio de Janeiro and São Paulo. A spatial inverse distance weights matrix with 174 urban areas has also been formed. These distance data were obtained from Brazil's Transportation Ministry using Brazil's highways, and from Aondefica (Guide of distances between Brazil's cities) websites.[132]

[132] These websites are, respectively: http://www.transportes.gov.br/bit/inrodo.htm; http://www.estradas.com.br/new/header_sites/mapas.asp, and http://www.aondefica.com/.

6.5. Empirical Results and Discussion

6.5.1. Cross-section Levels Models

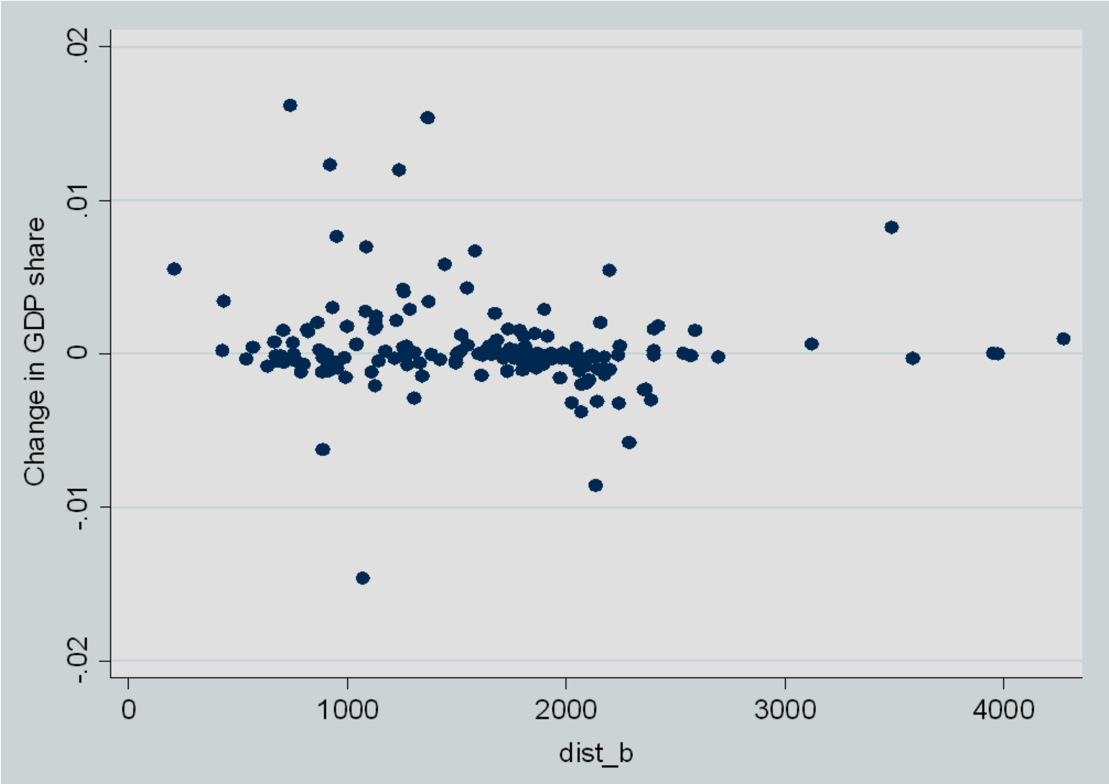
This subsection presents the results from equation (6.1). The results in Table 6.1 show that the initial conditions play an important role in determining GDP levels over time (Krugman, 1991a, 1991b). In every year, the level of GDP is positively correlated with its level in 1939. None of the parameters for distance is significant implying that proximity to any of the three cities - Brasilia, Rio de Janeiro and São Paulo - had no effect on regions' GDP levels. This finding is supported (in relation to Brasilia) by the pattern in Figure 6.5. This shows the change in GDP share for regions from 1939 to 2008. There is no apparent relationship between regions' GDP shares and their proximity to Brasilia.

Table 6.1: Cross-section Levels Models for GDP

	Constant	gdp1939	dist_b	dist_r	dist_s	Number of observations	R-Squared
1949	81609	1.3870***	-32.9958	-28.4221	5.5384	174	0.8987
	(1.24)	(37.75)	(-0.29)	(-1.08)	(0.27)		
1959	61680	3.1841***	60.0301	-55.2019	39.2006	174	0.9192
	(0.47)	(43.15)	(0.26)	(-1.04)	(0.96)		
1970	326800	6.2878***	-229.7687	-17.5688	11.5339	174	0.7969
	(0.74)	(25.30)	(-0.30)	(-0.10)	(0.08)		
1975	447647	11.4404***	166.3022	250.8030	-254.4666	174	0.7175
	(0.44)	(20.14)	(0.09)	(0.62)	(-0.81)		
1980	883067	17.9905***	-55.5763	330.7112	-222.3956	174	0.7025
	(0.54)	(19.54)	(-0.02)	(0.50)	(-0.43)		
1985	1362942	18.6287***	-148.5116	298.0742	-144.9836	174	0.6618
	(0.73)	(17.76)	(-0.05)	(0.40)	(-0.25)		
1996	2699707	22.2034***	-2798.6000	277.1090	-38.2558	174	0.6216
	(1.11)	(16.37)	(-0.66)	(0.29)	(-0.05)		
1999	2655332	21.0469***	-1351.8170	343.2896	-285.1544	174	0.5958
	(1.08)	(15.29)	(-0.31)	(0.35)	(-0.37)		
2000	2835591	22.0184***	-1148.9200	317.8972	-325.7367	174	0.5720
	(1.05)	(14.50)	(-0.24)	(0.29)	(-0.39)		
2001	2885838	22.2623***	-1356.2430	293.1258	-302.4045	174	0.5774
	(1.07)	(14.69)	(-0.29)	(0.27)	(-0.36)		
2002	3117183	23.4678***	-1692.4960	223.6048	-210.8496	174	0.6032
	(1.15)	(15.53)	(-0.36)	(0.21)	(-0.25)		
2003	2932727	23.0149***	-1239.9930	357.9345	-359.2086	174	0.5900
	(1.08)	(15.07)	(-0.26)	(0.33)	(-0.42)		
2004	3398735	24.2203***	-1528.2510	160.6457	-221.1182	174	0.5669
	(1.13)	(14.35)	(-0.29)	(0.13)	(-0.24)		
2005	3194668	25.8707***	-989.5803	-93.0791	48.1797	174	0.5826
	(1.03)	(14.87)	(-0.18)	(-0.07)	(0.05)		
2006	3954819	26.7364***	-1940.0180	-408.8236	272.5428	174	0.5739
	(1.21)	(14.58)	(-0.34)	(-0.31)	(0.27)		
2007	4033675	28.4285***	-1846.4090	-73.3931	23.1194	174	0.5708
	(1.15)	(14.48)	(-0.30)	(-0.05)	(0.02)		
2008	4078597	29.8128***	-1477.4430	-306.2765	173.8492	174	0.5679
	(1.10)	(14.38)	(-0.23)	(-0.21)	(0.15)		

Notes: Equations are estimated using SUR. Significance levels: *=10%, **=5%, and ***=1%. Values of z statistics are in brackets.

Figure 6.5: Change in GDP share, 1939 to 2008



Notes: dist_b is distance to Brasília in kilometres; GDP is in thousands of BR$ in 2000.

In addition to the tests for GDP, models for GDP per capita and population are also considered.[133] Table 6.2 shows the results for GDP per capita. As it can be seen, when considering a cut-off of 10 percent significance level, the initial conditions which are measured by GDP per capita in 1939 (gdp_pc1939) are important through to 2003. However, after 2002, none of the initial conditions is significant at the 5 percent level and three of the six parameters are not even significant at the 10 percent level. This suggests that while there is long-lasting persistence in regional GDP per capita levels (Krugman, 1991a, 1991b), the influence of initial conditions on GDP per capita does dissipate over time, possibly due to convergence forces (as examined explicitly in the next formulation). The parameters for distance to Brasilia and to Rio de Janeiro were not significant in any year. However the coefficient of São Paulo's distance is negative and significant at the 5 percent level for every year other than 1959. This result indicates that regions nearest São Paulo gained in terms of GDP per capita (relative to their 1939 level) compared to those cities further from São Paulo, consistent with positive agglomeration impacts emanating from the dominant

[133] For descriptive statistics of the models for the three dependent variables, see Appendixes 6.2-6.3. The estimation method for these models is again Seemingly Unrelated Regression (SUR).

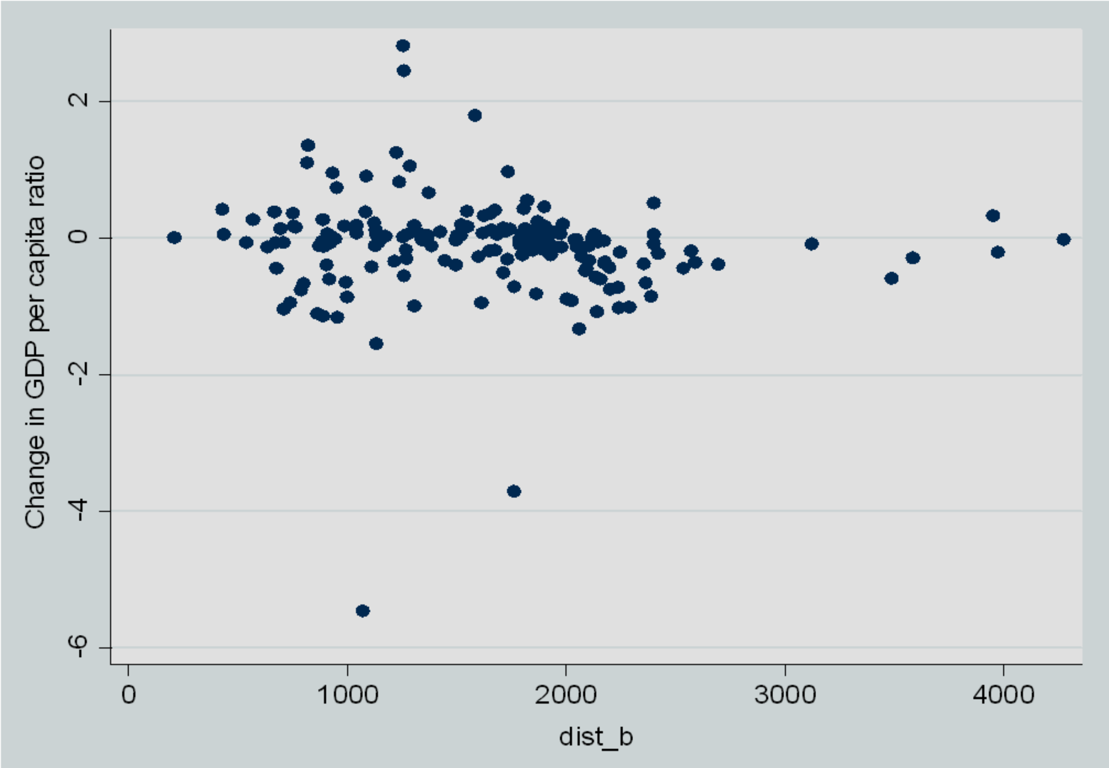
regional economy of Brazil. Conversely, the ‘Brasília shock’ had no effect on the spatial pattern of per capita incomes in the Brazilian economy, as shown graphically in Figure 6.6.

Table 6.2: Cross-section Levels Models for GDP per capita

	Constant	gdp_pc1939	dist_b	dist_r	dist_s	Number of observations	R-Squared
1949	1.4506***	0.6413***	-0.0004	0.0002	-0.0005***	174	0.6578
	(3.56)	(10.84)	(-0.54)	(1.14)	(-3.84)		
1959	1.7040**	1.0729***	-0.0004	-0.0001	-0.0003	174	0.5571
	(2.35)	(10.20)	(-0.36)	(-0.45)	(-1.24)		
1970	3.9228***	0.7603***	-0.0018	0.0004	-0.0014***	174	0.5811
	(4.59)	(6.14)	(-1.20)	(1.23)	(-4.97)		
1975	5.2482***	1.0838***	-0.0013	0.0009	-0.0026***	174	0.5513
	(3.55)	(5.05)	(-0.50)	(1.56)	(-5.37)		
1980	7.0368***	1.2757***	-0.0015	0.0013	-0.0034***	174	0.5127
	(3.58)	(4.47)	(-0.44)	(1.69)	(-5.21)		
1985	6.9948***	1.4687***	-0.0010	0.0008	-0.0030***	174	0.5466
	(3.74)	(5.41)	(-0.31)	(1.11)	(-4.77)		
1996	8.3192***	0.7329***	-0.0037	0.0008	-0.0028***	174	0.4872
	(4.81)	(2.92)	(-1.24)	(1.09)	(-4.84)		
1999	8.6982***	0.5326**	-0.0021	0.0007	-0.0030***	174	0.5382
	(5.27)	(2.22)	(-0.74)	(1.06)	(-5.47)		
2000	8.6316***	0.5260*	-0.0008	0.0004	-0.0030***	174	0.4885
	(4.47)	(1.88)	(-0.25)	(0.54)	(-4.60)		
2001	8.8501***	0.4390*	-0.0018	0.0005	-0.0031***	174	0.5062
	(4.87)	(1.67)	(-0.58)	(0.67)	(-5.05)		
2002	9.3481***	0.5658**	-0.0025	0.0002	-0.0028***	174	0.4914
	(4.88)	(2.04)	(-0.75)	(0.19)	(-4.37)		
2003	9.1138***	0.4625*	-0.0019	0.0005	-0.0032***	174	0.4951
	(4.75)	(1.66)	(-0.59)	(0.65)	(-4.97)		
2004	10.2821***	0.3576	-0.0030	0.0000	-0.0029***	174	0.4629
	(4.94)	(1.19)	(-0.84)	(0.04)	(-4.19)		
2005	10.1065***	0.4965	-0.0022	-0.0008	-0.0021***	174	0.3990
	(4.29)	(1.45)	(-0.54)	(-0.86)	(-2.64)		
2006	10.4587***	0.5378	-0.0022	-0.0014	-0.0017**	174	0.3638
	(3.99)	(1.41)	(-0.50)	(-1.34)	(-1.97)		
2007	10.5269***	0.6207*	-0.0025	-0.0007	-0.0023***	174	0.4185
	(4.40)	(1.79)	(-0.59)	(-0.75)	(-2.89)		
2008	10.9793***	0.6547*	-0.0025	-0.0011	-0.0022**	174	0.3766
	(4.01)	(1.65)	(-0.53)	(-1.01)	(-2.38)		

Notes: Equations are estimated using SUR. Significance levels: *=10%, **=5%, and ***=1%. Values of z statistics are in brackets.

Figure 6.6: Change in GDP per capita ratio, 1939 to 2008



Notes: dist_b is distance to Brasília in kilometres; GDP per capita is in thousands of BR$ in 2000.

Rio de Janeiro and São Paulo each had a prominent role on population developments in their nearest regions (Table 6.3). Since 1959, Rio's distance coefficient is positive and significant at least at the 5 percent level while since 1975, São Paulo's distance parameter is negative and significant[134] at least at the 5 percent level. These results show that proximity to Rio was associated with lower population levels relative to pre-war conditions while proximity to São Paulo was correlated with higher population. On the other hand, Brasilia's distance parameter was not significant, even at the 10 percent significance level, over the period of analysis. Over time, population levels were strongly positively related to their initial levels for regions, with the coefficient on population in 1939 (pop1939) significant at the 1 percent level in all cross-sections. Consistent with these findings, Figure 6.7 again shows no apparent relationship between distance to Brasilia and change in regions' population shares from 1939 to 2008.

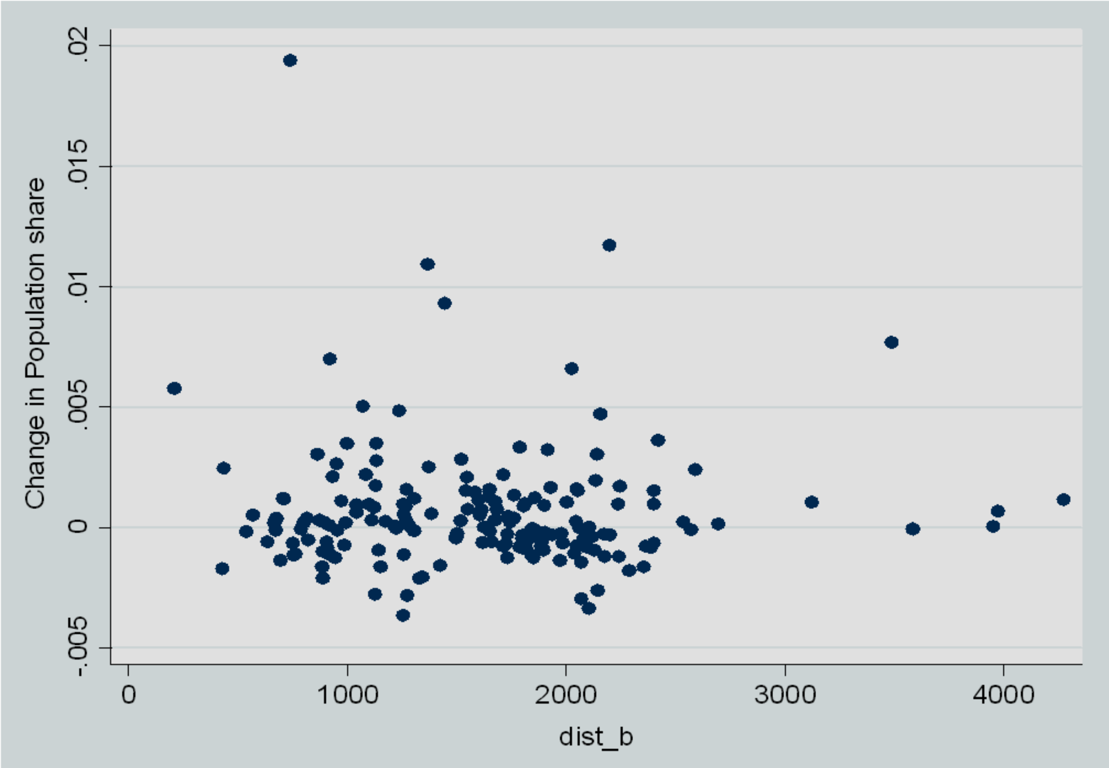
[134] And also for 1959, but at 10 percent level.

Table 6.3: Cross-section Levels Models for Population

	Constant	pop1939	dist_b	dist_r	dist_s	Number of observations	R-Squared
1949	2937.0030	1.0481***	7.5315	4.0342	-3.2015	182	0.9219
	(0.14)	(44.21)	(0.20)	(0.48)	(-0.48)		
1959	3845.0970	1.5678***	-19.2787	33.2737**	-22.5818*	182	0.8642
	(0.09)	(32.93)	(-0.26)	(1.97)	(-1.68)		
1970	42737.0300	2.3515***	-126.3791	78.3854**	-46.9132*	182	0.7727
	(0.48)	(24.27)	(-0.83)	(2.28)	(-1.71)		
1975	57404.5800	2.8755***	-162.8710	126.9734***	-76.5641**	182	0.7288
	(0.48)	(21.69)	(-0.78)	(2.69)	(-2.04)		
1980	72071.7500	3.3996***	-199.3631	175.5615***	-106.2150**	182	0.6939
	(0.46)	(19.96)	(-0.74)	(2.90)	(-2.21)		
1985	103533.7000	3.9392***	-258.6038	229.4861***	-135.4899**	182	0.6626
	(0.54)	(18.60)	(-0.78)	(3.05)	(-2.40)		
1996	41124.3700	4.7924***	-168.7809	324.0943***	-187.3099**	182	0.6281
	(0.16)	(17.34)	(-0.39)	(3.30)	(-2.40)		
1999	46932.2700	5.0199***	-175.1835	349.3047***	-201.4247**	182	0.6167
	(0.17)	(16.93)	(-0.38)	(3.31)	(-2.41)		
2000	47133.1100	5.0954***	-174.6887	358.1000***	-206.2423**	182	0.6136
	(0.17)	(16.82)	(-0.37)	(3.33)	(-2.41)		
2001	45627.0600	5.3405***	-175.4039	371.3084***	-214.8436**	182	0.6093
	(0.16)	(16.68)	(-0.35)	(3.26)	(-2.38)		
2002	45368.3300	5.4193***	-174.7792	378.6142***	-218.6772**	182	0.6064
	(0.15)	(16.58)	(-0.34)	(3.26)	(-2.37)		
2003	43456.7900	5.4999***	-170.6040	387.0505***	-223.3639**	182	0.6036
	(0.14)	(16.48)	(-0.32)	(3.26)	(-2.37)		
2004	42689.7400	5.6702***	-169.1815	403.7658***	-233.5861**	182	0.5984
	(0.13)	(16.31)	(-0.31)	(3.27)	(-2.38)		
2005	47815.6900	5.7625***	-178.7582	712.7590***	-238.4698**	182	0.5952
	(0.15)	(16.10)	(-0.32)	(3.27)	(-2.27)		
2006	49046.1000	5.8542***	-180.3142	421.6548***	-243.5557**	182	0.5921
	(0.15)	(16.10)	(-0.32)	(3.26)	(-2.27)		
2007	25493.1300	5.9592***	-136.7866	435.2751***	-250.6000**	182	0.5912
	(0.08)	(16.07)	(-0.23)	(3.30)	(-2.39)		
2008	26088.1500	6.0700***	-137.8748	441.5811***	-253.1335**	182	0.5894
	(0.08)	(16.01)	(-0.23)	(3.28)	(-2.36)		

Notes: Equations are estimated using SUR. Significance levels: *=10%, **=5%, and ***=1%. Values of z statistics are in brackets.

Figure 6.7: Changes in Population share, 1939 to 2008



Notes: dist_b is distance to Brasília in kilometres; Population is the number of inhabitants.

6.5.2. Cross-section Growth Models

Table 6.4 shows the results for specification (6.5) estimated using the SUR method. From 1970 to 1999, distance to Brasilia is positive and significant at least at the 10 percent level. This result indicates that, over this period, Brasilia had a crowding-out effect, impeding growth of its nearby regions. By 2008, Brasilia's crowding-out effect was not statistically different from zero. On the other hand, São Paulo's distance parameter is negative and statistically significant at the 5 percent level in all cross-sections. This result indicates that the closer to São Paulo a region *r* was, the higher was region *r*'s growth rate after controlling for initial conditions (population and GDP per capita).

Rio's role is found to be weak for most of the period; its distance coefficient is negative and significant at 5 percent only for 2008. Importantly, however, its coefficient for the 2008 cross-section is almost identical to that of São Paulo, and of similar significance level. The 2008 cross-section summarises the development process over the complete period for which data are available (1939-2008). The result for this cross-section indicates that agglomeration forces, which previously had benefitted regions close to São Paulo, also benefitted regions close to Rio de Janeiro in the 2000s. This result is consistent with theoretical contributions finding that agglomeration externalities have grown in importance since the 1990s (McCann 2009a).

The coefficient on region's *r* per capita income in *t-T* is negative and significant at the 1 percent level in all cross-sections, indicating a negative relationship between income growth and initial income, which is consistent with the neoclassical growth theory of beta convergence (Barro & Sala-i-Martin, 1992, 2004). The annual convergence rate, however, declined from about 3 percent in 1949 to around 1 percent in 2008. This reduction in the annual rate of convergence is not surprising given that the window for convergence is much longer for the 2008 cross-section than for the 1949 cross-section. Regional income convergence has been found in other previous studies (see, for instance, this thesis' section 2.4; Ferreira, 2000; Azzoni, 2001; Silveira-Neto & Azzoni, 2006; Gondim, Barreto & Carvalho, 2007; Resende, 2011). However, none of these studies examined the role of Brasilia's creation or examined the potential for agglomeration externalities associated with Rio de Janeiro or São Paulo to supplement the convergence pattern.

Table 6.4: Cross-section Growth Models for Growth of per capita GDP

Growth	Constant	yrt_T	iprt	dist_b	dist_r	dist_s	Number of observations	R-Squared
1939-49	0.1799*** (3.46)	-0.0302*** (-5.80)	0.0003 (0.48)	0.0055 (0.57)	-0.0116 (-1.56)	-0.0158** (-2.18)	174	0.2350
1939-59	0.1010*** (3.88)	-0.0218*** (-8.44)	0.0007** (2.19)	0.0044 (0.91)	-0.0067* (-1.82)	-0.0091** (-2.49)	174	0.3652
1939-70	0.0573*** (3.33)	-0.0167*** (-9.78)	0.0006** (1.96)	0.0070** (2.19)	-0.0012 (-0.47)	-0.0111*** (-4.62)	174	0.3937
1939-80	0.0662*** (4.47)	-0.0143*** (-10.40)	-0.0005 (-0.90)	0.0075*** (2.92)	-0.0003 (-0.16)	-0.0107*** (-5.54)	174	0.4184
1939-85	0.0612*** (3.55)	-0.0119*** (-9.93)	0.0002 (0.19)	0.0047** (2.12)	-0.0016 (-0.96)	-0.0072*** (-4.33)	174	0.4038
1939-99	0.0335* (1.89)	-0.0109*** (-11.47)	0.0016 (1.39)	0.0033* (1.89)	-0.0021 (-1.56)	-0.0053*** (-4.05)	174	0.4773
1939-08	0.0398** (2.19)	-0.0096*** (-11.05)	0.0013 (1.05)	0.0019 (1.19)	-0.0030** (-2.47)	-0.0033** (-2.78)	174	0.4807

Notes: Equations estimated using SUR. yrt_T is urban area (region) r income per capita in initial period (1939); iprt is urban area (region) r predicted population in period *t* from the first stage population regression model (see the fourth dependent variable in appendix 6.1). All independent variables are in logarithms. Significance levels: *=10%, **=5%, and ***=1%. Values of *z* are in brackets.

Overall, these conclusions imply that Brasilia, if anything, impeded growth of its neighbouring regions, while São Paulo's positive influence on neighbouring regions' growth remained strong throughout the period. The agglomeration influence on surrounding regions of Rio de Janeiro, the second ranked city, was minor throughout the period to 1999.

These results indicating the importance of São Paulo relative to Rio de Janeiro are consistent with other studies. For instance, states' manufacturing employment growth rates from 1981 to 2006 were already analysed in Chapter 3 where it was found that while the national and São Paulo annual average growth rates were 2.56 percent and 1.28 percent respectively, Rio de Janeiro's was -0.05 percent[135]. Azzoni (1997), after analysing the dynamics of the states' shares in national GDP, found an increase in São Paulo's share from about 30 to 35 percent and a decrease in Rio's share from 20 to 10 percent over 1939 to 1995.

However, the result for the full period through to 2008 indicates that Rio de Janeiro reasserted its influence through the 2000s. By 2008, its impact on neighbouring regions' per capita income growth (since 1939) was of similar

[135] Rio was the sole state with a negative annual average growth rate.

magnitude to São Paulo. Thus Brazil's two major cities exerted a positive impact on neighbouring regions' per capita income growth over the 70 year period to 2008 after convergence forces are taken into account. By contrast, the major regional development shock associated with Brasilia's creation, had no impact.

6.5.3. Cross-section Spatial Growth Models

The tests of whether regional development policy or agglomeration forces dominate in Brazil are extended by accounting for spatial dependence in the convergence growth model. Prior to the estimation of the spatial models, it is useful to first test for spatial dependence. The results of these tests are in Table 6.5. The Moran's *I* statistic is not significant at even the 10 percent level in any year. In addition, only one of the Robust Lagrange Multiplier (LM) statistics for each of the spatial error and spatial lag models is significant at the 10 percent level, and neither of the LM statistics is significant for the full period (2008) model.[136]

For completeness, spatial error and spatial lag models for all cross-sections have been estimated despite the minimal level of spatial dependence implied by the statistics in Table 6.5. As expected, given these test statistics, none of the ρ or λ coefficients from equations (6.6.1) or (6.6.2) is significant at the 10 percent level for any year. Accordingly, these results are not presented, since the estimates are inefficient relative to those in Table 6.4 that exclude the insignificant spatial dependence parameters.

[136] Magalhães, Hewings & Azzoni (2005, pp. 16-17), citing Anselin and Rey (1991), point out that the criteria for comparison between the spatial error and spatial lag models is to look at the significance levels of their Robust Lagrange Multipliers. While Table 6.5 shows that the spatial error model out-performs the spatial lag model in all years except 1985, the clear picture is that spatial dependence of either type is not present.

Table 6.5: Tests for Spatial Dependence

Growth	Test	Moran's *I* (error)	Robust LM (error)	Robust LM (lag)
1939-49	Value	0.040	1.570	1.533
	p-Value	0.968	0.210	0.216
1939-59	Value	-0.089	0.115	0.086
	p-Value	1.000*	0.735	0.769
1939-70	Value	-0.723	0.848	0.629
	p-Value	1.000*	0.357	0.428
1939-80	Value	1.626	0.186	0.072
	p-Value	0.104	0.666	0.789
1939-85	Value	-1.067	1.262	1.663
	p-Value	1.000*	0.261	0.197
1939-99	Value	-0.380	4.146	3.826
	p-Value	1.000*	0.042	0.050
1939-08	Value	-0.368	1.036	0.868
	p-Value	1.000*	0.309	0.352

Notes: LM stands for Lagrange Multiplier. *These p-values were higher than one due to instability of inverse distance weights matrix that attributes 1 for coincident regions (http://resources.esri.com).

6.6. Conclusion

This chapter analysed the impact of the creation of Brasilia city on the growth of Brazilian regions from 1939 to 2008. The impacts of São Paulo and Rio de Janeiro on neighbouring regions' growth were also evaluated, given the importance of these two regions for Brazil's development process. The aim was to test whether the effects of agglomeration forces (associated with the economic dominance of the largest cities of São Paulo and Rio de Janeiro) more than offset the effect of regional development policy (the creation of Brasilia) on regional growth patterns in Brazil. Different model specifications have been employed to test robustness.

Some key conclusions can be drawn from estimations. First, cross-section Seemingly Unrelated Regression (SUR) models were estimated for the levels of three separate economic dependent variables, controlling for regions' pre-war (1939) characteristics. It was found that Brasilia had no significant impact in reversing concentration patterns for regional GDP, per capita income or population levels. Similarly, Rio de Janeiro had no effect on GDP or per capita income levels of its neighbours and had a negative impact on population outcomes for its neighbouring regions. By contrast, São Paulo has acted to increase per capita income and population of neighbouring regions, consistent with agglomeration forces acting to spread the benefits of São Paulo's economic dominance to nearby cities.

A second set of cross-section SUR models was estimated for regional growth of income per capita. These estimates confirmed the presence of (small) beta convergence of per capita incomes across Brazil's regions. In addition to this effect, São Paulo generated a positive spillover for per capita incomes in neighbouring regions across all years covered by the study. For much of the period, Rio de Janeiro had no significant effect on growth of its neighbours' per capita incomes, although there was a positive effect once the sample was extended to 2008. Furthermore, the magnitude of this effect is similar to that of São Paulo's. By contrast, Brasilia had either a crowding out effect or zero effect on neighbouring regions' per capita income growth.

These results indicate that, at least for the case of Brazil, standard convergence models for per capita incomes (Barro & Sala-i-Martin, 1992, 2004) need to be supplemented by other factors. In particular, the convergence model needs to be supplemented by incorporating the impact of agglomeration forces exerted by the largest cities (São Paulo and Rio de Janeiro). However, no modification to the convergence model is required to reflect the major regional development shock associated with the creation of Brasilia.

Tests for spatial dependence in the estimated convergence model found no evidence of spatial lags or spatial errors. Consistent with these tests, estimated cross-section spatial lag and spatial error models showed no significant effects of either a spatial lag or spatial error term. Thus the three distance variables that were included in the convergence equation are sufficient to cater for any structural spatial dependence within the convergence model.

Overall, the results indicate that standard convergence forces have been at work within Brazil over the period 1939-2008. The results also demonstrate that Brasilia's creation had no positive effect on growth of its neighbouring regions. Thus the creation of Brasilia did not meet policy makers' expectations as it had little or no impact on the growth of the country's interior regions. By contrast, the agglomeration forces associated with São Paulo (and, more recently, Rio de Janeiro) had a strong positive impact on neighbouring regions' per capita income growth, consistent with the modern economic geography literature (Krugman, 1994; Krugman & Venables, 1995; Skilling, 2001; Poot, 2004; Crawford, 2006; Silveira-Neto & Azzoni, 2006; McCann, 2009a). These agglomeration forces have benefited income growth of regions adjacent to the São Paulo and Rio de Janeiro urban areas, the two largest concentrations of economic activity and the two most globally linked cities in Brazil. It can be concluded therefore that in Brazil's case the fundamental forces associated with agglomeration have dominated spatial development patterns within the country and outweighed any impacts intended for the major regional development initiative associated with Brasilia's creation.

CHAPTER 7 – CONCLUSION

This thesis analysed the regional socio-economic transformation which Brazil underwent over the last century. Yet it demonstrated that from 1939 to 2008 the pattern of concentration of economic activity across Brazilian regions, states and in the two (economically) most important metropolitan areas, São Paulo and Rio de Janeiro, was quite stable[137]. Why did this happen? To answer this question, five conclusions can be drawn from the analysis. First, from Chapter 2, it was seen that although "first nature" geography is important, its advantages are more than offset by "second nature" geography. Most lagging states are also coastal states in Brazil, but in addition to their low local market potential, they lack market access, making their industries unable to take advantage of increasing returns to scale that arise in agglomeration economies. The lagging states (mainly located in the northeast region) had poor performance in a number of socio-economic indicators. They also lacked skilled labour due to the lowest literacy rates and schooling levels in the nation, which resulted in more specialisation in land-related sectors.

Secondly, the initial descriptive evidence was re-examined in Chapter 3 by means of the Glaeser *et al.* (1992) model which tested the importance of industrial clusters and evaluated whether knowledge externalities are important for regional growth. Although the results are conditional to the data used (for instance, the competition externality was only included in a cross-section model due to a lack of required data), there is considerable evidence that state growth is mostly explained by specialisation, consistent with both Marshall-Arrow-Romer and Porter externalities (Marshall, 1920; Arrow, 1962; Romer, 1986; Porter, 1990; De Groot, Poot and Smit, 2009) rather than with Jacobs' externalities (Jacobs, 1969). The latter externalities were confirmed in a single cross-section model only.

Thirdly, the importance of the effects of agglomeration forces, especially the role of market potential (or home market effect), was tested in Chapter 4. This thesis confirmed the heterogeneity of people's (and businesses') locational decisions for Brazil. The power law holds for the 100 largest urban areas over the entire 20th century. The power parameter (or the slope) from the traditional power

[137] Venables (2009, p. 331) mentions that, according to economic geography literature, "spatial unevenness in economic activity and income is an equilibrium outcome". The evidence provided in this thesis indicates that Brazil basically seems to have been in the same uneven equilibrium since the pre-Second World War period.

law and from the corrected version of this law (Gabaix & Ibragimov, 2006) resemble each other: they decrease from -0.63 and -0.60 in 1907 to -0.89 and -0.84 in 2008, respectively. Zipf's law does not hold because the absolute power law parameter is less than one. Except for 1939-1940, the city size distribution has become increasingly uneven. In other words, there is a city size divergence (or a convergence to Zipf's law). The behaviour of the power parameter from the traditional power law can be divided into three phases: Phase I: the slope is fairly static around -0.6 from 1907 to 1939. Phase II: the slope falls steeply from -0.60 to -0.86 from 1940 to 1983. Phase III: the slope is relatively stable, changing from -0.86 to -0.89 from 1984 to 2008. These phases were considered for tests of Gibrat's law of proportional city growth, confirmed after 1984 by panel unit root tests that tested the relationship between urban areas' growth and their initial sizes for the whole sample. For the previous sub-periods (or phases), Gibrat's law was rejected in favour to a weak mean reversion hypothesis (though the coefficient of lag of city size is very small). These findings were supported by both the Levin-Lin-Chu and Im-Pesaran-Shin panel unit root tests for the whole sample and, to a smaller extent, for individual urban areas across sub-periods.

In contrast, panel unit root tests for the 100 largest urban areas rejected Gibrat's law in each sub-period: the relationship between urban areas growth and their initial sizes was negative and significant at 10 percent from 1907 to 1939, and positive and significant at 1 percent from 1940 to 2008. Therefore, the direction of rejection of Gibrat's law since 1940, in line with the convergence to Zipf's law, shows that there has been increase in agglomeration of population in the 100 largest urban areas. This led to increase in size dispersion across all urban areas.

Fourthly, to understand the sources of regional growth, Chapter 5 employs exploratory spatial data analysis (ESDA), the classic and spatial shift-share methods (Selting & Loveridge, 1994; Cochrane & Poot, 2008; Nazara & Hewings, 2004) to analyse state-level employment changes from 1981 to 2006. The findings from the classic shift-share and those from ESDA resemble each other: the north and centre-west states grew fastest due to higher performance in two components, the industry-mix and competitive effects, and there were only a few states that went against the trends of the national economy during periods when the nation had little economic growth. Thus, although Glaeser *et al.'s*

(1992) model estimated in Chapter 3 explained regional employment growth by specialisation, this does not appear to be true for the fast growing states over the 1981-2006 period. The spatial shift-share enhances the understanding of regional growth. This method shows that, on the one hand, the states that grew fastest (the north and centre-west states) usually performed well in spatial components of the shift-share method, namely the potential spatial spillover and spatial competitive effects; on the other hand, the states with low growth had almost zero or negative rates for both spatial components.

Finally, Chapter 6 tests the effect of the creation of Brasilia city in the late 1950s. More specifically, it evaluates whether the effects of agglomeration forces (that historically caused concentration of income and population in the two largest cities of São Paulo and Rio de Janeiro)[138] confirmed by the power law in Chapter 4, more than offset the effects of an exogenous shock (the creation of Brasilia city) on the pattern of economic activity established before Brasilia's creation (Krugman, 1994; Krugman & Venables, 1995; Skilling, 2001; Poot, 2004; Crawford, 2006; McCann, 2009a). Using different modelling strategies for socio-economic variables at the levels (GDP, GDP per capita and population), Chapter 6 finds that, based on cross-section levels models, Brasilia had no effect on the spatial pattern of GDP, per capita income and population. From successive cross-section models for growth of GDP per capita since 1939, this thesis confirmed the neoclassical theory of beta convergence (Barro & Sala-i-Martin, 1992, 2004; Ozgen, Nijkamp & Poot, 2010), i. e. that the per capita income growth at time *t* is negatively correlated with per capita income at time *t*-1. This result was consistent with the one found after introduction of spatial autocorrelation, although the models with distance variables as means to capture spatial effects were more efficient. Models of growth of GDP per capita show that Brasilia either had no effect or had a crowding-out effect on the spatial pattern of economic activities in Brazil, with a negative impact on growth of its neighbouring regions. While Rio de Janeiro, from its positive effect on its neighbours in 1959, lost importance that was recovered by 2008 only, the largest city of São Paulo increased its importance (this is captured by an increase in significance level of distance parameter from 1949 to 1970, which had stable significance from 1970), benefiting growth of GDP per capita of its neighbouring regions. This finding, therefore, implies that

[138] Indeed, this concentration is also observed at the regional and state levels, as shown in Chapter 1, that it has been almost timeless.

the effects of agglomeration forces outweigh those of the large scale regional policy behind Brasilia's creation.

All these conclusions show that regional growth is strongly associated with location in Brazil. From these conclusions, a question is asked; i. e. which of the three theoretical frameworks dominates as the most suitable explanation for Brazil's development process? Is it the mainstream neoclassical economics, or is it regional science (modern economic geography) combined with spatial economic theory? The fact that a number of models confirmed a convergence process in which some of the lagging states grew faster than average and the richest ones showed at most modest growth is consistent with mainstream economics, because the marginal productivity of any production factor is much higher in poor and low income states compared to the rich states. The rich states reached levels of intensity in the use of production factors for which inputs improvements (infrastructure development, increase in education, etc) had small effects on their growth rates. Most features, however, are more consistent with the other two approaches used to explain regional disparities in Brazil, five of which are emphasised.

First, even though there is national convergence, its rate is very small[139] and the lagging states actually do not catch-up. Secondly, the key socio-economic indicators (GDP, GDP per capita, and population density) are the highest in the leading regions, and the thesis identifies two geographical timeless clusters (or clubs) associated with development of states: a cluster (or club) of contiguous rich states (southeast and south) - the core - and another of contiguous poor states (north and northeast) - the periphery[140]. Thirdly, an inspection of states' sectoral employment shares and states' Hirschman-Herfindahl Index for sectoral employment data shows consistence with predictions of the New Economic Geography literature (e.g. Brakman et al. 2001) (e.g. the nation's poorest states in the northeast region – not in north - are more agricultural and almost lack manufacturing while the nation's richest states and regions have a significant

[139] For instance, this thesis showed in Chapter 6 that, at the metropolitan area level, the convergence rate in GDP per capita fell from about 3 percent from 1939 to 1949 to just 1 percent from 1999 to 2008.

[140] The centre-west region does not clearly belong to any of these two clubs and its insertion in Brazil is ambiguous. For example, at the regional level, for regions' GDP shares and population density, this region is somewhere in between the north and northeast regions (the periphery) standards. Nevertheless, for data of GDP per capita, it resembles the core regions, especially since 1970.

share of manufacturing in their economies). Fourthly, the platforms of the global economy (Poot, 2004; McCann, 2008, 2009a, 2009b) are in the rich regions (e.g. the two cities and states of São Paulo and Rio de Janeiro in the southeast region), which are also the largest national markets [141] and have the highest market potential. Fifthly, as a consequence, despite only modest growth in the rich regions in recent decades, the gaps between rich and poor regions do not narrow because increasing returns to scale operate in favour of the rich regions which yields a scale of concentration of economic activities, as demonstrated in this thesis, that is almost timeless with a clear pattern of core-periphery. These five features provide more insights that explain the observed regional disparities better than the mainstream economics solely in Brazil.

From these conclusions, the first two hypotheses of this thesis shown in section 1.4 of Chapter 1 cannot be rejected. In other words, the essentially time-invariant pattern of concentration of economic activities in Brazil holds because the patterns of spatial interactions among regions have not significantly changed since the Second World War, which implies that to significantly reduce concentration in Brazil the economy needs strong simultaneous shocks (i.e., more place based policies). A single shock, as shown in Chapter 6 of this thesis, can be insufficient to significantly change spatial patterns of income and population. Place-based policies matter: as seen in Chapter 6, the creation of Brasilia obviously had major implications for Brasilia itself, but did not offset the agglomeration forces that led to the dominance of São Paulo. Similarly, investment in other lagging regions may not offset the advantages of São Paulo and Rio de Janeiro, but they may improve economic conditions in the lagging regions themselves. Obviously, simple income transfers or subsidies of inefficient industries are not going to help lagging regions[142]. The benefits may be more in

[141] Even if considering that the transport costs would be high due to distance, the fact that infrastructure is less developed makes the volume of inter- and intraregional trade very small in Brazil, which is harmful for the lagging and peripheral regions of north and northeast.

[142] Previous studies pointed out successful fiscal measures to attract economic activities and help in development of the lagging regions. These measures include, for instance, fiscal incentives that helped development of southern Goiás state even in municipalities off the highway network such as Araguaia Paraense (Katzman, 1975), as well as tax incentives for manufacturing and mining sectors (Amazonas state) and mining (Carajás mining project in Serra Norte in Pará state) (Gomes, 2002). Furthermore, Chapter 5 of this thesis provides some policy advice for lagging regions in Brazil. The overall picture showed by sectoral employment data over the period 1981-2006 is that all states followed national trend. However, policy advice has to simultaneously consider location quotients at the state level (see Table 5.3), the trends of sectoral shares at the national level and both state and nationwide sectoral average growth rates (a complete Table is available upon request). Overall, north region states had location quotients higher than 1 in almost all sectors and,

creating the precondition for endogenous growth: good infrastructure, a highly educated population and a promotion of innovation activities that exploit the localised advantages and potential of the lagging regions.

Consequently, the third hypothesis is clearly rejected, because the thesis demonstrated that the characteristics of the rich regions and those of the poor ones are consistent with the circular causality outlined by the economic geography literature (see, for instance, Krugman, 1991a, 1991b, 1995; Brakman, Garretsen & van Marrewijk, 2001; Fujita & Krugman, 2004; Bosker, 2008; Capello & Nijkamp, 2009). This circular causality is the reason why, as emphasised by the research questions of this thesis, there is a large concentration of economic activities and there is no clear sign of a fall in the disparity of the three key socio-economic indicators (GDP, GDP per capita, and population) in Brazil.

These findings have one key policy implication. The government of Brazil should focus more on state and municipality levels to alleviate the problem of peripherality for the lagging regions and improve regional prosperity according to the economic geography literature. For example, efforts to create agglomerations at the municipality level are more likely to succeed than at the state level. Most conditions for regional prosperity (e.g. improvement of market potential) can only be reached by combining both short and long-run policies[143]. The third hypothesis of this thesis states that the mechanisms of the growth process cannot be clearly detected, i. e. it is difficult to understand how they function in Brazil. This hypothesis has been rejected, which has policy implications because, in reality, the mechanisms for growth are basically driven by the dichotomy of agglomeration forces versus diffusion forces according to the economic geography literature. However, there is a caveat. Resources should be directed to the collection of more disaggregated data (i.e. data at the local level) from which

most importantly, grew faster in sectors that, at the national level, grew fastest and that individually had increasing shares in Brazil's total employment. These are Commerce, Services, and Transportation and Communication, which should be prioritised in terms of receiving fiscal incentives. Conversely, in the troublesome northeast region, states had a relative specialisation (LQ>1) in construction sector (which along with Electricity, Water & Gas, and Mining sectors was stable at the national level) and in agriculture & fishing sector (which had a smooth fall despite the fact that in 2006 was still accounting for 20 percent of the nation's total employment). Therefore, in the northeast region, it is urgent to not encourage further the agriculture & fishing sector and set up incentives for the most national promising sectors (Commerce, Services, and Transportation and Communication). Fortunately, Northern states had low LQs (LQ<0.45) in agriculture & fishing sector. The national trend of financial sector and manufacturing is inconclusive due to fluctuations in their shares.

[143] These include, for instance, improvements of infrastructure and human capital in lagging regions by Brazilian government, as emphasised above.

the analysis of growth patterns can be more refined hopefully resulting in more reliable models and possible policy prescriptions.

There are two possible avenues for future research. First, from Chapter 4, one research extension would be to evaluate the impact of births and deaths in urban areas on the city size distribution, using non-parametric alternative approaches such as Kernel density function (Black & Henderson, 2003; Stata, 1996-2012). Specifically, this research would evaluate the extent to which an introduction of a new city into the system changes de distribuition. This would help to better understand the regularity of market potential (or home market effect). Secondly, it can be seen from observation that the two clusters (or clubs) of regions have differences in infrastructure development in Brazil. Although Chapter 6 examined partially and indirectly the extent to which infrastructure related to the creation of Brasilia affected growth of regions (with São Paulo and Rio de Janeiro as drivers of Brazil's economic system), one potentially promising further investigation would be to more directly analyse the impact of infrastructure at the regional level and see whether infrastructural investment can break the spatial pattern (which is almost timeless, as was seen throughout the thesis) of activities and regional development. This could be achieved by means of building General Equilibrium models for the regions. Broadly, this line of research will directly evaluate how much differences in transport and transaction costs (features tied to "second nature" geography in the literature) harm development of lagging regions and/or benefit leading regions. To find adequate data for this task will be a huge challenge.

References

Abreu, M. P. (2008a). The Brazilian economy, 1930-1980. In L. Bethell (Ed.), *The Cambridge history of Latin America* (Vol. IX) (pp. 283-394). Cambridge, England: Cambridge University Press.

Abreu, M. P. (2008b). The Brazilian economy, 1980-1994. In L. Bethel (Ed.), *The Cambridge history of Latin America* (Vol. IX) (pp. 395-430). Cambridge, England: Cambridge University Press.

Abreu, M. P., & Werneck, R. L. F. (2008). The Brazilian economy, 1994-2004. In L. Bethel (Ed.), *The Cambridge history of Latin America* (Vol. X) (pp. 431-454). Cambridge, England: Cambridge University Press.

Almeida, P. R. de (2004). *A experiência brasileira em planejamento económico: Uma síntese histórica.* Available from http://www.pralmeida.org/05DocsPRA/1277HistorPlanejBrasil.pdf. <Accessed February 27, 2011>.

Amorim, M. A., & Da Mata, D. (2008). *The Cariri region of Ceará and the footwear cluster*. Paper presented at Conference "Trajetórias de Desenvolvimento Local e Regional: Uma comparação entre as Regiões do Nordeste Brasileiro e a Baixa Califórnia, México", Fortaleza, Brazil.

Anderson, G., & Ge, Y. (2005). The size distribution of Chinese cities. *Regional Science and Urban Economics, 35*(6), 756-776.

Andrade, T. A., & Serra, R. V. (2001). *Análise do desempenho produtivo dos centros urbanos brasileiros no período 1975/96*. Available from Núcleo de Estudos e Modelos Espaciais Sistêmicos website: http://www.nemesis.org.br/

Andrikopoulos, A., Brox, J., & Carvalho, E. (1990). Shift-share analysis and the potential for predicting regional growth patterns: Some evidence for the region of Quebec, Canada. *Growth and Change, 21*(1), 1-10.

Anselin, L. (1989). *What is special about spatial data? Alternative perspectives on spatial data analysis.* Paper prepared for presentation at the Spring 1989 Symposium on Spatial Statistics, Past, Present and Future, Department of Geography, Syracuse University.

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

Arcelus, F. (1984). An extension of shift-share analysis. *Growth and Change, 15*(1), 3-8.

Arrow, K. J. (1962). The economics implications of learning by doing. *Review of Economic Studies*, *29*(3), 155-173.

Autant-Bernard, C., Mairesse, J., & Massard, N. (2007). Spatial knowledge diffusion through collaborative networks. *Papers in Regional Science*, *86*(3), 341-350. doi:10.1111/j.1435-5957.2007.00134.x

Azzoni, C. R. (1997a). *Distribuição pessoal de renda nos estados e desigualdade de renda entre estados no Brasil: 1960, 70, 80 e 91*. Available from Núcleo de Estudos e Modelos Espaciais Sistêmicos website: http://www.nemesis.org.br/

Azzoni, C. R. (1997b). *Concentração regional e dispersão das rendas per capita estaduais: Análise a partir de séries históricas estaduais de PIB, 1939-1995.* Available from Núcleo de Estudos e Modelos Espaciais Sistêmicos website: http://www.nemesis.org.br/

Azzoni, C. R. (1999). Regional income inequalities by income class in Brazil: Is inequality higher among rich or poor people? *Working Paper Series.* doi: 10.2139/ssrn.202688.

Azzoni, C. R. (2001). Economic growth and regional income inequality in Brazil. *The Annals of Regional Science, 35*, 133-152.

Azzoni, C. R. (2003). Sobre a necessidade da política regional. *Nereus Working Paper #08.* Available from The University of São Paulo Regional and Urban Economics Lab website: www.econ.fea.usp.br/nereus.

Azzoni, C. R., Menezes-Filho, N., Menezes, T., & Silveira-Neto, R. (2000). *Geography and income convergence among Brazilian states*. Retrieved July 21, 2008 from Inter-American Development Bank Research Department website: http://www.iadb.org/res/pub_desc.cfm?pub_id=R-395.

Baer, W. (2008). *The Brazilian economy: Growth and development* (6th Ed.). Boulder, CO: Lynne Rienner.

Barff, R. D., & Knight, P. L. (1988). Dynamic shift-share analysis. *Growth and Change, 19*(2), 1-10.

Barossi-Filho, M., & Azzoni, C. R. (2003). A time series analysis of regional income convergence in Brazil. *Nereus Working Paper #09.* Available from The University of São Paulo Regional and Urban Economics Lab website: www.econ.fea.usp.br/nereus.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy, 100*(2), 223-251.

Barro, R. J., & Sala-i-Martin, X. (2004). *Economic growth*. Cambridge, MA: MIT Press.

Baum, C. F. (2003). *Notes section 9: Panel unit root tests* [PowerPoint Slides]. Available at: http://fmwww.bc.edu/ec-c/s2003/821/ec821.s2003.php

Bavaud, F. (1998). Models for spatial weights: A systematic look. *Geographical Analysis, 30*(2), 153-171.

Becker, B. K. (1977). *A implantação da rodovia Belém-Brasília e o desenvolvimento regional*. Retrieved March 1, 2011, from Annual Record of the Federal University of Rio de Janeiro Geosciences Institute website: www.anuario.igeo.ufrj.br/anuario_1977/vol_01_32_46.pdf

Bertaud, A. (2001). *The costs of utopia: Brasilia, Johannesburg, and Moscow.* Available from Alain Bertaud's website: http://alain-bertaud.com. <Accessed throughout February, 2011>.

Bértola, L. (2002). *Income distribution and the Kuznets curve: Argentina and Uruguay since the 1870s.* Paper presented at the 7th Annual Meetings of the LACEA - Latin American and Caribbean Economic Association (11-13, October), Madrid, Spain.

Berzeg, K. (1978). The empirical content of shift-share analysis. *Journal of Regional Science, 18*(3), 463-469.

Berzeg, K. (1984). A note on statistical approaches to shift-share analysis. *Journal of Regional Science, 24*(2), 277-285.

Bethel, L. (2008a). Politics in Brazil under Vargas, 1930-1945. In L. Bethell (Ed.), *The Cambridge history of Latin America* (Vol. IX) (pp. 3-86). Cambridge, England: Cambridge University Press.

Bethell, L. (2008b). Politics in Brazil under the Liberal Republic, 1945-1964. In L. Bethell (Ed.), *The Cambridge history of Latin America* (Vol. IX) (pp. 87-164). Cambridge, England: Cambridge University Press.

Bishop, K. C., & Simpson, C. E. (1972). Components of change analysis: Problems of alternative approaches to industrial structure. *Regional Studies*, *6*(1), 59–68.

Black, D., & Henderson, J. V. (1999). A theory of urban growth. *Journal of Political Economy*, *107*(2), 252-284.

Black, D., & Henderson, J. V. (2003). Urban evolution in the US. *Journal of Economic Geography*, *3*(4), 343-372.

Blien, U., & Wolf, K. (2002). Regional development of employment in Eastern Germany: An analysis with an econometric analogue to shift-share techniques. *Papers in Regional Science*, *81*(3), 391–414.

Bosker, M. (2008). *The empirical relevance of geographical economics*. Unpublished PhD thesis, Utrecht University, Utrecht, The Netherlands.

Bowling, K. R. (1991*). The creation of Washington, D.C.: The idea and location of the American capital.* Fairfax, VA, USA: George Mason University Press.

Boyce, R. R. (2004). Geographers and the Tennessee Valley Authority. *The Geographical Review, 94*(1), 23-42.

Brakman, S., Garretsen, H., van Marrewijk, C., & van den Berg, M. (1999). The return of Zipf: Towards a further understanding of the rank-size distribution. *Journal of Regional Science*, *39*(1), 739–767.

Brakman, S., Garretsen, H., & van Marrewijk, C. (2001). *An introduction to geographical economics: Trade, location and growth.* Cambridge, England: Cambridge University Press.

Brakman, S., Garretsen, H., & van Marrewijk, C. (2009). *The new introduction to geographical economics.* Cambridge, England: Cambridge University Press.

Brauch, M. D., & Monasterio, L. M. (2007). *Spatial dynamics of economic growth in Brazil (1970-2000): An empirical approach.* Available from Núcleo de Estudos e Modelos Espaciais Sistêmicos website: http://www.nemesis.org.br/

Brazil's Ministry of Foreign Affairs, www.mre.gov.br.

Brown University (n.d). *Spatial analysis with GeoDa: Spatial autocorrelation.* Available from Brown University S4 - Spatial Structure in the Social Science website: http://www.s4.brown.edu/

Campbell, R. H., Skinner, A. S., & Todd, W. B. (Eds.) (1976). *An inquiry into the nature and causes of the wealth of nations.* Oxford, England: Clarendon Press.

Capello, R., & Nijkamp, P. (2009). (Eds.). *Handbook of regional growth and development theories*. Cheltenham, England: Edward Elgar Publishing.

Centro de Pesquisa e Documentação de História Contemporânea do Brasil (CPDOC), Fundação Getúlio Vargas (FGV) de São Paulo, http://cpdoc.fgv.br/

Chahad, J. P. Z., Comune, A. E., & Haddad, E. A. (2002). *Shifts in regional employment: An application of the Haynes-Dinc model for Brazil.* Paper presented at the 42nd Congress of the European Regional Science Association (27-31, August), Dortmund, Germany.

Ciccone, A., & Hall, R.E. (1996). Productivity and the density of economic activity. *American Economic Review, 86*(1), 54–70.

Clark, G. L., Feldman, M. P., & Gertler, M. S. (Eds.) (2003). *The Oxford Handbook of Economic Geography*. New York: Oxford University Press.

Cochrane, W., & Poot, J. (2008). Forces of change: A dynamic shift-share and spatial analysis of employment change in New Zealand labour market areas. *Studies in Regional Science*, *38*(1), 51-78.

Coelho, R. L. P., & Figueiredo, L. (2007). Uma análise da hipótese da convergência para os municípios Brasileiros. *Revista Brasileira de Economia, 61*(3), 331-352.

Combes, P. P. (2000). Economic structure and local growth: France, 1984–1993. *Journal of Urban Economics, 47*, 329-355.

Córdoba, J-C. (2008). On the distribution of city sizes. *Journal of Urban Economics, 63*(1), 177–197.

CORECON - Conselho Regional de Economia de São Paulo (or, São Paulo’s Regional Council of Economics). Retrieved throughout 2009, from http://www.coreconsp.org.br/indicadores/dip/planos_economicos.html.

Cornish, E. (1991, July-Aug.). Building utopia: Lesson from Brasília. *The Futurist*, 29-32.

Couto, R. C. (2001). *Brasília Kubitschek de Oliveira*. Rio de Janeiro, Brasil: Editora Record.

Crawford, R. (2006). The effects of agglomeration on economic activity: The empirical evidence on mechanisms and magnitudes. *Ministry of Economic Development Occasional Paper #3*, Wellington, New Zealand.

Dantas, A., Ribeiro, K., & Yamashita, Y. (2006). *Assessing spatial-temporal impacts of transport infrastructure policy in Brazil*. Paper presented at the Transportation Research Board 85th Annual Meeting (22-26, January), Washington DC, the United States.

Daumal, M., & Zignago, S. (2010). Measure and determinants of border effects of Brazilian states. *Papers in Regional Science*, *89*(4). doi:10.1111/j.1435-5957.2009.00265.x

Dedina-Wagner (1961?). *Brasília, a nova capital do Brasil*. São Paulo, Brasil: Edições Zenit.

De Groot, H. L. F., Poot, J., & Smit, M. J. (2009), Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories* (pp. 256-281). Cheltenham, England: Edward Elgar Publishing.

Delgado, A. P., & Godinho, I. M. (2006). Long term evolution of the size distribution of Portuguese cities. *Working Paper #230*, Faculty of Economics, University of Porto.

del Rio, V. (2009a). Preface: What can we learn from Brazilian urbanization? In V. del Rio & W. Siembieda (Eds.), *Contemporary urbanism in Brazil: Beyond Brasília* (pp. xvii-xxx). Florida A&M University, Tallahassee, FL, USA: University Press of Florida.

del Rio, V. (2009b). Introduction: Historical background. In V. del Rio & W. Siembieda (Eds.), *Contemporary urbanism in Brazil: Beyond Brasília* (pp. 1-36). Florida A&M University, Tallahassee, FL, USA: University Press of Florida.

del Rio, V. & Siembieda, W. (2009)(Eds.). *Contemporary urbanism in Brazil: Beyond Brasília*. Florida A&M University, Tallahassee, FL, USA: University Press of Florida.

Dinc, M., & Haynes, K. E. (1999). Regional efficiency in the manufacturing sector: Integrated shift share and data development analysis. *Economic Development Quarterly*, *13*(2), 183-199.

Dinc, M., & Haynes, K. (2005). Productivity, international trade and reference area interactions in shift-share analysis: Some operational notes. *Growth and Change, 36*(3), 374-394.

Dinc, M., Haynes, K., & Qiangsheng, L. (1998). A comparative evaluation of shift-share models and their extensions. *Australasian Journal of Regional Studies, 4*(2), 275-302.

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *Review of Economics and Statistics, 90*(1), 37–48.

Distrito Federal. www.distritofederal.df.gov.br. <Accessed February, 2011>.

DNIT (2011). Departamento Nacional de Infraestructura de Transportes. Retrieved February 2011, from http://www.dnit.gov.br/menu/rodovias.

Doll, C., & Muller, J-P (2000). *Mapping Anthropogenic activities from Earth observation data* [PowerPoint Slides]. The Center for International Earth Science Information Network (CIESIN), Columbia University.

Dowall, D. E., & Monkkonen, P. (2007). Consequences of the plano piloto: The urban development and land markets of Brasília. *Urban Studies*, *44*(10), 1871–1887. doi: 10.1080/00420980701560018.

Dunn, E.S. (1960). A statistical and analytical technique for regional analysis. *Papers of the Regional Science Association, 6*(1), 97-112.

Duranton, G. (2002). *City size distributions as a consequence of the growth process*. Available from ideas.repec.org/p/cep/cepdps/0550.html.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic Review*, *97*(1), 197-221.

The Economist (1998). Who says Brasília doesn’t work? *The Economist, 347*(8073), 37-38.

The Economist (1999). Dreaming up Brasília. *The Economist, 353*(8151), 57-57.

The Economist (2011, May 19). Brazil's north-east - Catching up in a hurry: The country’s poorest region is narrowing the gap with the prosperous south. *The Economist*, http://www.economist.com/node/18712379 <Accessed June 13, 2011>.

The Economist (2009, November 12). Brazil takes off: Now the risk for Latin America’s big success story is hubris. *The Economist*. http://www.economist.com/node/14845197 <Accessed August, 2011>.

Eeckhout, J. (2004). Gibrat’s law for (all) cities. *American Economic Review*, *94*(5), 1429–1451.

Eeckhout, J. (2009). Gibrat's law for (all) cities: Reply. *American Economic Review, 99*(4), 1676–1683.

El-Dahdah, F. (2003). Brasilia – the project of Brasilia. In R. El-Khoury and E. Robbins (Ed.), *Shaping the city: Studies in history, theory, and urban design* (pp. 41-56). New York: Routledge Taylor & Francis Group.

Encyclopaedia Britannica website: www.britannica.com <Accessed June, 2010>.

Enders, W. T. (1980). Regional disparities in industrial growth in Brazil. *Economic Geography, 56*(4), 300-310.

Epstein, D. G. (1973). *Brasília, plan and reality: A study of planned and spontaneous urban development.* Berkeley, CA: University of California Press.

Esteban-Marquillas, J. M. (1972). A reinterpretation of shift-share analysis. *Regional and Urban Economics, 2*(3), 249-255.

Farret, R. (2001). *Land and urban development policies in a planned city: Achievements and challenges in Brasília, Brazil, network city and landscape.* Retrieved February 21, 2011 from http://www.nsl.ethz.ch/displus/147/pdf/farret.pdf.

Fausto, B. (1999). *A concise history of Brazil.* (A. Brakel, Trans.). Cambridge, England: Cambridge University Press. (Original work published 1998).

Ferreira, A. (2000). Convergence in Brazil: Recent trends and long-run prospects. *Applied Economics, 32*, 479-489.

Fotopoulos, G., & Spence, N. (1999). Spatial variations of net entry rates of establishments in Greek manufacturing industries: An application of a shift-share ANOVA model. *Environment and Planning A, 31*(10), 1731-1755.

Fotopoulos, G., Kallioras, D., & Petrakos, G. (2010). Spatial variations of Greek manufacturing employment growth: The effects of specialization and international trade. *Papers in Regional Science, 89*(1), 109-134.

Fujita, M., & Krugman, P. (2004). The new economic geography: Past, present and the future. *Papers in Regional Science, 83*(1), 139–164. doi: 10.1007/s10110-003-0180-0.

Gabaix, X. (1999a). Zipf's law and the growth of cities. *American Economic Review* (Papers and Proceedings), *89*(2), 129–132.

Gabaix, X. (1999b). Zipf's law for cities: An explanation. *The Quarterly Journal of Economics, 114*(3), 739-767.

Gabaix, X., & Ioannides, Y. M. (2003). The evolution of city size distributions. *Working Paper #310*, Department of Economics, Tufts University.

Gabaix, X., & Ibragimov, R. (2006). Log(Rank-1/2): A simple way to improve the OLS estimation of tail exponent. *Working Paper #2106,* Harvard Institute of Economic Research.

Galanternick, M. (2003). Trade lanes: Brasília, open city – Brasília newest dry port brings the remote capital's exporters closer to the world. *Latin Trade,* 56-56. www.latintrade.com.

Gangopadhyay, K., & Basu, B. (2009). City size distributions for India and China. *Physica A: Statistical Mechanics and its Applications, 388*(13), 2682-2688.

Garmestani, A. S., Allen, C. R., Gallagher, C. M., & Mittelstaedt, J. D. (2007). Departures from Gibrat's law, discontinuities and city size distributions. *Urban Studies, 44*(10), 1997-2007.

Garmestani, A. S., Allen, C. R., & Gallagher, C. M. (2008). Power laws, discontinuities and regional city size distributions. *Journal of Economic Behavior & Organization, 68*(1), 209-216.

Garrison, C. B. (1974). Industrial growth in the Tennessee Valley region, 1959 to 1968. *American Journal of Agricultural Economics, 56*(1), 50-60. Available at: http://www.jstor.org/stable/1239346 .

Gazel, C. R., & Schwer, R. K. (1998). Growth of international exports among the states: Can a modified shift-share analysis explain it? *International Regional Science Review, 21*(2), 185-204. doi: 10.1177/016001769802100205.

Getis, A. (1991). Spatial interaction and spatial autocorrelation: A cross-product approach. *Environment and Planning A 23*(9), 1269-1277. doi:10.1068/a231269. Reprinted with permission of PION Limited, London.

Gibrat, R. (1931). *Les inégalités économiques; applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel.* Paris, France: Librairie du Recueil Sirey.

Giesen, K., & Suedekum, J. (2009). Zipf's law for cities in the regions and the country. *Working Paper #3928*, The Institute for the Study of Labor (IZA).

Giesen, K., Zimmerman, A., & Suedekum, J. (2009). *The size distribution across all cities – Zipf, lognormal, double Pareto, or what?* Paper presented at the 56th Annual North American Meetings of the Regional Science Association International (19-21, November), San Francisco, California, the United States.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy, 100*(6), 1126-1152.

Glaeser, E.L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography, 1*(1), 27–50.

Glaeser, E. L., Ponzetto, G. A. M., & Tobio, K. (2011). Cities, skills, and regional change. *Working Paper #16934*, National Bureau of Economic Research. Available at: http://www.nber.org/papers/w16934

Gomes, G. M. (2002). *Regional development strategies in Brazil*. International Conference on Regional Development and Foreign Direct Investment in Fortaleza, Ceará, Brazil. Available from Organisation for Economic Cooperation and Development (OECD) website: http://www.oecd.org.

Gomes, E. (2009). *Terra de colonizador, gado, soja, algodão...* Diário de Cuiabá. http://www.diariodecuiaba.com.br/detalhe.php?cod=58833.

Gonçalves, P. R. (2002). *As cidades satélites de Brasília: Registro histórico*. Retrieved March 7, 2011 from the University of Brasília website: http://vsites.unb.br/ics/sol/itinerancias/bsb/historico/c_satelites.pdf.

Gondim, J. L. B., Barreto, F. A., & Carvalho, J. R. (2007). Condicionantes de clubes de convergência no Brasil. *Estudos Econômicos, 37*(1), 71-100.

González-Val, R., Ramos-Gutiérrez, A., & Sanz-Gracia, F. (2011). *Size Distributions for All Cities: Lognormal and q-exponential functions.* Paper presented at the 51st European Regional Science Association (30th August - 3rd September), Barcelona, Spain.

Google maps website. http://maps.google.co.nz/ <Accessed through the second semester of 2009>.

Hanham, R. Q., & Banasick, S. (2000). Shift-share analysis and changes in Japanese manufacturing employment. *Growth and Change, 31*(1), 108-123.

Hanson, G. H. (1998). Regional adjustment to trade liberalization. *Regional Science and Urban economics, 28*(4), 419-444.

Hanson, G. H. (2003), Firms, workers, and the geographic concentration of economic activity. In G. L. Clark, M. P. Feldman, & M. S. Gertler (Eds.), *The Oxford Handbook of Economic Geography* (pp. 477-494). New York: Oxford University Press.

Haynes, K. E., & Dinc, M. (1997). Productivity change in manufacturing regions: A multifactor/shift-share approach. *Growth and Change, 28*(2), 201-221.

Haynes, K., & Machunda, Z. B. (1987). Considerations in extending shift-share analysis: Note. *Growth and Change, 18*(2), 69-78.

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy, 103*(5), 1067-1090.

Hill, B. M. (1975). A simple approach to inference about the tail of a distribution. *Annals of Statistics*, 3, 1163-1174.

Holanda, F. de., Ribeiro, R., & Medeiros, V. (2008). *Brasília, Brazil: Economic and social costs of dispersion.* Paper presented at the 44th International Society of City and Regional Planners (*ISOCARP*) Congress "Sustainable City - Developing World" (16-19, September), Dalian, China.

Holmes, T. J., & Lee, S. (2010). Cities as six-by-six-mile squares: Zipf's law? In E. L. Glaeser (Ed.), *Agglomeration Economics* (pp. 105-131). Cambridge, MA: University of Chicago Press.

Hoppes, R. B. (1991). Regional versus industrial shift-share analysis—with help from the Lotus spreadsheet. *Economic Development Quarterly, 5*(3), 258-267.

IBGE (2006). *Estatísticas do século XX*. Rio de Janeiro, Brasil: IBGE. www.ibge.gov.br.

IBGE. Brazilian Institute of Geography and Statistics. www.ibge.gov.br.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics,* 115, 53-74. doi:10.1016/S0304-4076(03)00092-7

Ioannides, Y., & Overman, H. G. (2003). Zipf's law for cities: An empirical examination. *Regional Science and Urban Economics*, *33*(2), 127–137.

IPEA. Institute of Applied Economic Research. www.ipea.gov.br.

IPEA (2010a). Institute of Applied Economic Research. www.ipea.gov.br.

IPEA (2010b). Brasilia: Impactos econômicos da capital no centro-oeste e no País. *Working Paper #44*, Brasilia, Brazil: IPEA.

Issitt, M. L. (2010). *Brasília.* Our World: Brazil, 2-2.

Jacobs, J. (1969). *The economy of cities*. New York: Vintage.

Kamecke, U. (1990). Testing the rank size rule hypothesis with an efficient estimator. *Journal of Urban Economic*s, *27*(2), 222-231.

Katzman, M. T. (1975). Regional development policy in Brazil: The role of growth poles and development highways in Goiás. *Economic Development and Cultural Change, 24*(1), 75-107. Available at: http://www.jstor.org/stable/1152782

Keil, S. R. (1992). On the value of homotheticity in the shift-share framework. *Growth and Change*, *23*(4), 469-493.

Kinoshita, T., Kato, E., Iwao, K., & Yamagata, Y. (2008). Investigating the rank-size relationship of urban areas using land cover maps. *Geophysical Research Letters, 35*, L17405.

Kironde, J. M. L. (1993). Will Dodoma ever be the new capital of Tanzania? *Geoforum, 24*(4), 435-453.

Knudsen, D. C. (2000). Shift-share analysis: Further examination of models for the description of economic change. *Socio-Economic Planning Sciences*, *34*(3), 177-198.

Knudsen, D. C., & Barff, R. (1991). Shift-share analysis as a linear model. *Environment and Planning A*, *23*(3), 421–431.

Kohlsdorf, M. E., Kohlsdorf, G., & Holanda, F. de (2009). Brasília: Permanence and transformations. In V. del Rio & W. Siembieda (Eds.), *Contemporary urbanism in Brazil: Beyond Brasília* (pp. 42-64). Florida A&M University, Tallahassee, FL, USA: University Press of Florida.

Krugman, P. (1991a). *Geography and trade.* Cambridge, MA: MIT Press.

Krugman, P. (1991b). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), 483-499.

Krugman, P. (1994). *Urban concentration: The role of increasing returns and transport costs.* Proceedings of the World Bank Annual Conference on Development Economics.

Krugman, P. (1995). *Development, geography, and economic theory.* Cambridge, MA: MIT Press.

Krugman, P. (1996). Confronting the mystery of urban hierarchy. *Journal of the Japanese and International Economies, 10*(4), 399–418.

Krugman, P. (1998). Space: The final frontier. *Journal of Economic Perspectives, 12*(2), 161-174.

Krugman, P., & Venables, A. J. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics, 110*(4), 857–880.

Kubitschek, J. (1978). *Cinquenta anos em cinco.* 3º Volume da Coleção Meu Caminho para Brasília. Rio de Janeiro, Brasil: Bloch Editores S. A.

Kubitschek, J. (2000). *Por que construí Brasília*. Brasília, Brasil: Senado Federal. Retrieved February, 2011, from Biblioteca Digital do Senado Federal website: http://www2.senado.gov.br/bdsf/item/id/1039.

Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, *45*(1), 1-28.

Le Gallo, J., & Kamarianakis, Y. (2010). The evolution of regional productivity disparities in the European Union from 1975 to 2002: A combination of shift–share and spatial econometrics. *Regional Studies*, 1–17, iFirst article. doi: 10.1080/00343400903234662.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. London: CRC Press.

Levin, A., Lin, C-F., & Chu, C-S. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108, 1-24.

Levy, M. (2009). Gibrat's law for (all) cities: Comment. *American Economic Review, 99*(4), 1672–1675.

Lobo, E. M. L. (1996). Brazil since 1889. In B. A. Tenenbaum (Ed.), *Encyclopedia of Latin American history and culture* (pp. 424-432). New York: Simon & Schuster Macmillan.

Lovejoy, D. (1985). Islamabad, Dodoma and Brasilia: A tale of three cities. In A. Evin (Ed.), *The Expanding Metropolis: Coping with the urban growth of Cairo* (pp. 222-225). Singapore: Concept Media/the Aga Khan Award for Architecture. Retrieved July 26, 2011 from http://archnet.org/library/documents/one-document.jsp?document_id=2750

Loveridge, S. (1995). A practical approach to shift-share analysis. *Journal of the Community Development Society*, *26*(1), 111–124.

Loveridge, S., & Selting, A. C. (1998). A review and comparison of shift-share identities. *International Regional Science Review. 21*(1), 37-58.

Macalester College, Saint Paul, Minnesota, the Unites States. http://www.macalester.edu/

Macedo, R., & Barbosa, F. (1997). Brazil: Instability and macroeconomic policies. In M. J. F. Willumsen & E. G. da Fonseca (Eds.), *The Brazilian economy: Structure and performance in recent decades* (pp. 1-23). Miami, FL: North-South Center Press at the University of Miami.

Magalhães, A., Hewings, G. J. D., & Azzoni, C. R. (2005). Spatial dependence and regional convergence in Brazil. *Investigaciones Regionales*, *6*, 5-20.

Markusen, A. R., Noponen, H., & Driessen, K. (1991). International trade, productivity, and U.S. regional job growth: A shift-share interpretation. *International Regional Science Review*, *14*(1), 15-39.

Marshall, A. (1890). *Principles of economics*. London: Macmillan.

Matlaba, V. J. et al. (2010). *Agglomeration externalities and 1981-2006 regional growth in Brazil*. Paper presented at the 57th Annual North American Meetings of the Regional Science Association International (10-13, November), Denver, Colorado, the United States. **[Chapter 3 of this thesis]**

Matos, R. (2000). *Aglomerações urbanas, rede de cidades e desconcentração demográfica no Brasil.* Paper presented at 12nd Encontro Nacional de Estudos Populacionais, Caxambu, Minas Gerais, Brasil. http://www.abep.org.br.

Mayor, M., & López, A. J. (2008). Spatial shift-share analysis versus spatial filtering: An application to Spanish employment data. *Empirical Economics, 34*, 123–142. doi: 10.1007/s00181-007-0167-9.

McCann, P. (2001). *Urban and regional economics*. New York: Oxford University Press.

McCann, P. (2003). Geography, trade and growth: Problems and possibilities for the New Zealand economy. *Working Paper #03/03*, New Zealand Treasury, Wellington, New Zealand.

McCann, P. (2008). Globalization and economic geography: The world is curved, not flat. *Cambridge Journal of Regions, Economy and Society, 1,* 351–370. doi:10.1093/cjres/rsn002.

McCann, P. (2009a). Economic geography, globalisation and New Zealand's productivity paradox. *New Zealand Economic Papers*, *43*(3), 279-314.

McCann, P. (2009b). *Economic geography, globalisation and New Zealand's productivity paradox* [PowerPoint Slides]. Motu Public Policy Seminar, December. Retrieved July 19, 2011 from http://www.motu.org.nz/publications/detail/economic_geography_globalisation_and_new_zealands_productivity_paradox.

McCann, P., Mameli, F. M., & Faggian, A. (2008). Employment growth in Italian local labour systems: Issues of model specification and sectoral aggregation. *Spatial Economics Analysis*, *3*(3). 1-18.

McDonough, C. C., & Sihag, B. S. (1991). The incorporation of multiple bases into shift-share analysis. *Growth and Change, 22*(1), 1-9.

Mellinger, A. D., Sachs, J. D., & Gallup, J. L. (2003). Development and underdevelopment. In G. L. Clark, M. P. Feldman, & M. S. Gertler (Eds.), *The Oxford Handbook of Economic Geography* (pp. 169-194). New York: Oxford University Press.

Mello, E. K. de., Wilheim, J., & Costa, L. (1960). *History, city planning, architecture, building* (2nd ed.) (Parallel text in Portuguese and English). São Paulo, Brazil: M. Gruenwald.

Mitchell, W., Myers, J., & Juniper, J. (2005). Extending shift-share analysis to account for spatial effects: A study using Australian census data. *Working Paper #05-19*, Centre of Full Employment and Equity, University of Newcastle, Callaghan, Australia.

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, *37*(1/2), 17-23. Available at: http://www.jstor.org.ezproxy.waikato.ac.nz/stable/pdfplus/2332142.pdf

Mori, T., Nishikimi, K., & Smith, T. E. (2005). A divergence statistic for industrial localization. *The Review of Economics and Statistics, 87*(4), 635-651.

Mori, T., Nishikimi, K., & Smith, T. E. (2008). The number-average size rule: A new empirical relationship between industrial location and city size. *Journal of Regional Science, 48*(1), 165–211.

Moser, S. (2010). City profile: Putrajaya: Malaysia's new federal administrative capital. *Cities, 27*(4), 285-297. doi:10.1016/j.cities.2009.11.002

Mossi, M. B., Aroca, P., Fernández, I. J., & Azzoni, C. R. (2003). Growth dynamics and space in Brazil. *International Regional Science Review, 26*(3), 393-418.

Moura, N. J., & Ribeiro, M. B. (2006). Zipf law for Brazilian cities. *Physica A,* 367, 441-448.

Nazara, S., & Hewings, G. J. D. (2004). Spatial structure and taxonomy decomposition in shift-share analysis. *Growth and Change, 35*(4), 476-490.

The New School (2010). *Vilfredo Pareto, 1848-1923*. History of Economic Thought, The New School. http://homepage.newschool.edu/het/

Newton, H. J. (Ed.) (2001). *Stata technical Bulletin: A publication to promote communication among Stata users*. Texas, USA: Department of Statistics, Texas A & M University.

Nishivama, Y., Osada, S., & Morimune, K. (2008). OLS estimation and the *t* test revisited in rank-size rule regression. *Journal of Regional Science, 48*(4), 691–715.

Nitsch, V. (2005). Zipf zipped. *Journal of Urban Economics, 57*(1), 86–100.

Nogueira, C. A. G., & Lopes, D. A. F. (2008). *Employment growth in Ceará: A shift-share analysis (2000-2005).* Paper presented at the 8th Congress of the Regional Science Association International (16-19, March), São Paulo, Brazil.

Noponen, H., Markusen, A. R., & Driessen, K. (1998). International trade and shift-share analysis: A response to Dinc and Haynes. *Economic Development Quarterly, 11*(4), 344-350.

Oliveira, M. de (2006). O ISEB e a construção de Brasília: Correspondências míticas. *Sociedade e Estado*, *21*(2), 487-512. Available from http://www.scielo.br/pdf/se/v21n2/a08v21n2.pdf. <Accessed February 23, 2011>.

Oliveira, R. A. de (2008). *Brasília e o paradigma modernista: Planejamento urbano no moderno atraso*. Unpublished MA dissertation, Faculdade de Arquitectura e Urbanismo, Universidade de São Paulo, São Paulo, Brazil.

Oliveira, L. L. (2009). *O Brasil de JK - Brasília, a meta-síntese.* Available from http://cpdoc.fgv.br/producao/dossies/JK/artigos/Brasilia/Meta-sintese. <Accessed March 10, 2011>.

Overman, H. G., & Ioannides, Y. M. (2001). Cross-sectional evolution of the US city size distribution. *Journal of Urban Economics*, *49*(3), 543-566.

Ozgen, C., Nijkamp, P., & Poot, J. (2010). The effect of migration on income growth and convergence: Meta-analytic evidence. *Papers in Regional Science*, *89*(3), pp. 537-561.

Parr, J. B. (1985). A note on the size distribution of cities over time. *Journal of Urban Economics*, *18*(2), 199-212.

Pisati, M. (2010). Spatreg, spatlsa, spatwmat: Stata module to estimate the spatial regression model, local indicators of spatial association, and spatial weight matrix. www.stata.com.

Poot, J. (2004). Peripherality in the global economy. In J. Poot (Ed.), *On the edge of the global economy* (Vol. IX) (pp. 3-26). Cheltenham, England: Edward Elgar Publishing Limited.

Porter, M. (1990). *The competitive advantage of nations.* New York: Free Press.

Portnov, B. A., & Benguigui, L. (2010). *Does Zipf's Law hold for primate cities? Some evidence from a discriminant analysis of world countries.* Paper presented at the 50th Anniversary European Congress of the Regional Science Association International "Sustainable Regional Growth and Development in the Creative Knowledge Economy" (19-23, August), Jonkoping, Sweden.

Prideaux, J. (2009, November 14). Getting it together at last: A special report on business and finance in Brazil. *The Economist*, 2-16.

Pumain, D., & Moriconi-Ebrard, F. (1997). City size distributions and metropolisation. *Geojournal*, *43*(4), 307-314.

Rabelo, R. F. (2003). Plano de metas e consolidação do capitalismo industrial no Brasil. *E & G Economia e Gestão, 2 & 3* (4 & 5), 44-55.

Rangel, L. A., Andrade, J., & Divino, J. A. (2008). *Economic growth and income inequality in Brazil: Analyzing the comparable minimum areas.* Retrieved March 21, 2009 from European Economics and Finance Society Prague Papers website: www.eefs.eu/conf/Prague/Papers/S602_Andrade_Divino_Rangel.pdf.

Ray, M.A., & Harvey, J. T. (1995). Employment changes in the European Economic Community: A shift-share analysis. *Review of Regional Studies, 25*(1), 97-110.

Resende, G. M. (2011). Multiple dimensions of regional economic growth: The Brazilian case, 1991-2000. *Papers in Regional Science*, *90*(3). doi:10.1111/j.1435-5957.2010.00336.x

Resende, M. (2004). Gibrat's law and the growth of cities in Brazil: A panel data investigation. *Urban Studies*, *41*(8), 1537-1549.

Road Transportation Information website of the Ministry of Transportation of Brazil: http://www.transportes.gov.br/bit/inrodo.htm. <Accessed March, 2009>.

Robinson, A. J. (1973). A planned new town: The Canberra experiment. *Land Economics*, *49*(3), 361-366.

Rolim, C. (2008). *Decomposing the GDP growth of the great Brazilian regions: Dynamic shift-share on recent Brazilian economic cycles.* Paper presented at the 8th

Congress of the Regional Science Association International (16-19, March), São Paulo, Brazil.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy, 94*(5), 1002-1037.

Rosen, K. T., & Resnick, M. (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics, 8*(2), 165–186.

Rossi-Hansberg, E., & Wright, M. (2007). Urban structure and growth. *Review of Economic Studies, 74*(2), 597-624.

Sanders, T. G. (1973). Urban growth and public policy in Brasília's satellite cities. *East Coast South America Series, 17*(6), 1-9.

Sandoval, M. A. L. (n. d.). *Breve histórico sobre a evolução do planejamento nacional de transportes.* DNIT, Brasília, Brasil: Ministério dos Transportes. Available from http://www.dnit.gov.br/planejamento-e-pesquisa/historico-do-planejamento-de-transportes. <Accessed throughout February, 2011>.

Sandroni, P. (Ed.) (1994). *Novo Dicionário de Economia.* São Paulo, Brasil: Editora Best Seller.

Santos, M. A. dos (2008). *Brasília, o lago Paranoá e o tombamento: Natureza e especulação na cidade modernista*. Unpublished MA dissertation, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, Brasil.

Selting, A. C., & Loveridge, S. (1994). Testing dynamic shift-share. *Regional Science Perspectives*. *24*(1), 23-41.

Serra, M. V., Dowall, D. E., Motta, D., & Donovan, M. (2005). Urban land markets and urban land development - An examination of three Brazilian cities: Brasília, Curitiba and Recife. *Working Paper #2004-03*, Institute of Urban and Regional Development, University of California, Berkeley, CA: The United States.

Shields, M. (2003). *Using employment data to better understand your local economy: Shift share analysis helps identify local growth engines.* In Understanding Economic Change in Your Community Series, Center for Economic and Community Development at Penn State. http://cecd.aers.psu.edu/ <Accessed August, 2010>.

Silva, M. V. B., & Silveira-Neto, R. M. (2007). Determinantes da localização industrial no Brasil e geografia econômica: Evidências para o período pós-real. *Revista Economia, 8,* 269-288.

Silva, N. do V. (2008). Brazilian society: Continuity and change, 1930-2000. In L. Bethell (Ed.), *The Cambridge history of Latin America* (Vol. IX) (pp. 455-544). Cambridge, England: Cambridge University Press.

Silveira-Neto, R., & Azzoni, C. R. (2006). Location and regional income disparity dynamics: The Brazilian case. *Papers in Regional Science, 85*(4), 599-613.

Simon, H. (1955). On a class of skew distribution functions. *Biometrika, 42*(3/4), 425-440.

Skidmore, T. (1967). *Politics in Brazil, 1930-64: An experiment in democracy.* New York: Oxford University Press.

Skidmore, T. (2010). *Brazil: Five centuries of change* (2nd Ed.). New York: Oxford University Press.

Skilling, D. (2001). *The importance of being enormous: Towards an understanding of the New Zealand economy.* Economic Transformation Project, New Zealand Treasury, Wellington, New Zealand.

Skouras, S., & Ioannides, Y. (2010). *Lognormality for most cities and power law for most of the population.* Available at: http://ase.tufts.edu/econ/papers/index.html

Snyder, D. E. (1964). Alternate perspectives on Brasilia. *Economic Geography, 40*(1), 34-45. http://www.jstor.org/stable/142172. <Accessed February 18, 2011>.

Song, S., & Zhang, K. H. (2002). Urbanisation and city size distribution in China. *Urban Studies, 39*(12), 2317–2327.

Soo, K. T. (2005). Zipf's Law for cities: A cross-country investigation. *Regional Science and Urban Economics, 35*(3), 239–263.

Soo, K. T. (2007). Zipf's law and urban growth in Malaysia. *Urban Studies, 44*(1), 1–14.

Stata (1996-2012). StataCorp LP. http://www.stata.com/

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association, 69*(347), 730-737. Available at: http://www.jstor.org/stable/2286009

Story, E. F. (2006). *Constructing development: Brasilia and the making of modern Brazil.* Unpublished PhD Thesis, Vanderbilt University, Nashville, Tennessee, the United States.

Tenenbaum, B. A. (1996) (Ed.). *Encyclopedia of Latin American history and culture*. New York: Simon & Schuster Macmillan.

Tranter, P. (1990). The role of the national capital spirit in Canberra's development. *Focus, 40*(4), 18-26.

Varian, H. R. (1992). *Microeconomic analysis* (3rd Ed.). New York: W. W. Norton & Company, Inc.

Veiga, J. E. (2003). Cidades imaginárias: O Brasil é menos urbano do que se calcula (2nd ed.). Campinas, São Paulo, Brasil: Editora Autores Associados.

Venables, A. J. (2009). Rethinking economic growth in a globalizing world: An economic geography lens. *African Development Review, 21*(2), 331–351. doi: 10.1111/j.1467-8268.2009.00212.x

Wang, J-F, Liu, X-H, & Chen, H-Y. (2010). *Construction of a cities evolution tree, with applications.* Paper presented at the 50th Anniversary European Congress of the Regional Science Association International "Sustainable Regional Growth and Development in the Creative Knowledge Economy" (19-23, August), Jonkoping, Sweden.

Williams, R. J. (2007). Brasília after Brasília. *Progress in Planning, 67,* 301–366.

Wooldridge, J. M. (2002). *Solutions manual and supplementary materials for econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

www.citybrazil.com.br

Xu, Z., & Zhu, N. (2009). City size distribution in China: Are large cities dominant? *Urban Studies, 46*(10), 2159–2185.

Yasin, M., Alavi, J., Sobral, F., & Lisboa, J. (2010). A Shift-share analysis approach to understanding the dynamic of the Portuguese tourism market. *Journal of Travel & Tourism Marketing, 17*(4), 11-22.

Zellner, A. (1962). An efficient method of estimating Seemingly Unrelated Regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*(298), pp. 348-368. http://www.jstor.org/stable/2281644?seq=2

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology.* Cambridge, MA: Addison-Wesley.

Zipser, T., Mlek, M., & Zipser, W. (2010). *Zipf's law in hierarchically ordered open system.* Paper presented at the 50th Anniversary European Congress of the Regional Science Association International "Sustainable Regional Growth and Development in the Creative Knowledge Economy" (19-23, August), Jonkoping, Sweden.

Appendices

Appendix 3.1: Fluctuation in the Logarithm of Manufacturing Industry Employment Across Brazilian Regions

Figure 3.1a: Fluctuation in the Logarithm of Manufacturing Industry Employment, North Region

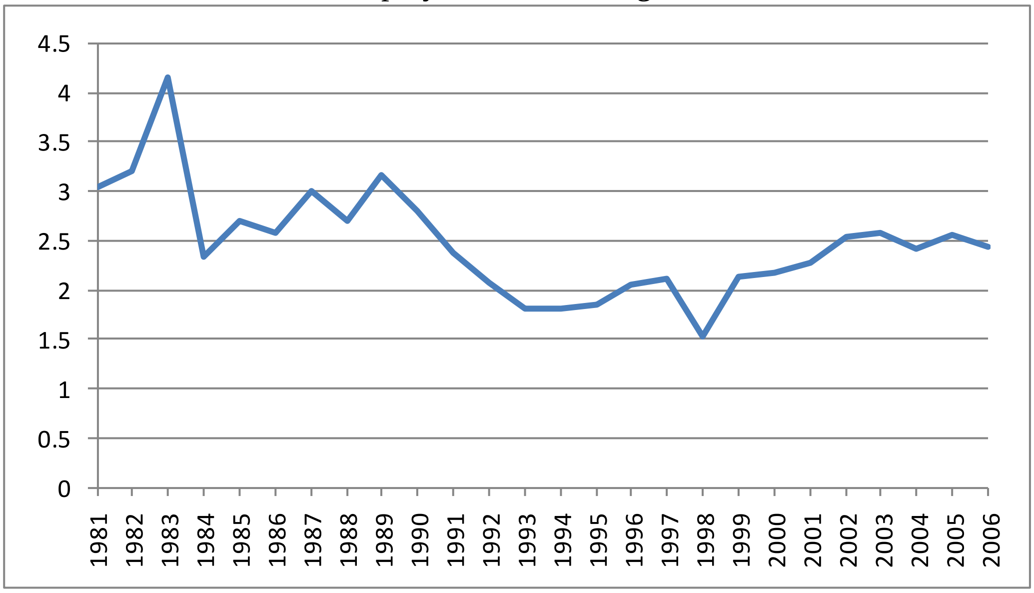


Figure 3.1b: Fluctuation in the Logarithm of Manufacturing Industry Employment, Northeast Region

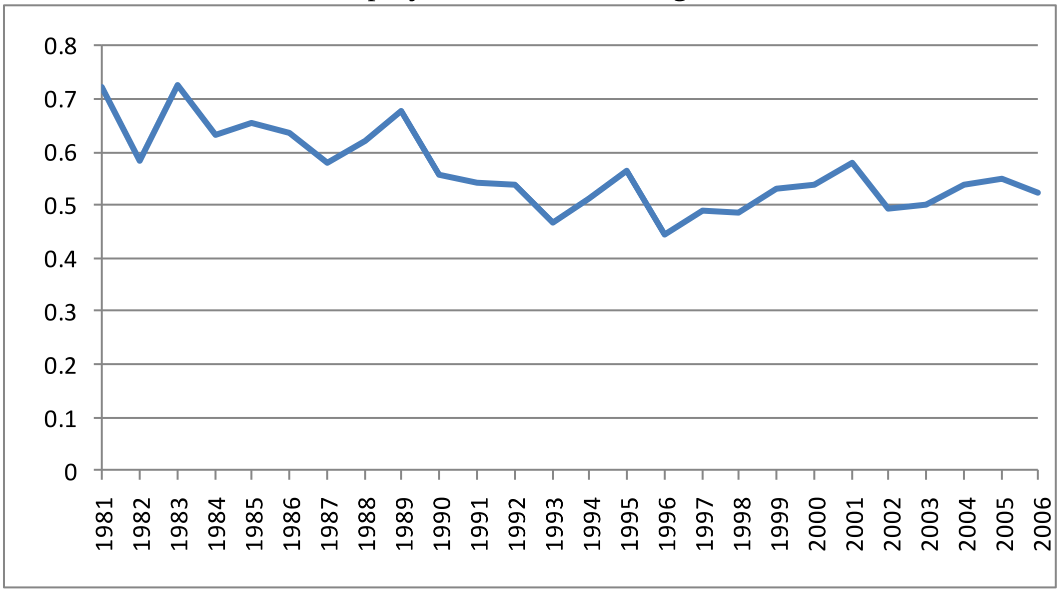


Figure 3.1c: Fluctuation in the Logarithm of Manufacturing Industry Employment, Middle-West Region

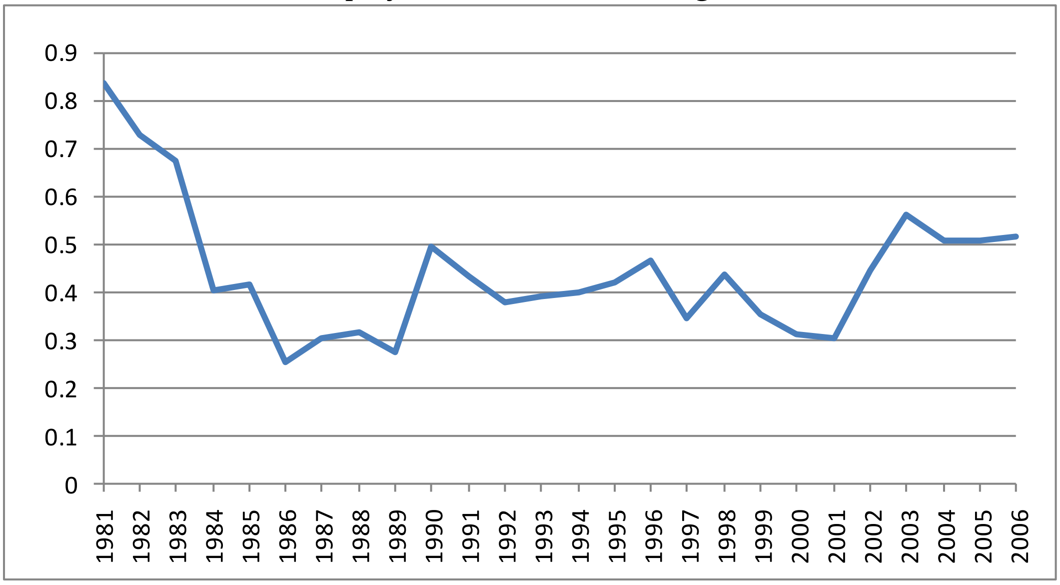


Figure 3.1d: Fluctuation in the Logarithm of Manufacturing Industry Employment, Southeast Region

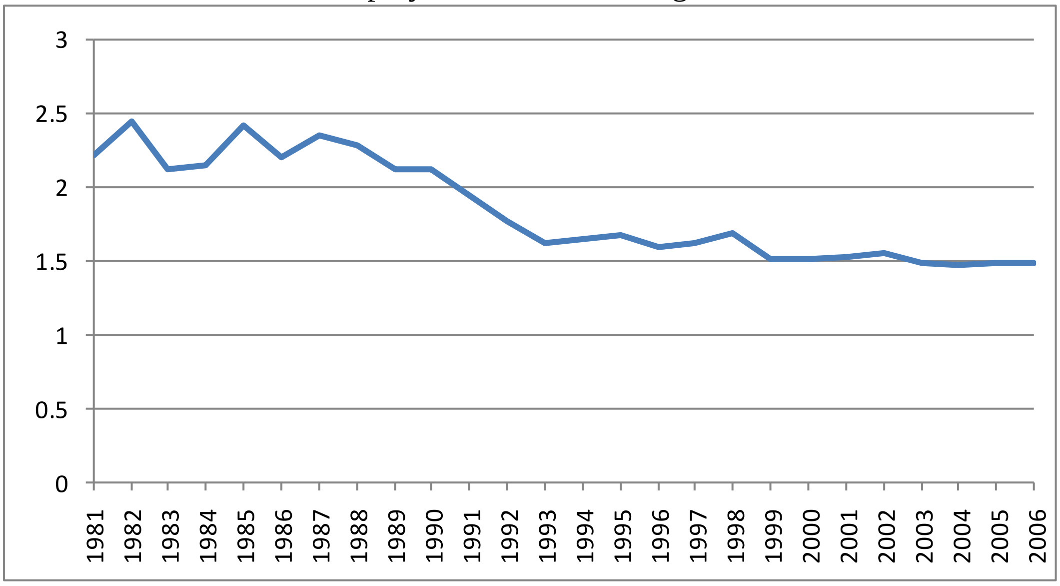
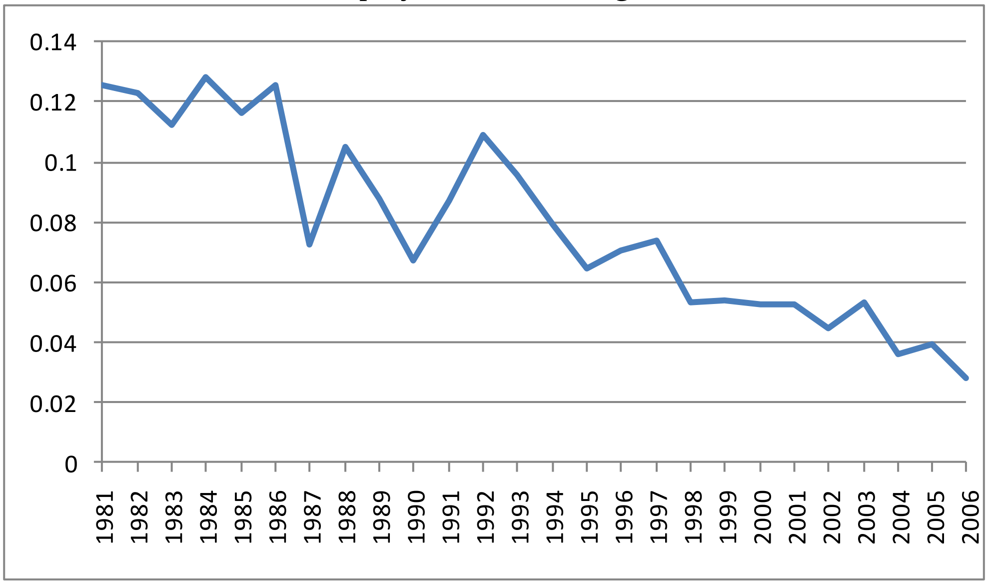


Figure 3.1e: Fluctuation in the Logarithm of Manufacturing Industry Employment, South Region



Appendix 3.2: Variables used in the analysis

Variable	Expression	Explanation	Cross-section Models	Annual Panel Models	Pooled Models
Dependent Variables					
Growth of Employment	$\mathrm{VDEP} = \mathrm{Growth} = \mathrm{Log}(\frac{E_t}{E_{t-1}})$ (as in Glaeser et. al. (1992))	The natural logarithm of states' manufacturing industry employment in *t* divided by states' manufacturing industry employment in *t*-1.	*t*=1995; *t*-1=1985.	*t* varies from 1981 to 2006.	---
Growth of Employment	$\mathrm{VDEP2} = y_{z,s} = \log\left(\frac{emp_{z,s,t}}{emp_{z,s,t-1}}\right) - \log\left(\frac{emp_{s,t}}{emp_{s,t-1}}\right)$ (as in Combes (2000))	The difference between the manufacturing industry *s* employment growth in the state *z* between *t*-1 and *t* and the national manufacturing industry employment growth in Brazil during the same period.	*t*=1995; *t*-1=1985.	*t* varies from 1981 to 2006.	---
Annual Average Growth of Employment	$Avgrwth_i$=ln(E1990/E1981)/9 $Avgrwth_i$=ln(E1998/E1990)/8 $Avgrwth_i$=ln(E2006/E1998)/8	The average annual growth of manufacturing industry employment, *E*, in each periods for each state *i*.	---	---	The periods are: 1981-90; 1990-98 and 1998-06; state *i*=1,...,26.
Independent Variables					
Specialisation	$\mathrm{Spec} = \frac{E_{z,s}/E_z}{E_s/E}$ (as in Glaeser et. al. (1992))	The ratio of manufacturing industry s state z employment share to Brazil manufacturing industry s employment share, in a given *t*.	*t*=1985.	*t* varies from 1981 to 2006.	*t* varies from 1981 to 2006.
Diversity	$\mathrm{Div}_{z,s} = \frac{1/\sum_{s'=1, s'\neq s}^{S}[\frac{emp_{z,s'}}{emp_z - emp_{z,s}}]^2}{1/\sum_{s'=1, s'\neq s}^{S}[\frac{emp_{s'}}{emp - emp_s}]^2}$ (as in Combes (2000))	Diversity is, for a given *t*, the inverse of an Herfindahl index of sectoral concentration based on the share of all sectors, except the one considered, where *S* is the total number of sectors and z is state. The 9 sectors *s* are: agriculture & fishing; commerce; construction; electricity, water & gas; finance; manufacturing; mining; services; and transportation & communications.	*t*=1985.	*t* varies from 1981 to 2006.	*t* varies from 1981 to 2006.

Competition	$\text{Comp} = \frac{N_{z,s}/E_{z,s}}{N_s/E_s}$ (as in Glaeser et. al. (1992))	The ratio of State z manufacturing industry establishment s to state z manufacturing industry employment s. This ratio is divided by its national indicator for a given *t*.	*t*=1985.	---	---
Varlag	$\text{Varlag}_t = \text{VDEP}_{t-1} = [\log\left(\frac{E_t}{E_{t-1}}\right)]_{t-1}$	The growth of manufacturing industry employment in the previous period in each year *t*, thereby starting in 1983.	---	*t* varies from 1981 to 2006.	---
Varlag1	$\text{Varlag1} = \text{VDEP2}_{t-1} = (y_{z,s})_{t-1}$	The difference between the employment growth of manufacturing sector *s* in state z between *t*-1 and *t* and the national manufacturing industry employment growth in Brazil during the same period, for the previous period in each year *t*, thereby starting in 1983.	---	*t* varies from 1981 to 2006.	---
Education	Education = LIT_t	The percentage of literate people of 15 years old or more.	*t*=1985.	*t* varies from 1981 to 2006.	*t* varies from 1981 to 2006.
Change in Education	$CHANGELIT_{t-1,t}=\ln(LIT_t/LIT_{t-1})$	The change of LIT between t-1 and t.	*t*=1995; *t*-1=1985.	*t* varies from 1981 to 2006.	*t* varies from 1981 to 2006.
Annual Change in Education	$ACHANGELIT_{t-1,t}=\ln(LIT_t/LIT_{t-1})$	The Annual Change of LIT.	---	*t* varies from 1981 to 2006.	---
Education Change within the Pools	$PCHANGELIT_{t-1,t}=\ln(LIT1990/LIT1981)$ $PCHANGELIT_{t-1,t}=\ln(LIT1998/LIT1990)$ $PCHANGELIT_{t-1,t}=\ln(LIT2006/LIT1998)$	The change of LIT within the time pools.	---	---	The pools are: Pool 1: 1981-90; Pool 2: 1990-98 and Pool 3: 1998-06.
Time dummies	Dummy = $dyear_t$	A dummy variable which assumes 1 in each *t* and 0 otherwise.	---	*t* varies from 1981 to 2006.	
Time dummies within the pools	Dummy=td_i	A dummy variable which assumes 1 in each time pool *i* (*i*=1, 2, 3), and 0 otherwise.	---	---	Pool 1: 1981-90; Pool 2: 1991-98 and Pool 3: 1999-06.
States' dummies	Dummy=$state_i$	A dummy variable which assumes 1 in each	---	*i* varies from 1	*i* varies from 1 to

		state *i* for all of its *t*, and 0 otherwise.		to 26; *t* from 1981 to 2006.	26; *t* is pooled as aforementioned.
Distance	Distance=$D_{i,SP}$ (as in Henderson et. al. (1992))	The distance in kilometres, which is constant over the years in each state, from centres of each *i* state's capital to São Paulo's city centre (the major country's market)[144].	*i* varies from 1 to 26.	*i* varies from 1 to 26.	*i* varies from 1 to 26.
Number of Observations	---	---	26	624	78

[144]The distance is based on the shortest ways with today's Federal, State and Municipal roads infrastructure (for São Paulo state itself distance is set equal to 1 km). That the infrastructure was similar at the beginning of the period considered compared with the end is assumed. This assumption is important especially for annual models because they deal with a long period in which Brazil moved through a process of long-run development.

Appendix 3.3: Descriptive Statistics

Variable	Model	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
VDEP	Cross-section	26	0.2898	0.3260	-0.1545	1.1083
SPECIALISATION	Cross-section	26	0.7153	0.3626	0.2441	1.8263
COMPETITION	Cross-section	26	1.5799	0.7378	0.5394	4.0913
DIVERSITY	Cross-section	26	1.0048	0.4577	0.4941	2.8391
DISTANCE	Cross-section	26	2065	1258.01	1	4756
LITERATE	Cross-section	26	76	14.08	49.46	92.98
CHANGELIT	Cross-section	26	0.0864	0.0768	-0.0233	0.2745
VDEP2	Cross-section	26	0.2187	0.3260	-0.2256	1.0372
VDEP	Panel	650	0.0457	0.1726	-0.9991	1.8803
VARLAG	Panel	624	0.0470	0.1748	-0.9991	1.8803
SPECIALISATION	Panel	676	0.7535	0.3474	0.1668	1.8484
DIVERSITY	Panel	676	0.9411	0.2181	0.4780	2.8391
DISTANCE	Panel	676	2065	1234.49	1	4756
LITERATE	Panel	676	80.73	11.81	45.42	96.24
ACHANGELIT	Panel	650	0.0073	0.0211	-0.1004	0.0935
VDEP2	Panel	650	0.0201	0.1660	-0.9624	1.8628
VARLAG1	Panel	624	0.0208	0.1681	-0.9624	1.8628
LOGSPECIALISATION	Panel	676	-0.3801	0.4373	-1.7912	0.6143
LOGDIVERSITY	Panel	676	-0.0844	0.2157	-0.7381	1.0435
LOGDISTANCE	Panel	676	7.1773	1.5950	0	8.4672
LOGLITERATE	Panel	676	4.3791	0.1597	3.8160	4.5668
AVGRWTH	Pooled	78	0.0453	0.0476	-0.0865	0.1822
SPECIALISATION	Pooled	78	0.7664	0.3372	0.3063	1.7732
DIVERSITY	Pooled	78	0.9205	0.1272	0.5430	1.1974
PCHANGELIT	Pooled	78	0.0607	0.0458	-0.0368	0.1786
DISTANCE	Pooled	78	2065	1241.57	1	4756
LOGSPECIALISATION	Pooled	78	-0.3536	0.4176	-1.1831	0.5728
LOGDIVERSITY	Pooled	78	-0.0933	0.1500	-0.6107	0.1801
LOGDISTANCE	Pooled	78	7.18	1.60	0	8.47

Appendix 4.1: Geographical Definition of Urban Areas in 2008

State (s)	UA - Urban Area	Municipalities in 2008
Acre	UA1 - Rio Branco	Rio Branco
Amazonas	UA2 - Manaus	Manaus; Manacapuru; Careiro; Rio Preto da Eva; Careiro da Várzea; Manaquiri; Iranduba
Amazonas	UA3 - Itacoatiara	Itacoatiara; Silves; Urucurá; Urucurituba; Itapiranga; São Sebastião de Uatumã
Amazonas	UA4 - Parintins	Parintins; Barreirinha
Amazonas	UA5 - Tefé (former Teffé)	Tefé (former Teffé); Uarini; Alvarães
Amapá	UA6 - Macapá	Macapá; Mazagão (formerly in Pará State); Santana
Pará	UA7 - Belém	Belém
Pará	UA8 - Goianésia do Pará	Goianésia do Pará; Tucuruí; Jacundá; Tailândia; Breu Branco
Pará	UA9 - Abaetetuba (former Abaete)	Abaetetuba (former Abaete); Moju; Igarapé-Miri (former Igarapé-miry); Barcarena
Pará	UA10 - Ananindeua	Ananindeua; Igarapé-Açu (former João Pessoa); Santa Isabel do Pará (former Santa Isabel); Castanhal; Bujaru; São Francisco do Pará; Nova Timboteua; Benevides; Santo António do Tauá; Santa Maria do Pará; Peixe-Boi; Bonito; Terra Alta; Santa Bárbara do Pará; Marituba
Pará	UA11 - Breves	Breves
Pará	UA12 - Cametá (misspelled as Caneta in 1907-1912 IBGE database)	Cametá; Limoeiro do Ajuru
Pará	UA13 - Itaituba	Itaituba; Senador José Porfirio (former Souzel); Altamira (former Xingú); Rurópolis; Pacajá; Uruará; Medicilândia; Brasil Novo; Vitória do Xingu; Anapu; Placas
Pará	UA14 - Marabá	Marabá; São João do Araguaia; Itupiranga; Parauapebas; Brejo Grande do Araguaia; Bom Jesus do Tocantins; Eldorado dos Carajás; Eldorado dos Carajás; São Domingos do Araguaia; Palestina do Pará; Abel Figueiredo; Nova Ipixuna
Pará	UA15 - Paragominas	Paragominas; Dom Eliseu; Ulianópolis; Ipixuna do Pará; Nova Esperança do Piriá
Pará	UA16 - Santarém	Santarém; Belterra
Rondónia	UA17 - Porto Velho	Porto Velho
Rondónia	UA18 - Guajará-Mirim	Guajará-Mirim
Rondónia	UA19 - Ji-Paraná	Ji-Paraná; Alvorada D'Oeste; Presidente Médici; Ouro Preto do Oeste; Machadinho D'Oeste; Ministro Andreazza; Theobroma; Urupá; Vale do Paraíso; Mirante da Serra; Teixeirópolis; Vale do Anari; Nova União
Roraima	UA20 - Boavista (former Bôa Vista do Rio Branco in Amazonas state)	Boavista (former Bôa Vista do Rio Branco in Amazonas state)
Roraima	UA21 - Caracaraí	Caracaraí (former Moura in former Amazonas state)
Tocantins	UA22 - Palmas	Palmas
Tocantins	UA23 - Araguaína (formerly in Goiaz State)	Araguaína (formerly in Goiaz State); Filadélfia (formerly in Goiaz State); Babaçulândia (formerly in Goiaz State); Arapoema (formerly in Goiaz State); Wanderlândia (formerly in Goiaz State); Nova Olinda (formerly in Goiaz State); Piraquê; Carmolândia; Aragominas; Muricilândia; Santa Fé do Araguaia; Pau D'Arco; Bandeirantes do Tocantins
Tocantins	UA24 - Gurupi (formerly in Goiaz State)	Gurupi (formerly in Goiaz State); Peixe (former Santa Terezinha in former Goiaz State); Dueré (formerly in Goiaz State); Formoso do Araguaia (formerly in Goiaz State); Figueirópolis (formerly in Goiaz State); Aliança do Tocantins; Cariri do Tocantins; Sucupira; Crixás do Tocantins
Alagoas	UA25 - Maceió	Maceió; Rio Largo (former Santa Luzia do Norte); Satuba; Santa Luzia do Norte; Coqueiro Seco
Alagoas	UA26 - Atalaia	Atalaia; Pilar; Marechal Deodoro (former Alagôas)
Alagoas	UA27 - São Miguel dos Campos	São Miguel dos Campos
Alagoas	UA28 - Arapiraca	Arapiraca; Junqueiro; Limoeiro de Anadia (former Limoeiro); Girau do Ponciano; Campo Grande; Feira Grande; Jaramataia; Lagoa da Canoa; Taquarana; Coité do Nóia; Craíbas
Alagoas	UA29 - Palmeira dos Índios	Palmeira dos Índios; Quebrangulo (former Victoria); Igaci; Paulo Jacinto; Belém; Tanque d'Arca; Mar Vermelho; Estrela de Alagoas

Alagoas	UA30 - Penedo	Penedo; Piaçabuçu (former Piassabussú); Igreja Nova (former Triumpho); Feliz Deserto
Bahia	UA31 - Salvador	Salvador; Camaçari (former Abrantes); Itaparica; São Francisco do Conde (former São Francisco); Candeias; Lauro de Freitas; Simões Filho; Vera Cruz; Salinas da Margarida; Dias D'Avila; Madre de Deus; Saubara
Bahia	UA32 - Porto Seguro	Porto Seguro; Santa Cruz Cabrália (former Santa Cruz); Itagimirim; Guaratinga; Eunápolis; Itabela
Bahia	UA33 - Alagoinhas	Alagoinhas; Catu (former Sant'Anna do Catú); Entre Rios; Inhambupe; Ouriçangas; Aramari; Teodoro Sampaio (former Theodoro Sampaio); Araçás
Bahia	UA34 - Tucano	Tucano; Nova Soure (former Soure); Ribeira do Pombal (former Pombal); Cipó; Araci (former Aracy); Ribeira do Amparo (former Amparo)
Bahia	UA35 - Bom Jesus da Lapa	Bom Jesus da Lapa; Paratinga (former Rio Branco); Carinhanha; Riacho de Santana; Malhada; Sítio do Mato
Bahia	UA36 - Brumado (former Bom Jesus dos Meiras)	Brumado (former Bom Jesus dos Meiras); Ituaçu; Livramento de Nossa Senhora (former Livramento); Aracatu; Dom Basílio; Presidente Jânio Quadros; Tanhaçu; Maetinga
Bahia	UA37 - Conceição do Coité	Conceição do Coité; Riachão do Jacuípe (former Riachão de Jacuhype); Santaluz; Valente; Retirolândia; Ichu; Teofilândia; São Domingos; Nova Fátima; Barrocas
Bahia	UA38 - Feira de Santana (former Feira)	Feira de Santana (former Feira); Irará; Coração de Maria; Santo Amaro; Serra Preta; Tanquinho; António Cardoso; Anguera; Candeal; Santa Bárbara; Santanópolis; Conceição do Jacuipe; Amélia Rodrigues
Bahia	UA39 - Guanambi	Guanambi; Caetité (former Caitité); Palmas de Monte Alto (former Monte Alto)
Bahia	UA40 - Ilhéus (former Ilheios)	Ilhéus (former Ilheios); Itabuna; Una; Ubaitaba; Coaraci; Uruçuca; Itajuipe; Ibicaraí; Floresta Azul; Aurelino Leal; Itapé; Buerarema; Arataca; São José da Vitória
Bahia	UA41 - Jacobina	Jacobina; Saúde; Miguel Calmon; Várzea do Poço; Serrolândia; Caldeirão Grande; Caém; Mirangaba; Várzea Nova; Capim Grosso; Ourolândia; Quixabeira
Bahia	UA42 - Jequié	Jequié; Maracás; Itiruçu (former Itirussú); Jaguaquara; Manoel Vitorino; Itagi; Jitaúna; Lajedo do Tabocal
Bahia	UA43 - Juazeiro (former Joazeiro)	Juazeiro (former Joazeiro); Curaçá; Jaguarari; Sobradinho
Bahia	UA44 - Santo Antônio de Jesus (former Santo Antônio)	Santo Antônio de Jesus (former Santo Antônio); Nazaré (former Nazareth); Aratuipe (former Aratuhype); Laje; São Miguel das Matas (former São Miguel); Conceição do Almeida (former Affonso Pernna); São Felipe; Muniz Ferreira; Varzedo
Bahia	UA45 - Teixeira de Freitas	Teixeira de Freitas; Alcobaça; Prado; Itanhém; Medeiros Neto; Lajedão; Ibirapuã; Itamaraju
Bahia	UA46 - Vitória da Conquista (former Conquista)	Vitória da Conquista (former Conquista); Itambé; Barra do Choça; Planalto; Anagé; Cândido Sales
Ceará	UA47 - Fortaleza	Fortaleza; Caucaia (former Soure); Aquiraz; Maranguape; Pacatuba; Maracanaú; Eusébio; Itaitinga
Ceará	UA48 - Sobral	Sobral; Massapé; Meruoca; Cariré; Alcântaras; Groaíras; Forquilha
Ceará	UA49 - Crato	Crato; Barbalha; Caririaçu (former São Pedro do Crato, then São Pedro do Carirí); Santana do Cariri (former Sant'Anna do Cariry); Juazeiro do Norte (former Juazeiro); Farias Brito; Nova Olinda
Ceará	UA50 - Iguatu	Iguatu; Jucás (former São Mateus); Várzea Alegre; Icó; Acopiara (former Afonso Pena); Cedro; Cariús; Orós
Ceará	UA51 - Quixadá	Quixadá; Quixeramobim; Itapiúna; Banabuiú; Ibicuitinga; Ibaretama; Choró
Ceará	UA52 - Limoeiro do Norte (former Limoeiro)	Limoeiro do Norte (former Limoeiro); Quixeré (former Quixará); Russas (former São Bernardo das Russas); Morada Nova; São João do Jaguaribe
Ceará/Piaui	UA53 - Tianguá (former Tyanguá)	Tianguá (former Tyanguá); Viçosa do Ceará (former Viçosa); Ubajara; Frecheirinha; São João da Fronteira (in Piaui state)
Maranhão	UA54 - São Luis	São Luis; Alcantara; Paço do Lumiar; São Jose de Ribamar; Raposa
Maranhão	UA55 - Bacabal	Bacabal; Pedreiras; São Luis Gonzaga do Maranhão (former São Luis Gonzaga); Vitorino Freire; Lago da Pedra; Lima Campos; São Mateus do Maranhão; Olho D'Agua das Cunhãs; Paulo Ramos; Lago do Junco; Igarapé Grande; Bernardo do Mearim; Trizidela do Vale; Lago dos Rodrigues; Bom Lugar; Alto Alegre do Maranhão

Maranhão	UA56 - Balsas (former Santo Antonio de Balsas)	Balsas (former Santo Antonio de Balsas); Riachão; Carolina
Maranhão	UA57 - Presidente Dutra	Presidente Dutra; Barra do Corda; Dom Pedro; Gonçalves Dias; São Domingos do Maranhão; Governador Eugênio Barros
Maranhão	UA58 - Codó	Codó; Caxias; Coroatá; Coelho Neto; Timbiras (former Monte Alegre); Santo António dos Lopes; Governador Archer; Peritoró; Capinzal do Norte
Maranhão	UA59 - Imperatriz	Imperatriz; João Lisboa; Açailândia; Governador Edison Lobão; São Francisco do Brejão; Itinga do Maranhão; Bom Jesus das Selvas
Maranhão	UA60 - Pinheiro	Pinheiro; São Bento (former São Bento dos Perizes, then São Bento dos Peris); Peri Mirim (former Macapá); Bequimão; Palmeirândia
Paraiba	UA61 - João Pessoa	João Pessoa; Santa Rita; Cabedelo; Bayeux
Paraiba	UA62 - Campina Grande	Campina Grande
Paraiba	UA63 - Patos	Patos; São Mamede; Malta; São José de Espinharas; Santa Teresinha; São José do Bonfim; Cacimba de Areia; Quixabá
Pernambuco	UA64 - Recife	Recife; Olinda; Igarassú (former Iguarassú); Joboatão dos Guararapes (former Joboatão); São Lourenço da Mata (former São Lourenço); Cabo de Santo Agostinho (former Cabo); Paulista; Moreno (former Morenos); Ilha de Itamaracá; Itapissuma; Abreu e Lima; Camaragibe
Pernambuco	UA65 - Caruaru	Caruaru; Altinho; Bezerros; São Caitano; Agrestina; Santa Cruz do Capibaribe; Toritama; Riacho das Almas
Pernambuco	UA66 - Arcoverde (former Rio Branco)	Arcoverde (former Rio Branco); Sertânia (former Alagoa de Baixo); Buíque; Pedra; Pesqueira (former Cimbres); Alagoinha; Tupanatinga; Venturosa
Pernambuco	UA67 - Belo Jardim	Belo Jardim; São Bento do Una (former São Bento); Sanharó
Pernambuco	UA68 - Carpina (former Floresta dos Leões)	Carpina (former Floresta dos Leões); Paudalho (former Pão d'Alho); Limoeiro; Nazaré da Mata (former Nazareth); Lagoa do Itaenga; Tracunhaém; Lagoa do Carro; Araçoiaba
Pernambuco	UA69 - Ouricuri (former Ouricury)	Ouricuri (former Ouricury); Parnamirim (former Leopoldina); Bodocó (former Granito); Exu (former Novo Exú); Araripina (former São Gonçalo); Trindade; Ipubi; Santa Cruz
Pernambuco/Bahia	UA70 - Petrolina	Petrolina; Casa Nova (former São José da Casa Nova) (in Bahia state); Lagoa Grande
Pernambuco/Ceará	UA71 - Salgueiro	Salgueiro; Cabrobó; Serrita (former Serrinha); Belém de São Francisco (former Belém); Penaforte (in Ceará state); Jati (in Ceará state); Terra Nova; Verdejante
Pernambuco	UA72 - Serra Talhada (former Villa Bella)	Serra Talhada (former Villa Bella); São José do Belmonte (former Belmonte); Flores; Floresta; Triunfo (former Triumpho); Mirandiba; Betânia; Calumbi
Pernambuco	UA73 - Vitória de Santo Antão (former Victoria)	Vitória de Santo Antão (former Victoria); Escada; Amaraji (former Amaragy); Gravatá; Glória do Goitá (former Gloria de Goytá); Pombos; Primavera; Chá Grande
Piaui/Maranhão	UA74 - Teresina	Teresina; Timon (Former Flôres) (in Maranhão state)
Piaui	UA75 - Parnaiba	Parnaiba; Ilha Grande
Rio Grande do Norte	UA76 - Natal	Natal; Macaiba; São Gonçalo do Amarante (former São Gonçalo); Parnamirim; Extremoz
Rio Grande do Norte	UA77 - Caicó	Caicó; Serra Negra do Norte (former Serra Negra); Florânia (former Flôres); Jardim do Seridó; Jucurutu (former São Miguel do Jucurutú); São João do Sabugi; Jardim de Piranhas; Cruzeta; Ouro Branco; São Fernando; Ipueira; Timbaúba dos Batistas; São José do Seridó
Rio Grande do Norte	UA78 - Ceará-Mirim	Ceará-Mirim; Taipu; Maxaranguape; Ielmo Marinho
Rio Grande do Norte	UA79 - Mossoró	Mossoró; Açu (former Assú); Areia Branca; Grossos; Governador Dix. Sept Rosado; Carnaubais; Baraúna; Tibau
Sergipe	UA80 - Aracajú	Aracajú; São Cristóvão; Nossa Senhora do Socorro (former Soccorro); Laranjeiras; Riachuelo; Maruim (former Maroim); Santo Amaro das Brotas (former Santo Amaro); Itaporanga D'Ajuda (former Itaporanga); Barra dos Coqueiros
Sergipe	UA81 - Nossa Senhora das Dôres	Nossa Senhora das Dôres; Capela (former Capella); Siriri (former Sriry); Cumbe; Feira Nova
Sergipe	UA82 - Itabaiana	Itabaiana; Frei Paulo (former São Paulo); Campo do Brito; Ribeirópolis; Macambira; Malhador; Moita Bonita; Areia Branca

Sergipe	UA83 - Lagarto	Lagarto; Riachão do Dantas (former Riachão); Boquim (former Buquim); Simão Dias (former Annapolis); Salgado; São Domingos
Distrito Federal/Goiás	UA84 - Brasília	Brasília; Planaltina; Formosa; Luziânia; Cidade Ocidental; Novo Gama; Valparaiso de Goiás
Goiás	UA85 - Goiânia	Goiânia; Aragoiânia
Mato Grosso	UA86 - Cuiabá	Cuiabá; Santo Antônio do Leverger (former Santo Antonio do Rio Abaixo); Nossa Senhora do Livramento (former Livramento); Poconé; Varzea Grande
Mato Grosso do Sul	UA87 - Campo Grande	Campo Grande
Espirito Santo	UA88 - Vitória	Vitória; Cariacica; Viana (former Vianna); Serra; Vila Velha (former Espirito Santo); Domingos Martins (former Santa Izabel); Santa Leopoldina (former Porto do Cachoeiro de Santa Leopoldina); Guarapari; Fundão (former Nova Almeida); Aracruz (former Santa Cruz); Ibiraçu (former Pau Gigante); Marechal Floriano
Espirito Santo	UA89 - Cachoeiro de Itapemirim	Cachoeiro de Itapemirim; Rio Novo do Sul (former Rio Novo); Alegre; Muqui (former São João do Muquí); Itapemirim; Castelo; Jerônimo Monteiro; Atilio Vivacqua; Vargem Alta; Marataizes
Espirito Santo	UA90 - Colatina (former Linhares)	Colatina (former Linhares); Itaguaçu (former Itaguassú); Pancas; Marilândia; João Neiva; São Domingos do Norte; São Roque do Canaã; Governador Lindenberg
Espirito Santo	UA91 - Linhares	Linhares; Rio Bananal; Sooretama
Espirito Santo	UA92 - Nova Venécia	Nova Venécia; São Mateus; Conceição da Barra; Vila Pavão
Minas Gerais	UA93 - Belo Horizonte	Belo Horizonte; Sabará; Caeté (former Caethé); Nova Lima (former Villa Nova de Lima); Esmeraldas (former Santa Quitéria until 1943); Santa Luzia (former Santa Luzia do Rio das Velhas); Contagem; Pedro Leopoldo; Brumadinho; Betim; Lagoa Santa; Mateus Leme; Rio Acima; Vespasiano; Matozinhos; Ribeirão das Neves; Igarapé; Taquaraçu de Minas; Nova União; Florestal; Raposos; Ibirité; Juatuba; São José da Lapa; Sarzedo; Mario Campos; São Joaquim de Bicas; Confins
Minas Gerais	UA94 - Ouro Preto	Ouro Preto; Mariana; Itabirito; Congonhas (former Congonhas do Campo); Belo Vale; Ouro Branco; Moeda
Minas Gerais/São Paulo	UA95 - Uberaba	Uberaba; Igarapava (in São Paulo state); Delta
Minas Gerais	UA96 - Uberlândia	Uberlândia (former Uberabinha); Araguari
Minas Gerais	UA97 - Juiz de Fora	Juiz de Fora; Matias Barbosa; Coronel Pacheco; Simão Pereira
Minas Gerais	UA98 - São João del Rei (Former São João Reis)	São João del Rei (Former São João Reis); Prados; Ritápolis; Dores de Campos; Tiradentes; Barroso; Coronel Xavier Chaves; Santa Cruz de Minas
Minas Gerais	UA99 - Montes Claros	Montes Claros; Bocaíuva (former Bocayuva); Coração de Jesus (former Inconfidência); Francisco Sá; São João da Ponte; Juramento; Francisco Dummont; Engenheiro Navarro; Capitao Enéas; Mirabela; Claro dos Poções; Guaraciama; Glaucilândia; Patis; São João da Lagoa
Minas Gerais	UA100 - Viçosa	Viçosa; Guaraciaba (Sant'Anna dos Ferras); Ervália (former Erval); Teixeiras; Coimbra; São Geraldo; São Miguel do Anta; Paula Candido; Porto Firme; Presidente Bernardes; Cajuri
Minas Gerais	UA101 – Varginha	Varginha; Campanha; Três Corãçoes (former Três Corações do Rio Verde); Três Pontas; Elói Mendes; Paraguaçu (former Paraguassú); Carmo da Cachoeira; Monsenhor Paulo; São Bento Abade
Minas Gerais	UA102 – Pouso Alegre	Pouso Alegre; Santa Rita do Sapucai (former Santa Rita do Sapucahy); Silvianópolis; Borda da Mata; Cachoeira de Minas (former Cachoeiras); Estiva
		Conceição dos Ouros; Congonhal; Ipuiúna; São Sebastião da Bela Vista; Espírito Santo do Dourado; Senador José Bento; Tocos do Moji
Minas Gerais/São Paulo	UA103 - Poços de Caldas	Poços de Caldas; São João da Boa Vista (in São Paulo state); Andradas (former Caracol); Caldas (former Parreiras); Campestre; Botelhos (former São José dos Botelhos); Aguas da Prata (in São Paulo state); São Sebastião da Grama (in São Paulo state); Santa Rita de Caldas; Divinolândia (in São Paulo state); Santo Antonio do Jardim (in São Paulo state); Ibitiúra de Minas; Bandeira do Sul

Minas Gerais	UA104 - Patos de Minas (former Patos)	Patos de Minas (former Patos); Patrocínio; Carmo do Paranaíba (Carmo do Paranahyba); Coromandel; Presidente Olegário; Abadia dos Dourados; Serra do Salitre; Guimarânia; Cruzeiro da Fortaleza; Lagoa Formosa; Lagamar; Varjão de Minas
Minas Gerais	UA105 - Barbacena	Barbacena; Santos Dumont (former Palmyra); António Carlos; Paiva; Oliveira Fortes; Ressaquinha; Desterro do Melo; Santa Bárbara do Tugúrio; Alfredo Vasconcelos
Minas Gerais	UA106 - Ipatinga	Ipatinga; António Dias (former Antonio Dias Abaixo); Coronel Fabriciano; Jaguaraçu; Timóteo; Santana do Paraíso
Minas Gerais	UA107 - Divinópolis	Divinópolis; Santo António do Monte; Itapecerica; Itaúna; Cláudio (former Apparecida do Claudio); Carmo do Cajuru; São Gonçalo do Pará; Nova Serrana; Perdigão; Araújos; Pedra do Indaiá; São Sebastião do Oeste; Igaratinga
Minas Gerais	UA108 - Governador Valadares	Governador Valadares; Galiléia; Alpercata; Divino das Laranjeiras; Mendes Pimentel; Frei Inocêncio; Mathias Lobato; Periquito
Minas Gerais	UA109 - Itabira	Itabira; Santa Maria de Itabira; Nova Era; Bom Jesus do Amparo; Bela Vista de Minas; João Monlevade; São Gonçalo do Rio Abaixo
Minas Gerais	UA110 - Passos	Passos; Cássia (former Santa Rita de Cássia); São Sebastião do Paraíso; São Tomás de Aquino; Alvinópolis; Capetinga; São João Batista do Glória; Itaú de Minas
Minas Gerais	UA111 - Sete Lagoas	Sete Lagoas; Paraopeba; Cordisburgo; Inhaúma; Jequitibá; Baldim; Caetanópolis; Cachoeira da Prata; Prudente de Morais; Funilândia
Rio de Janeiro	UA112 - Rio de Janeiro (former Distrito Federal prior to 1960)	Rio de Janeiro (former Distrito Federal prior to 1960); Niterói; São Gonçalo; Maricá; Itaborai; Magé; Nova Iguaçu (former Iguassú); Duque de Caxias; São João de Meriti; Nilópolis; Belford Roxo; Queimados; Mesquita
Rio de Janeiro	UA113 - Barra Mansa	Barra Mansa; Piraí; Rio Claro; Volta Redonda; Pinheiral
Rio de Janeiro	UA114 - Campos dos Goytacazes (former Campos)	Campos dos Goytacazes (former Campos); São João da Barra
Rio de Janeiro	UA115 - Angra dos Reis	Angra dos Reis; Itaguaí (former Itaguahy); Mangaratiba; Parati (former Paraty)
Rio de Janeiro	UA116 - Barra do Piraí (former Barra do Pirahy)	Barra do Piraí (former Barra do Pirahy); Valença; Vassouras; Mendes
Rio de Janeiro	UA117 - Cabo Frio	Cabo Frio; Araruama; Saquarema; São Pedro da Aldeia; Arraial do Cabo; Iguaba Grange; Armação dos Búzios
Rio de Janeiro	UA118 - Macaé (former Macahé)	Macaé (former Macahé); Casimiro de Abreu (former Barra de São João); Trajano de Morais (former São Francisco de Paula); Conceição de Macabu; Rio das Ostras; Carapebus
Rio de Janeiro	UA119 - Teresópolis (former Therezopolis)	Teresópolis (former Therezopolis); Petrópolis; Nova Friburgo; São José do Vale do Rio Preto; Guapimirim; Areal
São Paulo	UA120 - São Paulo	São Paulo; Guarulhos; Mogi das Cruzes; São Bernardo do Campo (former São Bernardo); Itapecerica da Serra (former Itapecerica); Cotia; Santana de Parnaíba (former Parnahyba); Franco da Rocha (former Juqueri); Santo André; Poá; Barueri; Mairiporã; Suzano; São Caetano do Sul; Itaquaquecetuba; Taboão da Serra; Embu; Itapevi; Pirapora do Bom Jesus; Caieiras; Arujá; Diadema; Mauá; Ferraz de Vasconcelos; Ribeirão Pires; Francisco Morato (former Belém); Carapicuiba; Osasco; Rio Grande da Serra; Jandira
São Paulo	UA121 - São Vicente	São Vicente; Santos; Guarujá; Itanhaém; Cubatão; Mongaguá; Peruíbe; Praia Grande; Bertioga
São Paulo	UA122 - Campinas	Campinas; Indaiatuba; Monte Mor; Valinhos; Sumaré; Paulinia; Hortolândia
São Paulo	UA123 - Piracicaba	Piracicaba; Rio das Pedras; São Pedro; Águas de São Pedro; Charqueada; Saltinho
São Paulo	UA124 - Sorocaba	Sorocaba; Porto Feliz; Itu (former Ytu); Salto (former Salto de Ytu); Boituva; Mairinque; Salto de Pirapora; Votorantim; Capela do Alto; Iperó; Alumínio
São Paulo	UA125 - Taubaté	Taubaté; Tremembé; Pindamonhangaba; Roseira; Santo António do Pinhal
São Paulo	UA126 - Ribeirão Preto	Ribeirão Preto; Sertãozinho; Cravinhos; Jardinapolis; Brodowski (former Brodósqui); Serrana; Dumont
São Paulo	UA127 - Jundiai (former Jundiahy)	Jundiai (former Jundiahy); Cabreúva; Itatiba; Jarinu (misspelled in IBGE data files as Jarim); Vinhedo; Várzea Paulista; Campo Limpo Paulista; Itupeva; Louveira
São Paulo	UA128 - Avaré	Avaré; Cerqueira César; Arandú

São Paulo	UA129 - Rio Claro	Rio Claro; Araras; Santa Gertrudes; Cordeirópolis; Corumbataí; Ipeúna
São Paulo	UA130 - Orlândia	Orlândia; São Joaquim da Barra (former São Joaquim); Morro Agudo; Nuporanga; Sales Oliveira
São Paulo	UA131 - Franca	Franca; Patrocínio Paulista (former Patrocínio do Sapucahy); Itirapuã; São José da Bela Vista; Cristais Paulista; Restinga; Ribeirão Corrente
São Paulo	UA132 - São José do Rio Preto (former Rio Preto)	São José do Rio Preto (former Rio Preto); Mirassol; Cedral; Bálsamo; Jaci; Bady Bassitt; Guapiaçu
São Paulo	UA133 - São Carlos	São Carlos; Analândia (former Annapolis); Ribeirão Bonito; Itirapina; Ibaté
São Paulo	UA134 - Araraquara	Araraquara; Rincão; Santa Lúcia; Américo Brasiliense; Motuca
São Paulo	UA135 - Bauru	Bauru; Piratininga; Duartina; Cabrália Paulista; Lucianópolis
São Paulo	UA136 - Marília	Marília; Gália; Garça; Vera Cruz; Pompéia; Echaporã (Former Bela Vista); Oriente; Oscar Bressane; Álvaro de Carvalho; Júlio Mesquita; Ocauçu; Lupércio; Alvinlândia; Guaimbé
São Paulo	UA137 – São José dos Campos	São José dos Campos; Caçapava; Joanópolis (former São João do Curralinho); Piracaia; Nazaré Paulista (former Nazareth); Santa Isabel; Guararema; Jacareí (former Jacarehy); Santa Branca; Paraibuna (former Parahybuna); Jambeiro; Monteiro Lobato (former Buquira) (joined with São José dos Campos between 1934 and 1948); Igaratá (former Iguaratá) (joined to Santa Isabel between 1934 and 1953)
São Paulo	UA138 – Limeira	Limeira; Santa Bárbara D'Oeste (former Santa Bárbara); Americana; Cosmópolis; Artur Nogueira; Iracemápolis; Nova Odessa; Engenheiro Coelho
Paraná	UA139 - Curitiba	Curitiba; Campo Largo; Araucaria; São José dos Pinhais; Piraquara (former Deodoro); Bocaiúva do Sul; Campina Grande do Sul (former Campina Grande); Cêrro Azul (former Serro Azul); Rio Branco do Sul (former Votuverava until 1908, and Rio Branco until 1947); Colombo; Almirante Tamandaré (former Tamandaré); Quatro Barras; Catanduvas; Mandirituba; Quitandinha; Agudos do Sul; Tijucas do Sul; Fazenda Rio Grande; Pinhais; Itaperuçu; Tunas do Paraná; Campo Magro
Paraná	UA140 - Londrina	Londrina; Cambé; Rolândia; Ibiporã; Jataizinho
Paraná	UA141 - Apucarana	Apucarana; Arapongas; Mandaguari; Jandaia do Sul; Califórnia; Cambira
Paraná	UA142 - Ponta Grossa	Ponta Grossa; Imbituva (former Santo António de Imbituva); Ipiranga (former Ypiranga); Carambeí
Paraná/Santa Catarina	UA143 - União da Vitória	União da Vitória; Porto União (in Santa Catarina state); Porto Vitória
Santa Catarina	UA144 - Florianópolis	Florianópolis; Biguaçu; São José; Palhoça; Santo Amaro da Imperatriz; Governador Celso Ramos; António Carlos; Aguas Mornas; São Pedro de Alcântara
Santa Catarina	UA145 - Itajaí	Itajaí; Camboriú; Brusque; Penha; Ilhota; Navegantes; Balneário Camboriú; Guabiruba; Itapema; Barra Velha; Botuverá
Santa Catarina	UA146 - Blumenau	Blumenau; Indaial (former Indaiá); Gaspar; Timbó; Pomerode; Luiz Alves; Rio dos Cedros
Santa Catarina	UA147 - Jaraguá do Sul	Jaraguá do Sul; São Bento do Sul (former São Bento); Campo Alegre; Joinville; Araquari (former Paraty, misspelled as Araquad in 1985 IBGE's database); Guaramirim; Corupá; Massaranduba; Schroeder; São João do Itaperiú; Balneário Barra do Sul
Santa Catarina	UA148 - Araranguá	Araranguá; Turvo; Jacinto Machado; Sombrio; Maracajá; Meleiro; Santa Rosa do Sul; Ermo; Balneário Gaivota; Balneário Arroio do Silva
Santa Catarina	UA149 - Bombinhas	Bombinhas; Porto Belo; Nova Trento; São João Batista; Canelinha
Santa Catarina	UA150 - Chapecó (former Xapecó)	Chapecó (former Xapecó); Seara; Xaxim; Coronel Freitas; Nova Erechim; Aguas de Chapecó; Caxambu do Sul; Arvoredo; Cordilheira Alta; Nova Itaberaba; Aguas Frias; Planalto Alegre; Guatambú; Paial
Santa Catarina/Rio Grande do Sul	UA151 - Concórdia	Concórdia; Itá; Mariano Moro (in Rio Grande do Sul state); Severiano de Almeida (in Rio Grande do Sul state); Peritiba; Presidente Castelo Branco; Jaborá; Irani; Ipumirim; Lindóia do Sul; Arabutã

Santa Catarina/Rio Grande do Sul	UA152 - Palmitos	Palmitos; Iraí (in Rio Grande do Sul state); São Carlos; Mondaí; Frederico Westphalen (in Rio Grande do Sul state); Caibi; Vicente Dutra (in Rio Grande do Sul state); Caiçara (in Rio Grande do Sul state); Riqueza
Santa Catarina	UA153 - Criciúma (former Cresciúma)	Criciúma (former Cresciúma); Siderópolis; Nova Veneza; Içará; Morro da Fumaça; Forquilhinha; Cocal do Sul; Treviso
Santa Catarina/Paraná	UA154 - Dionísio Cerqueira	Dionísio Cerqueira; Barracão (in Paraná state); São José do Cedro; Palma Sola; Guarujá do Sul; Flor da Serra do Sul (in Paraná state); Princesa
Santa Catarina	UA155 - Fraiburgo	Fraiburgo; Caçador; Videira; Lebon Régis; Rio das Antas; Tangará; Pinheiro Preto; Monte Carlo; Frei Rogério
Santa Catarina	UA156 - Herval d'Oeste	Herval d'Oeste; Joaçaba (former Cruzeiro do Sul then Cruzeiro); Lacerdópolis; Ouro; Erval Velho; Ibicaré; Treze Tilias; Luzerna; Ibiam
Santa Catarina	UA157 - Lages	Lages; São Joaquim (former São Joaquim da Costa da Serra); Campo Belo do Sul; Ponte Alta; São José do Cerrito; Otacilio Costa; Correia Pinto; Capão Alto; Painel; Bocaina do Sul; Palmeira
Santa Catarina/Paraná	UA158 - Mafra	Mafra; Rio Negro (in Paraná state); Itaiópolis; Canoinhas; Rio Negrinho; Papanduva; Três Barras; Major Vieira; Monte Castelo
Santa Catarina	UA159 - Orleans	Orleans; Urussanga; Braço do Norte; Grão Pará; Lauro Muller; São Ludgero; Bom Jardim da Serra
Santa Catarina	UA160 - Tubarão	Tubarão; Jaguaruna; Laguna; Pedras Grandes; Treze de Maio; Gravatal; Sangão; Capivari de Baixo
Santa Catarina	UA161 - Alfredo Wagner	Alfredo Wagner; Bom Retiro; Anitápolis; Rancho Queimado; Leoberto Leal; Imbuia; Chapadão do Lageado
Rio Grande do Sul	UA162 - Porto Alegre	Porto Alegre; Viamão; Gravataí; São Jerônimo; General Câmara (former Santo Amaro); São Leopoldo; Guaiba; Novo Hamburgo; Canoas; Esteio; Alvorada; Cachoeirinha; Sapucaia do Sul; Arroio dos Ratos; Charqueadas; Eldorado do Sul; Glorinha; Nova Santa Rita
Rio Grande do Sul	UA163 - Santo António da Patrulha (former Santo António)	Santo António da Patrulha (former Santo António); Osório (former Conceição do Arroio); Tramandaí; Capão da Canoa; Imbé; Riozinho; Maquiné; Xangri-lá
Rio Grande do Sul	UA164 - Caxias do Sul	Caxias do Sul (former Caxias in Rio Grande do Sul State); Bento Gonçalves; Garibaldi; São Francisco de Paula (former São Francisco de Paula de Cima da Serra); Farroupilha; Flores da Cunha; Canela; Carlos Barbosa; Nova Petrópolis; Gramado; Nova Pádua
Rio Grande do Sul	UA165 - Rio Grande	Rio Grande; São José do Norte
Rio Grande do Sul	UA166 - Santa Maria (former Maria da Bocca do Monte)	Santa Maria (former Maria da Bocca do Monte); Júlio de Castilhos; São Sepé; São Pedro do Sul (former São Pedro); Restinga Seca; Formigueiro; Ivorá; Silveira Martins; São Martinho da Serra; São João do Polesine; Itaara; Dilermando de Aguiar
Rio Grande do Sul	UA167 - Pelotas	Pelotas; Canguçu (former Cangussú); São Lourenço do Sul (former São Lourenço in Rio Grande do Sul state); Piratini (former Piratiny); Pedro Osório; Capão do Leão; Morro Redondo; Cerrito; Turuçu; Arroio do Padre
Rio Grande do Sul	UA168 – Passo Fundo	Passo Fundo; Carazinho ; Marau; Ernestina; Pontão; Coqueiros do Sul; Coxilha; Santo António do Planalto; Mato Castelhano
Rio Grande do Sul	UA169 – Santa Rosa	Santa Rosa; Santo Cristo; Tuparendi; Giruá; Três de Maio; Horizontina; Tucunduva; Independência; Cândido Godoi; Campina das Missões; Porto Mauá; Ubiretama; Senador Salgado Filho
Rio Grande do Sul	UA170 – São Luiz Gonzaga	São Luiz Gonzaga; Santo António das Missões; Bossoroca; Caibaté; Roque Gonzales; São Nicolau; Dezesseis de Novembro; São Miguel das Missões; Rolador
Rio Grande do Sul	UA171 – Palmeira das Missões	Palmeira das Missões (former Palmeira); Chapada; Santo Augusto; Redentora; Coronel Bicaco; Condor; Novo Barreiro; São José das Missões; Boa Vista das Missões; Dois Irmãos das Missões; Nova Ramada
Rio Grande do Sul	UA172 - António Prado	António Prado; Veranópolis (former Alfredo Chaves); São Marcos; Ipê; Protásio Alves; Vila Flores; Nova Roma do Sul
Rio Grande do Sul	UA173 - Arroio do Meio	Arroio do Meio; Lajeado; Estrela; Encantado; Roca Sales; Colinas; Capitão; Travesseiro; Marquês de Souza
Rio Grande do Sul	UA174 - Bagé	Bagé; Pinheiro Machado (former Cacimbinhas); Caçapava do Sul (former Caçapava in Rio Grande do Sul state); Lavras do Sul (former Lavras in Rio Grande do Sul state); Dom Pedrito; Hulha Negra; Candiota; Aceguá; Pedras Altas

Rio Grande do Sul	UA175 - Camaquã (former São João de Camaquam then São João Baptista de Camaquam)	Camaquã (former São João de Camaquam then São João Baptista de Camaquam); Tapes (former Dôres de Camaquam. According to IBGE, it was "joined with Porto Alegre in 1911 and 1912"); Dom Feliciano; Cristal; Cerro Grande do Sul; Amaral Ferrador; Arambaré; Sentinela do Sul; Chuvisca
Rio Grande do Sul	UA176 - Cruz Alta	Cruz Alta; Ijuí (former Ijuhy); Tupanciretã (former Tupanchretã); Ibirubá; Santa Bárbara do Sul; Panambi; Pejuçara; Augusto Pestana; Fortaleza dos Valos; Quinze de Novembro; Bozano; Boa Vista do Cadeado; Boa Vista do Incra
Rio Grande do Sul	UA177 - Lagoa Vermelha	Lagoa Vermelha; São José do Ouro; Sananduva; Barracão; Ibiaçá; Ibiraiaras; Caseiros; São Jorge; Guabijú; André da Rocha; Santo Expedito do Sul; Tupanci do Sul; Capão Bonito do Sul
Rio Grande do Sul	UA178 - Rio Pardo	Rio Pardo; Cachoeira do Sul (former Cachoeira); Encruzilhada do Sul (former Encruzilhada); Santa Cruz do Sul (former Santa Cruz); Candelária; Vera Cruz; Butiá; Pantano Grande; Minas do Leão; Passo do Sobrado; Vale Verde; Novo Cabrais
Rio Grande do Sul	UA179 - Rosário do Sul (former Rosário)	Rosário do Sul (former Rosário); Santana do Livramento (former Livramento); São Gabriel; Cacequi; Vila Nova do Sul; Santa Margarida do Sul
Rio Grande do Sul	UA180 - Santa Vitória do Palmar (former Santa Vitória)	Santa Vitória do Palmar (former Santa Vitória); Chuí
Rio Grande do Sul	UA181 - Santiago (former Santiago do Boqueirão then São Thiago do Boqueirão)	Santiago (former Santiago do Boqueirão then São Thiago do Boqueirão); São Vicente do Sul (former São Vicente); São Francisco de Assis; Jaguari; Mata; Unistalda; Jari; Capão do Cipó
Rio Grande do Sul	UA182 - Taquara (former Taquara do Mundo Novo)	Taquara (former Taquara do Mundo Novo); Rolante; Três Coroas; Campo Bom; Igrejinha; Nova Hartz; Parobé; Araricá
Rio Grande do Sul	UA183 - Uruguaiana	Uruguaiana; Alegrete; Itaquí (former Itaquy); Quaraí (former Quarahy)
Rio Grande do Sul	UA184 - Vacaria	Vacaria; Bom Jesus; Esmeralda; Campestre da Serra; Monte Alegre dos Campos; Muitos Capões
Rio Grande do Sul	UA185 – Erechim	Erechim; Getúlio Vargas; São Valentim; Aratiba; Jacutinga; Sertão; Barão de Cotegipe; Viadutos; Ipiranga do Sul; Erebango; Estação; Áurea; Tres Arroios; Ponte Preta; Centenário; Carlos Gomes; Barra do Rio Azul; Floriano Peixoto; Quatro Irmãos

Note: The municipalities involved in the definition of urban areas in some cases change over time. Due to its size, the detailed appendix showing merges and splits of municipalities from some urban areas over time has not been included in this thesis, but it is available upon request.

Appendix 4.2: Changes in Population of Urban Areas Across Decades in Brazil

Figure 4.2.1: The Urban Population of Brazil, 1910

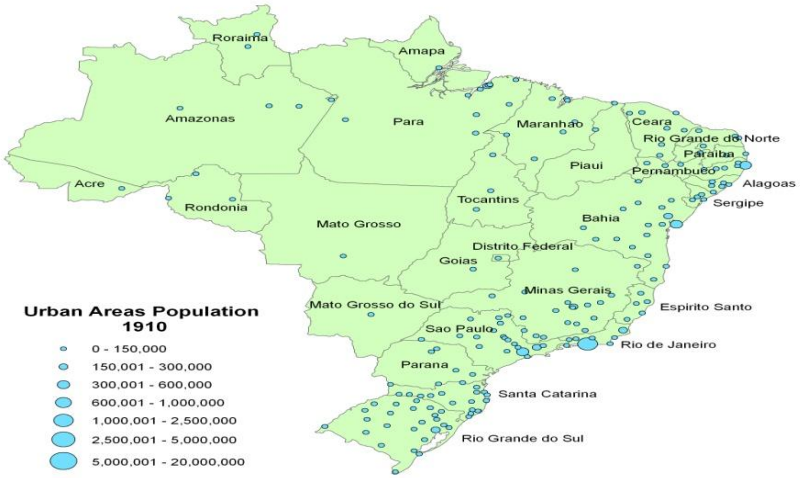


Figure 4.2.2: The Urban Population of Brazil, 1920

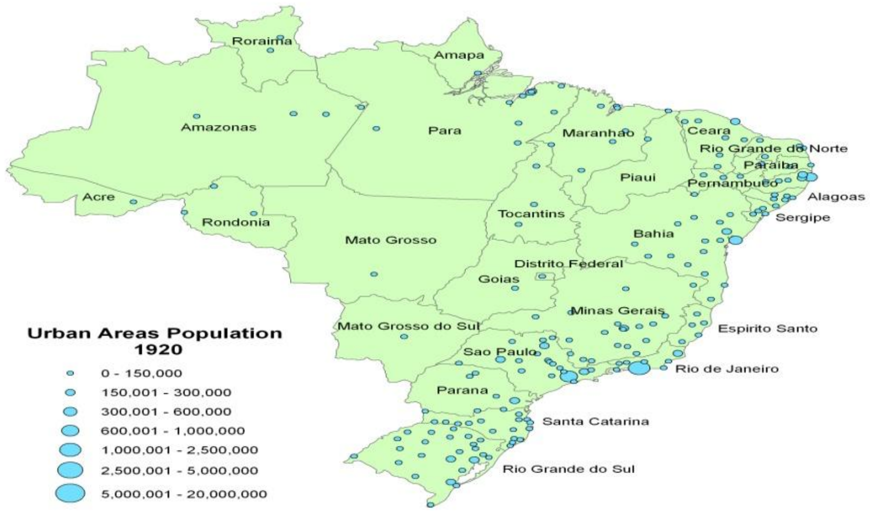


Figure 4.2.3: The Urban Population of Brazil, 1930

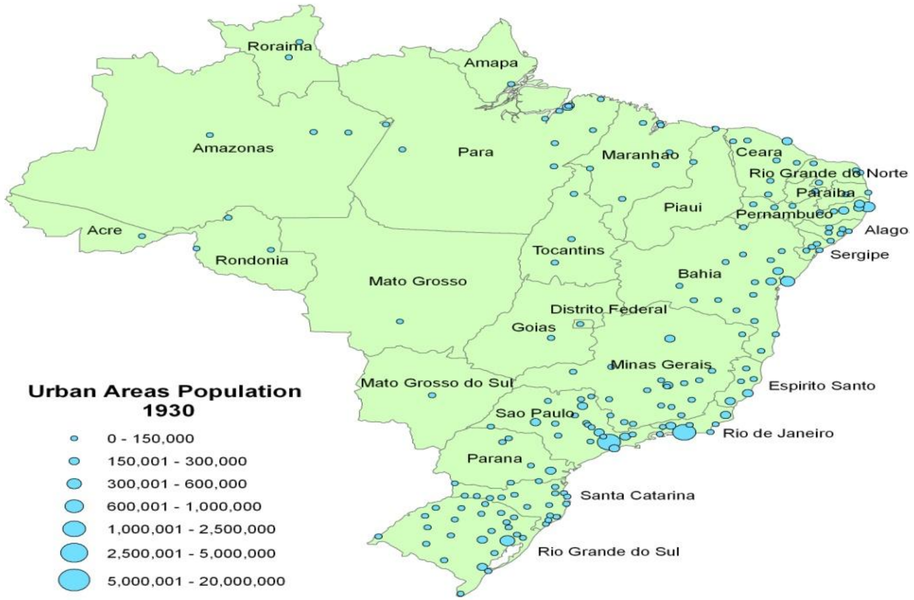


Figure 4.2.4: The Urban Population of Brazil, 1940

Figure 4.2.5: The Urban Population of Brazil, 1950

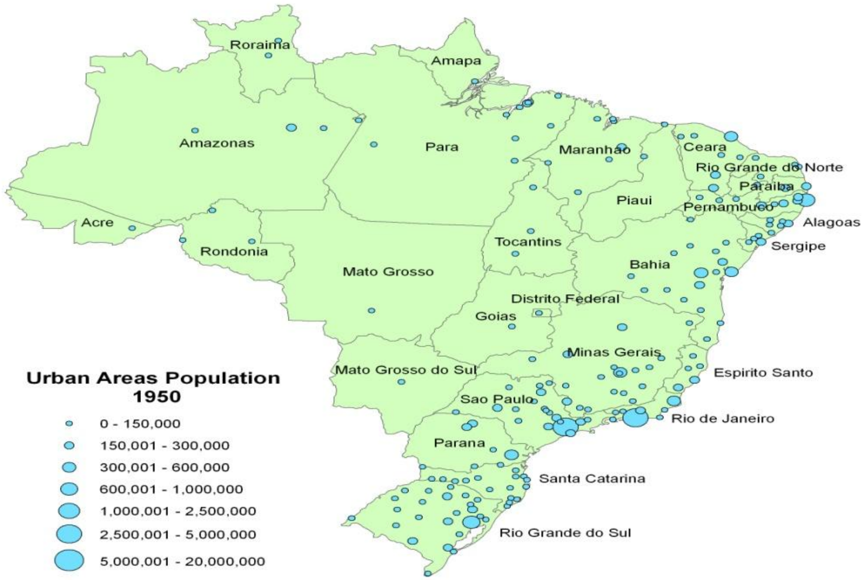


Figure 4.2.6: The Urban Population of Brazil, 1960

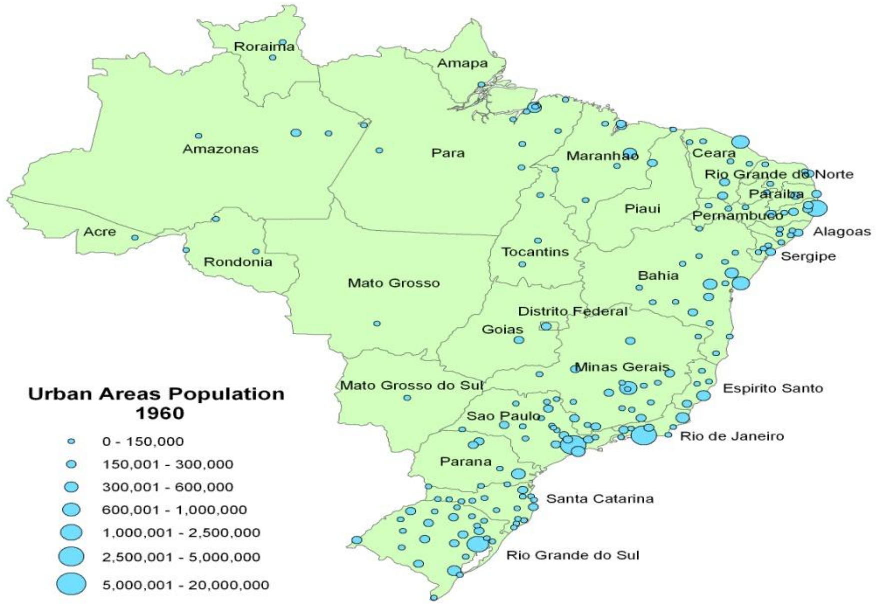


Figure 4.2.7: The Urban Population of Brazil, 1970

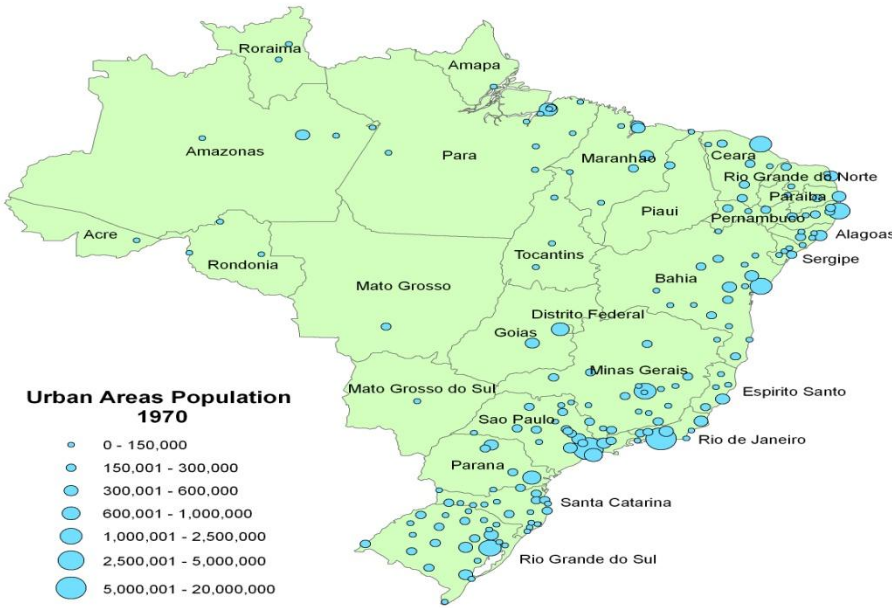


Figure 4.2.8: The Urban Population of Brazil, 1980

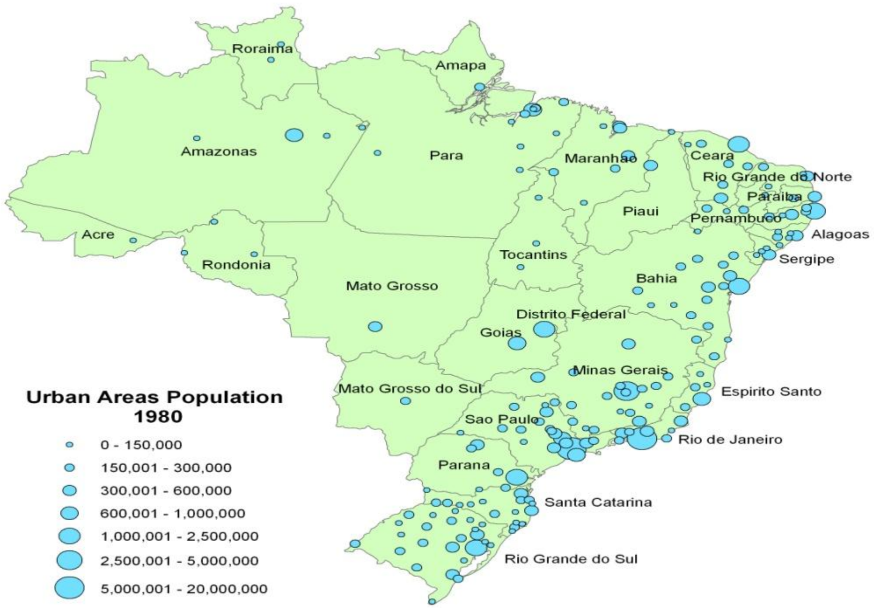


Figure 4.2.9: The Urban Population of Brazil, 1990

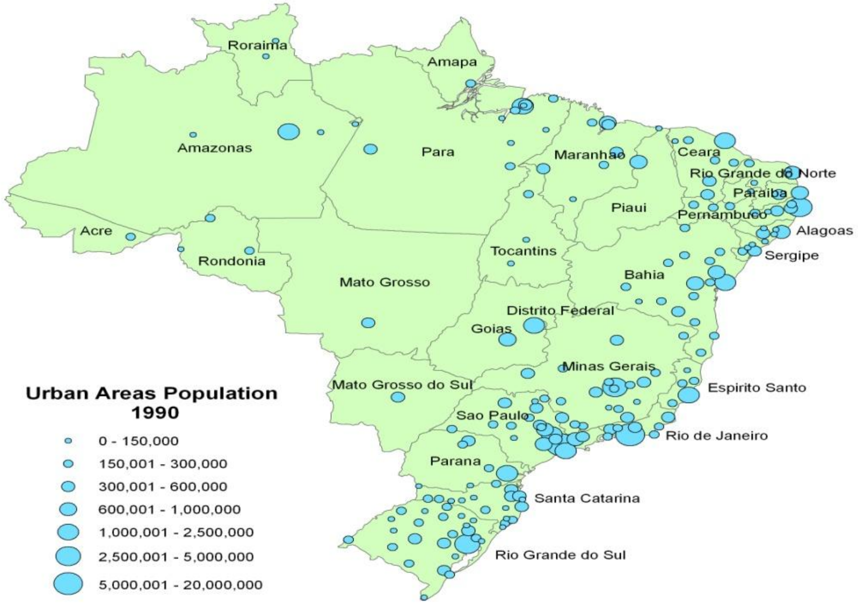


Figure 4.2.10: The Urban Population of Brazil, 2000

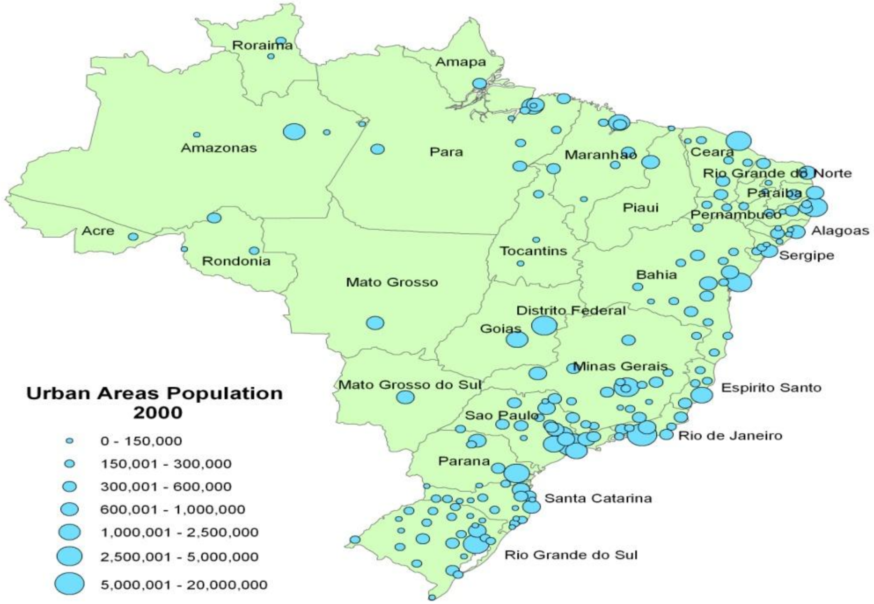
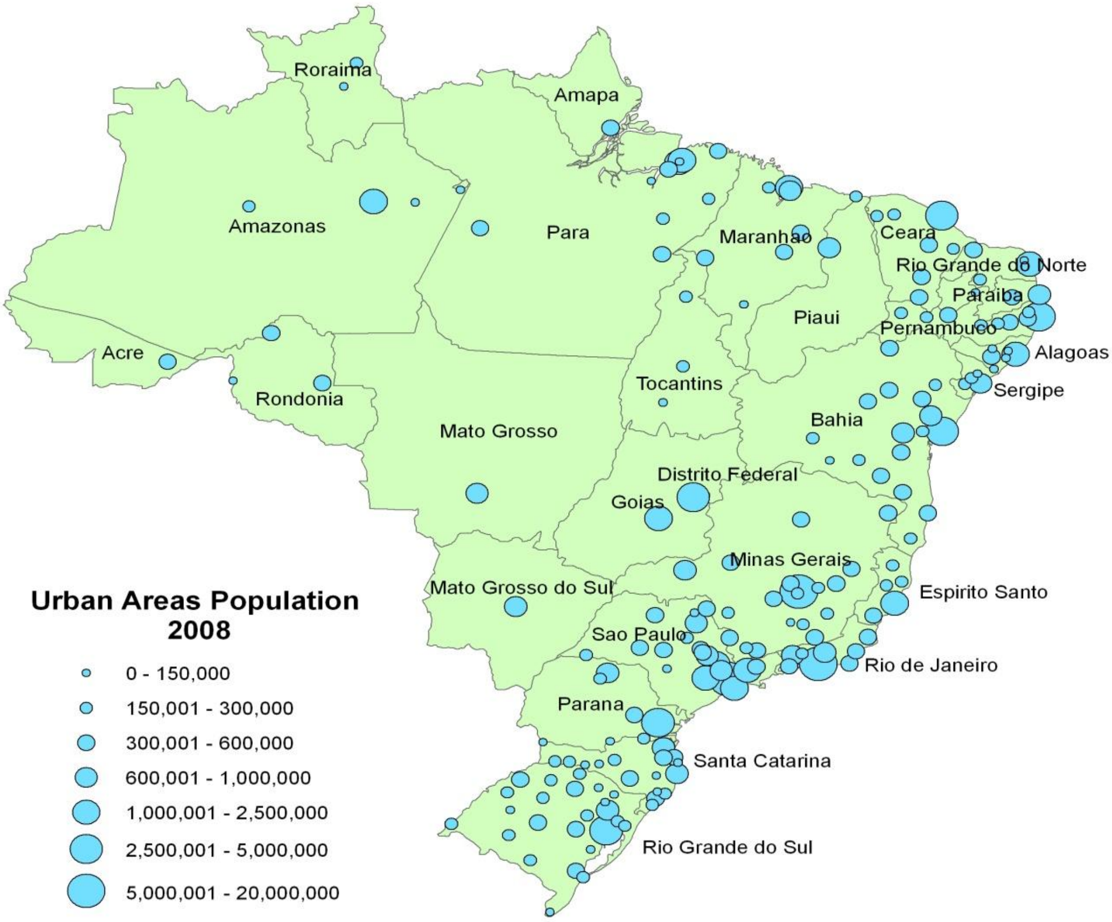
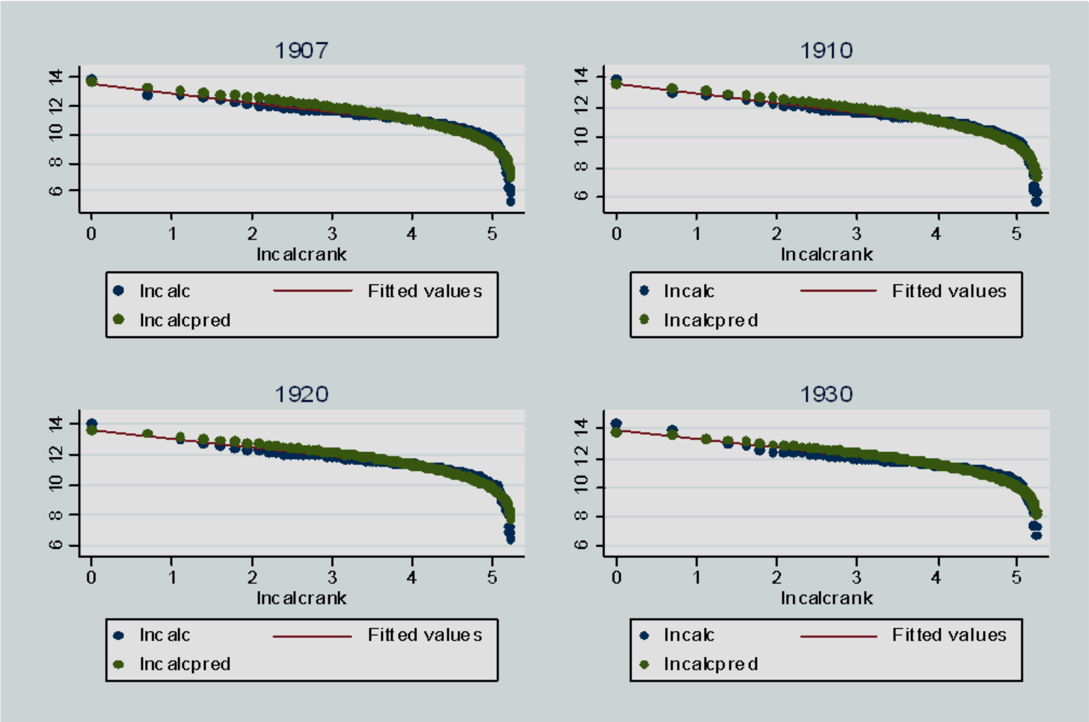


Figure 4.2.11: The Urban Population of Brazil, 2008



Appendix 4.3: Changes in the quality of the fit of both the Lognormal and Pareto curves in the city size distribution in Brazil

Figure 4.3.1: Lognormal and Pareto fit: Brazil, 1907 to 1930



Note for Figures 4.3.1 to 4.3.3: lncalc is the natural logarithm of population; lncalcrank is the natural logarithm of population rank; lncalcpred is the predicted logarithm of population from fitting a lognormal distribution; Fitted values line refers to the power law estimated on the 100 largest cities.

Figure 4.3.2: Lognormal and Pareto fit: Brazil, 1940 to 1970

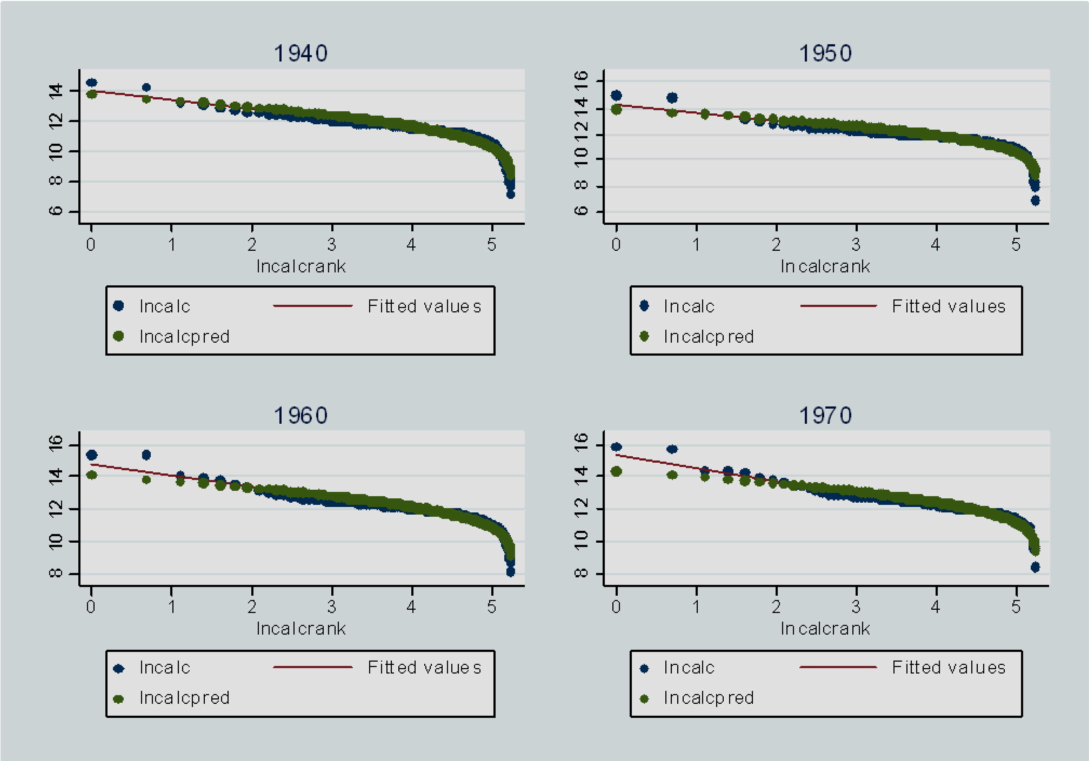
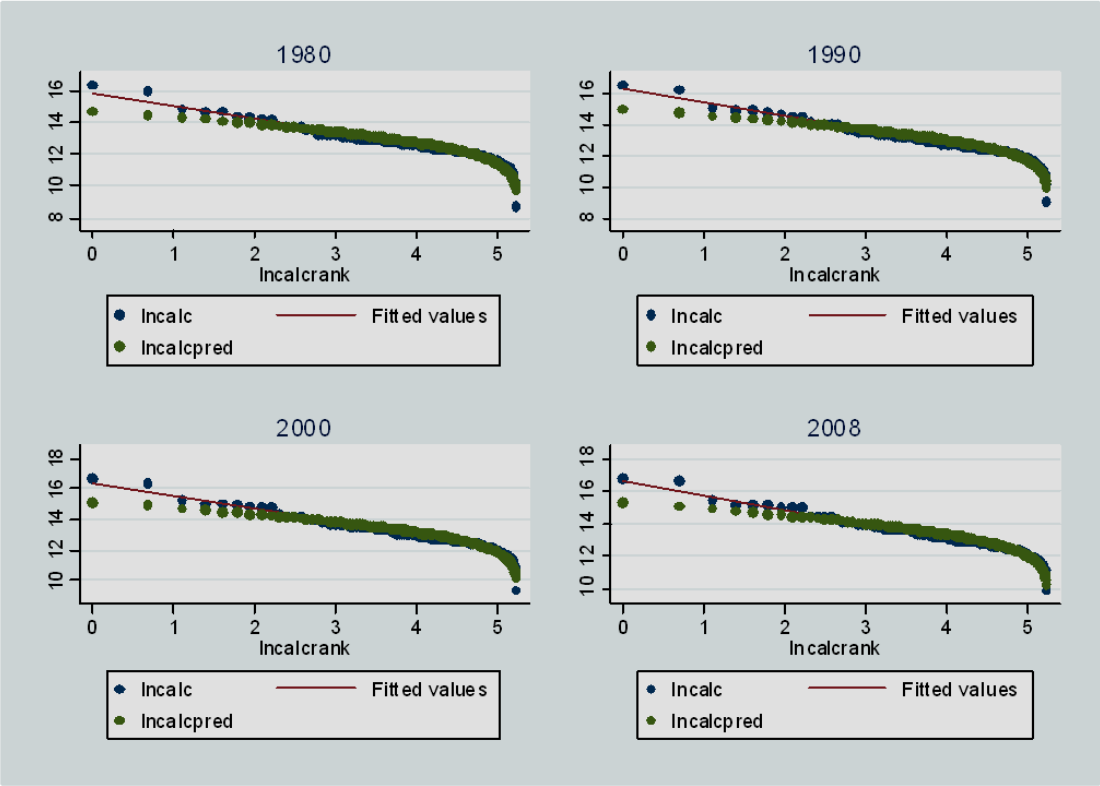


Figure 4.3.3: Lognormal and Pareto fit: Brazil, 1980 to 2008



Appendix 6.1: Variables used in the analysis

Variable	Brief description	Cross-section Levels Models	Cross-section Growth Models	Cross-section Spatial Growth Models
Dependent Variables				
$gdp_{r,t}$	Real GDP of urban area (region) r in period *t*	*t*=1949, 1959, 1970, 1975, 1980, 1985, 1996, and 1999 to 2008.	---	---
$gdp\_pc_{r,t}$	Real GDP per capita of urban area (region) r in period *t*	*t*=1949, 1959, 1970, 1975, 1980, 1985, 1996, and 1999 to 2008.	---	---
$pop_{r,t}$	Urban area (region) r population in period *t*	*t*=1949, 1959, 1970, 1975, 1980, 1985, 1996, and 1999 to 2008.	---	---
$pop_{r,t}$	Population of urban area (region) *r* in period *t*. **Note:** This dependent variable was defined for the first stage panel regression model of growth for which its dependent variable is follows. The independent variables of the first stage model are: distance of the urban area (region) r to Rio de Janeiro and Sao Paulo, the 1939 population of the urban area (pop_{rt_T}) and year. This first stage model has 1274 (182 regions x 7 years) observations.	*r* varies from 1 to 182.	*r* varies from 1 to 182.	*r* varies from 1 to 182.
depvar_n = (1/*T*). $\log(y_{r,t}/y_{r,t-T})$	The average annual per capita GDP (y) growth from *t*-T to *t* for urban area (region) *r*.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008. The equivalent *T* for *t* is, respectively, 10, 20, 31, 41, 46, 60, and 69.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008. The equivalent *T* for *t* is, respectively, 10, 20, 31, 41, 46, 60, and 69.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008. The equivalent *T* for *t* is, respectively, 10, 20, 31, 41, 46, 60, and 69.
Independent Variables				
$gdp_{r,t}$	Real GDP of urban area (region) r in period *t*	*t*=1939	---	---

$gdp\_pc_{r,t}$	Real GDP per capita of urban area (region) r in period *t*	*t*=1939	---	---
$pop_{r,t}$	Urban area (region) r population in period *t*	*t*=1939	---	---
dist_b	Distance (in Kilometres) of the urban area *r* to Brasília	---	---	---
dist_r	Distance (in Kilometres) of the urban area *r* to the city of Rio de Janeiro	---	---	---
dist_s	Distance (in Kilometres) of the urban area *r* to the city of São Paulo	---	---	---
pop_{rt_T}	Population of the urban area (region) *r* in initial period (*t*-T). **Note:** this variable is one of controls for the first stage panel growth model to predict iprt that was included in per capita GDP growth models.	(*t*-T) is 1939 for each *t*.	(*t*-T) is 1939 for each *t*.	(*t*-T) is 1939 for each *t*.
$year_t$	Year of the cross-section observation. **Note:** this variable is one of controls for the first stage growth regression model to predict region r population, or iprt.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.
y_{rt_T}	Logarithm of the urban area (region) *r* per capita GDP (y) in initial period *t*-T.	(*t*-T) is 1939 for each *t*.	(*t*-T) is 1939 for each *t*.	(*t*-T) is 1939 for each *t*.
ip_{rt}	Logarithm of the urban area (region) r predicted population in period t from the first stage regression model (see the fourth dependent variable).	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.
pop_b	Logarithm of Brasilia's population in period t.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.	*t*=1949, 1959, 1970, 1980, 1985, 1999, and 2008.	---
Number of Observations	---	174	174	174

Appendix 6.2: Descriptive Statistics for Cross-section Levels Models for GDP

Year	Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
1949 to 2008*	dist_b	182	1610.874	664.11	209	4275
	dist_r	182	1652.725	1031.64	99.9	5159
	dist_s	182	1566.453	1069.14	58.8	4756
	$gdp1939_r$	174	106235	181235	343	1318890
	$gdp_{r,t}$	178	174807.8	267080.9	41	2228314
1949	$gdp_{r,t}$	178	174807.8	267080.9	41	2228314
1959	$gdp_{r,t}$	179	339059.5	598242	1467	4160752
1970	$gdp_{r,t}$	179	665297.2	1269851	10527	1.03e+07
1975	$gdp_{r,t}$	180	1202222	2455223	18205	1.82e+07
1980	$gdp_{r,t}$	181	1902802	3866116	56464	2.65e+07
1985	$gdp_{r,t}$	181	2224565	4133054	23069	2.67e+07
1996	$gdp_{r,t}$	182	2512748	5042268	22913	3.70e+07
1999	$gdp_{r,t}$	182	2768169	4950894	60312	3.13e+07
2000	$gdp_{r,t}$	182	2964565	5308203	69061	3.35e+07
2001	$gdp_{r,t}$	182	2987420	5330565	67091	3.51e+07
2002	$gdp_{r,t}$	182	3114411	5485191	65725	3.59e+07
2003	$gdp_{r,t}$	182	3139266	5451658	62667	3.58e+07
2004	$gdp_{r,t}$	182	3367115	5863059	60754	3.91e+07
2005	$gdp_{r,t}$	182	3504406	6156935	63676	4.00e+07
2006	$gdp_{r,t}$	182	3675091	6421713	69584	4.38e+07
2007	$gdp_{r,t}$	182	3883711	6848236	71981	4.75e+07
2008	$gdp_{r,t}$	182	4104419	7206790	74573	5.06e+07

***Note:** All the distance variables also apply for models of both GDP per capita and Population.

.

Appendix 6.3: Descriptive Statistics for Cross-section Levels Models for GDP per capita and Population

Year	Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
1949	gdp_pc1939_r*	174	0.9348	0.7795	0.1381	7.2154
1949	$pop1939_r$*	182	96101.25	93282.87	1100	776203
1949	$gdp\_pc_{r,t}$	178	1.3301	0.8940	0.0377	4.3401
1949	$pop_{r,t}$	182	111292.9	101676.8	1085	792715
1959	$gdp\_pc_{r,t}$	179	1.8069	1.3838	0.2296	9.9656
1959	$pop_{r,t}$	182	148953.3	154872.8	3076	1198360
1970	$gdp\_pc_{r,t}$	179	2.5476	1.6772	0.4186	8.4527
1970	$pop_{r,t}$	182	203554.5	243660.7	4421	1791322
1975	$gdp\_pc_{r,t}$	180	3.7664	2.8092	0.4132	19.2397
1975	$pop_{r,t}$	182	242307.4	305257	5211	2144008
1980	$gdp\_pc_{r,t}$	181	5.1614	3.5557	0.5855	20.7852
1980	$pop_{r,t}$	182	281059.8	368950.1	6000	2638078
1985	$gdp\_pc_{r,t}$	181	5.4605	3.5126	0.8016	18.4825
1985	$pop_{r,t}$	182	327133.1	437065.2	6965	3162213
1996	$gdp\_pc_{r,t}$	182	4.5873	3.0523	0.3676	14.6268
1996	$pop_{r,t}$	182	408431.5	543245.4	9814	3843638
1999	$gdp\_pc_{r,t}$	182	5.1749	3.0654	0.6707	16.5292
1999	$pop_{r,t}$	182	431812.9	574189.7	10267	4078223
2000	$gdp\_pc_{r,t}$	182	5.4529	3.4016	0.9067	19.6120
2000	$pop_{r,t}$	182	439602.4	584215.2	10457	4152109
2001	$gdp\_pc_{r,t}$	182	5.3434	3.2536	0.9118	17.4978
2001	$pop_{r,t}$	182	457576.8	614224.4	14910	4398039
2002	$gdp\_pc_{r,t}$	182	5.5043	3.3758	0.9341	20.8434
2002	$pop_{r,t}$	182	465254.9	624692.9	15448	4478720
2003	$gdp\_pc_{r,t}$	182	5.5657	3.3938	0.9027	21.0696
2003	$pop_{r,t}$	182	473202	635352.4	15987	4559881
2004	$gdp\_pc_{r,t}$	182	5.7504	3.5689	0.9743	21.8082
2004	$pop_{r,t}$	182	488541.2	657782.5	17259	4730210
2005	$gdp\_pc_{r,t}$	182	5.7916	3.8167	1.0333	26.5323
2005	$pop_{r,t}$	182	497827.9	670217.9	17746	4824469
2006	$gdp\_pc_{r,t}$	182	5.9909	4.1326	1.0778	30.8519
2006	$pop_{r,t}$	182	506953.7	682576	18367	4918039
2007	$gdp\_pc_{r,t}$	182	6.1687	3.9462	1.1960	23.3249
2007	$pop_{r,t}$	182	516000.8	695184.4	18578	5011609
2008	$gdp\_pc_{r,t}$	182	6.4368	4.3610	0.9737	31.3662
2008	$pop_{r,t}$	182	526739.8	709169.3	18789	5105179

***Note:** These two variables were included in all cross-section models.

Appendix 6.4: Descriptive Statistics for Cross-section Growth Models and Spatial Growth Models for Growth of GDP per capita

Year	Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
1949	dist_b	182	7.2942	0.4473	5.3423	8.3605
1949	dist_r	182	7.1685	0.7800	4.6042	8.5485
1949	dist_s	182	7.0424	0.9027	4.0741	8.4672
1949	poprt_T	182	96101.24	93282.87	1100	776203
1949	yrt_T	174	-0.3371	0.7530	-1.9801	1.9762
1949	poprt	182	111292.8	101676.8	1084.5	792715
1949	iprt	182	6.1296	5.9553	0	14.638
1949	depvar_n	174	0.0378	0.0473	-0.1428	0.2129
1949	pop_b	182	0	0	0	0
1959	poprt	182	148953.3	154872.8	3075.8	1198360
1959	iprt	182	8.4304	5.4012	0	14.6687
1959	depvar_n	174	0.0347	0.0259	-0.0627	0.1823
1959	pop_b	182	0	0	0	0
1970	poprt	182	203554.5	243660.7	4421	1791322
1970	iprt	182	10.5884	4.0024	0	14.7015
1970	depvar_n	174	0.0339	0.0175	-0.0312	0.1025
1970	pop_b	182	13.3182	0	13.3182	13.3182
1980	poprt	182	281059.8	368950.1	6000	2638078
1980	iprt	182	12.0767	1.8025	0	14.7303
1980	depvar_n	174	0.0426	0.0144	-0.0093	0.0998
1980	pop_b	182	14.1001	0	14.1001	14.1001
1985	poprt	182	327133.1	437065.2	6965	3162213
1985	iprt	182	12.3898	1.1936	0	14.7445
1985	depvar_n	174	0.0396	0.0122	0.0027	0.0791
1985	pop_b	182	14.3809	0	14.3809	14.3909
1999	poprt	182	431812.9	574189.7	10267	4078223
1999	iprt	182	12.8333	0.5252	11.3309	14.7830
1999	depvar_n	174	0.0298	0.0103	-0.0108	0.0609
1999	pop_b	182	14.7069	0	14.7069	14.7069
2008	poprt	182	526739.8	709169.3	18789	5105179
2008	iprt	182	13.0076	0.4485	11.9001	14.8070
2008	depvar_n	174	0.0288	0.0094	-0.0074	0.0569
2008	pop_b	182	14.9770	0	14.9770	14.9770

Notes: Distance variables are current distances; population (poprt_T) and GDP per capita (yrt_T) in urban area (region) r refer to their initial levels (1939). These three variables do not vary across cross-sections. Thus, they are presented for 1949 only.