The London School of Economics and Political Science

Essays on Spatial Scope of Regional Economic Development in Brazil

Guilherme Mendes Resende

Thesis submitted to the Department of Geography and Environment of the London School of Economics and Political Science for the degree of Doctor of Philosophy (Ph.D)

London, June 2012

Declaration

I, Guilherme Mendes Resende, certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 60,045 words.

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by Tara Jo-Anne Murphy (bandtglass@gmail.com).

Abstract

The aim of my thesis is to investigate the spatial scope of regional economic growth and regional economic development policy in Brazil. First, it reviews the theoretical background on the spatial scope of economic development and growth literature as well as sets this discussion for the Brazilian context. This part forms the basis for the following empirical investigations. Then, the thesis investigates how the determinants of economic growth in Brazil may have manifested themselves differently on various spatial scales during the period of 1991-2000. The analysis suggests a general framework for addressing multiple spatial scales, spatial autocorrelation, spatial heterogeneity and model uncertainty. The robustness tests identified variables that are simultaneously significant on different spatial scales – higher educational and health capital, and better local infrastructure were related to higher rates of economic growth, although their impact on growth may differ across spatial scales. Next, the thesis investigates the extent of spatial autocorrelation effects in the context of regional economic growth at different spatial scales from 1970-2000 using standard panel data models. Among other results, it shows that spatial autocorrelation appears negligible at the state level but shows positive and significant values at the other three spatial scales. Moreover, the panel data models that control for time invariant fixed effects do not completely eliminate the spatial autocorrelation in the residuals at different spatial scales. Finally, the thesis formulates a framework to measure the micro- and macro - impacts of regional development policies in Brazil and applies this framework to measure the impact of northeast regional fund (FNE) industrial loans on employment and labour productivity growth at the micro (firm) level and on GDP per capita growth at macro (municipalities, micro-regions and spatial clusters) levels for the 2000-2003 and 2000-2006 periods. The results show a positive and statistically significant impact of the FNE industrial loans on job creation at the micro level but no significant impacts on the GDP per capita growth at the macro level.

Acknowledgements

Apart from my own efforts, the success of this thesis depended largely on the encouragement and guidance of many others. I take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this thesis. I want to thank Stephen Gibbons, Giordano Mion, Frédéric Robert-Nicoud, Lízia de Figueirêdo and Daniel da Mata, all of whom have provided useful comments on earlier drafts and improved the thesis. I should express my gratitude to Henri de Groot and Vassilis Monastiriotis who were the examiners of my thesis and provided detailed comments and relevant suggestions to improve the quality of the thesis. The thesis has also benefited from discussions with professors Henry Overman, Olmo Silva, Ian Gordon and Andrés Rodríguez-Pose in two internal seminars at LSE. I am very grateful to five anonymous referees for constructive and helpful comments. I am also grateful to the staff of the Institute of Applied Economic Research (Instituto de Pesquisa Econômica Aplicada)/Government of Brazil (IPEA), the Ministry for National Integration (MI/Brazil), the Bank of the Northeast (BNB), the Bank of the Amazon (BASA) and the Bank of Brazil (BB), all of which were involved in the regional development funds evaluation project during 2005/2006. This thesis has benefited from that experience. The views expressed are those of the author and do not represent the views of IPEA, MI, BNB, BASA or BB. Remaining errors and misunderstandings are, of course, the responsibility of the author. I acknowledge financial support from IPEA and the Brazilian Federal Agency of Support and Evaluation of Postgraduate Education (CAPES). Not least, I should thank my wife, Josy, for her love, patience and forbearance while I have spent hundreds of days working on this thesis. In addition, I would like to express my heartfelt thanks to my beloved mother Cleuza and my little son José Guilherme for their love and moral support and to my friends and classmates for their help and wishes for the successful completion of this thesis. This thesis is dedicated to my father, Aloísio, who provided so much love, support, comfort, growth, leadership, and wisdom in my life and who passed away at the beginning of my Ph.D. course. I thank him for his truthfulness, warm understanding and spiritual support.

Contents

List of Tables

List of Figures

1. Introduction

This thesis is comprised of five main chapters beyond this introduction chapter and the concluding remarks chapter. First, it is important to highlight that the thesis aims to address the question of how economic growth, as well as regional development policies, manifest different forms at different spatial scales in Brazil. This investigation refers back to the Modifiable Areal Unit Problem (MAUP) and the Ecological Fallacy (EF), but it sheds new light on a core problem in the literature related to regional economic growth and to regional development policy evaluation. As discussed throughout the thesis, the choice of the spatial scale of analysis is a problematic issue in applied research. In this sense, the thesis seeks to investigate to what extent ambiguities about spatial scale undermine, or inform, our understanding of regional growth policies. Related to this investigation, it also sheds light on the potential theoretical reasons for different results found across the models estimated at different spatial scales.

The idea of this thesis is to investigate the distortions in the empirical results caused by the use of different levels of analysis in the study of regional economic growth and regional development policies, by systematically repeating a method originally developed to examine this phenomenon at a single scale across multiple scales. In this thesis, the term "scale" is defined as nested sets of spatial units of different spatial resolutions (e.g., municipalities nested within micro-regions, nested in turn within states). This empirical exercise is carried out using several Brazilian geographic stratifications (municipalities, minimum comparable areas, micro-regions, meso-regions, spatial clusters and states) commonly employed in the empirical literature about Brazil. In the case of the regional development fund evaluation, the study also employs micro (firm) level data. Rey and Janikas (2005) noted that a number of studies have examined the robustness of growth regression to various aspects of research design (Levine and Renelt, 1992; Sala-i-Martin, 1997; Sala-i-Martin et al., 2004), however changes in spatial unit of analysis have yet to be incorporated in this important line of research.

The thesis focuses on the investigation of the measurement issue that might cause variability in regional economic growth estimates due to the use of different spatial scales, likely due to the MAUP. According to Fotheringham et al. (2000: 237), the two components of the MAUP are: "*(i) the scale effect: different results can be obtained from the same statistical analysis on different levels of spatial resolution; (ii) the zoning effect: different results can be obtained owing to the regrouping of zones at a given scale*". This thesis only explores the scale effect of MAUP. Of note, Behrens and Thisse (2007) point out that from an empirical point of view, the concept of region one retains is often intrinsically linked to the availability of data. Behrens and Thisse (2007) discuss that the concept of region is problematic in theory. In this respect, they argue that:

> "*it is well known how poorly representative the so-called 'representative consumer' may be (Kirman, 1992). Likewise, the word 'industry' is still in search of a well-defined theoretical meaning (Triffin, 1940). Grouping locations within the same spatial entity, called a region, gives rise to similar difficulties. It is, therefore, probably hopeless to give a clear and precise answer to our first question (What is a region?), which is essentially an empirical one. When we talk about a region, we must be happy with the same theoretical vagueness that we encounter when using the concept of industry. Note that both involve some 'intermediate' level of aggregation between the macro and the micro*" (Behrens and Thisse, 2007: 459).

For this reason, Behrens and Thisse (2007) argue that the question of the spatial scale of analysis becomes a problematic issue in applied research. It is worth noting that most studies on economic growth and convergence process (see Chapters 2 and 3 for a detailed discussion) do not employ a rigorous analysis of spatial scale choice and do not make any comparison between results using alternative spatial scales. An exception, is the work originally developed by Cheshire and Hay (1989) that concentrated on obtaining functional regions that would be "geographically meaningful" in order to capture the economic sphere of influence of a group of smaller administrative units, so as to deal with the MAUP. However, even if all researchers in a specific field agree on a unique set of areal units (e.g., functional regions), statistical variations might occur when other levels of spatial aggregation in the data are used and these variations are "of interest in order to place a set of results into a meaningful statistical and spatial perspective" (Openshaw, 1979: 143). Furthermore, the discrepancy between economic growth studies may occur because they do not work with the same econometric specifications to evaluate growth determinants and convergence. For this reason, the same econometric specifications are employed at all spatial scales used in this thesis.

Finally, the main Brazilian regional development policy that has been enacted since 1989 by the Constitutional Financing Funds for the Northeast (FNE), the North (FNO), and the Central-West (FCO) is assessed. The objective of this policy is the economic and social development of lagging Brazilian regions through subsidised loans given mainly to small-scale farmers and small industrial firms. However, assessment of the impacts of regional policies in Brazil has rarely occurred over the years, as shown in Chapter 3. Thus, the goal of Chapter 6 is to contribute in filling – at least partially – this gap in the literature by measuring the micro- and macro-impacts of the FNE loans. Given the availability of data, the analysis includes only those FNE loans for firms in the industrial, commerce and services sectors, which represent roughly 40% of the FNE loans granted during the analysis period. The investigation involves two levels of analysis that are often implemented separately in the impact-evaluation literature. First, it measures whether the FNE loans create jobs and/or increase labour productivity at the firm level. Second, it assesses the impact of the FNE loans on regional inequalities by investigating whether the FNE loans reduce the regional GDP per capita gap. This combined approach is useful because evaluations can suggest, for instance, that the regional development funds create jobs and/or increase labour productivity at the firm (micro) level. However, it is still necessary to demonstrate that the program has solved, or at least reduced, regional inequalities. These micro- and macro-effects have been overlooked in the literature that deals with the impact of regional development funds. It is important to note once again that with regard to the macro-analysis, three levels of spatial aggregation of the observational units are employed (municipalities, micro-regions and spatial clusters) because, for instance, an aggregation problem might prevent us from identifying the effects of FNE loans on GDP per capita growth at the municipal level. In this sense, Chapter 6 seeks to provide a more complete picture of FNE performance during 2000-2006.

The rest of the thesis is organised as follows. Chapter 2 reviews the theoretical rationale and background of regional economic development policies and growth literature. Of note, the spatial scope of the economic growth theories is examined in detail. In addition, some recent econometric issues to investigate the determinants of regional economic growth are analysed, namely, the spatial econometrics literature and MAUP or Ecological Fallacy (EF).

Chapter 3 provides a descriptive analysis on the empirical literature and information for the Brazilian case. First, the spatial scales in Brazil for studying economic growth are described and literature on the role of spatial scales on Brazilian economic growth is discussed. Furthermore, some socioeconomic variables of the Brazilian economy at different spatial scales are described. Chapter 3 also reviews the justifications of regional economic development in Brazil as well as the evaluation literature regarding the Brazilian regional development funds. Finally, it discusses some issues related to policy process, policy objectives and the types of evaluation as well as the strategy of these funds.

Chapter 4 seeks to understand how the determinants of economic growth in Brazil may have manifested themselves differently on various spatial scales (states, municipalities, micro-regions and spatial clusters) between 1991 and 2000. Analysing this issue sheds light on the MAUP (a measurement issue). In addition, it also suggests potential explanations for the origin of this variability. This latter issue relates to the scale-dependent determinants of economic growth (a structural issue). The analysis reveals that the results change as the scale level changes, and suggests a general framework for dealing with multiple spatial scales, spatial autocorrelation, spatial heterogeneity and model uncertainty. Importantly, the robustness tests identified variables that are simultaneously significant on different spatial scales – higher educational and health capital, and better local infrastructure were related to higher rates of economic growth – although their impact on growth may differ across spatial scales.

Chapter 5 continues to investigate the spatial dimensions of regional economic growth in Brazil. It focuses on the extent of spatial autocorrelation effects by testing whether the residuals of traditional panel-data growth-model estimates are spatially autocorrelated at different spatial scales for the period between 1970 and 2000 in Brazil. In this sense, the idea is to explore whether alternative non-spatial panel data models (including those that control for time invariant fixed effects in panel data) eliminate the spatial autocorrelation across different spatial scales. Among other results, it shows that the spatial autocorrelation that is often detected in traditional regional growth regressions is dependent on the choice of spatial scale and the spatial weight (W) matrix. Indeed, spatial autocorrelation seems to be negligible at the state level. On the other hand, although the spatial autocorrelation in residuals at the other three spatial scales (minimum comparable areas, micro-regions and meso-regions) show positive and statistically significant values across distances of more than 1,500 kilometres, their levels are largely reduced at distances of more than 900 kilometres (in particular for the 1970s and 1980s). Interestingly, an increasing clustering of the regressions' residuals over time was demonstrated (in particular for the 1990s). Finally, the fixed-effect, first-difference and system-GMM approaches in this empirical exercise do not completely eliminate the spatial autocorrelation across different spatial scales.

Chapter 6 presents an effort to evaluate the regional development funds in Brazil. It develops a framework to measure the micro- and macro-impacts of regional development policies in Brazil using the first-differences method that controls for observable characteristics and unobserved fixed effects. Next, it applies this framework to measure the impact of the FNE industrial loans on employment and labour productivity growth at the micro (firm) level and on GDP per capita growth at macro (municipalities, micro-regions and spatial clusters) levels for the 2000-2003 and 2000-2006 periods. The results show a positive and statistically significant impact of the FNE industrial loans on job creation at the micro level but no significant impacts on the GDP per capita growth at the macro level. The last chapter presents the concluding remarks of this thesis.

2. The Spatial Scope of Regional Economic Growth: Theories and Methods

This chapter reviews the theoretical literature discussed throughout the thesis, thus it sets the stage for the empirical applications that are to come in Chapters 4, 5 and 6. Section 2.1 aims to review – in the mainstream of economics – the theoretical background of regional economic development and growth literature that is the basis for the empirical analysis (i) of regional economic growth in Brazil developed in Chapters 4 and 5 and (ii) of a specific Brazilian regional economic development policy that is assessed in Chapter 6. Section 2.1.3 aims to discuss the spatial scope of the economic growth theories examined in Sections 2.1.1 (Neoclassical growth models and their predecessors) and 2.1.2 (Endogenous growth models and recent developments) to shed some light on the potential theoretical reasons for different results found across regional economic growth models estimated at different spatial scales. In this sense, Section 2.1.3 focuses on the discussion of the structural (theory based) issues underlying economic growth at different scales.

Moreover, Section 2.2 discusses some recent econometric issues in studying the determinants of regional economic growth. The focus is on (i) the spatial econometrics literature that seeks to measure the effects of spatial interactions among neighbouring regions (Section 2.2.1), (ii) the parameter/spatial heterogeneity (Section 2.2.2) as well as (ii) the discussion of a measurement issue that may cause variability in econometric estimates due to the use of different spatial scales (Section 2.2.3), namely, MAUP or Ecological Fallacy (EF) and is empirically investigated in Chapters 4, 5 and 6. Note that, complementarily to the structural reasons discussed earlier, in Section 2.2.3 this phenomenon is examined as a measurement issue.

Herein it is important to note that the meaning of development and growth is treated as synonymous. Note that growth is mainly referred as increases in per capita income or increases in output per capita. On the other hand, development can be viewed as a multi-dimensional process that comprehends progressive changes in the socio-economic structure and indicators of a country, such as human development (better educational and health conditions), poverty reduction, improvements in infrastructures, institutional organization of production, human rights as well as a better distribution pattern of income among people. Although, there has been a long discussion of the differences between economic development and economic growth, which is relevant in my view, the thesis does not distinguish between the two concepts[1]. I take this approach because the study carried out in the thesis always looks to growth of income per capita conditioning on several socio-economic factors that are related to the concept of development. Thus, it is possible to argue that the literature on growth and development are intrinsically connected. In this sense, Lucas (1988) explains why to use levels and rates of growth of per capita income as meaning development:

> "*By the problem of economic development I mean simply the problem of accounting for the observed pattern, across countries and across time, in levels and rates of growth of per capita income. This may seem too narrow a definition, and perhaps it is, but thinking about income patterns will necessarily involve us in thinking about many other aspects of societies too, so I would suggest that we withhold judgment on the scope of this definition until we have a clearer idea of where it leads us*" (Lucas, 1988: 3).

As well noted by Ray (1998: 9), "*[n]either Lucas nor any intelligent person believes that per capita income is development. What's hidden in these words is actually an approach, not a definition*". Furthermore, Islam (1995) argue that when the panel data framework is adopted in economic growth analyses, it creates a bridge between development economics and the neoclassical empirics of growth. When differences in the aggregate production function are allowed by means of individual (fixed) effects, they are called upon to focus attention on all the tangible and intangible fixed effects that underlie much of the discussion of development economics. This issue is discussed in more detail in Chapter 5.

[1] To read more about the differences between growth and development, see, for instance, Ray (1998) and Sen (2000).

2.1. Theories of Economic Growth and Their Implications for Growth Policy

There has been a long and intense debate about the rationale for regional economic development policies among academics, specialists and policy makers. This section, rather than providing a complete review of the justifications for all economic perspectives, briefly summarises—in the mainstream of economics—the theoretical justifications for regional economic development policy. In this context this section provides background discussion of theories of economic growth, their spatial scope, and their implications for growth policy that underlie the investigations of the determinants of Brazilian regional economic growth in the next chapters.

Regional policies are justified by the existence of market failures, such as credit market imperfections, externalities and imperfect information. Given these failures, regional development agencies around the world have designed policies to mitigate these failures. Furthermore, “new economic geography” (NEG) models shed light on the possible trade-off between equity and efficiency when regional policies are carried out. Such issues are discussed below.

2.1.1. Neoclassical Growth Models and Their Predecessors

Before Solow’s growth model (Solow, 1956), the discussion concerning the role of the state in promoting economic growth and industrialisation was based on two basic ideas: (i) the concept that higher growth in output per capita was due to higher investment rates, as highlighted by the Harrod-Domar model (Harrod, 1948; Domar, 1946); and (ii) the concept of the “big push” (Murphy et al., 1989), which emphasizes that the government can establish the correct rate of investment across many sectors of the economy, thus creating backward and forward linkages that would make industrialization profitable and self-sustainable. This idea was formerly introduced by Rosenstein-Rodan (1943) and developed by many others (Nurkse, 1953; Scitovsky, 1954; Fleming, 1955; Hirschman, 1958).

However, in light of the neoclassical growth models [introduced by Solow (1956)], the role of the state in reducing regional per capita income disparities weakened. These models—also called exogenous growth models based on constant returns to scale, diminishing marginal returns, perfect competition, and no transaction costs—predict that, due to the diminishing returns to capital, regional disparities are only temporary and should decrease over time. This is the case of the absolute β-convergence, which assumes that because economies are structurally similar and because the production function is the same, there is convergence in both per capita income levels and growth rates at equilibrium. On the other hand, when regions differ in the parameters that determine their steady state (e.g., structural characteristics such as saving rates, schooling, infrastructure)[2], each region should converge toward its own steady state level of per capita income rather than toward a common level. In this case, which is called conditional β-convergence, there is only convergence in growth rates, and it is compatible with the persistence of large differences in levels of development between regions (Islam, 2003).

Indeed, the debate about factors that affect long run economic growth came with Solow's (1956) growth model. However, this model was augmented by many others conditioning variables. For instance, the exogenous growth model has been augmented by the inclusion of education capital (Mankiw et al., 1992), health capital[3] (Bloom et al. 2004; McDonald & Roberts, 2002), migration (Barro & Sala-i-Martin, 2003), and growth externalities (López-Bazo et al., 2004; Ertur & Koch, 2007). Herein, it is worth noting that most of the studies examine convergence process based on

[2] Barro and Sala-i-Martin (2003: 43) explain that "*[t]he key difference between improvements in knowledge and increases in the saving rate is that improvements in knowledge are not bounded. That is, the production function can shift over and over again because, in principle, there are no limits to human knowledge. The saving rate, however, is physically bounded by one. It follows that, if we want to generate growth in the long-run per capita income and consumption within the neoclassical framework, growth must come from technological progress rather than from physical capital accumulation.*" In this sense, neoclassical growth models predict that improvements in, for instance, schooling, infrastructure or saving rate only lead to a higher transitional growth rate.

[3] McDonald & Roberts (2002) develop an augmented Solow model that incorporates both health and education capital since human capital is a complex input that consists of more than the knowledge capital suggested by Mankiw et al. (1992).

aggregate labour productivity, using GDP per worker (or per capita) as the measure[4]. Of note, I should explain that when the analysis is conducted at the regional level where specialization patterns are relatively pronounced, the sectoral dimension of the growth and convergence processes might be a key factor to be examined. However, sometimes data unavailability at the regional level creates constrains for an investigation on this issue (which is the case of this current thesis that aims to investigate economic growth at multiple spatial scales). In a seminal paper, Bernard and Jones (1996) show evidence that convergence of aggregate productivity may mask substantial differences at the sectoral level by examining convergence by sectors for 14 OECD countries during 1970-1987. In sum, the results show that many sectors, such as services, show evidence of convergence that may be driving the aggregate convergence result. However, manufacturing sector shows little evidence of convergence in labour and technological productivity. One possible explanation given by Bernard and Jones (1996) is the distinction between tradable and nontradable goods in a world with specialization and spillovers as modelled by Krugman (1987)[5]. Recently, Mulder and De Groot (2007) provide a systematic comparison of cross-country differences in energy- and labour productivity at a detailed sectoral level for 14 OECD countries over the years 1970 to 1997. It is also important to note that sectoral analysis is not without problems as pointed out by Sørensen (2001), who argues that the applied purchasing-power parities (PPPs) are not appropriate as conversion factors for productivity measurement in manufacturing [e.g., as converted in Bernard and Jones (1996)], showing evidence that convergence in manufacturing depends heavily on the choice of the base year. In a reply, Bernard and Jones (2001: 1169) concluded that "*future research is needed to construct conversion factors appropriate to each*

[4] As noted by Bernard and Jones (1996), the exceptions are Dollar and Wolff (1988, 1993).

[5] "*The nontradable-goods sectors will behave very much like an aggregate growth model, and technological productivity levels will converge in these sectors as the technology for producing similar goods diffuses over time. (…) On the other hand, in the tradable-goods sectors, comparative advantage leads to specialization, and to the extent that countries are producing different goods, there is no a priori reason to expect the technologies of production to be the same or to converge over time*" (Bernard and Jones, 1996: 1237).

sector and that research relying on international comparisons of sectoral productivity and income should proceed with caution until these conversion factors are available".

2.1.2. Endogenous Growth Models and Recent Developments

From the 1990s, using the so-called endogenous growth models (also called "new growth theory") as a base, regional development agencies around the world have implemented policies to carry out a more active regional policy. This wave of research on economics, pioneered by Romer (1986) and Lucas (1988), seeks to explain why differences in per capita income arise and persist over time. These models endogenize economic growth by introducing externalities in production function that are created by investments in human capital and in technology. Other examples of endogenous growth models are Romer (1990), Barro (1990) and Alesina & Rodrik (1994). Romer's (1990) model shows that an economy with a larger total stock of human capital (that is devoted to research sector) will experience faster growth. Barro (1990) relates a high level of productive government spending (e.g., infrastructure) to high rates of economic growth. Alesina & Rodrik's (1994) growth model shows an inverse relationship between income inequality and economic growth.

The general conclusion from new growth models is that government interventions in activities such as R&D (research and development) and education could improve economic growth performance[6]. Moreover, models such as Ravallion and Jalan's (1996) have stressed the importance of geographic variables (for instance, local infrastructure, access to public utilities, knowledge of local environment and local institutions) on affecting the marginal productivity of capital and labour. Ravallion and Jalan (1996) show that community capital has positive external effects in the micro-level growth process just as the externalities of knowledge do in Romer's (1986) model. Using farm-household-level data for rural southern China, the authors reveal a strong

[6] Barros (1993: 543) points out that "the vehicle for such policies could be either subsidy or government direct engagement in externality producing economic activities".

external effect of community capital on the productivity of private investment. In recent years, the role of spatial externalities' effects in the economic growth dynamics has been examined using the appropriate spatial statistics and econometric methods (for the Brazilian case, see Lall and Shalizi, 2003; Silveira-Neto and Azzoni, 2006; Resende, 2011). Given the existence of spatial externalities' effects, it is possible to increase the size of the jurisdiction in order to deal with such spillover, thereby internalizing all of the benefits and costs (Oates, 1999). This evidence sheds light on the relevance of coordination of government policies among jurisdictions and among public and private agents to foster growth, especially the coordination of public investments in lagging regions.

Moreover, in poor areas—like in Brazil's lagging macro-regions—where the ratio of poverty to wealth is high, the existence of credit market failures might imply that this high level of personal income inequality will persist over time. Thus, subsidies to small and medium entrepreneurs—for instance, the regional development funds in Brazil—might resolve the problems of these environments that have imperfect information in the credit market or enforcement difficulties. In this instance, Banerjee and Newman (1993) develop a model whereby the occupational choices made by individuals depend on the distribution of wealth. Based on initial conditions of high personal income inequality and enforcement imperfections, lenders will limit borrowing and require collateral to ensure the re-payment. Thus, anyone with initial wealth below a quantifiable threshold cannot qualify for a loan to finance self-employment or to become an entrepreneur. In this situation, regional subsidies can be an important funding alternative to resolve both personal and regional income inequalities. Using similar arguments, Galor and Zeira (1993) show that divergence between economies can be attributed to differences in investment in human capital due to credit market imperfection as well. If borrowing is difficult or costly and investment in human capital is indivisible (there is a technological non-convexity which imply multiple long-run equilibria), the initial distribution of wealth will determine the long-run equilibrium in the economy. More precisely, in an economy with very unequal distribution of income, the poor dynasties – which individuals inherit less, work as unskilled, and leave less to their children – will have difficulty in borrowing to investment in human capital compared to those rich dynasties who inherit large amounts and have easier access to investment in human capital. The authors conclude that an economy that is initially poor or has a large initial amount of wealth unequally distributed, ends up poor in the long run. On the other hand, "*an economy which is initially rich and its wealth is distributed among many, ends up rich*" (Galor and Zeira, 1993: 42).

However, it is worth noting that these new growth models conclude that this sort of policy (for instance, subsidies to entrepreneurs) may actually generate high incentive distortions in the economy, which outweigh their benefits. For this reason, these models are more cautious in this kind of policy prescription. In fact, Krueger (1990) points out that in many countries government failure significantly outweighed market failure. Government failures can be grouped in terms of (i) commission and (ii) omission. The first set of failures consists of, for instance, "*high-cost public sector enterprises not traditionally associated with the public sector; government investment programs highly inefficient and wasteful; and government public sector deficits, fuelled by public sector enterprise deficits, excessive investment programs, and other government expenditures, led to high rates of inflation, with their attendant consequences for resource allocation, savings behavior, and the allocation of private investment*" (Krueger, 1990: 10). The latter group of government failures can be defined as the "*deterioration of transport and communications facilities, which raised costs for many private (and public) sector activities; maintenance of fixed nominal exchange rates in the face of rapid domestic inflation, buttressed by exchange controls and import licensing; insistence upon nominal rates of interest well below the rate of inflation with credit rationing so that governments could supervise credit allocation among competing claimants; and failure to maintain existing infrastructure facilities. As by-products of these failures, large-scale and visible corruption often emerged*"

(Krueger, 1990: 10). Moreover, it is relevant the notion of the free rider problem when discussing subsidies as an instrument to attract firms to lagging regions in a regional development policy context. The question, in this case, is how to select the firms or individuals who really need the subsidy to start or increase a business in a specific target region. In other words, it is hard to avoid spending money on firms or individuals that would invest in regional development policies (e.g., become an entrepreneur in a lagging region) also without the subsidy. How to face and minimize this problem is still a practical issue to be dealt by policy makers. In an environment of scarcity of resources to subsidize all applicants, it is necessary a careful analysis (sometimes spending more time and money) by the responsible agency (e.g, bank) in choosing and attracting the ones that most need the subsidy to develop a business in a given region. Indeed, as pointed out by Page and Tassier (2010), designing mechanisms that overcome the free rider problem and result in the efficient allocation of, for instance, public goods has long been discussed in mechanism design literature which creatively constructed incentives that induce individuals to truthfully reveal their preferences (Groves and Ledyard, 1977).

During the 1990s, another economic field called "new economic geography" (NEG) focused on developing a formal abstract model of spatial agglomeration. These models have focused on the role that increasing returns combined with transport costs play in generating a concentration of economic activity in a limited number of agglomerations (Krugman, 1991; Fujita et al., 1999; Fujita and Thisse, 2002). In recent years, pioneered by Baldwin et al. (2003), NEG models have discussed implications for policy, including the trade-off between national growth and regional economic equality. In other words, these models suggest that spatial agglomeration (regional inequality) might raise national growth as a whole[7]. Martin (2008: 7) discusses this trade-off and points out that a key implication of these models is that "*policies to stem spatial*

[7] An illustration of this trade-off can be drawn from Baldwin et al. (2003: 430): "*An income transfer to the poor region lowers income inequality and spatial concentration but lowers the growth rate of the whole economy*".

agglomeration, or that seek to reduce it, in an effort to close inter-regional (or intra-regional) economic disparities, may be economically inefficient from a growth point of view". However, the empirical validation of this trade-off is still an open question[8]. Finally, one important discussion that has emerged is the space-neutral versus the place-based approaches that are analysed in Section 3.4 (Chapter 3).

The next subsection seeks to provide a theoretical background for understanding the spatial scope of the determinants of economic growth discussed so far. The growth models discussed above do not distinguish between alternative spatial scales; in other words, the region in these theoretical models can refer to any spatial aggregation. Nonetheless, these models are used next to provide some insights on how explanatory variables may impact economic growth at different spatial scales providing the rationale for their inclusion in the econometric specifications in Chapters 4 and 5.

2.1.3. The Spatial Scope of Economic Growth Determinants: Potential Theoretical Reasons for Different Results Across Models Estimated at Different Spatial Scales

Several factors may have been responsible for driving the regional performance that can be grouped in proximate sources of growth and wider influences (Temple, 1999). The first set of variables consists of the production factors that directly influence regional economic growth, such as physical and human capital. The latter group comprises all other variables that indirectly affect growth by improving knowledge or technology transfer or the efficiency of input allocation via infrastructure, population density, income inequality and spatial externalities, for example. Herein, I provided some theoretical reasons to explain how these explanatory variables may impact economic growth at different spatial scales or, in other words, I discuss the scale-dependent determinants of economic growth (or the structural issue).

[8] See Martin (2008) for a cautionary note on this trade-off.

Proximate sources of growth:

- Physical capital (and the convergence hypotheses). Information on physical capital is often unavailable at subnational levels and, and thus this variable is excluded from the set of explanatory variables of regional growth regressions. The lack of availability of physical capital measures at finer regional level is not restricted to Brazil, as noted in Lesage and Fisher (2008) in a study for Europe. This fact is problematic because it causes the omitted variable problem that may bias the regressions estimates. Some panel data approaches partially deal with this problem by including fixed effects that might control for this kind of omitted variable (Islam, 1995). Despite such omission, the neoclassical growth framework (see Solow, 1956) provides a simple rationale for the convergence hypothesis. The convergence property comes from the law of decreasing returns to capital accumulation [i.e., capital tends to accumulate slower (faster) in regions where it is relatively abundant (scarce)]. As discussed earlier, to test this convergence hypothesis, it is estimated initial levels of regional per capita income on subsequent growth rates as known absolute β-convergence (because economies are structurally similar and because the production function is the same, there is convergence in both per capita income levels and growth rates at equilibrium). However, if regional differences are introduced in the regression by means of structural variables (that determine regions' steady state) such as saving rates, schooling and infrastructure then each region should converge toward its own steady state level of per capita income rather than toward a common level. This is the case of conditional β-convergence, where there is only convergence in growth rates and it is compatible with the persistence of large differences in levels of development between regions (Islam, 2003). For this reason, the initial level of income is introduced to control for decreasing returns to capital accumulation (Ottaviano and Pinelli, 2006). One interpretation of the initial level of income per capita coefficient suggests that if the coefficient for initial per capita income is inversely related to the per capita income growth, the β-convergence prediction of Solow's (1956) model cannot be rejected. However, some authors, such as Friedman (1992) and Quah (1993), highlight that a negative coefficient for initial per capita income can be due to the more general phenomenon of mean reversion[9], and, by reading convergence into this scenario, growth researchers are falling into Galton's fallacy (Islam, 2003). In other words, Bode and Rey (2006: 175) explain that, "*regression towards the mean is actually an 'incidental parameter' or 'errors in variables' problem that biases the convergence parameter systematically towards 'convergence' in cross-section regressions.*" For this reason, instead of indirectly testing and perhaps erroneously finding a convergence phenomenon, an alternative approach is to directly test the convergence of per capita income by evaluating the dynamics of dispersion of this variable. In the results section, I also determine whether the dispersion of per capita income between regions decreases over time, that is, if so-called σ-convergence occurs[10]. Finally, it is important to explain why it is expected to verify differences in the magnitudes of the convergence coefficient at different spatial scales. In the regressions estimates, the sizes of the initial income per capita coefficients are expected to be larger at finer spatial scales because, for instance, municipalities resemble to the notion of an open economy with perfect capital mobility. Barro et al.'s (1995) neoclassical model of the open economy with perfect capital mobility predicts the possibility that economies will jump instantaneously to a steady state of income per capita; this can be understood by a higher rate of convergence. The assumption of a more open economy is not difficult to justify in the municipal level context in Brazil, where the intensity of flows of capital, trade and people across municipal borders are

[9] There is a tendency for a stochastic process to remain close or tend to return to a long-run average value over time.
[10] Islam (2003: 314) points out that despite the limitations of β-convergence results, research interest in this concept has continued, in part because the, "*methodologies associated with investigation of β-convergence also provide information regarding structural parameters of growth models, while research along the distribution approach (σ -convergence) usually do not provide such information.*"

higher than states borders which can be viewed as more closed economies than municipalities (the empirical investigation on spatial dependencies across different spatial scales in Chapters 4 and 5 will shed light on this issue).

- Human capital. Because human capital is a complex input that consists of more than educational capital (McDonald and Roberts, 2002), I decompose the stock of human capital into two parts: educational and health capital. Theoretical and empirical papers have shown that increases in educational capital positively impact the growth rate of per capita income (Mankiw et al., 1992). I measure a region's stock of educational capital according to the residents' average level of schooling (in years), a factor that may raise productivity and ultimately foster economic growth[11]. The central reasons to include health capital in growth equations are its importance in measuring human capital composition and its influence on economic growth. As highlighted by Bloom et al. (2004), healthier workers have several key characteristics: they are physically and mentally more energetic and robust; they are more productive and earn higher wages; and they are less likely to be absent from work because of illness (or illness in their family). I use infant mortality rate as a proxy for health capital to test whether there is an aggregated influence of population health on economic growth. Glaeser et al. (2003) show that the presence of positive spillovers or strategic complementarities creates a "social multiplier" where aggregate coefficients of human capital (proxied, for instance, by years of schooling) will be greater than individual coefficients. In the context of this current study, we can think municipalities as being the micro (individual) level of analysis. For this reason, it is possible to argue that at the aggregate level (e.g., at micro-regional or state level), the coefficient of human capital may be inflated by externalities. Glaeser et al. (2003) point out that the coefficients may rise with the level of aggregation due to the existence of a social multiplier, which also supports

[11] There is an important and controversial discussion in the literature about the influence of schooling on economic growth. See, for example, Pritchett (1996) and Temple (1999).

the idea that there are human capital spillovers, as suggested by a wide body of literature (e.g. Lucas 1988 and Rauch 1993).

Wider influences on growth:

- Local infrastructure. The availability of local infrastructure is captured by an index that accounts for several dimensions of housing services and utilities, such as electricity, sewage, water provision and garbage collection[12]. Theoretical and empirical evidence (Aschauer, 1989; Barro, 1990; Easterly and Rebelo, 1993) has shown that infrastructure spending is likely to raise economic growth rates by improving the productivity of the private sector. Indeed, improvements in the infrastructure in a specific region are likely to increase growth performance in that region; however, the neighbouring regions may be affected. In this sense, the way we aggregate the data and analyse the results of these effects is very relevant. For instance, the results using different spatial aggregation may show varying influence of the flows of mobile capital and skilled labour across different spatial scales that investments in public infrastructure can generate in attracting capital and labour force and ultimately affecting economic growth. At the state level, we cannot observe the flows and effects across local areas within a specific state; we only have the aggregate effect that might represent a blend of individual and contextual effects (Manski, 1993; Anselin, 2002) (See Section 2.2.3 for details). Lall and Shalizi (2003: 679) argue that the effects of investments in public infrastructure in one location can draw production away from other locations, thereby having a detrimental effect on growth performance in neighbouring regions. On the other hand, it is also possible to obverse positive effects in neighbouring regions when central city investments provide benefits to the suburbs[13]; *"[i]n any case, spatial*

[12] See Da Mata et al. (2007b) for further details on how this index was constructed.
[13] Lall and Shalizi (2003: 679) exemplify noting that "*Boarnet (1998) shows that highway projects in California counties provide benefits to investing counties at the expense of other counties within the state, sugesting possibilities of negative output spillovers from public investments. When production factors are mobile, public investments in one location can draw production away from other locations. In another study of spatial spillovers, Haughwout (1998) argues that central city investments provide benefits to the suburbs, demonstrating the case for positive spatial externalities. While Boarnet's work suggests that*

externalities are likely to induce some distortions in economic behaviour" (Lall and Shalizi, 2003: 679).

- Population density. New economic geography models (Baldwin and Forslid, 2000) shed light on the positive impact of agglomeration externalities on economic growth rates. Population density within regions acts as the proxy for the agglomeration externalities. Local density is expected to capture the agglomeration effects within a region. The magnitude of these agglomeration effects may depend on the spatial scales of analysis, because, for instance, population density probably appears to be higher at a finer scale (e.g., municipalities) than population density at a spatial scale within larger regions (such as a state). Thus, centripetal effect of agglomerations might be operating at finer scales or, in other words, agglomeration-related centripetal forces may be much more relevant at the local than the state level.

- Income inequality. Using the political economy argument, the theoretical literature states that high inequality is harmful for growth. In simple terms, if inequality in income and wealth is high, then the median voter will choose a higher level of taxation that lowers economic growth (Alesina and Rodrik, 1994). Beyond this theory, there is a rich literature relating income inequality to growth that can be grouped into four broad categories: "*the endogenous fiscal policy approach, the socio-political instability approach, the borrowing and investment in education approach, and the joint education/fertility approach*" (Dominicis et al., 2008: 656). Benabou (1996), Perotti (1996) and Aghion et al. (1999) present a detailed literature review on this topic. Of note, Dominicis et al. (2008) provide a comprehensive literature review and conduct a meta-analysis approach to systematically describe, identify and analyse the variation in outcomes of empirical studies that relates income inequality to growth. Particularly interesting is the theory

individual counties will be tempted to overinvest, Haughwout's work suggests that individual cities will underinvest in public infrastructure".

developed by Galor and Moav (2004) showing how economic growth may vary depending on the stage of development in a country (or can be viewed as depending on different stage of development of regions within a country, i.e., less developed vs. developed regions). Dominics et al. (2008) neatly summarise the unified theory of Galor and Moav (2004) as follow:

> "*The positive impact of inequality upon growth reflects the situation of an economy during its early stage of industrialization. In this phase, the accumulation of physical capital is the principal engine of growth and it is promoted by disparities among individuals. Once the economy has passed over this initial phase, the accumulation of human capital becomes the prime engine of growth and a more equalitarian distribution of resources allows more people to invest in education. In this stage, in the presence of credit constraints, access to education is easier if wealth is evenly spread among individuals, and hence policy decisions have to be directed towards inequality-reducing strategies. Their conclusions are particularly relevant for less developed countries (LDCs). In contrast with the historical growth path of the currently developed countries, where physical capital was the prime engine of growth, human capital accumulation may be the prime engine of growth in some LDCs, even in the early stages of development. In some of the current LDCs, the strong presence of international capital inflows weakens the beneficial role of inequality in stimulating physical capital accumulation. In addition, the tendency to import skill-based technologies in LDCs increases the returns to human capital accumulation and, given credit constraints, strengthens the negative effect of inequality on human capital accumulation, and thus economic growth*" (Dominics et al., 2008: 656-7).

- Transportation costs. Theoretical models (Ottaviano and Puga, 1998; Lafourcade and Thisse, 2008) have shown that with decreasing transportation costs, regional inequalities will increase and then decrease. Other models integrate an endogenous growth model with the core-periphery model showing that a decrease in transportation costs may have non-linear effects on growth (Baldwin et al., 2003). With regard to Brazil, Da Mata et al. (2007a) have found that transportation costs are inversely related to the rate of economic growth. The non-linearity between growth and transportation costs can be addressed by allowing the impacts of transportation costs on growth to vary between two spatial regimes. This approach is used here; however, future work may use other empirical approaches to test a non-linear relationship between these variables. Moreover, as suggested earlier, the impact of transportation infrastructure on economic growth may vary as the scale of analysis changes. For instance, if this impact is analysed at the state level, the focus will be on the connectivity between these aggregate regions. On the other hand, at the municipal level, such an analysis might examine the impact of transportation cost reductions within the borders of states.

- Spatial externalities. The spatial growth model specifications discussed in the next section seek to capture the effects of spatial externalities. These models introduce a geographical component, which enables measurement of geographic spillovers among neighbouring regions. However, it is important to note that the extent and strength of these externalities may depend on the level of aggregation of the spatial units. For instance, spatial autocorrelation might be higher at the municipal level than at the state level, because, for instance, states are more self-contained than municipalities or, in other words, states are much more closed economic entities than municipalities. As noted by Oates (1999) it is possible to increase the size of the jurisdiction to deal with such spillovers, thereby internalising the benefits and costs. Corrado and Fingleton (2012) note that hierarchical models (also known as multilevel models) can be used in regional science and spatial economics to study a hierarchy of effects from cities, regions containing cities, and countries containing regions; thus, when these effects emanating from different hierarchical levels are not recognized, it can lead to incorrect inference. Herein, the adopted approach is to systematically replicate the regression specification chosen to examine the extent of spatial externalities at a single scale across multiple spatial scales. Lall and Shalizi (2003) enumerate some theoretical reasons why location effects and spatial externalities matter in examining determinants of growth that include: (i) agglomeration economies[14]; (ii) Marshallian externalities of knowledge

[14] *"Drawing on the central argument of the 'new economic geography' literature, growth in any region is influenced by its ability to access large markets (Krugman, 1995; Venables, 1998). These economies are not a function of the size of a specific industry but of the overall size of the agglomeration. Thus, competitive enterprises accessing larger markets can enhance productivity. In addition to market size,*

diffusion and labour market pooling[15]; (iii) common informal norms and institutions[16]; (iv) policy adoption[17]. Although much of these theoretical arguments on spatial externalities are about their positive effects, it is possible to argue some reasons why it can be observed negative spatial externalities effects. For instance, as regard the policy adoption argument discussed above, it could also "*be negative policy imitation where governments may not necessarily maximize growth but maximize rent-seeking and this behavior may be imitated by governments in neighboring regions*" (Lall and Shalizi, 2003: 665). Moreover, Lall and Shalizi (2003: 679) suggest that improvements in the structural variables (e.g., economic structure, workforce quality, and infrastructure quality) are likely to increase growth performance in the region; however, "*if growth in a particular region is higher than that of its neighbors, the region is likely to attract mobile capital and skilled labor from neighboring regions, thereby having a detrimental effect on growth performance in neighboring regions*".

2.2. Econometric Issues on Studying Regional Economic Growth

This section discusses – among a vast list of econometric issues – three econometric issues that are close related to the spatial scope of growth determinants that are examined in this thesis: (i) spatial dependence; (ii) parameter/spatial heterogeneity; and (ii) the Modifiable Areal Unit Problem (MAUP) / Ecological Fallacy (EF). The former deals with econometric issues that are spatial by definition and the last two (ii and iii)

---

agglomeration benefits potentially include access to specialized services (banking and finance), interindustry linkages, physical and economic infrastructure, and a larger medium for information exchange. Limiting the scope of the analysis to administrative units without considering the economic agglomeration (to which the region may belong) and the effects of market access are likely to limit the scope of the analysis" (Lall and Shalizi, 2003: 664).

[15] "*For technological externalities, innovations in one region are adopted in neighboring regions through diffusion, thereby creating convergence in production processes and linkages in development outcomes. In Marshall's (1920) terminology, diffusion is spatially localized and does not extend to all locations. The second source of Marshallian agglomeration is labor market pooling, where production units in one region can gain access to a shared pool of labor in the larger regional economy*" (Lall and Shalizi, 2003: 664).

[16] "*Neighboring regions are quite likely to share common informal norms and institutional structures making them react similarly to exogenous shocks (North, 1990)*" (Lall and Shalizi, 2003: 665).

[17] "*Growth rates could be correlated across space due to 'copy cat policy adoption' (Easterly and Levine, 1998). They suggest that policies leading to high growth may provide a model of the efficacy of public intervention to governments in neighboring regions*" (Lall and Shalizi, 2003: 665).

can be viewed as a separate field of spatial econometrics but as discussed below they can be also discussed in the spatial econometrics context.

The empirical study of economic growth determinants and convergence process is controversial and suffers from substantial drawbacks. For instance, Friedman (1992) and Quah (1993) highlight that a negative coefficient for initial per capita income can be due to the more general phenomenon of mean reversion[18], and by reading convergence into this scenario, growth researchers are falling into Galton's fallacy (Islam, 2003). In this case, the approach adopted in this thesis was to look to convergence using both concepts, the β–convergence and σ-convergence. In sum, the primary econometric problems with this growth literature are related to each other and include: (i) identification of β-convergence and economic divergence; (ii) endogeneity; (iii) outliers; (iv) missing data; (v) measurement error; (vi) robustness with respect to choice of explanatory variables; (vii) spatial correlation in errors; (viii) parameter/spatial heterogeneity; and the (ix) MAUP and EF. For a comprehensive discussion of issues from (i) to (vi), see Durlauf et al. (2005). The problems of spatial dependence (vii), parameter/spatial heterogeneity (viii) and the MAUP/EF (ix) are discussed next. Section 2.2.1 discussed some topics on spatial econometrics, namely, spatial dependence in general, spatial panel data models and spatial weight matrices. Then, parameter/spatial heterogeneity and MAUP/EF are examined, respectively, in Sections 2.2.2 and 2.2.3.

2.2.1. Spatial Econometrics

In this section, some spatial econometric issues are discussed. First, a general discussion on spatial dependence is conducted. Note that, spatial dependence results from (a) the existence of spillover effects; (b) spatially correlated variables that have been omitted; or (c) measurement error or misspecification of the functional form. The

[18] There is a tendency for a stochastic process to remain close or tend to return to a long-run average value over time.

discussion of these issues is important to inform the spatial analysis conducted in Chapter 4 and the exploratory spatial analysis carried out in Chapter 5. Despite spatial panel data models are not estimated in this thesis, a brief discussion on their recent developments is conducted in section 2.2.1.2 as well as some drawbacks of spatial econometrics as whole is highlighted.

2.2.1.1. Spatial Dependence

During the last two decades, an increasing dissemination of spatial econometrics techniques has been observed among regional scientists, economists and researchers in several fields (Anselin, 1988; LeSage, 1999; Conley, 1999). Well known is the vast research of applied spatial econometrics on the interdependencies among spatial units and their effects on, among others, regional economic growth, trade flows, knowledge spillovers, migration, housing prices, tax interactions, city's growth controls (e.g., López-Bazo et al., 2004; Gamboa, 2010; Fischer et al., 2009; LeSage and Pace, 2008, Jeanty et al., 2010; Gérard et al., 2010; Brueckner, 1998).

The idea of spatial dependence, basically, comes from the fact that that observation at one location depends on other observations at a number of other locations (Wilhelmsson, 2002). Indeed, there is a close resemblance to (time-) serial dependence (De Graaff et al., 2001). Corrado and Fingleton (2012) describe a general single equation spatial econometric model specification – the spatially autoregressive model with autoregressive disturbances (SARAR) model – as follow:

$$Y = \rho W_1 Y + X\beta + W_2 X \rho_x + \varepsilon \quad , \tag{2.1}$$

$$\varepsilon = \lambda W_3 \varepsilon + u \, , \tag{2.2}$$

$$u \sim iid(0, \sigma^2). \, , \tag{2.3}$$

In Eq. (2.1) Y is an $N \times 1$ column vector with observations for the dependent variable, X is an $N \times K$ matrix of observations on exogenous variables (for the sake of simplicity X includes the constant term), and ε and u are vectors of error terms.

The spatial matrices W_1 and W_2 allow, respectively, endogenous and exogenous spatial lags and the spatial matrix W_3 represents a spatial error process. LeSage (1999) note that, W_1, W_2 and W_3 can be equal, but there may be identification problems in this case. Thus, ρ and λ are the spatial autoregressive parameters; and β and ρ_x are $K \times 1$ vectors of coefficients. Note that, for the error process, there is a scalar λ and an $N \times 1$ vector of innovations u drawn from an *iid* distribution with variance σ^2.

The SARAR model nests the most common spatial econometric models usually employed in the empirical literature. For instance, as discussed in LeSage and Fischer (2008), imposing that $\lambda = 0$ leads to the Spatial Durbin Model (SDM). On the other hand, imposing that $\lambda = 0$ and $\rho_x = 0$, we have the Spatial Autoregressive (SAR) model. Assuming that $\rho = 0$ and $\rho_x = 0$ leads to the Spatial Error Model (SEM). Alternatively, assuming that $\lambda = 0$ and $\rho = 0$ leads to a spatially lagged X regression model (SLX) that assumes interactions between exogenous characteristics of nearby observations (WX) that directly affect Y. Finally, imposing the restriction that $\rho = 0$, $\rho_x = 0$ and $\lambda = 0$ leads to the standard non-spatial least squares (OLS) regression model. It is important to note that if ρ and ρ_x are significantly different from zero, their omissions in a regression give us biased estimates of β coefficients. These omissions will cause the residuals to be spatially correlated. Moreover, the regression may have spatially correlated residuals because of measurement error or misspecification of the functional form; in this case, using OLS in the presence of non-spherical errors yields unbiased estimates for the estimated parameters (β) but a biased estimate of the parameters' variance. Lee (2004) shows that (quasi) ML estimation provides consistent estimators for these spatial models conditional on the assumption that the spatial econometric model estimated is the true data generating process (Gibbons and

Overman, 2012). Nevertheless, Gibbons and Overman (2012) demonstrate how these models are related to each other. These authors show that the SAR and SLX models are nested within the SDM. Moreover, the SEM model can be rearranged to give the SDM. In this sense, Abreu et al. (2005) stress the over-reliance of the literature on the SEM and SAR models has tended to obscure other models available to capture spatial effects and the failure to consider the model in its reduced form can lead to problems in interpreting the estimated coefficients (in the form of direct, indirect and induced effects). Gibbons and Overman (2012) criticise the spatial models by arguing that distinguishing which of these spatial models generates the data that the researcher has at hand is very difficult in applied research. For this reason, they highlight that "*there are lessons to be learnt from the spatial econometrics literature but for most applied economic researchers the appropriate strategy should be based on the experimental paradigm which puts issues of identification and causality at center stage*" (Gibbons and Overman, 2012: 188). Partridge et al. (2012: 170-171) summarize the current discussion on the usefulness of spatial econometrics that ranges from the view of urban economists that propose to abandon spatial econometrics and follow the experimental route (e.g., Gibbons and Overman, 2012) to the understanding that standard spatial econometrics need to be refined with more careful theoretical treatment, "*constructing better W matrices, and using hierarchical approaches to achieve better identification*" (e.g., Corrado and Fingleton, 2012).

Finally, it is worth noting that there is a line of literature – which is not consensual among researchers – that proposes formal tests of model selection to help practitioners to choose the most appropriate spatial model(s). In a literature review, Abreu et al. (2005) observe that most spatial econometric studies on economic growth chose the appropriate empirical model on the basis of diagnostics tests carried out on the data, rather than, an empirical model derived from a theory. In sum, Abreu et al. (2005) describe the procedure: (i) a model is estimated using Ordinary Least Squares (OLS); (ii) spatial diagnostics is computed; (iii) these diagnostics indicate whether there is spatial autocorrelation in the residuals; (iv) in addition, Lagrange Multiplier tests indicate whether a spatial lag (SAR) or spatial error model (SEM) of spatial dependence is the most appropriate [following the decision rule suggested by Anselin and Florax (1995) and Florax et al. (2003)].

Ertur et al. (2006) describe in detail the LM tests to discriminate between the two forms of spatial dependence, spatial error or spatial lag (respectively LMERR and LMLAG and their robust versions):

> "*A classical 'specific to general' specification search approach outlined in Anselin and Rey (1991) or Anselin and Florax (1995) in the context of spatial econometric modeling can then be applied to decide which spatial specification is the more appropriate. If LMLAG is more significant than LMERR and R-LMLAG is significant but R-LMERR is not, then the appropriate model is the spatial autoregressive model. Conversely, if LMERR is more significant than LMLAG and R-LMERR is significant but R-LMLAG is not, then the appropriate specification is the spatial error model. The performance of such an approach is experimentally investigated in Florax and Folmer (1992). Furthermore, Florax, Folmer, and Rey (2003) showed by means of Monte Carlo simulation that this classical approach outperforms Hendry's (1979) 'general to specific' approach*" (Ertur et al., 2006: 18-19).

Nevertheless, Ertur et al. (2006) stress that this classical approach has drawbacks: (i) the significance levels of the sequence of tests are unknown; every test is conditional on arbitrary assumptions and it does not always lead to the 'best model'; (ii) some studies such as Getis and Griffith (2002) and Cravo and Resende (2012) prefer to filter the variables to get rid of spatial autocorrelation; and (iii) Conley (1999) proposed an alternative approach based on nonparametric estimation of covariance matrices yielding standard error estimates for coefficients that are robust versus spatial autocorrelation and heteroscedasticity, following the idea in time-series (Newey and West, 1987; or Andrews, 1991).

2.2.1.2. Spatial Panel Data Models

Specifically to the panel data methods, Baltagi and Pirotte (2010) examine the standard panel data estimators under spatial dependence using Monte Carlo experiments and show that when the spatial coefficients are large, test of hypothesis based on the standard panel data estimators that ignore spatial dependence can lead to misleading inference. Recently, new developments on spatial panel data models have been emerged in the spatial econometrics literature that propose alternative spatio-temporal models to investigate convergence and growth of regions, regional markets (Keller and Shiue, 2007), labour economics (Foote, 2007), among other fields. Anselin et al. (2008) provide a list of alternative spatial panel data models and the respective likelihood functions, but properties of estimation methods are not analysed. Elhorst (2012: 5) also examines a collection of spatial dynamic panel data (SDPD) models that include one or more of the following variables and/or error terms: "*a dependent variable lagged in time, a dependent variable lagged in space, a dependent variable lagged in both space and time, independent variables lagged in time, independent variables lagged in space, serial error autocorrelation, spatial error autocorrelation, spatial-specific and time-period-specific effects*".

Lee and Yu (2010a) examine some recent developments of spatial panel data models for both static and dynamic cases which consider the fixed effects, spatial lags and spatial disturbances specifications [for another surveys, see also Elhorst (2010a, 2012)]. Specifically, these spatial dynamic panel data models can be applied to investigate economic growth and convergence processes of regions which employ income per capita growth rates versus lagged levels of the explanatory variables (X) including the respective spatial lags (WX). Lee and Yu (2010a) established the asymptotic properties of estimators and provided consistent estimators for spatial panel models. However, spatial dynamic panel data models are not without problems as pointed out by Elhorst (2012: 25):

"(…) *not every method is able to tackle the potential bias in the coefficient of the variable WY_t adequately. Second, some estimators underperform when T is small; treating the initial observations endogenous instead of exogenous may be beneficial under these circumstances. Third, not every estimator is able to deal with endogenous explanatory variables other than the dependent variables lagged in space and/or time. Fourth, the stationarity conditions that need to be imposed on the parameters of the model are not always implemented correctly. A final problem is the treatment of spatial-specific effects; many studies adopt a random effects specification, where a fixed effects specification might be more appropriate*".

Finally, as highlighted by Elhorst (2012) a relevant development in the spatial econometrics literature is the increasing investigation of direct, indirect, and induced effects of the independent variables[19]; both for cross-sectional and to spatial panel data models. This point can be better explored in LeSage and Pace (2009) and Autant-Bernard and LeSage (2011) and it is also discussed in Abreu et al. (2005) and Partridge et al. (2012).

Of note, more than ten years after the first paper of Elhorst (2001) dealing with dynamic models in space and time, Elhorst (2012) discusses the most popular forms of dynamic panel data models stressing that each form appeared to have certain shortcomings and these forms are dependent on the purpose of a particular empirical study and the structure of the data, the researcher might determine which form is most appropriate and which estimator to apply. As highlighted earlier, although it is not

[19] Abreu et al. (2005) provide an explanation of the correct interpretation of a marginal effect of X on Y using the spatial lag model. The marginal effect on an increase in X on Y is $\partial Y / \partial X = I\beta + \rho W\beta + \rho^2 W^2\beta + \rho^3 W^3 \beta ...$, where "*the first term on the right hand side is a matrix with direct effects on the diagonal (the effects on y_i of a marginal change in x_i, where i refers to a spatial unit), and zeros in off-diagonal positions. The second term is a matrix with zeros on the diagonal and indirect effects in the off-diagonal positions for the regions j, defined as the neighbors of I in the spatial weights matrix. These indirect effects are spillovers of the direct effects, and both effects are local in the sense that only the regions in the sense that only regions in which there has been an exogenous shock and their neighbors are affected. The third and higher-order terms refer to spatial spillovers induced by the direct and indirect changes in the first and second terms, and therefore be referred to as induced effects. (…) [T]his implies the spatial lag model links all the regions in the system, so that the spatial effects in the model are global in nature. (…) It is therefore incorrect to compare the coefficient β of the spatial lag model with the coefficient β in an OLS model, since the first represents only the direct marginal effect of an increase in X, while the second represents the total marginal effect*" (Abreu et al., 2005: 31-2). In addition, Partridge et al. (2012: 169) point out that it is inappropriate to estimate the indirect effects from a spatial lag model when, for instance, the spatially lagged X regression model (SLX) is the corrected one and conclude that "*[t]he upshot is that supporting theory and more careful specification tests are needed before calculating such indirect effects*".

consensual among researchers, the LM test is one option that can be used to choose the most appropriate spatial panel data model as explained in the previous section.

2.2.1.3. Spatial Weight Matrices

After discussing the issue of spatial dependence and the alternative spatial econometric models, this subsection briefly describes the most commons spatial weight matrices (W matrices) used in applied work and discusses pros and cons of using each one. The W matrix is used in the spatial econometrics literature to model the spatial interactions among spatial units. Ertur et al. (2006) point out that, usually, each region is connected to a group of neighbouring regions by means of a purely spatial pattern introduced exogenously, by assumption, in this W matrix. On the other hand, weights based on "social distance" or "economic distance[20]" should be used with more care because of endogeneity issues. It is worth noting that the regional definitions have not been randomly created, as they depend on, e.g., history and population. However, the purely spatial weight matrix presents fewer endogeneity problems than those that use "social distance" or "economic distance" as a measure of interaction. W is the row standardised spatial weights matrix where w_{ij} are elements of a spatial weighting matrix such that the elements w_{ij} in each row sum to 1. The elements w_{ij} indicate the spatial connectivity between region *i* and region *j*; and w_{ii} on the diagonal are set to zero.

Several W matrices can be used: a simple binary contiguity matrix[21], a binary spatial weight matrix with a distance-based critical cutoff, above which spatial interactions are assumed negligible, more sophisticated generalized distance-based spatial weight matrices (based on distance decay as such, inverse distance or inverse squared distance) with or without a critical cutoff; yet, "*[t]he critical cutoff can be the*

[20] Corrado and Fingleton (2012) discuss some examples based on "economic distance" weights.
[21] For instance, the standardised first-order contiguity matrix (also called the queen contiguity matrix) in which the element w_{ij} in the matrix is 1 if areas *i* and *j* share borders or vertices, and 0 otherwise.

same for all regions or can be defined to be specific to each region leading in the latter case, for example, to k-nearest neighbors weight matrices when the critical cutoff for each region is determined so that each region has the same number of neighbors" (Ertur et al., 2006: 12).

For instance, in Chapter 4 I consider pure geographical distance, i.e., the spatial weight matrix W is based on the k-nearest neighbours calculated from the great circle distance between region centroids. I conduct that analysis using a spatial weight matrix based on the ten nearest neighbours (k=10). In addition, the results were carried out using k = 5 and k = 15 to check for robustness. In Chapter 5, a binary spatial weight matrix with a distance-based critical cutoff, above which spatial interactions are set to zero, are used because the purpose of the analysis is to provide a measure of the extent of spatial autocorrelation in kilometres. For this reason, alternative cutoffs are employed in that analysis.

Finally, note that there are pros and cons in using alternative W matrices, namely the fact that k-nearest neighbours weight matrices (used in Chapter 4) keep fixed the number of neighbours (and thus accounting for large/remote regions in some parts of the country and very small/linked regions in other parts of the country when you have in the same sample of large and small regions) while the distance-based matrices (used in Chapter 5) fix the size of the spatial field, thus accounting for more than the relevant neighbouring regions in small/metropolitan regions. Abreu et al. (2005) suggest that whenever possible the choice of critical values for distance-based matrices should be based on theoretical considerations [see Corrado and Fingleton (2012) for a deeper discussion on the theoretical justifications for alternative W matrices]; however, they also note that if there is high heterogeneity in the size and spatial distribution of the regions under analysis, it might be difficult to choose a critical cutoff based on empirical considerations. It is important to keep in mind that analyses in Chapters 4 and 5 are conditional upon the choice of the spatial weight matrices. Although, the results of statistical inference depend on spatial weights, I used alternative k-nearest neighbours weight matrices and alternative distance-based critical cutoffs matrices to check for robustness of the results. At this point, it is important to note that apart the choice of a specific W matrix (for instance, k-nearest neighbours or distance-based) the most important issue is the analysis of the robustness of the results using k = 5, 10 and 15 (see Chapter 4) or alternative distance cutoffs (see Chapter 5)[22]. Moreover, it is important to bear in mind that there is a growing literature that calls for "*a stronger more theoretical basis for W to supplement the very significant atheoretical empirical foundations (…) from current work on games, network formation, dynamics and equilibria that is occurring within the social science, notably within the economics of networks*" (Corrado and Fingleton, 2012: 27). In this sense, Corrado and Fingleton (2012) argue that the concept of the W matrix is undeniably necessary in one form or another and is in any case almost inescapable.

2.2.2. Parameter/Spatial Heterogeneity

While spatial dependence has been the focus of empirical investigation [Rey and Montouri (1999), Fingleton (1999), López-Bazo et al. (2004) and Ertur and Koch (2007), for instance), parameter or spatial heterogeneity is less examined *per se*. The more general issue is coined as parameter heterogeneity and is associated to the idea of club convergence[23]which is examined – in non-spatial models using different statistical methodologies – in Durlauf and Johnson (1995), Desdoigts (1999) and Durlauf et al. (2001). The assumption is that coefficients are not stable across different regimes (or clubs). Durlauf et al. (2005) discuss the importance of testing for parameter heterogeneity. Indeed, Quah (1996, 1997) suggests analysis of the distribution of GDP per capita to identify different dynamics across economies and argues that traditional

[22] In general, results found in both chapters suggest that spatial dependence wanes as the number of neighbouring regions increases, suggesting that spatial interactions are bounded in space.
[23] As highlighted by Ertur et al. (2006: 8) "*the concept of club convergence is based on endogenous growth models that are characterized by the possibility of multiple, locally stable, steady state equilibria as in Azariadis and Drazen (1990). Which different equilibria an economy will be reaching depends on the range to which its initial conditions belong. In other words, economies converge to one another if their initial conditions are in the 'basin of attraction' of the same steady state equilibrium. When club convergence exists, the convergence equation should be estimated per club, corresponding to different regimes*".

growth regressions could be misleading because they are not able to capture different dynamics across regions. Ertur et al. (2006) note that similarities in legal and social institutions, culture and language may present economic structure that are locally uniform in a spatial context. This fact may create situations where rates of convergence and other growth determinants show similar estimates for observations located nearby in space. For this reason, Ertur and LeGallo (2008: 3) argue "*parameter heterogeneity is then spatial in nature and estimating a 'global' relationship between growth rate and initial per capita income, which applies in the same way over the whole study area, doesn't allow capturing the important convergence rate differences that might occur in space*". In this sense, the spatial club-convergence analysis [Ertur et al. (2006) and Fischer and Stirböck (2006), for instance] links the idea of club-convergence examined in Durlauf and Johnson (1995) to the notion of spatial heterogeneity.

Ertur and LeGallo (2008) note that, the instability in space of economic relationships (coined as spatial heterogeneity) can be observed at several spatial scales, for instance, behaviours and economic phenomena are not similar in the centre and in the periphery of a city, in an urban region and in a rural region, in the "West" of the enlarged European Union and in the "East", and in the "North" and in the "South" in Brazil. It is important to note that the problem is more serious in cross sections because information is not available over time and individual heterogeneity cannot be controlled.. In this sense, the problem described as spatial heterogeneity can be viewed as an omitted variable problem and one way to control this problem is to include location fixed effects (e.g., regional dummies) or other explanatory variables to account for omitted factors that vary at the regional or local level (e.g., environmental or amenities variables).

In addition, Ertur and LeGallo (2008) explain that spatial heterogeneity can be reflected by varying coefficients (structural instability), or by varying error variances across observations, (groupwise heteroskedasticity), or both:

> *"In an econometric regression, these differences may appear in two ways: with space-varying coefficients and/or space-varying variances. The first case is labeled structural instability of regression parameters, which vary systematically in space. The second case pertains to heteroscedasticity, which is a frequent problem in cross-sections"* (Ertur and LeGallo, 2008: 3).

Therefore, as proposed by Ertur et al. (2006) spatial convergence clubs can be detected using exploratory spatial data analysis, which relies on geographic criteria. This approach is applied in the spatial analyses conducted in Chapter 4.

Finally, it is important to note that the links between spatial autocorrelation and spatial heterogeneity are quite complex (Ertur et al., 2006). Indeed, Anselin (2001) points out that spatial heterogeneity often occurs jointly, with spatial autocorrelation in applied econometric studies. Indeed, Ertur et al. (2006) argue that, in cross section, spatial autocorrelation and spatial heterogeneity may be observationally equivalent; in other words, a spatial autocorrelation of residuals may simply indicate that the regression is misspecified[24]. De Graaff et al. (2001) present some reasons why spatial dependence and spatial heterogeneity should be handled jointly: (i) there may be no difference between dependency and heterogeneity in an empirical analysis; (ii) spatial dependency may create a particular form of heteroscedasticity, and (iii) it may be difficult to separate the two effects in an empirical study.

[24] "*For example, in polarization phenomena, a spatial cluster of extreme residuals in the center may be interpreted as heterogeneity between the center and the periphery or as spatial autocorrelation implied by a spatial stochastic process yielding clustered values in the center. Finally, spatial autocorrelation of the residuals may be implied by some spatial heterogeneity that is not correctly modelled in the regression (Brundson, Fotheringham, and Charlton (1999) provided such an example)*" (Ertur et al., 2006: 10). In sum, according to these definitions, one form of spatial heterogeneity (spatial heteroscedasticity) is just a specific version of the SEM model with a block diagonal spatial weights matrix. Another sort, spatial heterogeneity in the intercept is just an omitted variables problem. The only sort of spatial heterogeneity that is distinct is spatial variation in the coefficients (which may or may not be an omitted variables problem).

2.2.3. Modifiable Areal Unit Problem (MAUP) and Ecological Fallacy (EF)

This section treats, as a measurement issue, the variation in statistical results of regional economic growth estimates across different levels of spatial aggregation in the observational units. In other words, the cause of the variability in economic growth estimates can be, perhaps, due to the Modifiable Areal Unit Problem (MAUP) or the Ecological Fallacy (EF). Openshaw and Taylor (1981: 60) explain that MAUP is related to the fact that the same individual data can be aggregated to show alternative areal representations "*which yield data with little resemblance to the 'real' data that existed before any spatial aggregations occurred*". Moreover, EF is related to the fact that parameters estimated from macro-level data are not appropriate to make inferences about behavioural and socio-economic relations at a more disaggregate level (individual/micro-level). In this sense, both MAUP and EF indicate an aggregation bias or effect. Of note, the term EF is typically used in social sciences (e.g., Hannan, 1971) and it is similar to the MAUP in geography (e.g., Openshaw and Taylor, 1979, 1981; Openshaw, 1984) as discussed in Peeters and Chasco (2006). However, it is important to point out that that MAUP may be more complicated than EF because MAUP is associated to the uncertainties on the choice of alternative number of zones (or zoning systems) and the implications that this holds for spatial analysis (Openshaw and Taylor, 1981). As pointed out in Chapter 1, it is important to distinguish between the scale and the zoning effects that MAUP may present. As discussed in Openshaw and Taylor (1981), the scale of study is related to the selection of an appropriate number of zones; however, it is possible to produce alternative zoning systems by regrouping zones at a given scale. Openshaw and Taylor (1981) review the literature on MAUP and gives an example – using a small data set of 99 Iowa counties – that 12 alternative zone aggregations (scale levels) produce a robust regression coefficient that varied from -14.6 to +16.2 and a level of fit which is nearly perfect or incredibly poor; yet, similar zoning effects have also been observed. Openshaw and Taylor (1981: 62) noted that "*there are no longer grounds for believing that methods exist which are not affected, unless there is proof. It cannot be assumed as a matter of faith*".

Section 3.2 (in Chapter 3) shows that many studies in the context of growth regressions for the Brazilian case (and for growth regressions applied for other countries as well) do not employ a rigorous analysis of spatial scale choice and do not make any comparison between spatial scales. In a recent paper, Briant et al. (2010) evaluated, in the context of economic geography estimations, the magnitude of the distortions potentially induced by the choice of various French geographic stratifications. From this specific exercise, they concluded that the first source of MAUP (scale or number of zones) is prejudicial to economic geography estimations, whereas the second source (zoning) is less important. Furthermore, they found that distortions due to specification choices are much larger than variations due to scale or zoning choice. However, Briant et al. (2010) note that the French administrative zoning systems are less sensitive to the MAUP by definition[25]; thus, it is important that researchers replicate the exercises in the context of other countries. In addition, these authors pointed out that there are many other questions in empirical economic geography on which the magnitude of the MAUP should be assessed. Of note, to my knowledge Resende (2011) – which is a version of the paper presented in Chapter 3 – is the first study of regional economic growth exploring alternative spatial scale dimensions. Interestingly, Abreu et al. (2005) use a meta-analysis of around 600 estimates taken from a random sample of empirical growth studies published in peer-reviewed journals to investigate the variations in the magnitude of (beta) β-convergence rate. From this meta-analysis strategy, a series of factors that may create heterogeneity in the empirical convergence literature which include the spatial level of aggregation (countries or regions) have been identified. This evidence suggests that

[25] Briant et al. (2010) explain that the French economical and institutional design may be particularly well-designed to minimize MAUP problems because the size of the French Départements "*was chosen so that individuals from any point in the Département could make the round trip by horse to the capital city in no more than two days, which translated into a radius of 30–40 km. Hence, it might well be that the French administrative zoning systems are less sensitive to the MAUP by definition*" (Briant et al., 2010: 301).

the spatial unit of analysis is an important factor to be examined in empirical growth studies.

Openshaw and Taylor (1981) neatly state that the problem (MAUP) can be avoided simply by not studying spatially aggregated data. However, most empirical research on economic growth performance has used aggregated data because the common way to calculate income growth between two or more periods is to use aggregate data (from countries, regions, counties, etc.). Aggregate studies may also go in this direction because they need to use such obvious macroeconomic variables as inflation, investment, roads, amenities, etc., which are, by definition, aggregate variables. In this sense, Openshaw and Taylor (1981: 60) concluded that "*if the problem of studying data for modifiable units cannot be avoided then it is essential that the consequential limitations of such studies are clearly understood*".

As noted earlier, there is a similar issue that is related to the aggregation problem. It is referred to as ecological regression and is often criticised for yielding invalid inferences, the so-called Ecological Fallacy (EF) problem (Anselin, 2002). This subtle problem is linked to fundamental differences in the underlying economic process under study. More precisely, the ecological fallacy happens when behavioural and socio-economic relations are inferred for a disaggregate level (micro-level) of analysis using parameters estimated for an aggregate level (macro-level). Anselin (2002) observes that, even in very simple situations, the ecological approach creates problems of interpretation.

Anselin (2002: 21) provides a simple example in which a regression model is specified at the individual level where both individual-level variables [x_{ik} is a characteristic of individual *i* in group *k* (e.g., income for household *i* in municipality k] and group-wise aggregates [$\overline{x_k}$ is the group average for that characteristic[26] (e.g., municipal average income)] are included:

$$y_{ik} = \alpha + x_{ik}\beta + \overline{x_k}\gamma + \varepsilon_{ik} \tag{2.4}$$

In the "neighbourhood effects" literature (Manski, 1993, 2000), β corresponds to the *individual* effect and γ the *contextual* effect. The macro regression that relates the group averages to each other (where $\overline{y_k} = \sum_i y_{ik} / n_k$) is specified in Eq. (2.5).

$$\bar{y}_k = \alpha + \overline{x_k}(\beta + \gamma) + \bar{\varepsilon}_{ik} \tag{2.5}$$

The implications of this aggregated model are twofold. First, at this aggregate level, the error term will become heteroskedastic because the groups do not have the same number of members. Second, separate identification of the individual and contextual effects are no longer possible because the coefficient of the average $(\beta + \gamma)$ in the aggregate model represents a blend of individual and contextual effects[27]. Here, it is worth noting that even if we assume that the municipal level is the micro-level of analysis (instead of the household), the problem appears again when the study is carried out using another aggregate level.

Finally, as demonstrated by Anselin (2002), if the spatial dimension is added to the former example, some other complexities become evident. For example, in the SAR model where a spatially lagged dependent variable ($\rho\sum_{j=1}^{n} w_{ij} y_{jh}$) is added on the right-hand side of the Eq. (2.4). This specification is usually implemented to model a spatial reaction function for economic agents *i*. Formally, the neighbourhood rule is defined by the specification of the spatial weights matrix, which is an $n \times n$ positive

---

[26] Where, $\overline{x_k} = \sum_i x_{ik} / n_k$ (n_k is the group size).

[27] Anselin (2002: 22) points out that "*even when there is no within-group heterogeneity (all the groups have the same β and γ coefficients), the estimate from the aggregate model only corresponds with an individual-level coefficient when there is no contextual effect (γ = 0). Similarly, it only corresponds to a "pure" contextual effect when there is no individual effect (β = 0)*". See Greenland (2002) for more explanations.

matrix (W). In each row *i*, a non-zero element w_{ij} defines *j* as being a neighbour of *i*. The diagonal elements are zero $(w_{ii} = 0)$ because an observation cannot be a neighbour of itself. At the aggregate level, the spatial lag dependent variable for the groups g (g=1,…,G) would be $\lambda\sum_{g=1}^{G} w_{kg}\,\overline{y}_g$. However, Anselin (2002) shows that the aggregate of groups of individual-level spatial lag terms are not equal to the spatial lag of the aggregate values. Basically, if the individual spatial weights included non-zero elements for individuals in the same group, then the aggregate weights should show non-zero diagonal elements, $w_{kk} \neq 0$, a situation that is usually ruled out. Recently, Arbia and Petrarca (2011) present a general framework to investigate the effects of MAUP on spatial econometric models showing how the presence of spatial effects affects the classical results. Arbia and Petrarca (2011) concentrate on the loss in efficiency of the parameters' estimators due to aggregation.

At present, these are worrying problems for the empirical economic growth literature, which has seen in the last years an increasing dissemination of alternative spatial scales of analysis and a growing interest on spatial econometric models. Further efforts are need to be made in understanding economic growth performance at different scale levels. The lack of care in the treatment of aggregation problems on empirical research, is neatly put by Openshaw and Taylor (1981: 63) as follows:

> "*Despite the potential havoc that this may well cause to all manner of geographical studies, it is apparent that geographers have been very slow at recognizing the implications. They have been equally reticent about making clear their assumptions with respect to this problem. Indeed, the standard practice is to simply ignore it altogether in the development and application of methods of analysis that completely fail to take it into account. What is even worse, many geographers see no need to even try*".

What is surprising is that more than thirty years after this statement, it remains true. Moreover, I would extend this lack of attention on MAUP by geographers to researchers in other fields such as regional scientists, economists working on regional and urban issues, among others practitioners.

3. Regional Economic Growth and Regional Economic Development Policy Evaluation in the Brazilian Context

This chapter locates the theoretical discussion conducted in the previous chapter for the Brazilian context. This contextualization is important to give a better understanding of empirical investigation of economic growth determinants carried out in Chapters 4 and 5 and the assessment of the Brazilian economic development policy conducted in Chapter 6. Section 3.1 presents the Brazilian spatial scales employed throughout the thesis. In Section 3.2, the literature on economic growth in Brazil at different spatial scales is reviewed. Section 3.3 describes some socioeconomic information on Brazil at different spatial scales. Section 3.4 discusses the justification for regional economic development policies in the Brazilian context. Finally, in Section 3.5 some issues related to the evaluation of regional development policies in Brazil are discussed such as its policy process, its objectives and the types of evaluation as well as it reviews the strategy of the Brazilian regional development funds since 1989. The discussion in Sections 3.4 and 3.5 provide useful background information for the impact evaluation of the FNE industrial loans that is conducted in Chapter 6.

3.1. Spatial Scales in Brazil for Studying Economic Growth

When the economic growth process is analysed across multiple scales, it is possible to obtain a better understanding of how geography shapes economic growth. There are many types of regions in Brazil, ranging from densely settled urban centres to sparsely settled rural regions. Section 3.3 shows some descriptive data (such as population density, income per capita and schooling) on different spatial scales in Brazil. Brazil is roughly twice the size of the European Union (which consists of 27 countries) and is divided into 27 states[28] that are the main political-administrative units in the country. Municipalities represent the smallest administrative level, dealing with local policy

[28] More precisely, there are 26 states and one federal district.

implementation and management. The "Atlas do Desenvolvimento Humano no Brasil" (IPEA, PNUD and FJP, 2003) provides data from the Census of 1991 using the 5,507 municipalities that existed in 2000, rather than the 4,491 municipalities that existed in 1991. Thus, it is possible to use municipal data with constant borders over the 1991-2000 period. This is the period analysed in Chapter 3. Moreover, in Chapter 4, to extend the analysis over the 1970-2000 period, the study make use of 3,657 minimum comparable areas (MCAs)[29] which can be interpreted as municipalities with constant borders over the 1970-2000 period. In this sense, it was necessary to make some adjustments in the data because the number of municipalities increased from 3,920 municipalities in 1970 to 5,507 municipalities in 2000. To address this problem, municipalities were merged into 3,657 MCAs – defined by Reis et al. (2005) as sets of municipalities whose borders were constant from 1970 to 2000.

It is worth noting that the analyses carried out at the state level do not always provide sufficient detail to satisfactorily capture unobserved heterogeneity. This caveat may mask meaningful geographic variation relative to smaller units of analysis. On the other hand, the use of municipal-level data has a potential to provide spatial autocorrelation that can arise as artefacts of slicing up homogenous regions. Municipalities are territorial units for the production of regional statistics for Brazil and the municipal boundaries might not always approximate the functional borders of the regional economy.

One solution to this problem has been to define functional regions. An example of such functional regions is the 559 micro-regions (used in Chapter 4) defined by IBGE [Instituto Brasileiro de Geografia e Estatística (Brazilian Institute of Geography and Statistics)] in 1990 as being a group of contiguous municipalities within the same state. These micro-regions were grouped according to natural and production characteristics.

[29] The total number of MCAs is 3,659, but Chapter 4 uses 3,657. Fernando de Noronha (in the state of Pernambuco) and Ilhabela (in the state of São Paulo) were excluded because they are islands and do not adjust to the spatial weight matrices used in the analyses. These exclusions do not alter the results of the chapter.

In the analysis that follows in Chapter 5, it was necessary to make some adjustments and the 559 micro-regions were merged into 522 micro-regions defined by Reis et al. (2005) as sets of micro-regions whose borders were constant from 1970 to 2000. IBGE also defines 134 meso-regions that are lager areas than the micro-regions. Meso-regional scale is based on the following dimensions: the social aspects, the natural setting, and the communication and place network as an element of space articulation. According to IBGE (2011) the division of Brazil into micro- and meso-regions is "*relevant to formulate public policies; to subsidize the system of decisions relative to the localization of economic, social and tributary activities; to subsidize the planning, surveys and identification of space structures of metropolitan areas and other forms of urban and rural agglomerations*." Another example of functional region is the spatial clusters proposed by Carvalho et al. (2007). These authors defined 91 spatial clusters by employing an original cluster methodology in the form of an algorithm that grouped contiguous municipalities that share similar characteristics among the 46 variables reported in the Brazilian Census of 2000. Table 3.1 shows the four spatial scales and some statistics concerning their sizes (in square kilometres) used in the empirical analyses conducted in Chapter 4.

Figure 3.1

Spatial Scales in Brazil to Analyse the Period Between 1991 and 2000

States (n = 27)	**Micro-regions** (n = 559)	**Spatial Clusters** (n = 91)	**Municipalities** (n = 5,507)
			
Area Mean = 315,982 Km2 Area Min = 5,822 Km2 Area Max = 1,577,820 Km2 Area Standard Deviation = 378,718	Area Mean = 15,262 Km2 Area Min = 18 Km2 Area Max = 333,857 Km2 Area Standard Deviation = 29,659	Area Mean = 93,753 Km2 Area Min = 350 Km2 Area Max = 1,340,216 Km2 Area Standard Deviation = 196,110	Area Mean = 1,549 Km2 Area Min = 3 Km2 Area Max = 161,446 Km2 Area Standard Deviation = 5,738

Note: Own elaboration from data of IBGE and Carvalho et al. (2007).

Furthermore, Figure 3.2 shows the four spatial scales and some statistics concerning their sizes (in square kilometres) used in the panel data analysis of Chapter 5 over the 1970-2000 period.

Figure 3.2
Spatial Scales in Brazil to Analyse the Period Between 1970 and 2000

States (n = 27)	**Meso-regions** (n = 134)	**Micro-regions** (n = 522)	**MCAs*** (n = 3,657)
			
Area Mean = 312,994 Km^2 Area Min = 5,771 Km^2 Area Max 1,558,987 Km^2 Area Standard Deviation = 372,070 Km^2	Area Mean = 63,066 Km^2 Area Min = 2,937 Km^2 Area Max 650,338 Km^2 Area Standard Deviation = 103,804 Km^2	Area Mean = 16,189 Km^2 Area Min = 190 Km^2 Area Max 439,498 Km^2 Area Standard Deviation = 42,083 Km^2	Area Mean = 2,311 Km^2 Area Min = 8 Km^2 Area Max 367,284 Km^2 Area Standard Deviation = 14,157 Km^2

Note: Own elaboration from data of IBGE. * Minimum Comparable Areas (MCAs).

3.2. The Literature on the Role of Spatial Scales on Economic Growth in Brazil

There are several studies about regional economic growth in Brazil. However, these studies examine the processes of convergence and the determinants of the economic growth only at a single spatial scale. A sample of studies on convergence in Brazil includes Ferreira and Diniz (1994), Azzoni (2001), Azzoni et al. (2000), Vergolino et al. (2004), De Vreyer and Spielvogel (2005) and Silveira Neto and Azzoni (2006) among others that are discussed below. Surveying the Brazilian literature about the determinants of economic growth I have found plenty of papers discussing the theme using state level data, very few papers using micro-regions data and an increasing number of papers in recent years that employ municipal aggregation of data.

Most of papers use state level data to run growth regressions. Ferreira and Diniz (1994) find absolute β-convergence of per capita income among Brazilian states in the period 1970-1985. Similar results are found for the period 1948-1995 (Azzoni, 2001). Ferreira (1999) shows that the results about absolute β-convergence among states in Brazil are robust with regard to period variations. On the other hand, some papers test the prediction of conditional β-convergence including some exploratory variables in economic growth regressions. Using ten cohort means (for a given state in a given year), Azzoni et al. (2000) reveal the existence of conditional β-convergence and indicate that the geographical variables (climate, latitude and rain) seem to be important determinants of economic growth. Furthermore, the results show that schooling and infrastructure variables (sewerage system and piped water) are some of the main factors behind the differences in steady-state rate of income growth in Brazil between 1981 and 1996. Silvera Neto (2001) shows empirical evidence of growth spillovers among Brazilian states economies in the period 1985-1997 by using spatial econometric models. However, Silvera Neto and Azzoni (2006) show that after conditioning on the initial educational levels and manufacturing shares of the states, spatial dependence disappears over the period 1985–2001. Finally, Resende & Figueirêdo (2005) run two robustness tests[30] using 25 variables suggested by the literature for Brazilian states between 1960 and 2000. The estimations of panel data models show that urbanization, infant mortality rates, fertility rates, climate, tax burden and migration have a robust correlation with the growth rates of GDP per capita of the Brazilian states. Moreover, they do not reject the occurrence of conditional β-convergence for the Brazilian states.

Another spatial scale used to study Brazilian economic growth determinants is called micro-regions as shown in Figure 3.2. Vergolino et al. (2004) include initial income, regional dummies and education as exploratory variables to analyze the process of economic growth for the Brazilian micro-regions during the period 1970-96. They argue the existence of two clubs of convergence in Brazil: North/South and

[30] The first approach is the Extreme Bounds Analysis (EBA) test proposed by Levine and Renelt (1992). An alternative approach was considered by Sala-i-Martin (1997) where he argues that instead of analyzing the extremities of the coefficients estimates of a specific variable, it is necessary to make the analysis of the distribution of all coefficients of this variable.

Northeast/Southeast/Centre-West. In the former, it shows a high speed rate of convergence and in the latter there is not any signal of convergence process. Furthermore, the results support the hypothesis under which human capital plays an important role in the economic growth of Brazilian micro-regions. Cravo (2010) is another study on Brazilian economic growth using a panel of 508 Brazilian micro-regions for the period 1980–2004. The author shows the presence of SMEs is not positively correlated with economic growth at micro-regional level and SMEs' human capital is more important for growth in more developed regions.

Recently, growth regressions have been used to discuss economic growth among the Brazilian municipalities. Andrade et al. (2002) find evidence in favour of absolute and conditional β-convergence, for the period 1970-1996, using both OLS and quantile regressions[31]. When regional dummies are added to the estimation, results from OLS and quantile regression are not statistically different. The exceptions to this rule are the North and Northeast regions that present different results from OLS when using quantile regression. However, the conclusion in favour of convergence still remains (Andrade et al., 2002). Also, De Vreyer and Spielvogel (2005) employ municipal units to analyse Brazilian economic growth for the period 1970-1996. The main equation includes the per capita GDP in 1970 to test for conditional β-convergence, spatial lags of GDP per capita in 1970 and economic growth rates, a set of controlling variables, and regional dummies that could cause differences in the rate of technological progress and the steady state across municipalities. By using spatial econometric models they found spatial externality effects and conditional β-convergence at work among municipalities. Furthermore, the illiteracy rate, the primary sector (agriculture) share and the share of urban population are negatively correlated

[31] Coelho and Figueiredo (2007) employ another technique to analyse economic growth of Brazilian municipalities over the period 1970-2000: the regression tree approach proposed by Durlauf and Johnson (1995) and Johnson and Takeyama (2003) that allows testing the club convergence hypothesis. The results based on the regression tree method demonstrate the importance of initial conditions such as income per capita and human capital.

with economic growth. On the other hand, the mean size of households and the share of households with electricity[32] have a positive effect on municipal economic growth.

This variability in previous results should not be entirely attributed to the measurement issue (MAUP and EF, see Section 2.2.3) or interpreted as a statistical artefact. Statistical variations, such as those discussed above and later in this thesis, might be related to a structural issue according to the discontinuity perspective because, in the case of the economic growth debate, different relationships may exist between explanatory variables (including spatial externalities) and economic growth on different spatial scale levels. As the sociologist Hannan (1971: 3) has so neatly stated, "*Those who operate from discontinuity perspectives will certainly expect to find large and important differences in analogous models estimated at different levels of aggregation. However, to those who operate from continuity or homology assumptions, such effects should be quite disturbing. Since these effects would not have any direct theoretical meaning, the variations in estimates obtained at different levels must be considered a statistical artifact.*" The discontinuity thesis is analysed by Hannan (1971) who argues from a substantive point of view that it is necessary to develop cross-level theories to deal with scale problems. Such theories could connect micro- and macro-processes providing explanations for the relationships between micro- and macro-variables (e.g., for instance, individual level => neighbourhood level => municipal level => micro-regional level => state level). From the perspective of regional economic growth literature, such an approach might use structural arguments – which I coined in Section 2.1.3 as the scale-dependent determinants of economic growth – seeking to suggest some potential explanations to the differences in the results across the different spatial scales.

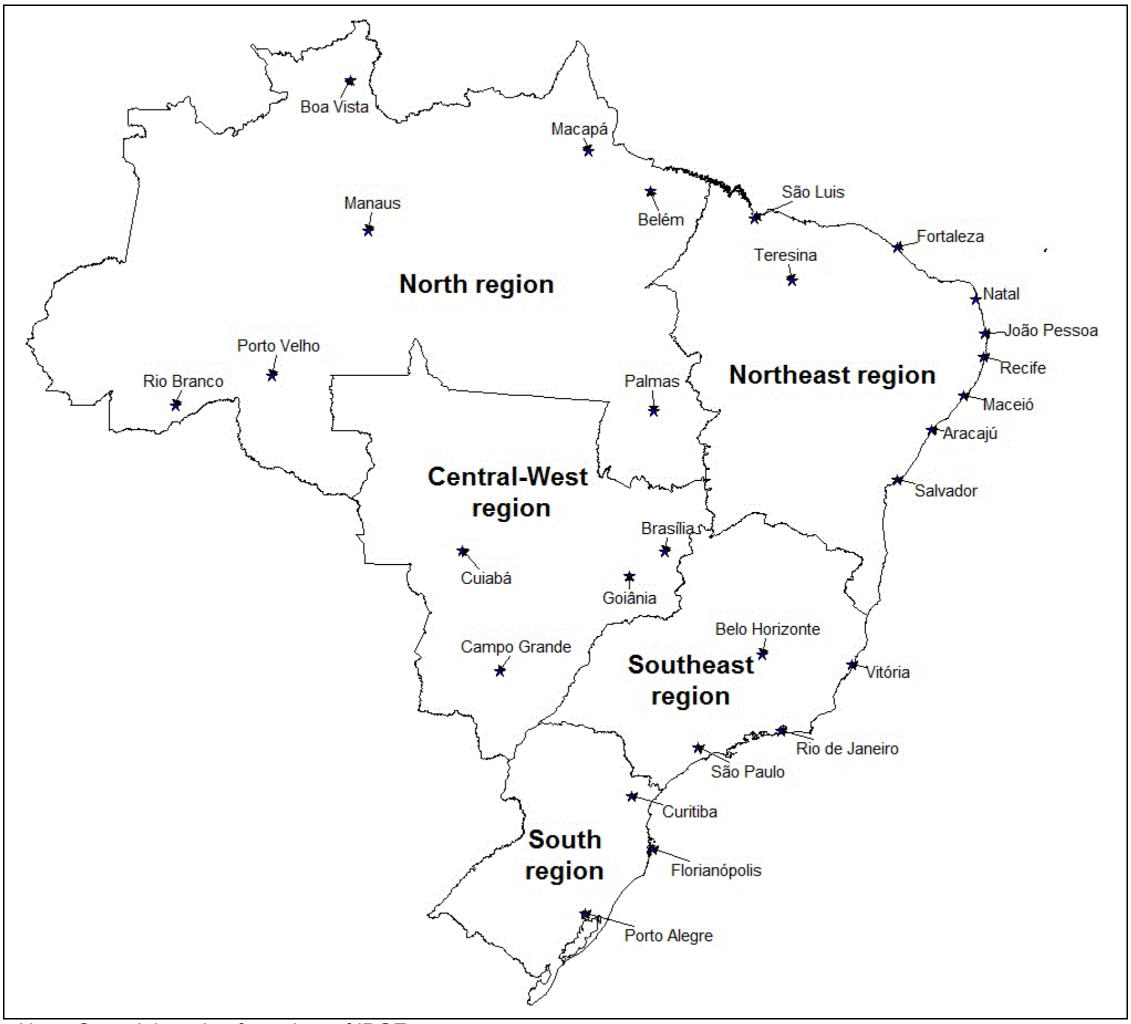
[32] All variables are measured in 1970.

3.3. Empirical Information on the Brazilian Case

This section provides an analysis of some variables of the dataset employed in the thesis. Firstly, the dataset used in Chapter 4 is examined at the four spatial scales, namely, 5,507 municipalities, 559 micro-regions, 91 spatial clusters and 27 states. This analysis covers the 1991-2000 period, where income per capita growth rates between 1991-2000 versus lagged levels of the explanatory variables (in 1991) are regressed. Secondly, the panel dataset at four spatial scales (3,657 minimum comparable areas, 522 micro-regions, 134 meso-regions and 27 states) investigated in Chapter 5 is described over the 1970-2000 period. This data is organised in intervals as close as possible to 10 years using Brazilian Censuses information. In this sense, the panel analysis is in income per capita growth rates versus lagged levels of the explanatory variables.

The aim of the maps below is to describe the spatial distribution of some variables used in the following chapters. Firstly, Figure 3.3 show the geographical distribution of the twenty-seven states capitals and the boundaries of the five Brazilian macro-regions, namely, South (which has three states), Southeast (four states), Central-West (four states), Northeast (nine states) and North (seven states) regions. Note that most of the state capitals are localised in the coastal areas (Macapá, Belém, São Luís, Fortaleza, Natal, João Pessoa, Recifce, Maceió, Aracajú, Salvador, Vitória, Rio de Janeiro, São Paulo, Florianópolis and Porto Alegre). For this reason, we verify that around 25% of the total Brazilian population have been living along the coastal municipalities (in 2010). In Brazil, the total population is around 190 million in 2010, and in 1970 it was 90 million (see www.ibge.gov.br).

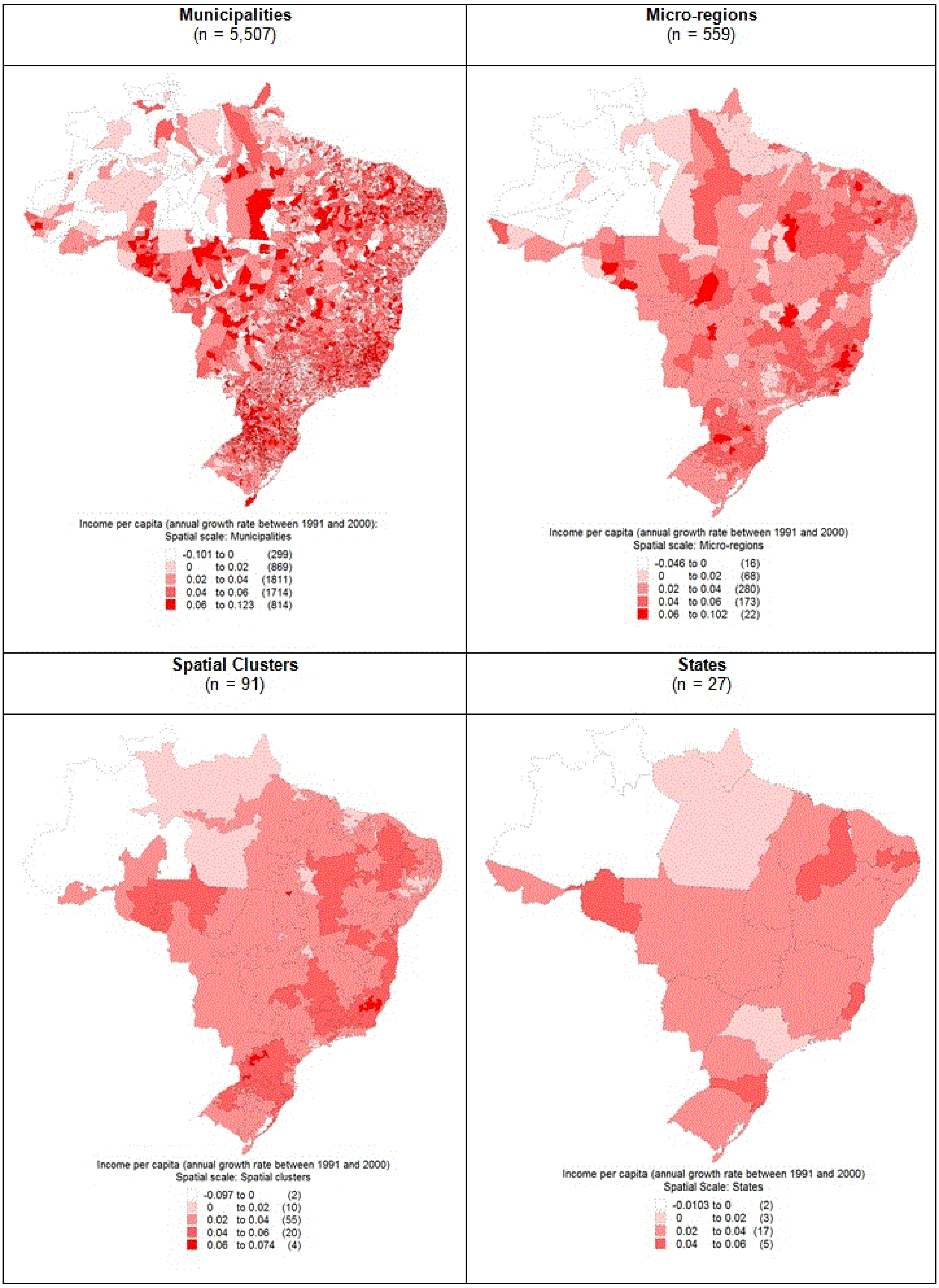
Figure 3.3 – The 5 Brazilian Macro-Regions and the 27 State Capitals



Note: Own elaboration from data of IBGE.

Figure 3.4 shows the average annual growth rates of income per capita between 1991 and 2000 across municipalities, micro-regions, spatial clusters and states in Brazil. In the figures, the intervals (the ranges represented by colours) are kept fixed across the spatial scales for the sake of comparability. Most regions had an average economic growth between 2% and 4% per annum over the 1990's across the four spatial scales under analysis (1811 municipalities, 280 micro-regions, 55 spatial clusters and 17 states are within this interval, 2% and 4%).

Figure 3.4 – Average Annual Growth Rates of Income Per Capita Between 1991 and 2000

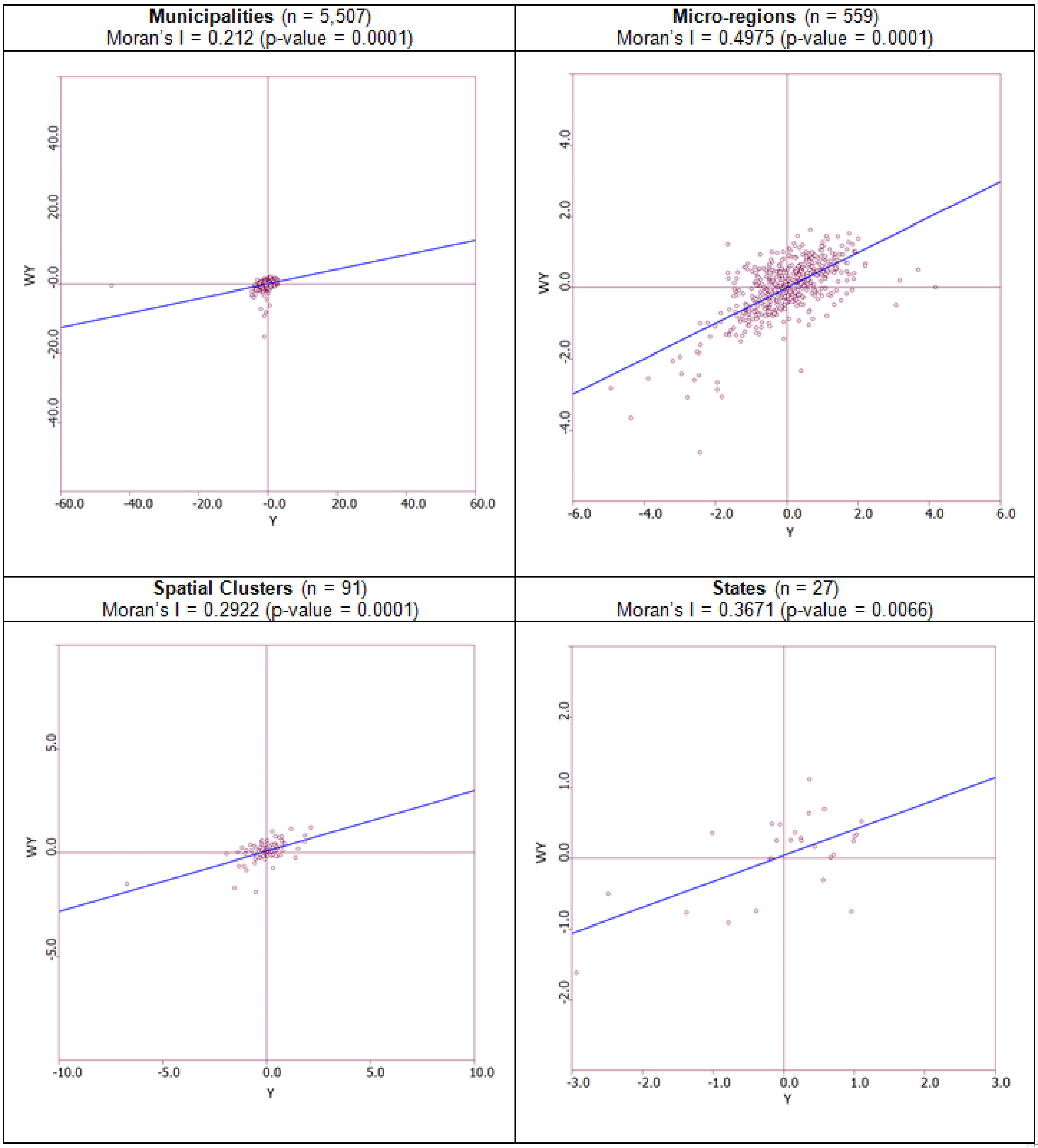


Note: Own elaboration from data of IBGE.

From the visualisation of these maps (Figure 3.4), it is not easy to verify a clear spatial clustering of high (or low) values of economic growth rates. It is possible to suggest a spatial concentration of high growth rates in the South and low growth rates of income per capita in the North region. To shed some light on the spatial dependencies in the growth rates of income per capita between 1991 and 2000 (the dependent variable used in Chapter 4), the global Moran's I statistics is calculated for each spatial scale using the standardised first-order contiguity matrix (also called the queen contiguity matrix, or the queen W matrix). Moran's I is a measure of global spatial autocorrelation (Cliff and Ord, 1981). See Section 5.3 (Chapter 5) for the details on the Moran's I formula.

The global indicator of spatial association (Moran's I) calculated to the growth rates of income per capita, captures the general pattern throughout the four Brazilian spatial scales. Figure 3.5 shows Moran's I statistic and the Moran scatterplot for the average annual growth rates of income per capita between 1991 and 2000. The Moran scatterplot is a useful tool for visualising the spatial interactions summarised by the global statistic. The scatterplot displays the "spatial lag" of the variable under analysis ("WY" is the W matrix times the variable of interest) for each area plotted against the variable (Y) for each area. In sum, the results show that Moran's I statistic is greater than zero at the four scale levels which means there is a positive spatial autocorrelation, i.e., areas with high (low) economic growth rates values tend to be close to areas with high (low) economic growth rates values. The statistical significance of Moran's I is calculated using 9,999 permutations (Anselin, 1995) for each one of the four spatial scales. The four Moran's I statistics (presented in Figure 3.5) are highly significant and means the null hypothesis of no spatial autocorrelation is rejected.

Figure 3.5 – Moran's I Scatterplot (Average Annual Growth Rates of Income Per Capita Between 1991 and 2000)



Note: Own elaboration using GEODA software.

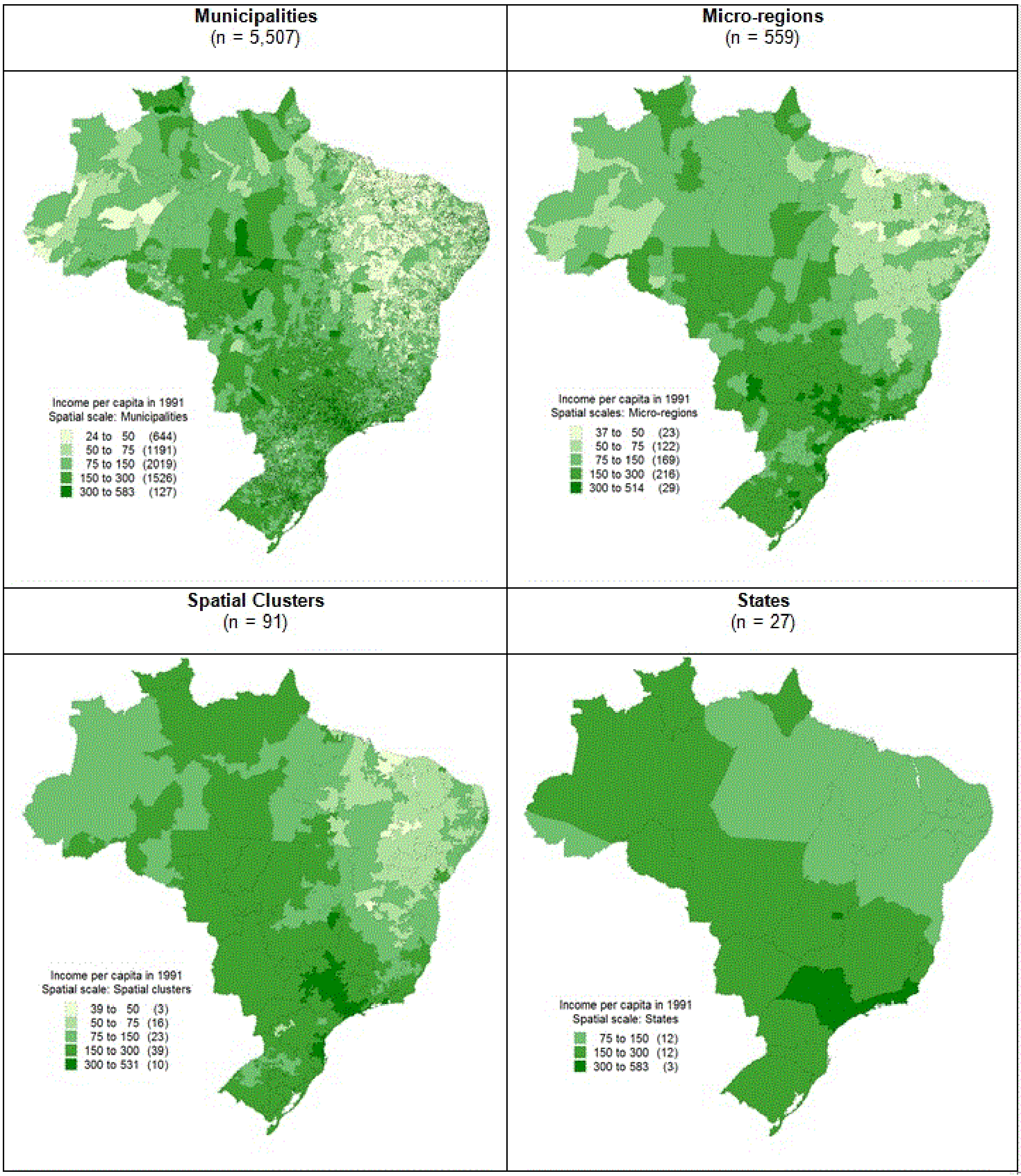
As explained earlier, the explanatory variables of the growth equations are in levels of the initial period. In this sense, some variables for the year 1991 – employed in Chapter 4 – are described next. For details about the source of these variables see Section 4.3.2 (in Chapter 4). Figure 3.6 maps the income per capita in 1991 (the values are in R$, the Brazilian currency) at the municipal, micro-regional, spatial cluster and state levels. The average income per capita in 1991 in Brazil was approximately

R$230. The level of income per capita in 1991 displays a clear spatial pattern of clustering, where high values are localised in the South, Southeast and Central-West regions, and low values are concentrated in the North and Northeast regions (see Figure I.D.1 and Figure I.D.2 in Appendix I). For this reason, the study of Moran's I was not necessary here to investigate the spatial dependencies of levels of income per capita. Moreover, the occurrence of spatial autocorrelation in levels of income per capita is already well documented in the literature about Brazil (Mossi et al. 2003; Silveira-Neto and Azzoni, 2006; Cravo and Resende, 2012).

Figures 3.7 and 3.8, respectively, show population density and average years of schooling in 1991 (across municipalities, micro-regions, spatial clusters and states in Brazil). In 1991, the average population density in Brazil was 17.2 inhabitants per squared kilometre (146,825,475 inhabitants/8,531,514 Km^2). As expected, when the spatial scale becomes finer, the population density appears to be higher in some areas. For instance, while the highest population density at the state level is 292 inhabitants per squared kilometre; there are 69 municipalities that have population density between 1,000 and 12,200. As discussed above, population is concentrated along the coastal areas and, for this reason, the population density is also higher in the coastal regions. This is reflected in all spatial resolutions under investigation. In 1991, the average years of schooling of the Brazilian population was 4.9, and, in 2000, 5.9. Figure 3.8 clearly displays that Northeast region is lagging behind in regards to schooling. For instance at the municipal level, apart from the state capitals of the Northeast region, most of the North-eastern municipalities have an indicator of less than 3 years of schooling.

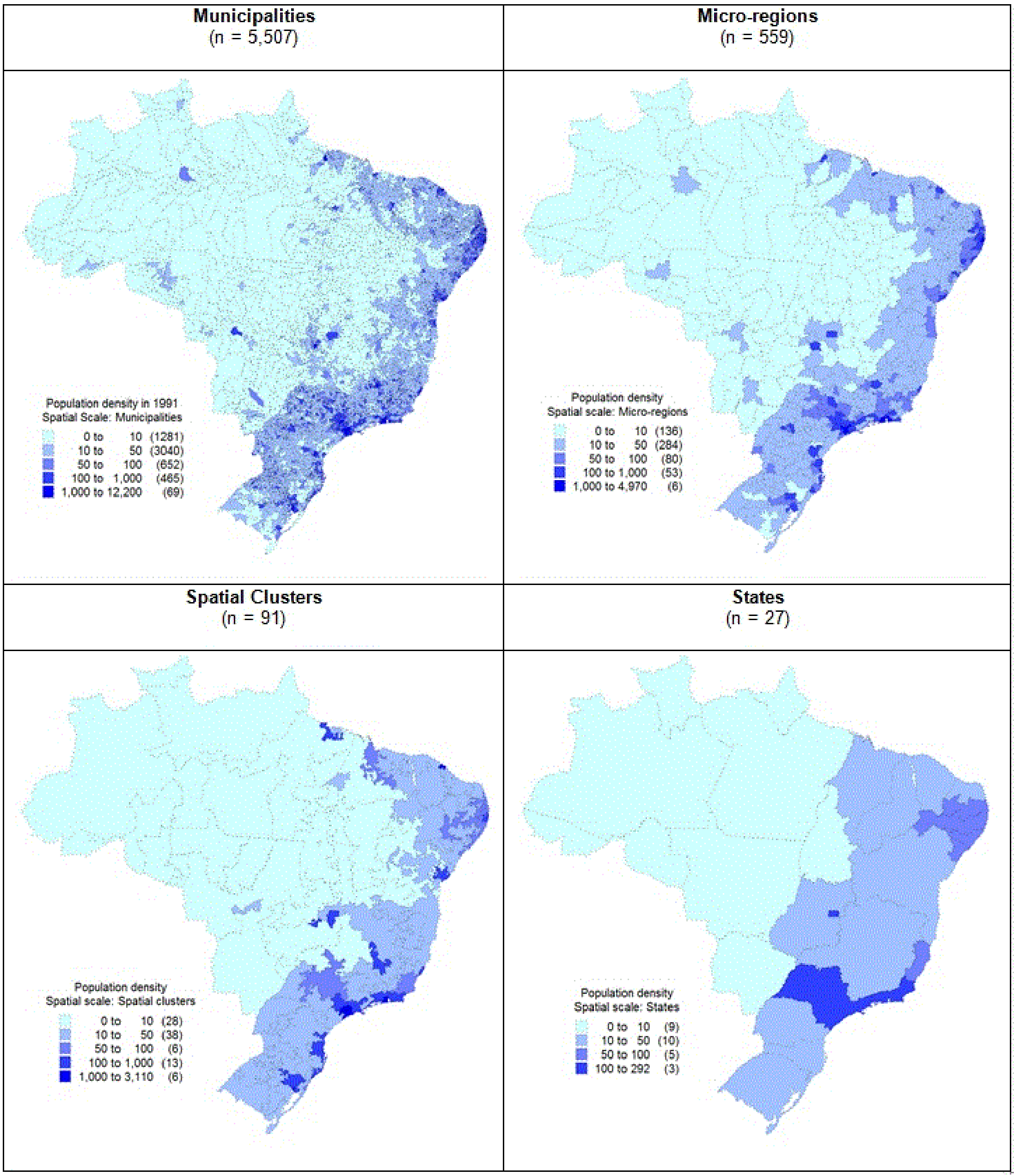
Finally, Figure 3.9 maps the income per capita in 1970, 1980, 1991 and 2000 across the spatial scales employed in Chapter 4, namely, minimum comparable areas (MCAs), micro-regions, meso-regions and states. These maps across time and scales are an interesting way to visualise the dynamics of the income per capita. Per capita income information is deflated to Real (R$) in 2000. Beyond the clear pattern of spatial concentration of income per capita across space in Brazil; it is possible to observe the drop of income per capita between 1980 and 1991. This is explained by a phenomenon known as the Brazilian "lost decade", which was a period of debt crises, hyperinflation and high rates of unemployment (Baer, 2003). The next subsection discusses the rationale for regional economic development policies in the Brazilian context.

Figure 3.6 – Income Per Capita in 1991



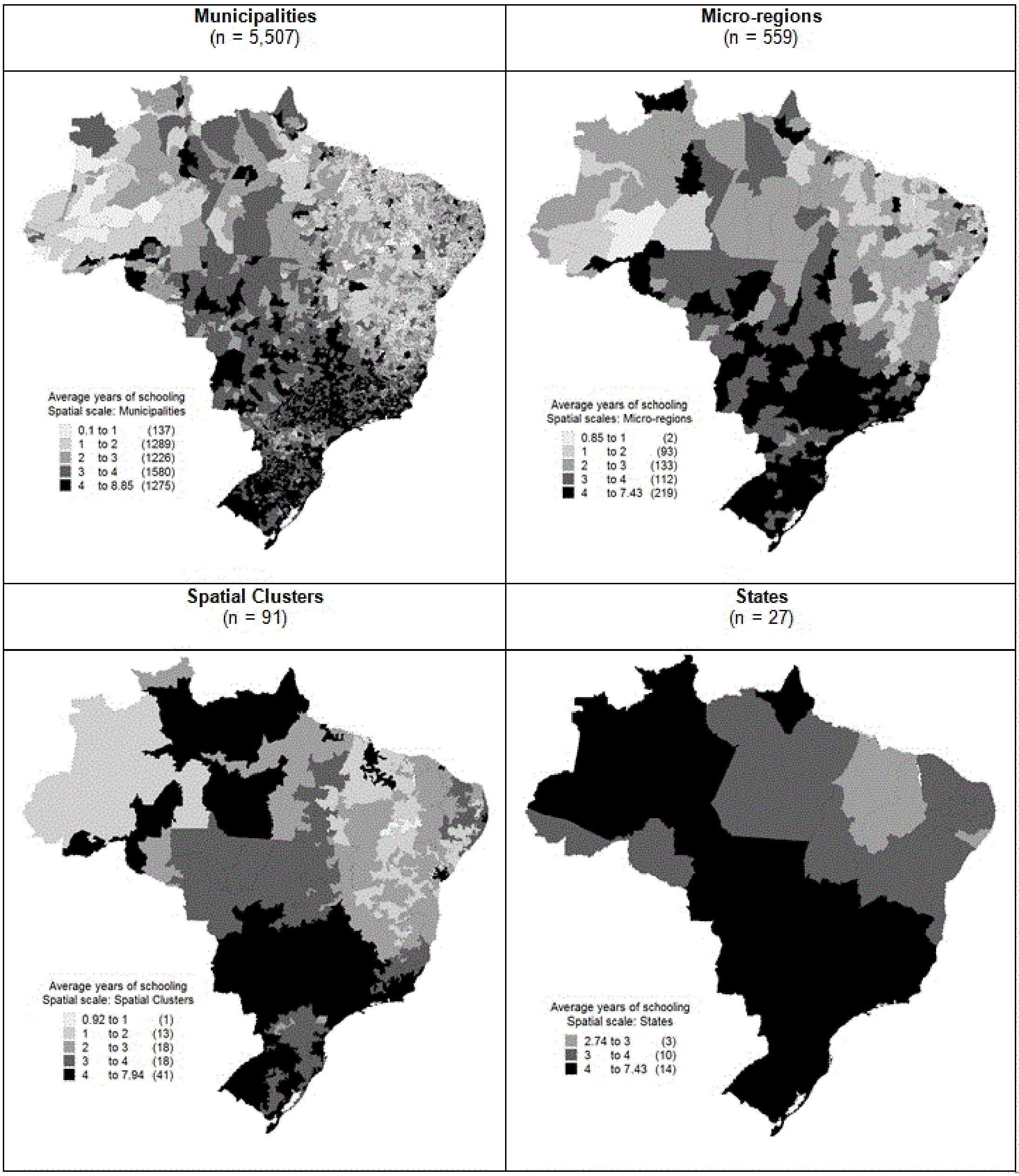
Note: Own elaboration from data of IBGE.

Figure 3.7 – Population Density (Population/Km^2) in 1991



Note: Own elaboration from data of IBGE.

Figure 3.8 – Average Years of Schooling in 1991

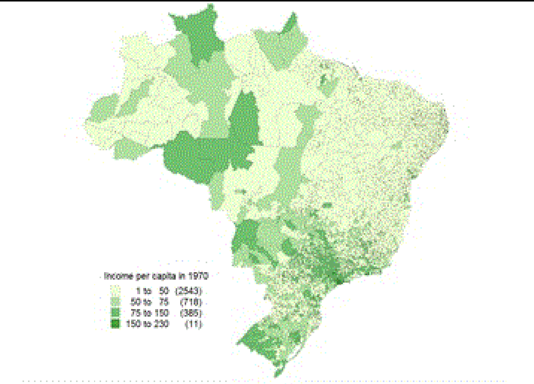


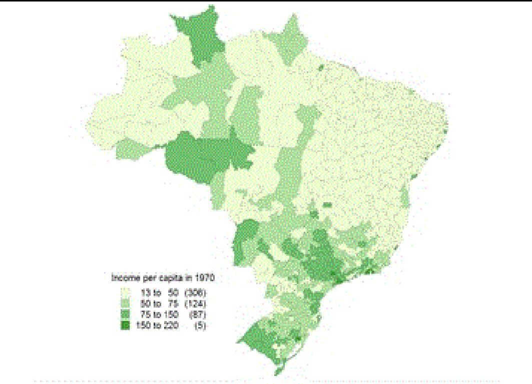
Note: Own elaboration from data of IBGE.

Figure 3.9 – Income Per Capita (in 1970, 1980, 1991 and 2000)

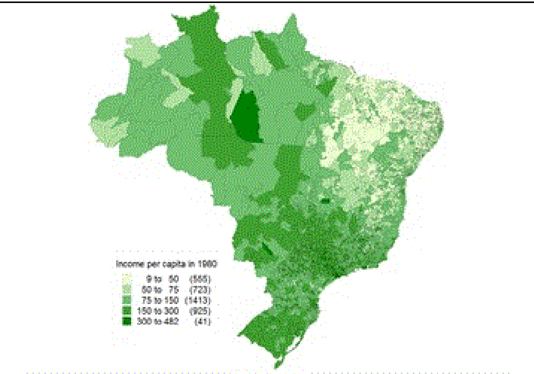
Minimum Comparable areas (MCAs) (n = 3,657) | **Micro-regions** (n = 522)

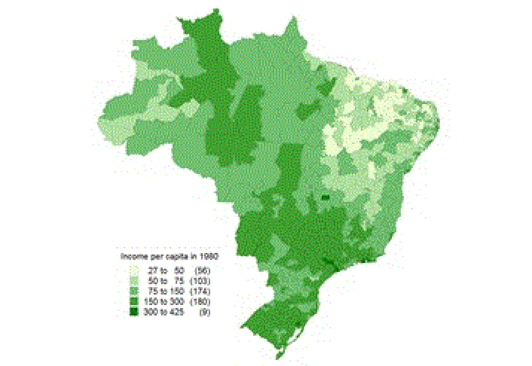
1970



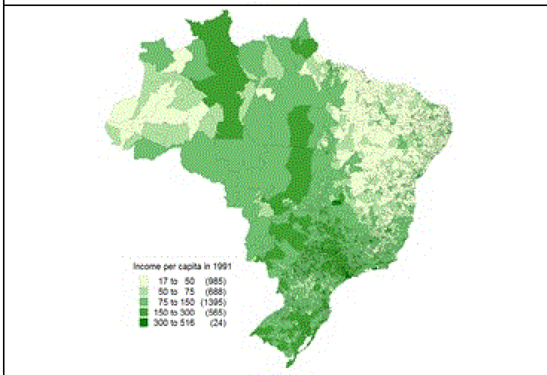


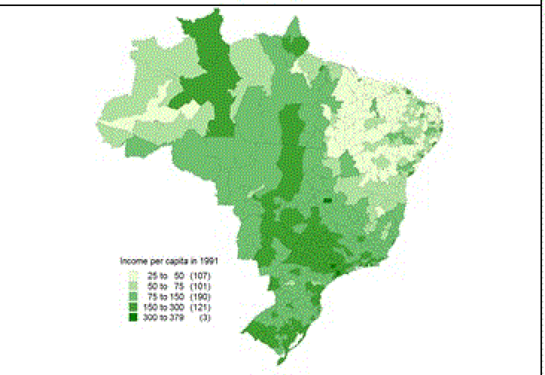
1980



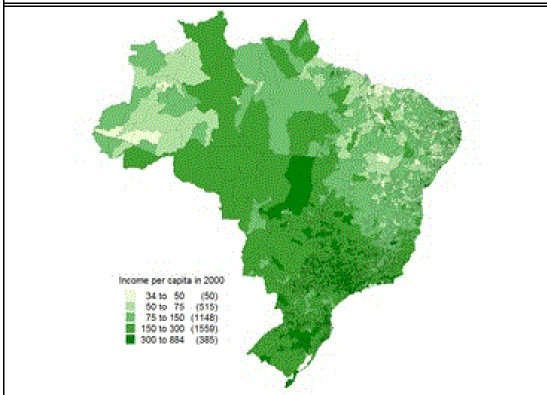


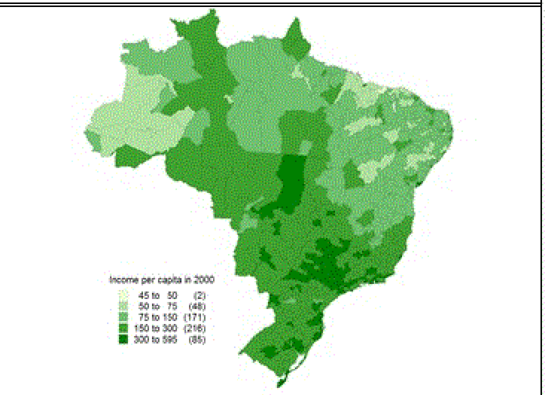
1991





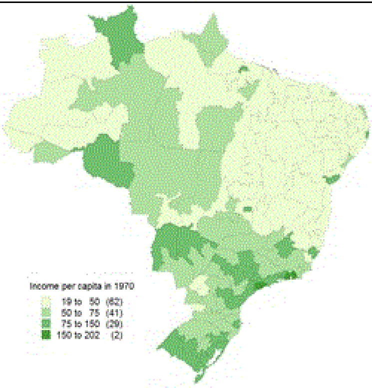
2000

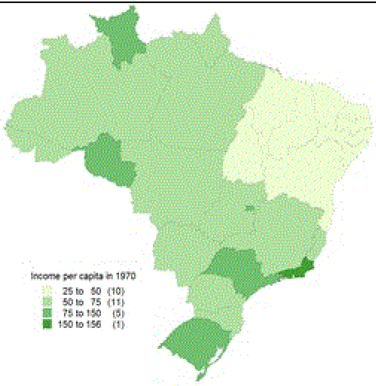


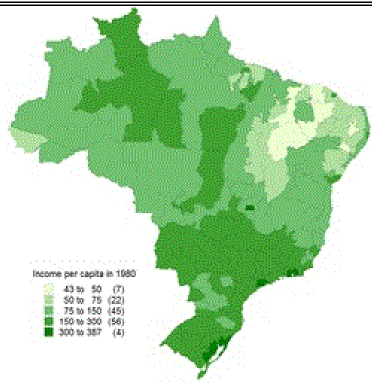


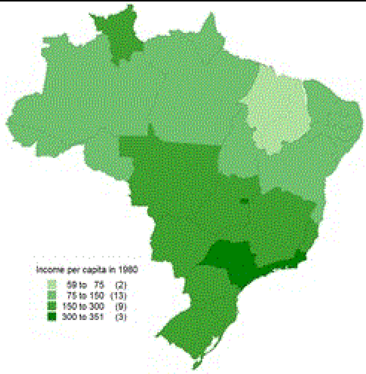
(Continued)

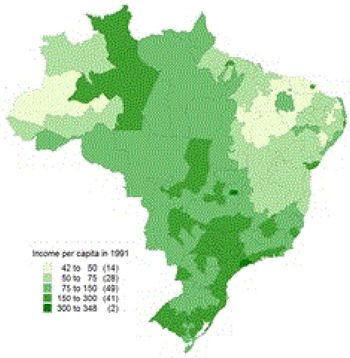
(Continued)

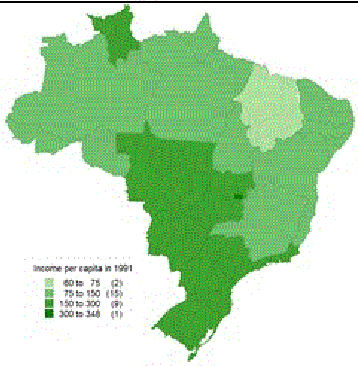
<table>
<tr><th>Meso-regions
(n = 134)</th><th>States
(n = 27)</th></tr>
<tr><td>1970</td><td>1970</td></tr>
<tr><td>


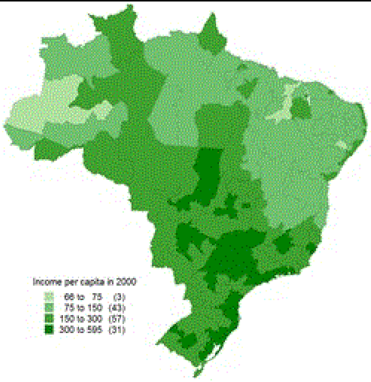
</td><td>


</td></tr>
<tr><td>1980</td><td>1980</td></tr>
<tr><td>


</td><td>


</td></tr>
<tr><td>1991</td><td>1991</td></tr>
<tr><td>


</td><td>


</td></tr>
<tr><td>2000</td><td>2000</td></tr>
<tr><td>


</td><td>


</td></tr>
</table>

Note: Own elaboration from data of IBGE.

3.4. Justifications for Regional Economic Development Policy in Brazil

The main justification for a regional development policy in Brazil dates to the 1950s and is based on the CEPAL's (Economic Commission for Latin America and Caribbean) centre-periphery arguments. As discussed in Ferreira (2004), the work written in 1958 by Celso Furtado in the GTDN[33] transposed the ideas of CEPAL—namely, the terms of trade disadvantage of the countries in Latin America—to the North–South imbalance within Brazil[34]. CEPAL also based their policy recommendations on the existence of market failures. However, the prevailing view was that the market failures should be corrected via relative price distortions—subsidies, for example—which would allocate resources more efficiently. Following Furtado's suggestion, the federal government created, in 1959, the Superintendency for the Development of the Northeast (SUDENE), which was responsible for coordinating all public interventions, such as tax and investment credits, infrastructure investments (mainly in energy and roads), long-term financing and tax incentives for firms in the Northeastern region. In 1974, SUDAM was created for the development of the Amazon region with the same objectives. However, after suspicions of corruption surrounding both organizations (SUDENE and SUDAM), they were both shut down in 2001[35]. Another regional policy created in 1989 is the regional development funds (FNE, FNO, FCO), which aims to promote the economic and social development of the Brazilian lagging macro-regions (Northeast, North, and Central-West) through subsidies to small agricultural and industrial producers seeking to reduce credit constraints. These regional development funds and the latest developments of Brazilian regional policy are discussed in Section 2.5. Recently, some Brazilian economists (for instance, Barros, 2011; Ferreira, 2004; and Pessôa, 2001) have criticized such policies, arguing that regional problems in Brazil are an issue of secondary importance when compared to the inequality among

[33] "Grupo de Trabalho para o Desenvolvimento do Nordeste", ("Working Group for the Development of the Northeast"). See Furtado (1997) for the reprinted document

[34] Castro (1971) and Cano (1976) are other references that also justify Brazilian inequalities between North and South based on the imbalance of exchanges between the two regions.

[35] SUDENE and SUDAM have been recreated in 2007.

households[36]. Pessôa (2001) argues that a subsidy policy to industry is not the best recommendation for solving inequalities that are embodied in the individual (skill level, for example)[37]. In the same way, Ferreira (2004) points out that it has been observed that once you have control of education and other relevant factors, the contribution of the region to inequality is relatively small as shown by Barros and Mendonça (1995) and Menezes-Filho (2001). These authors argue for a change in the focus of regional policy from subsidy of physical capital accumulation to mass investments in human capital (Ferreira, 2004). Recently, Barros (2011) measures the contribution of individual and local (area) factors to the observed income inequality between the Northeast (poor) and Southeast (rich) regions. The study shows that after controlling for differences in quantity (years of schooling) and quality of education and for cost of living, it appears that GDP per capita in the Northeast is the same as observed in the Southeast region.

The discussion that has been posed by Pessôa (2001), Ferreira (2004) and Barros (2011) is similar to the argument provided by Gibbons et al. (2010) on people versus place based policies in the UK context. First, Gibbons et al. (2010) show a picture of pronounced and very persistent disparities across local areas in Britain over the period 1998-2008. Then, they examine "*to what extent these disparities arise because of differences in the types of workers in different areas (sorting) versus different outcomes for the same types of workers in different areas (area effects)";* and conclude that area effects explain less than 1% of overall wage variation (Gibbons et al., 2010: 2)[38]. In this sense, "*who you are is much more important than where you live in determining earnings and other outcomes*" (Overman and Gibbons, 2011: 24). Duranton and Monastiriotis (2002) also suggest sorting as an explanation of spatial disparities in UK over the period 1982-1997. In sum, the studies suggest that disparities are driven by 'people' rather than 'place' (Gibbons et al., 2010).

[36] In 2007, personal income inequality, measured by the Gini index, was 0.53 in Brazil, one of the highest indexes in the world.
[37] See Pessôa (2001) for the discussion of regional problem vs. social problem.
[38] One caveat of this analysis is because it does not control for differences in costs of living and in access to amenities across places, thus, it focuses on nominal rather than real wages. This issue is relevant and is taken into account in Gibbons et al. (2011).

Recently, Barca et al. (2012) examine the rethinking of regional development policy intervention that has emerged, namely, the space-neutral versus the place-based approaches. These authors discuss the rethinking which has taken place by exploring a series of highly influential reports on the topic produced by the World Bank (2009), the European Commission (Barca, 2009), the OCDE (2009a, 2009b), and the Corporación Andina de Fomento (CAF, 2010) and an earlier report by Sapir et al. (2004). Barca et al. (2012) advocates in favour of place-based policies in contrast to the 2004 Sapir Report and the World Bank's (2009) World Development Report '*Reshaping Economic Geography*' saying that:

> "*[t]he place-based approach therefore argues that there are alternative pathways to development, which require attention to detail and the institutional context. Mega-urban growth at the top of the urban hierarchy, as advocated by the World Bank (2009), is just one such development option, an option which brings its own challenges with it, and an option which so far has not been demonstrated to be an optimal solution (Henderson, 2010). The World Bank (2009) has effectively given up on institutional reform as an essential part of the development process and substituted it with mega-urban growth, thereby foregoing all of the alternative pathways. In contrast, by acknowledging the limits of the central state to design good local development policies, place-based strategies recognize the need for intervention based on partnerships between different levels of governance, both as a means of institution-building and also of identifying and building on local knowledge (Pike et al., 2007)*" (Barca et al., 2012: 147).

3.5. Evaluation of Regional Economic Development Policy in Brazil

Evaluations aim to answer questions such as when and how interventions or treatments 'work' and seeks to inform decisions about improvements, expansions or modifications that can be made in a specific policy or program (Bartik and Bingham, 1995). This subsection discusses some issues related to evaluation process of regional economic development policies and describes the main regional policy in Brazil as well as the evaluation literature on this policy. In Brazil, the primary regional economic development policy has been in place since 1989. This policy seeks to facilitate the economic and social development of lagging macro-regions by offering loans below market interest rates, primarily, to small-scale farmers and small industrial firms. Such development is directed by the Constitutional Financing Funds for the Northeast (FNE), the North (FNO), and the Central-West (FCO) (henceforth referred to as the regional development funds or simply, regional funds). However, there have been very few evaluations of how these regional development funds are being used. A review of the literature carried out by the author reveals that regional development funds in Brazil are, indeed, rarely evaluated because during the period of 2000 to 2009, there are only two papers (out of 4,619) concerning Brazilian regional development funds evaluation that were published in the selected journals [namely, Silva et al. (2009) and. Soares et al. (2009)].[39] The investigation of the possible reasons for the scarcity of studies on regional development funds evaluation in Brazil is beyond the scope of this thesis[40]. In the next section, regional development policy process, its objectives and the types of evaluation are discussed. The strategy of the Brazilian regional development funds since 1989, is also reviewed.

[39] Amongst the Brazilian journals and leading regional science journals there were only two papers on this issue, and by comparison 20 papers on the EU in the same sample. The search was limited to a selected sample of top journals (the Brazilian journals are Economia e Sociedade, Estudos Econômicos, Pesquisa e Planejamento Econômico, Revista Brasileira de Economia, Revista de Econometria, Revista de Economia e Sociologia Rural, Revista de Economia Política and the top international regional science journals are Annals of Regional Science, International Regional Science Review, Journal of Regional Science, Papers in Regional Science, Regional Science and Urban Economics, Journal of Economic Geography and the Regional Studies journal). The papers on regional policy evaluation in EU countries are the following: Andersson (2005), Armstrong et al. (2001), Bradley (2006), Dall'erba and Le Gallo (2008), Dall'erba (2005), De la Fuente (2004), Esposti and Bussoletti (2008), Florio (2006), Frenkel et al. (2003), Greenbaum and Bondonio (2004), Harris andTrainor (2005), Lambrinidis et al. (2005), Leonardi (2006), Martin and Tyler (2006), Pereira and Andraz (2006), Pérez et al. (2009), Rodrguez-Pose and Fratesi (2004), Romero and Noble (2008), Romero (2009) and Skuras et al. (2006)]. The only problem with this approach would be if there were more papers on Brazil than the EU in the literature I did not review, which seems unlikely.

[40] It is worth noting that some authors, such as Bartik and Bingham (1995), have already tried to enumerate some reasons for the absence of more sophisticated evaluations of economic development programs (the focus of the work was the USA). In sum, they list six reasons: (i) evaluations with a comparable group require careful procedures to select the comparison group, including collection of extensive quantitative data over a period of time from both the firms participating in the economic development evaluation, and the comparison group; (ii) these data collection and design efforts may be expensive and time consuming; (iii) more rigorous evaluations will have a disproportionate part of their benefits going to groups other than those paying for the evaluation; (iv) administrators prefer a process evaluation as it would offer some clues as to how to improve the program, even if the evaluation by itself does not document what the program really accomplished; (v) state audit agencies frequently do not have staff who are trained in how to do studies that correct for selection bias due to a non-randomly selected comparison group; (vi) program administrators fear the political consequences of a negative evaluation. Hence, they avoid evaluations because with no evaluations, one can always claim success.

3.5.1. Policy Process: From Objectives to Evaluation

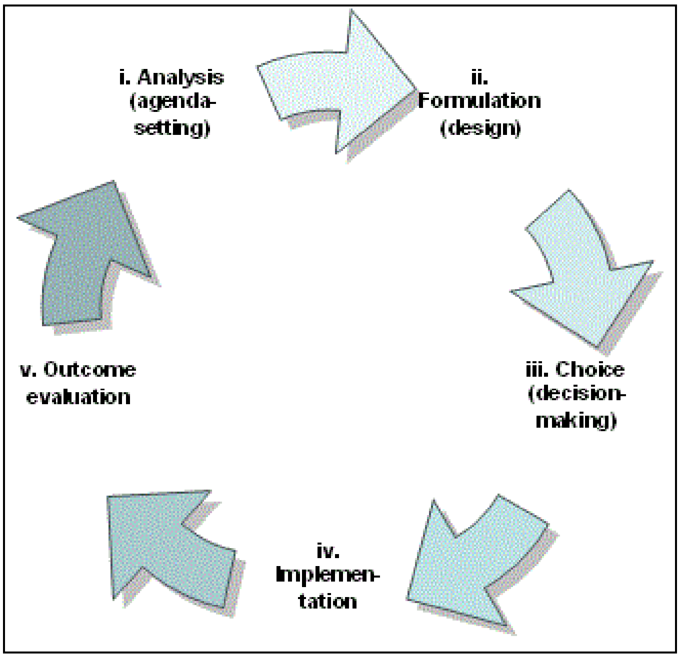
In general, how are policy objectives defined and evaluations carried out? Regional development policy follows the general public policy cycle, which is usually divided into five stages: (i) analysis (agenda setting), (ii) formulation (design), (iii) choice (decision making), (iv) implementation, and (v) outcome evaluation. In other words, first, "problems are defined and put on the agenda; next policies are developed, adopted and implemented; finally, these policies will be assessed against their effectiveness and efficiency and either terminated or restarted" (Jann and Wegrich, 2007: 44). It is worth noting that outcome evaluation is associated with the final stage in the cyclical model of policy process, but it is also closely related to the initial stages because the results given by the outcome evaluation will serve as input for the initial phases. Furthermore, evaluation studies form a separate sub-discipline as outcome evaluation is only one type of different perspective for evaluating research in terms of time (e.g., ex ante, ex post) and complexity (e.g., monitoring daily tasks or assessing impact on the problem)[41].

It is worth noting that the stages perspective has created different lines of research that have focused on particular stages—which follow a distinct set of questions, analytical perspectives and methods—rather than on the whole cycle (Jann and Wegrich, 2007). Also, policy process does not follow this sequence of discrete stages; instead, the stages are constantly connected and entangled in an ongoing process. Despite the limitations of modelling the policy process in terms of stages, first introduced by Lasswell (1956), I employ this approach as an ideal type of rational planning to organize and systemize the discussion around policy evaluation. With the limitations in mind, the following paragraphs briefly sketch the five stages of the cycle framework (see Diagram 3.1) and highlight the main issues related to the Brazilian regional development policy.

[41] The types of evaluation are discussed in the next subsection.

Diagram 3.1

'Cycle model' of the Policy Process



Note: Own elaboration based on Jann and Wegrich (2007).

i. Analysis (agenda setting): The first stage of the policy process is the recognition and analysis of a policy problem that requires state intervention. Then, the recognized problem goes to the agenda for analysis (agenda setting). In this phase, as indicated by Birkland (2007), groups have to fight to earn their issues' places among all of the other issues sharing the limited space on the agenda, and at the same time, they need to fight to keep other issues off the agenda, blocking action of competing issues. In Brazil, the regional inequalities were recognized as a problem in the 1950s, and since then, governments have undertaken some policies to deal with those inequalities. In the new Constitution of 1988, the regional inequalities remained a problem and new instruments (e.g., the regional development funds) were defined to fight against these inequalities.

ii. Formulation (design): This stage includes the definition of policy objectives—what should be achieved with the policy—and the consideration of different action alternatives in preparation for the final policy decision (Jann and Wegrich, 2007). In sum, this stage aims at formulating the set of alternatives that include "*identifying a* *range of broad approaches to a problem, and then identifying and designing the specific sets of policy tools that constitute each approach*" (Sidney, 2007: 79). As discussed above, the justification for regional policy in Brazil was influenced by the theories of CEPAL, which argue that the market failures should be corrected via subsidies/incentives to industry and agriculture in the lagging regions. One of the stated objectives of the Brazilian Constitution of 1988 was to reduce inequalities across Brazilian regions[42] using subsidies to the agricultural and industrial sectors in the lagging regions as the main policy tool.

iii. Choice (decision making): It is not easy to define a clear-cut separation between formulation and decision making. Indeed, this distinction is often impossible in practice. Roughly, choice or decision making can be defined as the final adoption of a specific public policy, i.e., the formal decision to take on the policy (Jann and Wegrich, 2007). In 1989, federal law nº 7827 created the regional development funds for the Northeast (FNE), the Central-West (FCO) and the North (FNO) with the objective of reducing regional inequalities by financing the productive sectors in those macro-regions. It is worth noting that because this is not a precisely defined objective, it will negatively affect the outcome evaluation process, as it will be difficult to measure policy effectiveness.

iv. Implementation: In this stage, policy will be executed by the responsible institutions and organizations. The program details (e.g., definition of agencies, laws) are specified as well as the allocation of resources (e.g., budgets, human resources). Pülzl and Treib (2007) discuss the implementation stage of the policy process, comparing top-down, bottom-up and hybrid approaches[43]. Concerning regional policy in Brazil, law nº 7827

---

[42] Art. 3rd. The fundamental objectives of the Federative Republic of Brazil are the following:
III – (…) to reduce the regional and social inequalities. [This extract from the Brazilian Federal Constitution of 1988 (Brasil, 2008), was translated by the author.]

[43] Pülzl and Treib (2007: 90) describe the three approaches as "*(a) top-down models put their main emphasis on the ability of decision makers' to produce unequivocal policy objectives and on controlling the implementation stage; (b) bottom-up critiques view local bureaucrats as the main actors in policy delivery and conceive of implementation as negotiation processes within networks of implementers; (c) hybrid theories try to overcome the divide between the other two approaches by incorporating elements of top-down, bottom-up and other theoretical models*".

(1989) defines the source of funding and designates the regional banks as being the operators of the regional development funds. Essentially, this kind of policy can be defined as a top-down approach.

v. Outcome evaluation: Evaluation research can be applied to the whole policy-making process and exists in various forms. The next subsection will discuss the various forms of evaluation research. Outcome evaluation includes assessing effectiveness, conducting a cost-benefit analysis and verifying whether the policy solved or at least reduced the problem. Depending on the results of the outcome evaluation, the policy will be redesigned, modified or terminated. Furthermore, Jann and Wegrich (2007: 54) point out that the activities of the evaluation are exposed to the logic and the incentives of the political process in at least two major ways: "*First, the assessment of policy outputs and outcomes is biased according to the position and substantial interest, as well as the values, of a particular actor. In particular, the shifting of blame for poor performance is a regular part of politics. Second, flawed definition of policy aims and objectives presents a major obstacle for evaluations. Given the strong incentive of blame-avoidance, governments are encouraged to avoid the precise definition of goals because otherwise politicians would risk taking the blame for obvious failure*".

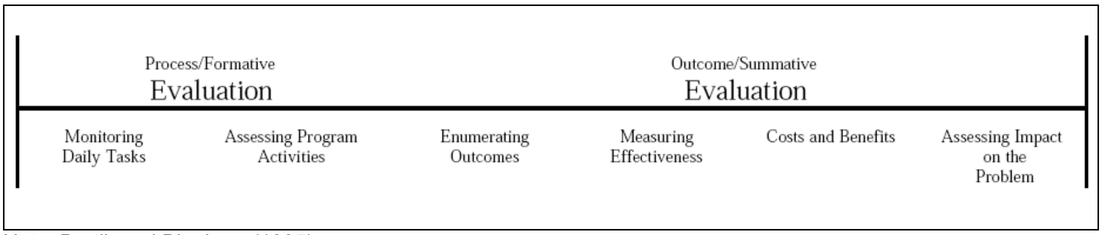
Regarding the Brazilian regional policy, it appears that the issue of blame-avoidance is one of the possible reasons for the infrequent evaluations of regional development funds over the years. Indeed, if there is no evaluation, how can governments be blamed for failures? In addition, even if evaluations are conducted, governments avoid the blame because the objectives of the Brazilian regional development funds are not precisely defined.

3.5.2. Types of Evaluation

As noted earlier, evaluation can be defined in several ways - in terms of time (e.g., ex ante, ex post), levels of complexity (e.g., monitoring daily tasks or assessing impact on the problem) or as an internal or external evaluation. Different from Brazil, the European Union, since the reform of the Structural Funds in 1988, has created a system of appraising, monitoring and evaluating all EU-funded regional development interventions. Bachtler and Wren (2006: 143) explain that the evaluation of Structural and Cohesion Funds programmes has to be conducted at defined points in the programming cycle: "*ex-ante to verify targets; at the mid-point to establish the need for corrective action; and ex-post to assess outcomes*". Although this can be a useful definition of types of evaluation, I prefer to discuss the types of evaluation by levels of complexity as the quality and the objectives of evaluation studies might be relatively uneven and diverse. Therefore, I follow the definition of Bartik and Bingham (1995) who look at evaluation as a continuum moving from the simplest form of evaluation, monitoring daily tasks, to the more complex, assessing the impact on the problem, as illustrated in Diagram 3.2.

Diagram 3.2

Types of Evaluation by Levels of Complexity



Note: Bartik and Bingham (1995).

Evaluation is divided into six levels ending with a judgment if the policy (or a specific program) works, i.e., solved the problem or at least reduced it. Bartik and Bingham (1995) point out that there is a tendency for governments to prefer process evaluation (monitoring daily tasks, assessing program activities and enumerating outcomes) as this lower level of evaluation only provides information about how to improve a program, rather than assess if the program is actually successful (e.g., creates jobs), which is the role of the outcome evaluation. Table 3.1 summarises the function of each type of evaluation by means of enumerating several questions that each type of evaluation has to answer[44].

Table 3.1

Function of each Type of Evaluation

Type of evaluation	Question that each type of evaluation has to answer
(i) Monitoring daily tasks	"*Are contractual obligations being met? Are staff members working where and when they should? Is the program administratively sound? Are daily tasks carried out efficiently? Are staff adequately trained for their jobs?*"
(ii) Assessing program activities	"*What activities are taking place? Who is the target of activity (businesses, cities, etc.), and with what problems or needs? How well is the program implemented?*"
(iii) Enumerating outcomes	"*What is the result of the activities described in the process evaluation? What happened to the target population? How is it different from before? Have unanticipated outcomes occurred and are they desirable? Have program objectives been achieved? How are the program recipients different from the way they were before?*"
(iv) Measuring program effectiveness	"*What would have happened in the absence of the program? Does the program work? What are the other factors that may have contributed to changes in the recipients? To answer these questions a cause and effect relationship must be established between the program and the outcome. Did the tax abatement 'cause' an increase in employment in the target company?*"
(v) Costs and Benefits	"*Do costs of the program outweigh the benefits of the program?*"
(vi) Assessing the impact on the problem	"*What changes are evident in the problem? Has the problem been reduced as a result of the program? What new knowledge has been generated for society about the problem or the ways to solve it?*"

Note: Bartik and Bingham (1995: 2-3).

As pointed out by Bartik and Bingham (1995), these six levels of evaluation provide a framework for assessing the quality of evaluations. To demonstrate that a program (or policy) accomplishes its targets, the evaluation must be at the highest levels: measuring effectiveness (e.g., it actually does create jobs) or assessing impact (e.g., there has been an improvement in the problem situation). Furthermore, simply because a program has been shown to be both substantively effective and has solved the problem, that does not mean that it should have ever been implemented. A cost-benefit analysis needs to be carried out to show that the program benefits outweigh its costs. Regarding the Brazilian regional policy, evaluations could suggest, for instance,

[44] These questions were extracted from Bartik and Bingham (1995: 2-3).

that the regional development funds create jobs and ultimately reduce regional inequalities. However, it is still necessary to demonstrate that the program is cost effective.

3.5.3. Brazilian Regional Development Funds (FNE, FNO, FCO)

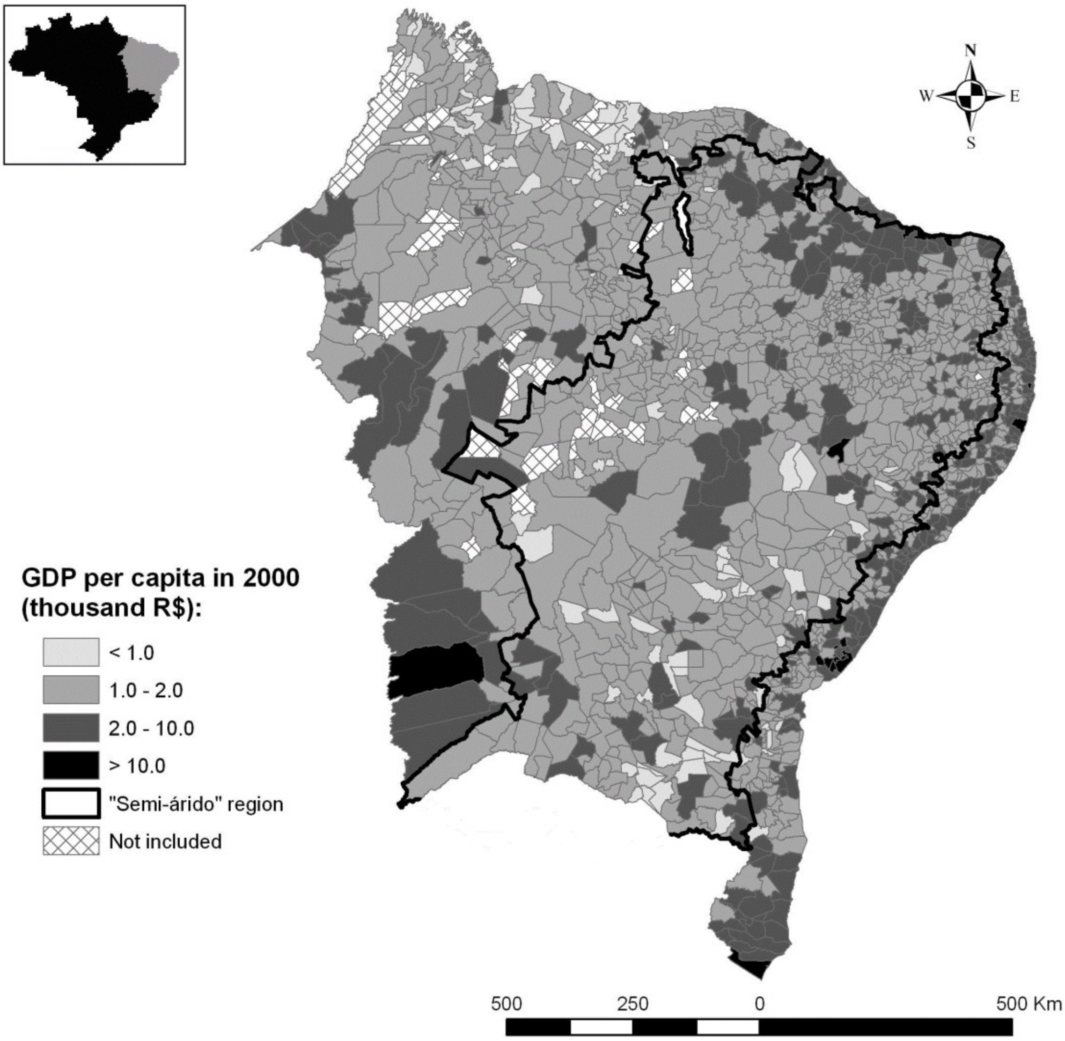
The regional development funds (FNE, FNO, and FCO) were created by federal law nº 7827 in 1989, based on article 159.I.c of the Federal Constitution of 1988. An equal portion (3%) of income taxes (from individuals and firms—"IR") and of the tax on industrialised goods ("IPI") represents the transfer of resources from the National Treasury to the regional development funds. It is important to note that the goal of the FNE, FNO and FCO defined by the federal law is to reduce regional inequalities through the financing of productive sectors in those macro-regions. As noted earlier, this imprecisely defined objective (or broad objective) is the major obstacle to outcome evaluations.

The total resources allocated to these funds each year is divided as follows: 60% goes to the FNE; 20%, to the FNO; and 20%, to the FCO. These resources are transferred from the National Treasury to the operating bank via the Ministry for National Integration ("Ministério da Integração Nacional"). Beyond the 3% IR and IPI taxes, the revenues for these funds come from the repayment of the loans (principal + interest). In this way, law nº 7827 (1989) defines the source of funding and designates the regional banks as being the operators of the regional development funds. The operator bank of FNE is the Bank of the Northeast (Banco do Nortedeste/BNB), and for the FNO, it is the Bank of Amazon (Banco da Amazônia/BASA), both of which were founded in the 1950s with the aim of fomenting and developing these lagging regions. The Central-West region does not have a regional bank, and the operator bank of FCO is the Bank of Brazil (Banco do Brasil/BB, a Brazilian federal bank).

Specifically, the operator banks of the regional development funds are the agents responsible for analysing and deciding whether to award the subsidised loans to applicants. The interest rates of the loans are fixed but vary depending on the size of the beneficiary and the sector. In rural FNE operations, the interest rates are between 6.00% and 10.75% per annum, and for the other operations, they are between 8.75% and 14.00% per annum (small businesses have the lower rates). It is worth noting that the average interest rates for the production sector in other banks were around 35% in 2000 (BANCO CENTRAL, 2000). Furthermore, good payers win compliance bonuses in the form of an interest rate reduction of approximately 15%. Applicants can be individuals, small businesses, enterprises or cooperatives/associations that want to finance a new business or an existing one located in the Northeast, North or Central-West region. There are some general guidelines that the banks follow when analysing applications: preference is given to (i) productive activities of individual and small farmers and (ii) small firms in other sectors, (iii) activities that intensively use raw materials and are labour-intensive and produce basic food for the population, and (iv) new centres, activities or clusters that can reduce the economic and social differences between regions. Moreover, by law, 50% of the FNE loans must be directed toward the "semi-árido" region. Figure 3.10 shows the boundaries of the "semi-árido" region and the GDP per capita in 2000 at the municipal level in the Northeast region (the FNE is the focus of an impact evaluation conducted in Chapter 6). Since the creation of the regional development funds, roughly 70% of the resources have been directed to the agricultural sector, 15% to the industrial sector and the remaining 15% has been devoted to commerce and service sectors.

Figure 3.10

Municipal GDP per capita in 2000 in the Northeast Region



Note: Own elaboration based on IBGE data.

Between 2000 and 2006, the regional development funds invested €10 (R$[45] 28) billion in lagging macro-regions (Northeast, North and Central-West) in Brazil. This fact represented 1.2% of the national GDP in 2006. It is interesting to note that, between 2000 and 2006, the European Union (EU 15 countries), which has been a paradigm of regional policy for the Brazilian governments[46], allocated €135 billion to regions with less than 75% of the average EU15 GDP per capita. Coincidentally, this expenditure also represented 1.2% of EU15 GDP in 2006. When comparing these

[45] Real (R$) is the Brazilian currency.

[46] For example, see the document on the European Union-Brazil dialogue on regional policy: <http://ec.europa.eu/regional_policy/international/pdf/eu_br_regint_en.pdf>.

numbers, it can be concluded that the Brazilian government has invested a significant amount of money in regional development policy.

Ferreira (2004) and Almeida Junior et al. (2007) conducted comprehensive studies of the resource allocation each year for these funds (FNE, FNO and FCO). Among other analyses, these authors show that the rate of non-performing FNE loans reached 31% in 2001. As pointed out by Ferreira (2004), before 2001 most bad credits were considered “under renegotiation” while, in fact, they were never paid back. This high default rate limited the Bank of Northeast from granting new loans during the 1998–2002 period. In 2002, a federal bailout plan capitalised the Bank of Northeast, and because of this, in the following years, it could increase the loans granted. Concerning FNO, the credit quality was also not good, reaching 13.2% default rate in 2002. On the other hand, FCO presents the lowest default rate at approximately 3% in 2002.

These regional development funds are not the only resources available from a public bank for lagging regions in Brazil. The Brazilian Development Bank (BNDES), a federal public bank established in 1952, also offers loans (interest rates are below market rates but are higher than those of the regional development funds) to companies of any size and sector in all Brazilian regions. While the focus of the regional development funds is the producers in the agricultural sector (60% of total loans), BNDES loans are directed toward large-scale industrial and infrastructure projects (75% of the total loans). However, unlike the operator banks of the regional development funds that work only in the lagging macro-regions, BNDES addresses the demand for funding in all Brazilian regions and does not have an explicit mandate regarding regional policy. Table 3.2 compares the regional development funds (FNE, FNO and FCO) loans and the BNDES loans by region for the period 2000 through 2007.

Table 3.2 shows that between 2000 and 2007, the average ratio between BNDES loans to the Northeast region (R$ 29.7 billion) and FNE loans (R$ 18.3 billion) was 1.6. Concerning FNO and FCO, the average ratios were 1.8 and 2.8, respectively. BNDES allocated R$ 69.8 billion in Northeast, North and Central-West regions between 2000 and 2007, which represents 22% of its total loans (R$ 322 billion) and twice the amount allocated by the regional development funds (FNE, FNO and FCO). The BNDES loans to the Southeast region (R$ 189.6 billion) represent almost 60% of the total BNDES loans during the period. This evidence suggests that BNDES loans respond to the demand for funding in the most dynamic regions (e.g., Southeast region).

Table 3.2

Regional Development Fund (FNE, FNO, FCO) and BNDES Loans by Region (2000–2007)

Region	Source of loans	2000	2001	2002	2003	2004	2005	2006	2007	Total
Northeast	BNDES	2,783	3,334	3,784	3,112	2,737	3,803	4,836	5,322	**29,712**
	FNE	569	302	254	1,019	3,209	4,174	4,588	4,247	**18,362**
	BNDES/FNE	4.9	11.0	14.9	3.1	0.9	0.9	1.1	1.3	**1.6**
North	BNDES	930	860	1,881	712	1,954	1,616	1,626	3,461	**13,039**
	FNO	697	454	605	1,075	1,321	976	986	1,110	**7,224**
	BNDES/FNO	1.3	1.9	3.1	0.7	1.5	1.7	1.6	3.1	**1.8**
Centre-West	BNDES	2,064	1,703	2,589	2,831	5,161	3,271	3,659	5,755	**27,032**
	FCO	292	979	1,439	920	1,172	1,468	1,444	1,974	**9,688**
	BNDES/FCO	7.1	1.7	1.8	3.1	4.4	2.2	2.5	2.9	**2.8**
Southeast	BNDES	13,008	14,494	23,074	20,036	21,299	28,740	31,415	37,581	**189,646**
South	BNDES	4,261	4,826	6,092	6,842	8,683	9,551	9,783	12,773	**62,809**
Total	BNDES all regions	**23,046**	**25,217**	**37,419**	**33,534**	**39,834**	**46,980**	**51,318**	**64,892**	**322,239**
	BNDES (Northeast) + (North) + (Central-West) regions (A)	**5,777**	**5,897**	**8,254**	**6,656**	**9,852**	**8,689**	**10,121**	**14,538**	**69,784**
	FNE+FNO+FCO (B)	**1,558**	**1,735**	**2,298**	**3,014**	**5,702**	**6,618**	**7,018**	**7,331**	**35,274**
	(A) / (B)	**3.7**	**3.4**	**3.6**	**2.2**	**1.7**	**1.3**	**1.4**	**2.0**	**2.0**

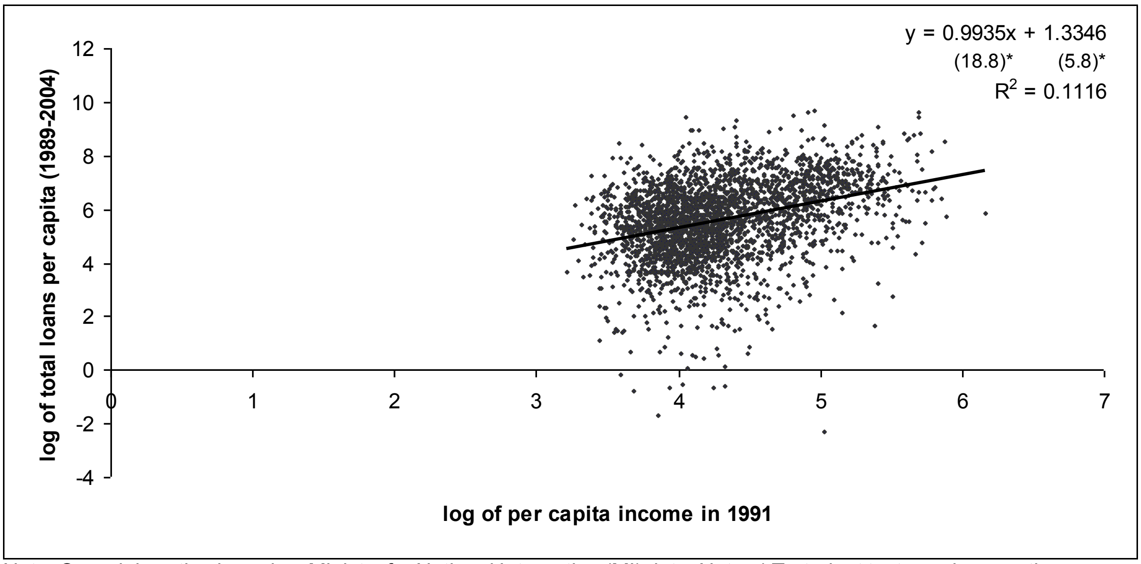
Note: Own elaboration based on BNDES and Ministry for National Integration (MI) data. Values are in R$ million, current prices.

Some authors, such as Almeida Junior et al. (2007) and Oliveira and Domingues (2005), argue that resource allocation of the regional development funds within each macro-region is guided by the demand side. In other words, only entrepreneurs within the prosperous areas have contracted these loans. Therefore, according to those authors, this fact may be generating an increase of intra-regional inequalities, i.e., the inequalities within the lagging macro-regions might be growing. Figure 3.11 aims to demonstrate this finding by plotting total regional development funds' loans per capita (between 1989 and 2004) against per capita income in 1991 at the municipal level. This simple correlation analysis shows what previous authors have already found using more sophisticated econometric methods: regional development funds' loans have been directed to the most prosperous areas (proxied by initial per capita income) within Northeast, North, and Central-West regions. Oliveira and Domingues (2005) suggest that the regional development funds are driven by the demand side. That is, they are requested by the local economic activities that fulfil the fund's requirements. Thus, it is likely that only the most developed activities, those located in municipalities with better access to information and banking infrastructure, have access to these funds. In the end, the consequence is that the impact of these regional development funds tends to be concentrated in the richest municipalities within the lagging macro-regions; therefore, having minimal impact on the economic development of the surrounding poor municipalities (Oliveira and Domingues, 2005).

Figure 3.11

Total Loans per capita (1989–2004) vs. Income per capita in 1991

(at municipal level)



Note: Own elaboration based on Ministry for National Integration (MI) data. Note: * T-student tests are in parentheses.

Silva et al. (2009) measure the effectiveness of regional development fund (FNE, FNO, and FCO) loans using propensity score estimates of firms that received loans (treatment group) and those that did not receive loans (control group) between 2000 and 2003. The results show that FNE has a positive impact on the growth rate for employment and no impact on the growth rate for wages. The study found that employment growth is approximately 60 percentage points higher for those firms that received loans than for those that did not receive loans over the period. With regard to FNO and FCO, there was no impact observed on the regional development funds on the two variables under study. This original research was sponsored by the Ministry for National Integration (MI/Government of Brazil) and generated policy reports using different time periods but essentially reported the same results[47]. Soares et al. (2009) employ the same propensity score method and expand the evaluation of FNE conducted by Silva et al. (2009), enlarging the time horizon under analysis. The results show significant impacts of FNE on employment growth for all periods between 1999

[47] See Almeida et al. (2007) and Silva et al. (2007).

and 2005; however no impact on the growth rate for wages was found[48]. Neither of the studies examines the loans granted to individuals in the agricultural sector, which represents roughly 60% of the total FNE during the period under study. For this reason, these results can be viewed as a partial evaluation of the regional development funds. Obviously, further evaluation and research is needed in this field.

Recently, regional development policy in Brazil has changed to target micro-regions (a group of contiguous municipalities) rated as stagnant or low income based on the National Regional Development Policy (PNDR) implemented by the Ministry of National Integration (MI) through Decree nº 6047 of 2007. The adoption of PNDR sub-regional types (namely, high income, growing, stagnant, and low income) aims at differentiating the micro-regional areas granted through regional development funding. This new approach for regional policy treats regional issues on a sub-regional scale, rather than as a macro-regional issue. This idea stems from the evidence that within the Northeast region, for instance, there are dynamic sub-regional areas (e.g., Petrolina/Juazeiro, Oeste Baiano) that have more capacity to attract private investments when compared with slow growing or low income sub-regional areas. Based on this concept, Araújo (1999) stresses the importance for focusing regional development policy (and public investments) in the stagnant or low income sub-regional areas, counterbalancing the natural tendency of the private investment to be directed to the most dynamic sub-regional areas. However, PNDR has at least three drawbacks. Firstly, micro-regional scale definition (groups of contiguous municipalities) may not represent a homogeneous set of municipalities that share similar characteristics (and problems) since economic shocks are not self-contained within micro-regions. Indeed, Resende (2011) suggests that micro-regions have externality effects that might spill over to the neighbouring micro-regions. The choice of a specific spatial scale to implement and evaluate the effectiveness of this regional policy should

[48] For instance, the impact of FNE on employment growth over the three-year period is 33 percentage points higher for financed firms.

be better justified. Secondly, the sub-regional typology employs income variables that are only available through the Census (1991 and 2000 are the most recent data) and municipal Gross Domestic Product (GDP) annual data that only have comparable data from 2002. These issues have a negative impact on both policy design and evaluation. Finally, the problem of low demand for loans in less developed areas will not be solved only by focusing on stagnant or low income micro-regions because, during the majority of the period, the regional development funds have not experienced an excess demand. For this reason, the relevant issue to address is how to create demand for funds in the stagnant or low income micro-regions.

In 2007, in addition to the PNDR, the development regional agencies (SUDENE and SUDAM) were recreated with the aim of identifying and defining regional development objectives; formulating regional development plans; and coordinating public interventions in the Northeastern region such as the regional development funds, sectoral plans (e.g., R&D and innovation system, environmental policy) and other incentives. Another programme, “Territories of Citizenship” (“Territórios da Cidadania”) launched in 2008, focused on selected poor territories over all Brazilian regions (mainly in the North and Northeast regions). This rural and local development action is coordinated by the Ministry of Agrarian Development (MDA) and encompasses 15 different ministries.

Despite some changes in Brazilian regional development policy, one issue seems to remain unchanged: the lack of outcome evaluation. This absence of empirical evaluation has limited the analysis of policy outcomes. Additionally, PNDR reliance on macro-data (e.g., GDP), where causation is difficult to prove and where counterfactual evidence is not developed, has prevented, and will continue to prevent, the debate from increasing our knowledge with its results, from discerning between good and bad practices, and from identifying the elements of the policy that should be improved.

4. The Spatial Scope of Regional Economic Growth Determinants: The Brazilian Case, 1991-2000

4.1. Introduction

The goal of this chapter is to analyse the spatial scope of regional economic growth in Brazil. With this aim, a regional dataset was prepared to examine the determinants of economic growth and convergence processes on four spatial scales[49] (states, municipalities, micro-regions and spatial clusters) between 1991 and 2000. Openshaw (1984) notes that many of the basic problems associated with the analysis of aggregated census data have long been recognised[50]. However, there is still little empirical evidence of how this aggregation problem has affected regional economic growth studies so far. Indeed, more research should be carried out in this field because what is true for a given spatial scale may not be applicable to another scale.

Despite the existence of a rich body of literature about the economic growth of various Brazilian regions, no studies have compared the process of economic growth across different scale levels. A survey of Brazilian literature about the determinants of economic growth revealed numerous papers discussing the theme using state-level data (Ferreira and Diniz, 1994; Ferreira, 1999; Azzoni et al., 2000; Azzoni, 2001; Resende and Figueirêdo, 2010), very few papers using data from the level of micro-regions (Vergolino et al., 2004) and an increasing number of papers in recent years that employ aggregated data at the municipal level (Andrade et al., 2002; De Vreyer and Spielvogel, 2005; Coelho and Figueirêdo, 2007). In addition, recent empirical studies (Magalhães et al., 2000; Silvera Neto, 2001; Lall and Shalizi, 2003; Silveira Neto and Azzoni, 2006) recognise the importance of spatial externalities, which can ultimately affect economic growth. Such spatial autocorrelation of economic growth

[49] In this article, the term "scale" is defined as nested sets of spatial units of different spatial resolutions (e.g., municipalities nested within functional regions, nested in turn within states).
[50] See, for instance, Gehlke and Biehl (1934), Robinson (1950) and Openshaw and Taylor (1979).

could manifest itself with varying intensity on different scale levels. Thus, multiple spatial dimensions of externality effects need to be analysed. Finally, Ertur et al. (2006) show that the estimated coefficients of growth regressions are not stable across space using regional data for the European Union. This phenomenon is known as spatial heterogeneity. This issue is taken into consideration to avoid bias in the growth regression estimates for Brazil.

While it is expected that this article will be informed by current discussions of spatial heterogeneity, spatial externalities and economic growth determinants (including convergence hypotheses), this article takes a distinct perspective on these three issues. The main question addressed by this chapter is as follows: How do the determinants of economic growth (including the intensity of spatial externalities) in Brazil vary according to the level of spatial aggregation used in the data? According to Rey and Janikas (2005), whereas a number of studies have examined the robustness of growth regression for measuring various aspects of research design (Levine and Renelt, 1992; Sala-i-Martin, 1997; Sala-i-Martin et al., 2004), changes in spatial scale have yet to be incorporated into this important line of research. To demonstrate how this phenomenon occurs, growth equations are systematically estimated – using the same time period and explanatory variables – across multiple scales. This strategy goes back to the exploration of the statistical literature, which proposes two approaches to analysing this measurement issue: the Modifiable Areal Unit Problem (MAUP) and the Ecological Regression Problem. MAUP refers to the variability in statistical results endemic to the selection of different area units (Openshaw and Taylor 1979, 1981)[51]. Another related problem discussed in the literature is the ecological fallacy. This problem appears when parameters estimated from macro-level data are used to make inferences about behavioural and socio-economic relations at a more disaggregate level (micro-level). Even if we assume that the municipal level is the micro-level of analysis (instead of the individual), the problem appears when the study is carried out

[51] MAUP has two components: the scale effect and the zoning effect.

using another aggregate level. Basically, these two concepts can be related to the measurement issue because both indicate an aggregation bias or effect[52]. Efforts to deal with MAUP have concentrated on obtaining functional regions that would be "geographically meaningful" to capture the economic sphere of influence of a group of smaller administrative units[53] (Cheshire and Hay, 1989).

However, Openshaw and Taylor (1979: 143) highlights that even if all researchers in a specific field agree on a unique set of areal units (e.g., functional regions), statistical variations might occur when other levels of spatial aggregation in the data are used and these variations "*will continue to be of interest in order to place a set of results into a meaningful statistical and spatial perspective.*" In addition to understanding how this variability comes to exist, it is important to understand why it occurs. Although understanding the reason for the existence of this variability is not the main objective of this chapter, the issue is addressed by an evaluation of the discontinuity thesis[54], examined in Hannan (1971)[55], using structural arguments (the scale-dependent determinants of economic growth). From the perspective of regional economic growth literature, such an approach supposes that each scale level could play a well-defined role in the economic growth process. The current work provides some potential explanations for the discontinuity thesis based on arguments of scale-dependent determinants of economic growth discussed in more detail in Chapter 2. However, in the context of regional growth literature, it is necessary to develop a cross-level theory (i.e., a theory linking scale levels) to provide a better understanding of such variability in the empirical results.

---

[52] Peeters and Chasco (2006) noted that the term ecological fallacy – typically used in social sciences – is isomorphic to the "modifiable areal unit problem" (MAUP) in geography.

[53] This strategy aims to identify the best scale for a particular problem (for example, to delineate the appropriate target boundaries or to choose the best spatial scale to implement and evaluate the effectiveness of public policies).

[54] "*Those who operate from discontinuity perspectives will certainly expect to find large and important differences in analogous models estimated at different levels of aggregation. However, to those who operate from continuity or homology assumptions, such effects should be quite disturbing. Since these effects would not have any direct theoretical meaning, the variations in estimates obtained at different levels must be considered a statistical artifact*" (Hannan, 1971:3).

[55] From a substantive perspective, the sociologist Hannan (1971) argued that it is necessary to develop cross-level theory to deal with scale problems.

The remainder of this chapter is organised as follows. The next section examines the role of spatial scales in the economic growth process in Brazil and the determinants of economic growth across different spatial scales. Section 4.3 discusses the econometric specifications of the study, the dataset and the spatial weight matrix. Section 4.4 presents the main results of the analyses. Section 4.5 provides a discussion of the results. The final section presents the main conclusions.

4.2. The Spatial Scope of Regional Economic Growth in Brazil

This section presents the four Brazilian spatial scales employed throughout the chapter. The roles of the various spatial scales, each of which can – in different ways – influence economic growth, are discussed. I conclude by revising the theoretical background of the economic growth determinants and how the explanatory variables may differently impact economic growth at different spatial scales (for more details on these see Chapter 2).

In Brazil, there are many types of regions, ranging from densely settled urban centres to sparsely settled rural regions as described in Section 3.3 (Chapter 3). As already described in Section 3.1, Brazil is divided into 27 states that are the main political-administrative units in the country and 5,507 municipalities that represent the smallest administrative level, dealing with local policy implementation and management. Also, there are micro-regions and spatial clusters aggregation levels that can be defined as functional regions. The micro-regions have been defined by IBGE [Instituto Brasileiro de Geografia e Estatística (Brazilian Institute of Geography and Statistics)] and group contiguous municipalities within the same state according to natural and production characteristics. The 91 spatial clusters proposed by Carvalho et al. (2007) were defined by an original cluster methodology in the form of an algorithm that grouped contiguous municipalities that share similar characteristics among the 46 variables reported in the Brazilian Census of 2000. Figure 4.1 (this is the Figure 3.1

shown in Section 3.1, Chapter 3) shows the four spatial scales employed here and some statistics concerning their sizes (in square kilometres).

Figure 4.1

Multiple Spatial Scales (Brazil)

States	Micro-regions	Spatial Clusters	Municipalities
			
n = 27 Area Mean = 315,982 Km^2 Area Min = 5,822 Km^2 Area Max = 1,577,820 Km^2 Area Standard Deviation = 378,718	n = 559 Area Mean = 15,262 Km^2 Area Min = 18 Km^2 Area Max = 333,857 Km^2 Area Standard Deviation = 29,659	n = 91 Area Mean = 93,753 Km^2 Area Min = 350 Km^2 Area Max = 1,340,216 Km^2 Area Standard Deviation = 196,110	n = 5,507 Area Mean = 1,549 Km^2 Area Min = 3 Km^2 Area Max = 161,446 Km^2 Area Standard Deviation = 5,738

Note: Own elaboration from data of IBGE and Carvalho et al. (2007).

Of note is that each empirical study about regional economic growth in Brazil examines the processes of convergence and the determinants of the economic growth using only one spatial scale described above. The major characteristics and spatial scale of analysis of a sample of studies on convergence in Brazil are summarised in Table 4.1.

Table 4.1

Sample of Studies on Convergence in Brazil

Paper	Scale level	Period	Convergence type	Model	Economic growth determinants
Ferreira and Diniz (1994)	State	1970-1985	Absolute β-convergence	Non-spatial model	Initial per capita income
Azzoni (2001)	State	1948-1995	Absolute β-convergence	Non-spatial model	Initial per capita income
Azzoni et al. (2000)	State	1981-1996	Conditional β-convergence	Non-spatial model	Initial per capita income, geographical variables (climate, latitude and rain), schooling and infrastructure variables (sewerage and running water).
Silvera Neto (2001)	State	1985-1997	Conditional β-convergence	Spatial lag model	Initial per capita income and schooling.
Resende and Figueirêdo (2010)	State	1960-2000	Conditional β-convergence	Non-spatial model	Initial per capita income, schooling, population growth, urbanisation, infant mortality rates, fertility rates, climate, taxation and migration.
Vergolino et al. (2004)	Micro-region	1970-1996	Conditional β-convergence	Non-spatial model	Initial per capita income, regional dummies and schooling.
De Vreyer and Spielvogel (2005)	Municipality	1970-1996	Conditional β-convergence	Spatial lag, spatial error and spatial Durbin models	Initial per capita income, illiteracy rate, the primary sector (agriculture) share, the share of urban population, the mean size of households and the share of households with electricity.
Silvera Neto and Azzoni (2006)	State	1985-2001	Absolute β-convergence Conditional β-convergence	Spatial error model Non-spatial model	Initial per capita income, schooling and share of manufacturing in the work force.

Note: Own elaboration.

Statistical variations, such as those illustrated in Table 4.1 might be related to a structural issue according to the discontinuity perspective because, in the case of the economic growth debate, relationships may exist between explanatory variables (including spatial externalities) and economic growth on different spatial scale levels as discussed in Section 2.1.3. The explanatory variables of economic growth used in this study are: initial level of income per capita (to test the convergence hypothesis), human capital (both educational and health capital), local infrastructure, population density, transportation costs, income inequality and the spatial externalities.

Although a cross-level growth model is not available to conclude that the results should change on different scale levels, I shed light on the fact that each spatial scale can play a role in terms of, for instance, the assignment of functions to different levels of government, which can differently influence economic growth at the four spatial scales discussed.. The responsibility of each level of government (federal, state and municipal) for the explanatory variables under study is as follows. Education in Brazil is financed and provided by the three levels of government as well as by the private sector. The latter charges tuition fees and is at liberty to be involved at all levels of education. The responsibility for public education is divided into (i) elementary education (states and municipalities), (ii) secondary education (states) and (iii) technical, technological and higher education (federal and states)[56]. According to the Constitution of 1988, Brazil's unified health system (Sistema Único de Saúde - SUS) was created to decentralise the provision of health services, increasing the autonomy of states and municipalities. In addition, there is a supplementary medical system, which includes the private plans and insurance companies[57]. Local infrastructure (housing infrastructure)[58], such as sewage, running water and the provision of electricity, is the joint responsibility of the three levels of government. Investments in transportation infrastructure, which are aimed at reducing transportation costs between municipalities in Brazil, are the responsibility of the federal and state governments.

Given the discussion above, I expect that the influence of all explanatory variables (including spatial externalities) on economic growth can be better captured at the municipal level. However, the aggregated influence of education, health and local infrastructure on economic growth will likely be observable at all spatial scales because these factors operate via public policies that extend across all scales. Although the coefficients of those three variables are expected to be statistically significant at all spatial scales, the magnitude of their impacts across scales may be different. Section 2.1.3 discusses the theoretical reasons to observe different results of convergence and impacts of human capital, infrastructure, population density, income inequality, transportation costs and spatial externalities on economic growth across different spatial scales. The theoretical arguments are related to the assumptions of an open economy (e.g., municipalities) versus a more closed one (states or functional regions) (Barro et al., 1995); the existence of 'social multipliers' (Glaeser et al., 2003); the

[56] For details, see Federal Law Nº 9.394/1996 and the Brazilian Federal Constitution of 1988 (Brasil, 2008) Art. 23, V. Art. 24, IX. Art. 30, VI. Art. 206, 208, 211, 212.
[57] For details, see the Brazilian Federal Constitution of 1988 (Brasil, 2008) Art. 23, II. Art. 24, XII. Art. 30, VII. Art. 195 par. 10. Art.196, 197, 198.
[58] See the Brazilian Federal Constitution of 1988 (Brasil, 2008) Art. 23, IX.

variability across space (because regions present different levels of development) of the effects of inequality on human capital accumulation, and ultimately on economic growth (Galor and Moav, 2004) and; agglomeration-related centripetal forces that might be more relevant at the local than the mega-regional levels, to cite some arguments. However, the impact of transportation infrastructure on economic growth may vary as the scale of analysis changes. If this impact is analysed at the functional regional or the state level, the focus will be on the connectivity between these aggregate regions. On the other hand, at the municipal level, such an analysis might examine the impact of transportation cost reductions within the borders of functional regions or states.

In regard to spatial autocorrelation, Oates (1999: 1130) highlights that, "*the existence and magnitude of spillover effects from localized public policies clearly depend on the geographical extent of the relevant jurisdiction.*" For this reason, spatial dependence is expected to be more evident at the municipal level than at the micro-regional, spatial cluster level or state level. As suggested by Oates (1999), it is possible to increase the size of the jurisdiction (municipalities, in the Brazilian case) to deal with such spillovers, thereby internalising the benefits and costs. Since 2005, Brazil has had some flexibility in terms of creating useful consortia[59] of municipalities to deal with particular issues (via inter-municipal coordinated decision-making), such as public utility services for water supply, sanitation, and health services. This chapter employs two spatial scales, which are the so-called functional regions (micro-regions and spatial clusters) that seek to keep such externalities within their boundaries. In addition, Lall and Shalizi (2003) provide some theoretical reasons why spatial externalities matter in examining determinants of growth that include agglomeration economies, Marshallian externalities of knowledge diffusion and labour market pooling, common informal norms/institutions and policy adoption. Finally, it is important to bear in mind that when using the same initial and final periods for all spatial scales, it may happen that, for instance, investments in education, health or local infrastructure impact quicker on

[59] For details, see Federal Law 11107 (2005).

municipal growth than they do on state level growth, which might be translated into different impacts across scale levels.

4.3. Econometric Specifications, Data and Spatial Weight Matrix

This section develops econometric specifications based on the determinants of economic growth discussed in Section 2 and on the spatial growth models proposed originally by Rey and Montouri (1999) and Fingleton (1999). In addition, this section discusses the dataset and the spatial weight matrix employed in the empirical strategy.

4.3.1. Econometric Specifications

The econometric specifications with a spatial component employed in this chapter derive from the spatial growth models initially suggested by Rey and Montouri (1999) and Fingleton (1999). The empirical models developed herein do not specify spatial scale choice; in other words, region, in these models, could refer to any spatial aggregation.

In empirical studies, the β-convergence hypothesis is traditionally tested by a simple linear regression model (for example, Barro and Sala-i-Martin, 1991, 1992) where the per capita income growth rate is estimated relative to the initial per capita income of the region by means of the Ordinary Least Squares (OLS) method. If that specification is modified to include other regional characteristics important to the dynamics of economic growth[60] (X is an $N \times K$ matrix of observations on other

[60] The dataset comprises of seven explanatory variables, which may be considered a small number of explanatory variables. As noted by one referee, for this reason, the estimates may suffer from missing variable bias in general, and the role of the latter variables may differ across spatial scales in particular. However, it is also fair to note that, in the context of regional growth studies, the current study uses an acceptable number of variables at a very disaggregate level as compared to other papers, such as Fingleton (1999) (three variables), Silveira Neto and Azzoni (2006) (three variables), Lall and Shalizi (2003) (six variables), De Vreyer and Spielvogel (2005) (eight variables), and López-Bazo et al. (2004) (nine variables).

exogenous variables), the absolute β-convergence gives way to the conditional β-convergence, which can be expressed by Eq. (4.1), the so-called Barro-regression[61]:

$$g = Y_0\beta_1 + X\beta_2 + \varepsilon \tag{4.1}$$

where g is an $N \times 1$ column vector with observations for per capita income growth for each region, Y_0 is an $N \times 2$ matrix including the constant term and the initial per capita income, and ε is the $N \times 1$ vector of error terms. A negative correlation between the growth rate and the initial per capita income (β_1<0) suggests that either mean reversion, conditional β-convergence, or both.

In Section 4.4, the first step is to run Eq. (4.1) using the OLS method to test for the existence of spatially auto-correlated errors at all scale levels. The next step is to run Eq. (4.2) or (4.3) to address the problem of spatial dependence in the growth regressions when necessary. In addition, the issue of spatial heterogeneity is investigated by dividing the samples into groups of regions with different spatial regimes [see Eq. (4.4)].

With regard to the spatial dependence issue, there is a large body of research that has investigated the determinants of regional economic growth and the convergence of regions in the presence of spatial externalities. The papers by Rey and Montouri (1999) and Fingleton (1999) seem to be the starting point for that line of research. More recent contributions to this field include López-Bazo et al. (2004) and Ertur and Koch (2007), who derived spatial economic growth models from the theory. Abreu et al. (2005) present a detailed literature review of this line of inquiry. Moreover, Section 2.2.1 (Chapter 2) discusses alternative spatial econometric specifications used in the growth literature.

[61] To mitigate the problem of endogeneity, all values for the explanatory variables are included at the start of the sampling period. Moreover, macro-regional dummies are included to control for macro-regional fixed effects, and when necessary, spatial models are used to deal with non-spherical errors. However, the study of causality in this type of economic growth estimate is not an easy task because it is hard to control for all omitted determinants. For this reason, causal findings should be interpreted with caution.

As discussed by Fingleton and López-Bazo (2006), most contributions to the study of spatial growth models have focused on spatial lag and spatial error models[62]. This chapter uses the approach proposed by Florax et al. (2003) for selecting the most appropriate econometric specification for the growth models in the presence of spatial dependence. This strategy consists of estimating standard OLS models [Eq. (4.1)] to check for spatial dependence while applying Lagrange Multiplier (LM) tests[63].

In the spatial lag model, the specification contains a spatially lagged dependent variable. Ignoring this spatial dependence will yield biased estimates of the coefficients. In the spatial error model, I follow the standard assumption that the error term in an OLS specification follows a first order spatial autoregressive process. As is well known, using OLS in the presence of non-spherical errors yields unbiased estimates for the estimated parameters but a biased estimate of the parameters' variance. Equations (4.2) and (4.3), respectively, present the spatial lag and spatial error models and can be estimated using maximum likelihood (ML) procedures.

$$g = \rho W g + X\beta + \varepsilon \qquad (4.2)$$

$$g = X\beta + \varepsilon, \quad \text{where} \quad \varepsilon = \lambda W \varepsilon + u \qquad (4.3)$$

W is the row standardised $N \times N$ spatial weights matrix. Thus, ρ and λ are the spatial autoregressive parameters, and ε and u are vectors of error terms. For the sake of simplicity, the X vector also includes the initial per capita income and the constant vector (Y_0). Similar to Eq. (4.1), to minimise the problem of endogeneity, values for all explanatory variables (X) are included at the start of the sampling period. Fingleton and López-Bazo (2006) noted that spatial lag (substantive) and spatial error (nuisance) cases produce rather different interpretations and policy implications with respect to the process of economic growth. In the former case, the assumption is that across-region externalities are due to knowledge diffusion and pecuniary externalities.

[62] I focus on the more commonly encountered forms of spatial models. Anselin (2003) provides a detailed discussion of several spatial processes.
[63] See Florax et al. (2003) for further details.

The latter can result from unobserved determinants that are correlated across regions and/or a mismatch between the spatial boundaries of the market process under study and the administrative boundaries used to organise the data (Rey and Montouri, 1999).

Finally, with regard to the spatial heterogeneity issue, the growth models estimated in Section 4.4.2.2 assume that coefficients are not stable across space, a fact that if not taken into account in the empirical analysis may be a source of bias. This issue was investigated by Ertur et al. (2006), who detected spatial convergence clubs in the estimation of growth regressions among 138 European regions over the period of 1980 to 1995[64]. The spatial club-convergence analysis relates the idea of club-convergence[65] examined in Durlauf and Johnson (1995) to the notion of spatial heterogeneity. Durlauf and Johnson (1995) rejected the hypothesis that the estimated coefficients in the cross-country regressions are the same in a different sub-sample of countries by grouping countries based on arbitrarily chosen cut off levels of initial income and literacy and also by using the regression-tree approach. To address this issue, the methodology suggested by Ertur et al. (2006) is employed in this study to group the Brazilian regions into different spatial regimes. Firstly, by means of exploratory spatial data analysis (ESDA) tools, such as Moran scatterplot[66], the spatial regimes of income per capita observed across Brazilian regions are identified: a cluster of rich regions (coined as spatial regime A) and a cluster of poor regions (spatial regime B). Secondly, if these spatial regimes are detected, Eq. (4.4) is estimated to test

[64] Fischer and Stirböck (2006) conducted a similar analysis using a sample of 256 European regions over the period of 1995 to 2000.

[65] As pointed out by Ertur et al. (2006: 8) "*the concept of club convergence is based on endogenous growth models that are characterized by the possibility of multiple, locally stable, steady state equilibria as in Azariadis and Drazen (1990). Which of these different equilibria an economy will be reaching depends on the range to which its initial conditions belong. In other words, economies converge to one another if their initial conditions are in the 'basin of attraction' of the same steady state equilibrium. When convergence clubs exist, one convergence equation should be estimated per club, corresponding to different regimes*".

[66] The Moran scatterplot is used to define spatial regimes in a way that is consistent across all spatial scales. As investigated by Ertur et al. (2006), the focus will be on the identification of regions in the first (High-High) and third (Low-Low) quadrants of the scatterplot, in other words, regions with high income per capita that tend to be surrounded by regions with high income per capita (HH, or spatial regime A) and regions with low income per capita surrounded by regions with low income per capita (LL, or spatial regime B).

whether convergence and the other economic growth determinants have statistically significant coefficients within each spatial club.

$$\begin{bmatrix} g_A \\ g_B \end{bmatrix} = \begin{bmatrix} Y_{0A} & X_A & 0 & 0 \\ 0 & 0 & Y_{0B} & X_B \end{bmatrix} \begin{bmatrix} \beta_{1A} \\ \beta_{2A} \\ \beta_{1B} \\ \beta_{2B} \end{bmatrix} + \begin{bmatrix} \varepsilon_A \\ \varepsilon_B \end{bmatrix} \tag{4.4}$$

The subscripts A and B indicate different spatial regimes. Eq. (4.4) is the counterpart of Eq. (4.1) and can be extended to include spatial effects, such as in Eq. (4.2) or Eq. (4.3), if spatial autocorrelation is detected.

The role of spatial scales in these economic growth models can be investigated by systematically repeating a method originally developed to examine this phenomenon at a single scale – using the same time period and explanatory variables – across multiple scales[67]. This empirical exercise is carried out in Section 4.4. In that section, I also combine multiple scale analysis with a variable uncertainty exercise, using the approach described by Levine and Renelt (1992) to verify the robustness of coefficients by including different sets of control variables.

4.3.2. Data and Spatial Weight Matrix

To investigate the determinants of economic growth on different scale levels in the context of growth regression estimates, I employ four Brazilian geographic stratifications discussed in section 3.1: 27 states, 559 micro-regions, 5,507 municipalities and 91 spatial clusters. The data are drawn from the municipal level and are then grouped to form other spatial scales.

[67] Yamamoto (2008) applied this approach to examine regional per capita income disparities in the USA on multiple spatial scales between 1955 and 2003. In that study, the focus is on methods such as inequality indices, kernel density estimation and spatial autocorrelation statistics. In addition, Briant et al. (2010) evaluate the magnitude of the distortions possibly induced by the choice of various French geographic stratifications in the context of economic geography estimations (wage equations).

The dependent variable is the income per capita growth rates between 1991 and 2000[68], and all explanatory variables are given in terms of 1991 values. Most of the socioeconomic data at the municipal level, such as logged per capita income, logged average years of schooling, logged infant mortality rate, logged Gini index, and logged population density, were obtained from the "Atlas do Desenvolvimento Humano no Brasil" (IPEA, PNUD and FJP, 2003). The Atlas provides data from the Census of 1991 using the 5,507 municipalities that existed in 2000, rather than the 4,491 municipalities that existed in 1991. Thus, it is possible to calculate per capita income growth between 1991 and 2000 at all scale levels. Indeed, the use of municipal data with constant borders limits the analysis to the 1991-2000 period. Logged transportation costs between all Brazilian municipalities and São Paulo are from IPEADATA. These transportation cost data are for the years 1980 and 1995. I estimated this variable for 1991 via interpolation. The cost of transportation to São Paulo is calculated through a linear program procedure as the minimum cost (given road and vehicle conditions) of travelling between a municipality's major headquarters and São Paulo. The local infrastructure index is constructed from the principal components analysis employed by Da Mata et al. (2007b). It takes into account several dimensions of housing services and utilities, such as electricity, sewage, water and garbage collection, and it is expected to capture the quantity of housing (or local) infrastructure in Brazilian municipalities[69]. Finally, the econometric specifications include regional dummies for the Brazilian macro-regions: Northeast, Southeast, South and Centre-West (the regional dummy for the North was excluded from the regressions to avoid perfect multicollinearity).

A spatial weight matrix is used to model spatial interdependence between regions. I consider pure geographical weights, which are exogenous by assumption, to

[68] The income per capita growth rates are averaged over ten years because municipal data are only available in the Brazilian population censuses conducted every ten years. Furthermore, given the presence of business cycle effects, the choice of ten-year growth averages seems to be a reasonable approach to avoid those influences. For instance, in 1994, Brazil launched the 'Plano Real' (Real Plan) – the stabilisation program – which ended a long period of high inflation rates that had started in the 1970s.
[69] I do not take the log of this variable because it has positive and negative values.

avoid endogeneity problems. The spatial weight matrix W used herein is based on the k-nearest neighbours calculated from the great circle distance between region centroids[70]. As pointed out by LeGallo and Ertur (2003), these matrices are preferable to the simple contiguity matrix [used, for example, by López-Bazo et al. (1999)] for various reasons. Two important reasons are that they (1) connect the islands of Ilhabela and Fernando de Noronha to continental Brazil and (2) force each unit to have the same number of neighbours, thus avoiding a situation in which rows and columns in W have only zero values[71]. In the next section, I show the results using a spatial weight matrix based on the ten nearest neighbours (k=10). In addition, a sensitivity analysis of the results was carried out using k = 5 and k = 15.

4.4. Results

The results of the baseline specification and diagnostics for spatial dependence are discussed with regard to the four spatial scales. Spatial econometric specifications and a spatial econometric framework for club-convergence testing are employed to correct for potential errors in the empirical strategy. Finally, a variable uncertainty exercise is conducted to investigate the robustness of the results[72]. At this point, some choices for the empirical investigation conducted in the next subsections should be explained. Firstly, I have followed the procedure suggested by Florax et al. (2003) to choose the appropriate spatial model. However, as pointed out by Anselin (2003: 158) "*equally valid and more accepted in the mainstream social science literature is the view that a substantive theoretical argument should suggest the nature of the externalities*". As

[70] The spatial weight matrices were calculated using the GeoDa software. In the menu option of GeoDa, centroids are central points. Central points are the average of the x and y coordinates of a polygon's vertices. Given the large size of the spatial units in the North region, the use of a different criterion for the definition of the centroids might result in important changes to the selection of neighbouring areas in that region.

[71] It is worth noting that these same properties can be obtained from a great circle distance based on a distance threshold above the minimum distance cut off. However, in a multiple scale context, this approach requires the definition of large distance cut offs given the large size of some regions (e.g., the Amazonas state).

[72] The Geoda software and the R software (version 2.10.1) with the package "spdep" were used to carry out all estimations.

noted in the introduction of this chapter the focal point of this chapter is the discussion of MAUP on growth regressions. In this sense, although the Florax et al. (2003) approach is not consensual, it provides a methodological guidance to choose the appropriate spatial model. Moreover, in Section 2.2 (Chapter 2) there is a discussion on the alternative spatial models and the problems in finding the "best model". The second choice made, was to handle jointly the spatial dependence and spatial heterogeneity. In cross sections, it is difficult to distinguish empirically between spatial autocorrelation and spatial heterogeneity; for instance, in some cases a spatial autocorrelation of residuals may simply indicate that the regression is misspecified[73] (Ertur et al., 2006). For this reason, I have chosen to follow Ertur et al. (2006) and Fischer and Stirböck (2006), for instance, in the empirical investigation conducted below and deal with both spatial dependence and heterogeneity. This issue is discussed in more detail in Section 2.2.2 of this thesis. Finally, robustness checks were conducted throughout the analyses because, in my view, this is a solid and well explained procedure to address the fragility of econometric inference regarding the choice of models as examined in Leamer (1978, 1983, 1985) and investigated in the growth regressions context by Levine and Renelt (1992), Sala-i-Martin (1997), Sala-i-Martin et al. (2004), for example.

4.4.1. Baseline Specification

The baseline specification [Eq. (4.1)] is estimated via OLS for the four spatial scales. Spatial dependence was assessed by applying the (robust) Lagrange Multiplier (LM) tests in the error terms. Table 4.2 shows results for conditional β-convergence. This specification recognises growth as a multivariate process.

[73] As explained in Section 2.2.2, one form of spatial heterogeneity (spatial heteroscedasticity) is just a specific version of the SEM model with a block diagonal spatial weights matrix. Another sort, spatial heterogeneity in the intercept is just an omitted variables problem. The only sort of spatial heterogeneity that is distinct is spatial variation in the coefficients (which may or may not be an omitted variables problem). Finally note that, it could happen that if the correct model is non-linear e.g. $g = b_0 + b_1 x + c_1 x^2 + \varepsilon$ then a model like $g = b_0 + b_2 x + c_2 Wx + \varepsilon$ may give you a significant c_2, simply because Wx is spuriously capturing something of x^2.

Table 4.2

OLS Baseline Estimation Results and Diagnostics for Spatial Dependence

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: OLS				
	political-administrative regions		functional regions	
Exploratory variables	(a) states	(b) municipalities	(c) micro-regions	(d) spatial clusters
ln(income per capita in 1991)	-0.0706***	-0.0608***	-0.0416***	-0.0677***
	(0.0209)	(0.0012)	(0.0035)	(0.0098)
ln(average years of schooling in 1991)	0.0112	0.0317***	0.0381***	0.0638***
	(0.0346)	(0.0013)	(0.0038)	(0.0153)
ln(Gini index in 1991)	0.1332**	-0.0070***	0.0076	-0.0343
	(0.0616)	(0.0027)	(0.0086)	(0.0341)
ln(infant mortality rate in 1991)	-0.0237*	-0.0127***	-0.0127***	-0.0113
	(0.0127)	(0.0010)	(0.0029)	(0.0110)
ln(transport cost to SP in 1991)	-0.0065	-0.0055***	-0.0018	-0.0033
	(0.0066)	(0.0007)	(0.0014)	(0.0039)
ln(population density in 1991)	0.0059**	-0.0002	-0.0003	0.000004
	(0.0024)	(0.0003)	(0.0006)	(0.0017)
local infra-structure in 1991	0.0015	0.0039***	-0.0006	0.0025
	(0.0068)	(0.0004)	(0.0011)	(0.0038)
Constant	0.5786***	0.3693***	0.2456***	0.3300***
	(0.1698)	(0.0109)	(0.0286)	(0.0925)
Regional dummies	yes	yes	yes	yes
Observations	27	5,507	559	91
Adjusted R-squared	0.7211	0.3948	0.3961	0.4333
Diagnostic for spatial dependence (W matirx: k=10):				
Lagrange Multiplier-Lag	0.6069	1057.4634***	217.7062***	2.1582
Robust Lagrange Multiplier-Lag	0.0049	46.9791***	19.9950***	0.0340
Lagrange Multiplier-Error	0.9813	2138.3393***	218.4065***	3.1039*
Robust Lagrange Multiplier-Error	0.3793	1127.8550***	20.6953***	0.9798

Note: Standard errors in parentheses; *** significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991].

First, the conditional convergence hypothesis cannot be rejected for all scales under analysis because the coefficients of initial income per capita are statistically significant. This finding means that each region at all scale levels converges towards its own steady state level of per capita income and not to a common level (such as in the absolute β-convergence case). Nevertheless, as discussed in Section 2.1.3, a more informative result for convergence is achieved by determining whether the cross-sectional dispersion of per capita income diminishes over time. Table 4.3 shows the results for σ-convergence between 1991 and 2000 on all spatial scales. Note that dispersion decreases (very slowly) on three of the spatial scales: states, spatial clusters and municipalities. Moreover, per capita income distribution at the micro-regional scale increases (or at least remains constant) from σ=0.574 in 1991 to σ=0.577 in 2000. The discrepancy in the β and σ-convergence results is due to the fact that these concepts capture two different aspects of per capita income across regions. As highlighted by Sala-i-Martin (1996), σ-convergence relates to whether or not the per capita income distribution across regions diminishes over time. On the other hand, β-convergence relates to the mobility of different individual regions within the given distribution of Brazilian per capita income.

Table 4.3

σ (Sigma)-Convergence

Scale level	N	Standard Deviation of (log of per capita income 1991) (a)	Standard Deviation of (log of per capita income 2000) (b)	Variation= (b-a)/a
States	27	0.426	0.410	-3.65%
Municipalities	5,507	0.583	0.579	-0.62%
Micro-region	559	0.574	0.577	0.55%
Spatial cluster	91	0.616	0.608	-1.30%

Note: Own elaboration.

Similar to existing empirical literature on Brazil, the results presented in Table 4.2 imply that control variables play a role in the performance of per capita income growth, as some of the estimated coefficients are statistically significant. It is useful to observe that explanatory variables seem to manifest differently in each of the four spatial scales, as indicated by the different magnitudes and significance of the coefficients between the spatial scales. Moreover, as the number of units increases, the model's explanatory power (adjusted R-squared) decreases. For example, at the municipal resolution (5,507 units), the adjusted R-squared term is 0.39, whereas at the level of the state (27 units)[74], the adjusted R-squared term climbs to 0.72. It is particularly important to analyse the diagnostics for spatial dependence because, in the

[74] It is important to note that the high R-squared for the state level may be a symptom of micronumerosity, which simply means a small sample size.

presence of spatial autocorrelation, the OLS coefficient parameters can be biased or inefficient, depending on the type of observed spatial dependence.

Table 4.2 shows the results of the test proposed by Florax et al. (2003) to identify the presence of spatial dependence across spatial units and to choose the best spatial econometric specification (spatial lag or spatial error). This strategy consists of estimating the standard OLS model and checking for spatial dependence by applying the (robust) LM tests[75]. For the conditional β-convergence equations, the LM statistics are not significant at the 5% level in the specifications for state-level data (column a) or spatial cluster-level data (column d), suggesting that the explanatory variables are able to deal with the spatial autocorrelation[76]. Indeed, Silvera Neto and Azzoni (2006) found that after conditioning their models on variables with very strong regional or geographic patterns across Brazilian states over the period 1985–2001, spatial dependence was not observed. These authors suggest that such significant explanatory variables reveal the potential channels through which strong spatial dependence in the process of income convergence occurs. Thus, the next subsection shows spatial correction only for the municipal and micro-regional levels. Following the approach presented by Florax et al. (2003), the spatial error model discussed in section 4.3.1 should be chosen because both LM-error and LM-lag reject the null hypothesis of no spatial dependence, and LM-error is greater than LM-lag in the specifications for municipalities and micro-regions. The robust versions of the LM tests support this option.

4.4.2. Spatial Analysis

4.4.2.1. Spatial Dependence

In this subsection, I report estimation results for the spatial models. Table 4.4 shows the results of the conditional β-convergence for the municipal and micro-regional levels using the spatial error model. First, the conditional β-convergence evidence cannot be

[75] LM-lag and LM-error tests consider the null hypothesis of no spatial dependency.
[76] For the absolute β-convergence estimations, all spatial scales suffer from spatial autocorrelation because the LM statistics are statistically significant. These results are available upon request.

rejected. Second, for the two spatial scales – municipal and micro-regional – the spatial autoregressive parameters (λ) are positive and significantly different from zero. Furthermore, the Likelihood Ratio tests[77] are highly significant (at the 1% level), indicating that the spatial error model specifications are appropriate.

Table 4.4

Spatial Error Model Results

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: ML		
	political-administrative regions	functional regions
Exploratory variables	(a) Municipalities	(b) micro-regions
ln(income per capita in 1991)	-0.0661***	-0.0352***
	(0.0013)	(0.0033)
ln(average years of schooling in 1991)	0.0358***	0.0336***
	(0.0014)	(0.0039)
ln(Gini index in 1991)	-0.0130***	-0.0134
	(0.0028)	(0.0083)
ln(infant mortality rate in 1991)	-0.0093***	-0.0122***
	(0.0011)	(0.0031)
ln(transport cost to SP in 1991)	-0.0055***	-0.0028
	(0.0009)	(0.0024)
ln(population density in 1991)	0.0010***	-0.0008
	(0.0003)	(0.0007)
local infra-structure in 1991	0.0046***	-0.0003
	(0.0004)	(0.0013)
lambda (λ)	0.6492***	0.7476***
	(0.0158)	(0.0412)
Constant	0.3723***	0.2193***
	(0.0108)	(0.0295)
Regional dummies	Yes	yes
Observations	5,507	559
Log-likelihood	14645.51	1732.72
Likelihood Ratio test (LR)	1175.90***	143.68***

Note: Asymptotic standard errors in parentheses; ***significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991]. The spatial weight matrix W is based on the 10-nearest neighbours.

From the measurement point of view, the results of Tables 4.2 and 4.4 indicate that MAUP jeopardises Brazilian economic growth estimates. The significance and magnitude of the coefficients vary at the four scale levels. For example, there is a positive and significant impact of educational capital (years of schooling) on economic

[77] The Likelihood Ratio test for spatial error dependence corresponds to twice the difference between the log likelihood in the spatial error model specification (Eq. 4.3) and the log likelihood in the specification given by Eq. (4.1); it follows a chi-square distribution with one degree of freedom.

growth at the municipal, micro-regional and spatial cluster scale levels. On the other hand, the coefficient for the local infrastructure variable seems to have a positive impact only at the municipal level.

The previous section focused on the conventional measure of σ-convergence. Recently, Rey and Dev (2006) and Egger and Pfaffermayr (2006) investigated σ-convergence in the presence of spatial effects. With regard the Brazilian regional dataset, the magnitude of spatial dependence of the logged income per capita was found to have similar magnitudes in 1991 and 2000 by use of the Moran's *I* statistic. Table 4.5 shows that apart from the results for the state level, the variations of Moran's *I* between 1991 and 2000 at the other spatial scales were quite small. For this reason, decomposition of the spatial effects for the measure of σ-convergence would not change the main conclusion of the previous section. A detailed analysis of the spatial σ-convergence is left for future research when a longer period of study is available.

Table 4.5

Moran's *I* of Log of Income Per Capita

Scale level	N	Moran's *I* of log of income per capita 1991 (a)	Moran's *I* of log of income per capita 2000 (b)	Variation= (b-a)/a
States	27	0.480***	0.559***	16.43%
Municipalities	5,507	0.746***	0.758***	1.65%
Micro-region	559	0.778***	0.807***	3.73%
Spatial cluster	91	0.442***	0.474***	7.13%

Note: *** Significant at 1% based on the permutation approach with ten thousand permutations. The spatial weight matrix W is based on the 10-nearest neighbours.

Next, a spatial econometric framework is employed to test the spatial club-convergence hypothesis. This step is important to determine whether the estimated coefficients are biased due to the spatial homogeneity assumption at each scale level, which underlies the presented analysis.

4.4.2.2. Spatial Heterogeneity

This subsection presents the results of the spatial club-convergence analysis, which assumes that coefficients are not stable across space. As explained earlier, the first step of this analysis is to detect the spatial regimes (or clubs) by means of the Moran scatterplot. Following the approach of Ertur et al. (2006), the logarithm of the initial level of income per capita is used to define the spatial regimes. In this case, the variable is the logged income per capita in 1991.

A Moran scatterplot was prepared for each spatial scale under analysis. It displays the spatial lag of the logged income per capita in 1991 against logged income per capita in 1991, both of which are standardised. Based on these Moran scatterplots, the Brazilian regions are split into four types. This exercise reveals a high proportion of High-High (HH) and Low-Low (LL) clustering types on the four spatial scales[78]. The Moran scatterplots corroborate the pattern of a clear polarisation of the Brazilian regions: a cluster of rich regions located in the south (spatial regime A) and a cluster of poor regions located in the north (spatial regime B). This fact indicates spatial heterogeneity suggesting that the convergence process and the impact of other explanatory variables on growth may be different across spatial regimes.

The new samples include only those regions located in spatial regimes A and B. Given the small number of observations, the regions located in the Low-High (LH) quadrant, which indicates poor regions surrounded on average by rich regions, were not included in the dataset for the growth regression estimates; conversely, those located in the High-Low (HL) quadrant were not included either. Furthermore, the spatial club-convergence regressions were not estimated for state and spatial cluster levels. The state level is not investigated here due to the lack of degrees of freedom for

[78] Specifically, at the state level, 11 states (41%) are of type HH and 11 states (41%) are of type LL; at the municipal level, 2558 municipalities (46%) are of type HH and 2338 municipalities (42%) are of type LL; at the micro-regional level, 268 micro-regions (48%) are of type HH and 227 micro-regions (41%) are of type LL; and, at the spatial cluster level, 42 spatial clusters (46%) are of type HH and 30 spatial clusters (33%) are of type LL. The spatial weight matrix W used here is based on the 10-nearest neighbours (k=10). However, it is worth noting that Moran scatterplots computed with the other spatial weight matrices, k=5 and k=15, lead to similar spatial regimes.

the second step of the analysis. Similarly, the spatial cluster scale is not examined due to the micronumerosity problem, which could result in misleading conclusions given the very large standard errors of the estimations. When interpreting the results, we should have in mind that LH and HL regions are not included in the samples. Therefore, the analysis is conditional upon the choice of the spatial weight matrix excluding those regions. It means that the spatial weight (W) matrix only connects the HH and LL regions. This is exactly the same procedure conducted by Ertur et al. (2006: 30) that was needed to make the similar exclusions and explain that: "*[t]he spatial clubs (LH) and (HL) containing only two regions and one region, respectively, are omitted due to the small number of observations in each and lack of degrees of freedom for the second step of our analysis*." Although somewhat problematic, this option should not invalidate the results of the HH and LL regions. Figures I.D.1 and I.D.2 (in the Appendix I.D) map the spatial regimes A and B at the municipal and micro-regional levels, respectively.

Table 4.6 presents the results for the spatial growth regression assuming two spatial regimes at the municipal and micro-regional levels. Note that these estimations also address spatially autocorrelated errors implying that all of the regions are spatially linked through the spatial weight matrix (W).

Table 4.6

Two Regimes Spatial Error Model Results

Spatial Regime	Explanatory variables	political-administrative region (a) Municipalities	functional region (b) micro-regions
	Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: ML		
Spatial Regime A (High-High)	ln(income per capita in 1991)	-0.0662***	-0.0501***
		(0.0021)	(0.0047)
	ln(average years of schooling in 1991)	0.0634***	0.0436***
		(0.0031)	(0.0078)
	ln(Gini index in 1991)	-0.0148***	-0.0099
		(0.0042)	(0.0129)
	ln(infant mortality rate in 1991)	-0.0076***	-0.0110***
		(0.0014)	(0.0038)
	ln(transport cost to SP in 1991)	-0.0020*	0.0013
		(0.0012)	(0.0022)
	ln(population density in 1991)	-0.0007	0.0008
		(0.0005)	(0.0009)
	local infra-structure in 1991	0.0017***	0.0006
		(0.0006)	(0.0018)
	Constant	0.3018***	0.2494***
		(0.0177)	(0.0354)
Spatial Regime B (Low-Low)	ln(income per capita in 1991)	-0.0765***	-0.0375***
		(0.0021)	(0.0050)
	ln(average years of schooling in 1991)	0.0274***	0.0256***
		(0.0017)	(0.0043)
	ln(Gini index in 1991)	-0.0041	0.0142
		(0.0042)	(0.0123)
	ln(infant mortality rate in 1991)	-0.0084***	-0.0061
		(0.0018)	(0.0045)
	ln(transport cost to SP in 1991)	-0.0164***	-0.0168***
		(0.0024)	(0.0042)
	ln(population density in 1991)	-0.0002	-0.0010
		(0.0005)	(0.0011)
	local infra-structure in 1991	0.0075***	0.0018
		(0.0007)	(0.0017)
	Constant	0.5170***	0.3426***
		(0.0235)	(0.0488)
	lambda (λ)	0.5946***	0.5588***
		(0.0186)	(0.0615)
	Regional dummies	Yes	yes
	Observations	4,896	495
	Log-likelihood	13172.27	1598.15
	Likelihood Ratio test (LR)	759.14***	52.60***

Note: Asymptotic standard errors in parentheses; ***significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991]. The spatial weight matrix W is based on the 10-nearest neighbours.

With regard to convergence processes, the models at both spatial scales present highly significant and negative coefficients for the initial income per capita level. Furthermore, at the municipal and micro-regional levels, the null hypotheses on the joint equality of coefficients across the two spatial regimes are rejected by the spatial version of the Chow–Wald test (Anselin, 1990). At the municipal level, the value of the test statistic is 232.17 (p-value=0.0000) and at micro-regional level the value is 22.30 (p-value=0.0343). Indeed, these results are consistent with the hypothesis of two club-convergence in Brazil.

The result presented in Table 4.6 indicate that explanatory variables manifest differently in each spatial regime and in each spatial scale under analysis, as demonstrated by the differences in the magnitude and significance levels of the coefficients between the spatial regimes as well as the spatial scales. However, the statistically significant coefficients are similar to those presented in Table 4.4. For instance, when spatial heterogeneity is not allowed, all coefficients are statistically significant at the municipal level (see Table 4.4). When the two spatial regime model is estimated, only the population density (in spatial regimes A and B) and the Gini index (in spatial regime B) coefficients become statistically insignificant. However, the noteworthy difference amongst the two spatial regime results is the magnitude of the estimated coefficients. For instance, the impact of schooling on economic growth at the municipal level within spatial regime A is more than double that within spatial regime B.

Also, it is important to determine whether the dispersion of per capita income between regions decreases over time at each spatial regime. Table 4.7 presents the results of σ-convergence at the municipal and micro-regional levels for both regimes and using the whole dataset for the sake of comparability. Looking at the municipal level, the σ-convergence results show opposite processes between the spatial regimes. Within spatial regime A, σ-convergence holds with a decrease of approximately 6% in the dispersion of logged income per capita between 1991 and 2000. By contrast, within spatial regime B, there is an increase in the dispersion of logged income per capita of approximate 10%. A similar pattern can be observed at the micro-regional level. Altogether, these findings demonstrate a very different process of convergence/divergence across the spatial clubs in Brazil. These findings also explain why the regions converged slowly (at the municipal level, a variation of -0.62% in the dispersion) or presented a small divergence (at the micro-regional level, a variation of +0.55%) over the 1990s, when all of the regions were analysed together.

Table 4.7

σ (Sigma)-Convergence in Two Spatial Regimes

Scale level	N	Spatial regime	Standard Deviation of (log of per capita income 1991) (a)	Standard Deviation of (log of per capita income 2000) (b)	Variation= (b-a)/a
Municipalities	2,558	A (High-High)	0.311	0.292	-5.96%
Municipalities	2,338	B (Low-Low)	0.284	0.312	9.68%
Municipalities	5,507	All units	0.583	0.579	-0.62%
Micro-region	268	A (High-High)	0.291	0.252	-13.21%
Micro-region	227	B (Low-Low)	0.261	0.269	3.20%
Micro-region	559	All units	0.574	0.577	0.55%

Note: Own elaboration.

Before I further analyse the results shown in Tables 4.2, 4.4 and 4.6, it is important to check the robustness of these results in two ways: (1) determine whether each coefficient remains significant and maintains its sign across a set of combinations of variables; and (2) conduct a sensitivity analysis based on 5- and 15-nearest neighbour spatial weight matrices. These robustness checks help to identify robust structural factors that drive economic growth on different scale levels in Brazil.

4.4.3. Robustness Checks

Herein, it is worth noting the relevance of robustness tests on the discussion of the determinants of economic growth[79]. I use the approach proposed by Levine and Renelt (1992) to test the robustness of coefficient estimates by including different sets of

[79] The literature review about robustness tests discussed herein is mainly based on Resende and Figueirêdo (2010) which conducted robustness tests for growth determinants at state level in Brazil.

control variables[80], which is a variant of Leamer's (1983) extreme-bounds analysis (EBA). The basic framework employed in this test aims at dealing with model uncertainty[81]. Brock and Durlauf (2001) highlight the challenge of studying the determinants of economic growth due to the "open-endedness" of the empirical literature on this subject, that is, at a conceptual level, one theory suggests determinants for growth without necessarily excluding determinants proposed by other theories. The cross-country regressions are extreme examples of this fact, as stressed by Sala-i-Martin et al. (2004: 814): "*the number of proposed regressors exceeds the number of countries in the world, rendering the all-inclusive regression computationally impossible*". Moreover, as highlighted by Sala-i-Martin et al. (2004), some empirical economists have simply "tried" combinations of variables that could be potentially important determinants of growth, and they then report the results of their preferred specification. As indicated by Sala-i-Martin et al. (2004: 814), "*such 'data-mining' could lead to spurious inference*".

To address the fragility of econometric inference regarding the choice of models in general, Leamer (1978, 1983, 1985) proposes a sensitivity analysis, more specifically, the extreme-bounds analysis (EBA) to identify robust empirical relationships. The robustness analysis of economic growth determinants appears in three seminal papers published in the American Economic Review (Levine, Renelt, 1992; Sala-i-Martin, 1997; Sala-i-Martin et al., 2004)[82]. Levine and Renelt (1992) verify the robustness of the determinants of growth in countries by applying a version of the so called "extreme bounds analysis" proposed by Leamer (1983). The authors conclude that very few regressors are statistically robust. In response to this pessimistic result, Sala-i-Martin (1997) employs a less severe test of the explanatory variables in growth regressions analysing the entire distribution of the variables and

[80] Following this idea, some authors have suggested other approaches (Sala-i-Martin, 1997; Fernandez et al., 2001; Brock and Durlauf, 2001; Sala-i-Martin et al., 2004).
[81] See Temple (2000) and Brock et al. (2003) for further discussions of model uncertainty.
[82] Other articles on robustness tests applied to growth regressions have been published, for example, Fernandez et al. (2001), Brock et al. (2003), Ley and Steel (2007) and Eicher et al. (2007).

checking their statistical significance. Then, Sala-i-Martin (1997) identifies a relatively large number of statistically significant variables associated to economic growth. In later work, Sala-i-Martin et al. (2004) showed that the approach proposed by Sala-i-Martin (1997) is a particular case of "Bayesian averaging models" and suggested a technique called 'Bayesian Averaging of Classical Estimates' (BACE), to determine the importance of a variable in the regression of economic growth. The results using the Bayesian approach confirm those found in Sala-i-Martin (1997)[83]. It is noteworthy that, Durlauf et al. (2004), in a survey of econometric techniques that are used to study economic growth, point out that one of the reasons to be optimistic about the line of empirical research on growth is the potential contribution that recent developments in methods of dealing with uncertainty models can bring to the discussions on growth issues.

Due to the lack of evidence of spatial autocorrelation at the state and spatial cluster scales, I employed Eq. (4.1), described in Section 4.3.1, to determine robustness. To test for robustness at municipal and micro-regional scales, where spatial dependence exists, the spatial error model [Eq. (4.3)] was employed. Furthermore, the spatial error model of Eq. (4.4) was used to test the robustness of the results in the presence of spatial heterogeneity. Temple (2000: 184) explains that, "*the central idea of EBA is to report an upper and lower bound for parameter estimates, thereby indicating sensitivity to the choice of specification.*" Basically, the EBA approach states that a coefficient is "robust" if it remains significant (at the 5% level) and does not change its sign across a set of combinations of other variables. Otherwise, the variable is determined to be "fragile". Instead of presenting only the upper and lower bounds of the coefficients, I follow Temple's (2000) suggestion and present information from a variety of models.

[83] In fact, the methods are quite similar. An advantage of the method proposed by Sala-i-Martin et al. (2004) is that it considers models of all sizes and no variable is kept "fixed" and therefore "untested."

In Appendix I.B, Tables I.B.1, I.B.2, I.B.3 and I.B.4 show the results of the robustness checks for states, municipalities, micro-regions and spatial clusters, respectively. Furthermore, in Appendix I.D, Tables I.D.1 and I.D.2 show the results of the robustness tests for the two spatial regimes models at the municipal and micro-regional levels. Given the multicollinearity problem and the high level of correlation between variables (see Appendix I.A), I tried to run models that eliminated variables with high correlation coefficients. To minimise the omission of relevant variables, regional dummies and initial per capita income are included in all of the regressions. For each spatial scale, I run eight models that are similar across the scale levels; that is, they include the same combination of variables.

When spatial heterogeneity, or two spatial regimes, is not allowed in the estimation framework, the results show that conditional β-convergence cannot be rejected for all spatial scales because the coefficients of initial income per capita are negative and significant for all specifications. On the municipal and micro-regional scales, the spatial autoregressive parameters (λ) can be deemed robust parameters because they appear highly significant in all specifications. This result suggests that spatial externalities are mainly operating within smaller-scale units of measurement, such as municipalities and micro-regions. Contrary to previous results (Tables 4.2 and 4.4) that show distinct results across spatial scales, I find common structural factors operating on different spatial scales when robustness tests are carried out (Tables I.B.1 to I.B.4 in Appendix I.B). Indeed, three factors suggested by the growth theories (higher educational and health capital[84] and better local infrastructure) seem to result in higher rates of economic growth at three levels of spatial scale: municipalities, micro-regions and spatial clusters. Furthermore, the analysis of their coefficients indicates that the magnitudes of such coefficients vary on different spatial scales.

[84] At the spatial cluster level, health capital (infant mortality rate) is statistically significant in specification (7), Table I.B.4. However, it seems that health capital is a robust variable because it appears statistically significant on the other scale levels (state, municipalities and micro-regions) in all specifications.

In addition, at the municipal level, more variables appear to be robust in their influence on growth. It is worth noting that transportation costs and population density coefficients are statistically significant at the 1% level in all model specifications. Finally, state-level results should be viewed with caution due the small sample size of that particular regression analysis. At the state level, higher population density and health capital correlate with higher economic growth rates. Although EBA is a useful method for communicating uncertainty surrounding parameter estimates, future research may propose other approaches to model selection and tests of robustness for multiple scale levels.

Sensitivity analysis of the results were carried out using 5- (k = 5) and 15-nearest neighbour (k=15) spatial weight matrices. This sensitivity analysis arrived at results similar to those discussed above. One interesting result of the robustness check comes from the investigation of the extent of the effects of spatial externalities, where I calculated Moran's *I* for the residuals of the municipal and micro-regional specifications in Table 4.2 using k= 5, 10, 15, 30, 60 and 120. Moran's *I* is a measure of global spatial autocorrelation, and it is used here to show that spatial dependence (at both the municipal and micro-regional levels) wanes as the number of neighbouring regions increases, suggesting that spatial externalities are bounded in space[85]. Table I.C.1 (in Appendix I.C) shows the values of Moran's *I*. This result is in line with the first law of geography, which states that, "*everything is related to everything else, but near things are more related than distant things*" (Tobler, 1970: 236).

Finally, robustness tests were conducted for the spatial club-convergence specifications, which allow coefficients to vary across two spatial regimes at the municipal (Table I.D.1) and micro-regional levels (Table I.D.2). All of the initial income per capita parameter estimates are negative and statistically significant for both spatial

[85]It is worth noting that the choice to use the same number of k-neighbouring regions across multiple scales is intended to give a clear rule of neighbourhood across spatial scales. On the other hand, it is important to realise that the extent of spillover differs from one spatial scale to the other when using a specific k-neighbourhood matrix. To address this gap, at least partially, the sensitivity analysis conducted in Table I.C.1 using other k-nearest neighbours was intended to indicate the extent of spillover effects at different spatial scales.

regimes. The results are consistent with an inference of two club-convergence in Brazil. Moreover, these robustness checks confirm that three structural factors are operating on municipal and micro-regional levels, namely, educational and health capital and local infrastructure. The population density coefficient does not seem to be robust because it is not statistically significant across the two spatial regimes at the municipal and micro-regional level in some specifications. More discussion on the reasons of the differences found in the estimates across different spatial scales and regimes is provided in the next section.

4.5. Discussion

The multiple scale analysis conducted in this chapter aimed to examine the variability of the economic growth determinant estimates using a Brazilian regional dataset grouped into four different spatial scales. This section discusses how explanatory variables could differently impact economic growth on the various spatial scales. These results are robust in the face of alterations in the conditioning set of variables. The first noteworthy result of this analysis is that when spatial heterogeneity is not allowed in the estimation framework, the conditional β-convergence (or mean-reversion process) is not rejected on all four of the spatial scales, although the dispersion of per capita income (i.e., σ-convergence) decreased too slowly on three spatial scales between 1991 and 2000 and increased at the micro-regional level. The test of the spatial club-convergence hypothesis shed new light on this previously reported result. First, the econometric estimates of the two spatial regime models at the municipal and micro-regional levels suggest the existence of two club-convergence in Brazil. This result is in line with the findings of studies that used Brazilian municipal data and suggest the formation of convergence clubs (Andrade et al., 2004; Coelho and Figueirêdo, 2007). If we compare the robustness results of spatial club-convergence at the municipal and micro-regional, it is possible to verify higher rates of convergence at the municipal level suggesting that municipalities are more open economies (Barro et al., 1995) than the micro-regions. For this reason, municipalities present faster convergence rates to their own state-steady levels of income per capita. Of note, this tendency of higher (slower) convergence rates at finer (larger) spatial scales do not hold when results at all spatial scales are compared if spatial heterogeneity is not allowed in the estimation framework (see for instance Table 4.2 and the respective robustness tests in Tables I.B.1, I.B.2, I.B.3 and I.B.4). Moreover, the σ-convergence analysis that accounts for the two spatial regimes reveals that within the spatial regime of rich regions there is a decrease in the dispersion of per capita income between 1991 and 2000, while within the spatial regime of poor regions an increase in the dispersion of per capita income can be observed. Altogether, these findings suggest that two different processes might be occurring across the spatial clubs in Brazil.

Local infrastructure, educational capital and health capital are robust predictors of economic growth at three scale levels (municipalities, micro-regions and spatial clusters). Tests for robustness were used to identify these scale-dependent determinants of economic growth. These findings suggest that variability in the statistical significance of the coefficients due to the selection of different spatial scales, (i.e., aggregation bias/effect) can be mitigated by carrying out this kind of robustness test (EBA). This result is in line with the analysis presented in section 4.2, in which it was suggested that the aggregated influence of education, health and local infrastructure would positively impact economic growth on all spatial scales because each level administers public policies related to these factors. In addition, it is worth noting that the magnitude of such impacts varies across spatial scales as well as across spatial regimes. Interestingly, in regards to the educational capital coefficients, the estimates at the spatial cluster level (that aims to represent functional regions, self-contained) are always higher than those at the municipal or micro-regional level. This fact corroborates the argument that the coefficients size may rise with the level of aggregation due to the existence of a social multiplier supporting the idea that there are human capital spillovers, One evidence that is related to this finding can be found in

Cravo and Resende (2012) who investigate the role of space to GDP per capita growth in Brazil between 1980 and 2004 at the micro-regional level by applying a spatial filter that eliminates the spatial dependence of the data. Maza and Villaverde (2009) argue that the filtered variables can be understood as the part of the data that is not explained by spillover effects from neighbouring economies. Cravo and Resende (2012) show that human capital coefficient became insignificant after removing the spatial autocorrelation from the data, a clear indication that its statistical significance is intrinsically linked to the spatial structure of the country. In other words, when the social multipliers suggested by Glaeser et al. (2003) are removed from the data the coefficients may decrease and eventually may become statistically insignificant.

When spatial heterogeneity is not allowed, the transportation cost coefficient is statistically significant (and negative) only at the municipal scale. This finding indicates that reductions in transportation costs have a positive impact on municipal economic growth. The analysis of multiple scale levels reveals that, despite finding no impact of transportation costs on economic growth in the functional regions (micro-region and spatial cluster) and states, reductions in transportation costs within the borders of a functional region or state might have affect economic growth. In addition, the observed statistical significance for the population density coefficient shows that agglomeration-related centripetal forces may be much more relevant at the local than the state level; in other words, the most densely populated municipalities in Brazil are growing faster. Observing Figure 3.7 in Section 3.3, it possible to conclude that municipalities localised at coastal areas (with all else constant) are growing faster. These two findings highlight the structural argument, showing that the influence of transportation costs and population density on growth might be operating only at the municipal level.

When the two spatial regime model results are examined, the main conclusions on the statistical significance of economic growth determinants remain practically unchanged. Of note, the coefficients of transportation costs are only robust on the spatial regime of poor regions at the municipal level and micro-regional level.

Therefore, reductions in transportation costs only within the borders of a poor spatial cluster or a poor state would affect economic growth. At the municipal level, the negative and statistically significant estimated coefficients of the Gini index in all specifications suggest that high personal income inequality is harmful for growth within the spatial club of rich municipalities. On the other hand, within the spatial club of poor municipalities some specifications in Table I.D.1 show that high personal income inequality foster economic growth at the municipal level. These different impacts of income inequality depending on the level of development of the regions are related to the theory developed by Galor and Moav (2004). These authors argue that in the early stage of development, economic growth is driven by physical capital accumulation that is promoted by disparities among individuals. In this sense, it is possible to justify the positive association between income inequality and growth within less developed areas in Brazil. However, in more developed areas (e.g., within the spatial club of rich municipalities), human capital accumulation is the main engine of economic growth. Galor and Moav (2004) explain that people can invest more in education if wealth is equally distributed among individuals, and hence income inequality will be negatively related to economic growth.

Furthermore, the population density coefficient in both spatial regimes is not robust at the municipal and micro-regional level because in some econometric specifications the coefficients are statistically insignificant. It is important to point out that although at the micro-regional level most of the coefficients of population density are not statistically significant (only one out of eight specifications in Table I.D.2); at the municipal level, most specifications show a positive and statistically coefficient of population density (see Table I.D1). These results suggest that what has been discussed in Section 2.2 may hold for the Brazilian case, that is, agglomeration effects may be operating only at a finer scale level (municipalities) and their effects seem to be higher in the spatial regime of rich regions.

Finally, spatial autocorrelation appear as geographic areas get smaller because only the error terms of the econometric specifications at the micro-regional and municipal levels are spatially autocorrelated. To effectively distinguish between the spatial error (SEM) and spatial lag (SAR) specifications of the growth models, this study followed the strategy outlined in Florax et al. (2003). The LM tests applied to the estimates at both spatial scales indicated spatial error specification to be the appropriate model. Equation (4.3) shows that a random shock introduced into a specific municipality (or micro-region) will not only affect the economic growth rate in that municipality (or micro-region) but also the growth rates of other municipalities (or micro-regions). As noted by Finglenton and López-Bazo (2006), the interpretation of random shocks in a cross-section of growth rates averaged over several years is more closely related to unobserved determinants and measurement errors that are correlated across regions than to, for example, shocks originating from business cycles. Possible unobserved determinants of economic growth not included in these models include cultural and institutional factors, which might be correlated across municipal and micro-regional units. It is also likely that there is a mismatch between the spatial boundaries of the economic growth processes under study and the municipal (and the micro-regional) boundaries used to group the data. Both interpretations are plausible because spatial dependence was not detected at the spatial cluster and state levels. Specifically, the spatial cluster level was constructed to capture the economic sphere of influence of a group of contiguous municipalities. In this sense, spatial externality effects may be bounded within each spatial cluster unit. Furthermore, some public policies with spatial components[86] may be better delivered if coordination existed between contiguous jurisdictions (e.g., municipalities). This possibility explains the proposal to work with groups of municipalities to deal with spillovers, thereby internalising all benefits and costs (Oates, 1999). Lastly, it is important to note that the

[86] For instance, adoption of a training program in a specific municipality to improve the education of workers or an improvement in local infrastructure might affect people's and firms' decisions about whether to migrate from neighbouring regions.

spatial process in the error terms may be translated to alternative spatial econometric specifications as discussed in Section 2.2.1.1. Indeed, the spatial Durbin model (SDM) nests most spatial models often used in the regional growth literature, namely, SAR, SEM and SLX models. LeSage and Fischer (2008) argue that SDM specification is a natural choice over competing alternatives and it provides a substantive justification for the spatial externalities by the inclusion of spatial lags of both the dependent and explanatory variables. Of course, exploring alternative spatial econometric specifications at different spatial scales can provide more insights on the spatial externalities across space, but it is beyond the scope of this study.

4.6. Conclusions

The aim of this chapter was to analyse Brazilian economic growth on different spatial scales, ranging from municipalities to state regions, between the years of 1991 and 2000. This study suggests a general framework for dealing with multiple spatial scales, spatial autocorrelation, spatial heterogeneity and model uncertainty.

This chapter showed that if a single regression is estimated on different scale levels, the results change with the scale level. Importantly, the robustness tests identified variables that were simultaneously significant on different spatial scales: higher educational and health capital and better local infrastructure were related to higher rates of economic growth. This result suggests that because public policies (education, health and local infrastructure) are operating across all scale levels, these factors will be significant across spatial scales, although their impacts on growth may differ across spatial scales as well as spatial regimes. Moreover, the significance of the coefficients for the transportation costs of the spatial regime of poor regions at the municipal and micro-regional levels suggests that reductions in transportation costs had an impact only within the borders of the poor spatial clusters and the poor states in Brazil over the 1990s.

The hypothesis of spatial club-convergence cannot be rejected, indicating different processes of convergence in Brazil. Furthermore, I directly tested whether the dispersion of per capita income between regions decreased over time in two spatial regimes; that is, whether σ-convergence exists within two different clusters, a cluster of rich regions and a cluster of poor regions. This analysis showed that per capita income distribution decreased within the spatial club of rich regions in Brazil between 1991 and 2000. For instance, per capita income dispersion fell by 13.2% at the micro-regional level during the study period. This dispersion increased for the spatial club of poor regions. For the club of poor municipalities, per capita income dispersion increased by 9.7% over the study period.

Spatial autocorrelation appeared at finer scales (the municipal and micro-regional levels). The underlying reasons for the spatial autocorrelation may involve externalities operating through the error term or can be related to unobserved determinants and measurement errors that are correlated across municipalities and micro-regions. This finding does not exclude the possibility that externalities may be operating through the spatial lags of the dependent and explanatory variables. I argue that public policies with a spatial component could be better implemented if coordination between contiguous municipalities was initiated, thereby internalising the benefits and costs of such policies.

One area requiring further research is the development of a cross-level theory linking spatial scales to improve our understanding of why economic growth differs from one scale to another. Furthermore, it is important to note that the results presented here are specific to the study period. Additional explanatory variables could be included in the analysis if relevant data were available. These issues will be left for future research because they are beyond the scope of this chapter.

5. Spatial Dimensions of Economic Growth in Brazil, 1970-2000: What is the Extent of Spatial Autocorrelation?

5.1. Introduction

This chapter extends the analysis of previous chapter by focusing on a descriptive analysis of the extent of the spatial autocorrelation effects in the context of regional growth literature, by testing whether the residuals of traditional growth model estimates are spatially autocorrelated at different spatial scales[87] using standard panel data models between 1970 and 2000 in Brazil. This approach allows us to investigate whether alternative non-spatial panel data models (that control for time invariant fixed effects) eliminate or, at least, mitigate the spatial autocorrelation. One of the advantages of using panel data framework is that it allows, as noted by Islam (1995), for differences in aggregate production functions for all the regions in the form of unobservable individual fixed effects, for instance correcting for omitted variable bias. The contribution of this current analysis is to explore the space and time dimensions of economic growth in Brazil using alternative panel data techniques to provide a measure of the extent of spatial autocorrelation (in kilometres) over three decades (1970-2000). To my knowledge, this is the first study of regional economic growth exploring both time and spatial scale dimensions. Some studies so far, such as Elhorst et al. (2010), advances the growth literature by using spatial econometric techniques to focus on time-space models; but they only examine the process of economic growth at one single spatial scale. However, as suggested by Resende (2011)[88] (and also discussed in Chapter 4) the need for spatial models (and therefore, the use of W matrices) depends on the level of spatial aggregation used in the data. In other words, the empirical exercise conducted by Resende (2011) shows that spatial autocorrelation

[87] In this thesis, the term 'scale' is defined as nested sets of spatial units of different spatial resolutions (e.g., municipalities nested within micro-regions, nested in turn within states).
[88] Resende (2011) focuses on the examination of the determinants of Brazilian regional economic growth at multiple levels of spatial aggregation using a cross-sectional data set only over the 1990s period.

appears only at finer scales. Except from this latter work, the studies thus far have only examined the existence of spatial autocorrelation in the process of economic growth at a single spatial scale to infer the consistency of spatial growth models with reality (e.g., Rey and Montouri, 1999; Fingleton, 1999; López-Bazo et al., 2004; Ertur and Koch, 2007; Fischer, 2011).

The term I use throughout this chapter is "spatial autocorrelation" because it takes a more conservative and descriptive view of the process of spatial interactions found in the empirical literature. Words such as spatial externalities, spatial effects, spatial spillovers, spatial dependence and interaction effects often suggest a relationship of causality between variables. However, as discussed in Chapter 2 given the problems to determine causality in the literature of economic growth (Temple, 1999 and Durlauf et al., 2005) and of spatial econometrics in general (McMillen, 2010; Gibbons and Overman, 2012), I prefer to use the term spatial autocorrelation in the current analysis.

As stated in the beginning of this chapter, herein it is conducted a descriptive analysis of regressions' residuals. Therefore, this study can be viewed as a first step for further investigations of spatial spillovers in the process of economic growth across multiple spatial scales using (spatial) panel data techniques. In this sense, the empirical exercise obtains estimates of the extent of spatial autocorrelation in the regressions' residuals by using alternative estimation techniques as well as spatial weight matrix specifications across different spatial scales (minimum comparable areas, micro-regions, meso-regions and states). Moreover, it provides information on the spatial autocorrelation effects for each decade under analysis, allowing for a comparison of the strength of such autocorrelation effects across time. It is also discussed throughout the chapter how spatial autocorrelation in the residuals can be interpreted in light of the alternative available theories. Among other results, this chapter shows that, at the state level in the Brazilian case, there is not spatial autocorrelation in the residuals of non-spatial panel data models demonstrating that these models are appropriate to investigate growth determinants and convergence process. However, for the other spatial scales under analysis, the results show that non-spatial panel data techniques are not able to deal completely with spatially autocorrelated residuals across those spatial scales.

The remainder of this chapter is organised as follows. Section 5.2 briefly reviews the literature on economic growth, focusing on the problems found in this literature. Section 5.3 discusses the empirical strategy of the study. Section 5.4 describes the dataset and the W matrices used in the analysis. In Section 5.5, the results are reported and discussed. Section 5.6 presents the main conclusions.

5.2. Literature Review

This literature review section aims to situate the discussion of the role of spatial scales in the growth literature as well as the issue of spatial interactions across regions, which can ultimately affect economic growth. First, theoretical and empirical literature is discussed. Then, some econometric issues in the growth literature are listed, and the modifiable areal unit problem (MAUP) and ecological fallacy (EF) are discussed.

The presence of spatially autocorrelated residuals in the non-spatial growth regressions motivates the estimation of the spatially augmented Solow models – as presented in Rey and Montouri (1999), López-Bazo et al. (2004) and Ertur and Koch (2007), for instance – to deal with such spatial autocorrelation. However, it is worth noting that there are alternative explanations for the existence of spatial autocorrelation in the residuals of the growth equations. Nuisance spatial dependence (spatial error) is one potential justification, as explained by Magrini (2004: 2763), it "*may result from measurement problems such as a mismatch between the spatial pattern of the process under study and the boundaries of the observational units.*" It is also likely that regions that are geographically close together may experience random shocks that affect both simultaneously. Another explanation is related to unobserved determinants that are correlated across regions (Finglenton and López-Bazo, 2006). Possible unobserved determinants of economic growth not included in these models include cultural, institutional and technological factors, which might be correlated across spatial units. One possible advantage of panel data framework is that if these unobserved determinants are supposed to be constant over time, the fixed-effect (FE), first-difference-over-time (FD) and system GMM (SYS-GMM) methods might eradicate them, at least partially.

Moreover, Fingleton and López-Bazo (2006) note that substantive dependence (spatial lag and/or spatial cross-regressive) assumes that across-region externalities are due to knowledge diffusion and pecuniary externalities. López-Bazo et al. (2004) discuss in some detail the substantive arguments for spatial dependence across regions. These authors built a spatially augmented growth model based on Mankiw et al. (1992) demonstrating that economic growth and initial productivity in the other regions boost growth in a given region, which is explained by regional spillovers of the diffusion of technology from other regions, caused by investments in physical and human capital. However, López-Bazo et al. (2004) recognise that it is also plausible that these externalities across economies might be caused by pecuniary externalities other than knowledge spillovers – such as those created by a specialised market for labour or output, or forward and backward linkages drawn from trade in intermediate goods – which are related with increasing returns at the firm level, as noted by contributors from the so-called new economic geography (Fujita et al., 1999). In this sense, López-Bazo et al. (2004: 44-45) highlight that:

> "*The role of distance could be more important for pecuniary externalities than for technological diffusion, because neighbors might take advantage of contiguity to a saturated economy; that is, it could be profitable for some suppliers to be located in a neighboring region with a lower degree of agglomeration, and still take advantage of proximity (Puga and Venables, 1996). Contiguous regions may also share markets for labor and final goods, and have similar capital or managerial talent at their disposal. When this is the case, pecuniary externalities could lead to the concentration of firms in macro-areas spanning several regions, thereby transferring externalities at the firm level to the aggregate regional level.*"

The debate about the correct specification of spatial models and especially the choice of the W matrices has increased[89]. In Chapter 2 these alternative spatial models and W matrices are discussed in more detail. In this context, some researchers have formulated some guidelines to specify the correct spatial model and the respective W matrix (Stakhovych and Bijmolt, 2008). Other authors have suggested a weighted spillover variable that directly enters into the model, avoiding the need for a W matrix (Harris and Kravtsova, 2009). Herein, it is important to note that this chapter will not use spatial econometric techniques to identify the correct argument among those discussed above, given the identification problems in the spatial econometric literature discussed in more detail by Gibbons and Overman (2012) and in Chapter 2 of this thesis. Finding an appropriate spatial econometric specification to identify the correct argument(s) of spatially autocorrelated residuals is beyond the scope of this chapter. It is worth noting that the issue of the correct spatial specification is a crucial point because each spatial specification (substantive or nuisance) gives alternative interpretations and policy implications for the process of economic growth (Fingleton and López-Bazo, 2006: 179). A final remark is that it is possible that nuisance and substantive arguments may be operating at the same time in the process of economic growth.

As already discussed in Chapter 2, in mainstream economic theory, the debate about factors that affect long-run economic growth began with Solow’s (1956) growth model. This model, also called the exogenous growth model, has been augmented by the inclusion of variables for educational capital (Mankiw et al., 1992), health capital (Bloom et al., 2004; McDonald and Roberts, 2002), migration (Barro and Sala-i-Martin, 2003), and knowledge spillovers (López-Bazo et al., 2004; Ertur and Koch, 2007; Fischer, 2011). These theoretical models predict conditional β-convergence, which

[89] The use of the W matrix and the respective spatial econometric specification has incited some debate in relation to the choice of the best W matrix to model the phenomenon under study (Stakhovych and Bijmolt, 2008; Folmer and Oud, 2008; Harris and Kravtsova, 2009), the best spatial model (Florax et al., 2003), the most appropriate estimation technique [e.g., maximum likelihood and instrumental variable (IV)/generalised method-of-moments (GMM) procedures] and whether this technique is able to determine causality (McMillen, 2010; Gibbons and Overman, 2012).

means that if countries (or regions) differ in the parameters that determine their steady-state (structural characteristics such as saving rates, human capital and infrastructure), each country (or region) should be converging towards its own steady-state level of per capita income and not towards a common level. After the Solow model, an alternative set of growth theories was developed, the so-called endogenous growth models. For instance, Romer (1986) stresses the externalities of knowledge investment, and Lucas (1988) shows the positive externalities of human capital accumulation. Other examples of the endogenous growth model are Romer (1990), Barro (1990) and Alesina and Rodrik (1994). These models are based on the presence of constant or increasing returns to capital, which breaks down the neoclassical model's prediction of convergence, leading to the conclusion that economies can diverge over time.

The interest of the empirical investigation of the convergence hypothesis is derived from the seminal paper of Baumol (1986), which tests the prediction of convergence based on a simple linear regression model where the per capita income growth rate of 72 countries is regressed on their initial per capita incomes by means of the Ordinary Least Squares (OLS) method. Barro (1991) has followed the same empirical approach, but his study is original because it links cross-country growth regressions to alternative growth theories and determinants[90]. Durluaf et al. (2005) point out that while Baumol (1986), Abramovitz (1986) and many others view convergence as the process of follower countries "catching up" to leader countries by adopting their technologies; Barro (1991), Mankiw et al. (1992), and others adopt the neoclassical model view that convergence is driven by diminishing returns to factors of production. Robinson (1971), Grier and Tullock (1989) and Kormendi and Meguire (1985) employed such regressions earlier, but these studies seem to have received less attention due to their appearance before endogenous growth theory emerged as a

[90] Durlauf et al. (2005: 580) make this point and also note that the standard regressions used in empirical growth research "*are sometimes known as Barro regressions, given Barro's extensive use of such regressions to study alternative growth determinants.*"

primary area of macroeconomic research, which, since then, have attracted interest on the empirical evaluation of growth theories (Durluaf et al., 2005).

After a first wave of cross-country regressions, there was a widespread interest in testing the convergence hypothesis and other growth determinants among regions within countries or groups of countries (e.g., US states and European Union regions). For example, Barro and Sala-i-Martin (1991) study convergence among US states and regions as well as European regions; Armstrong (1995) examines convergence and growth determinants using European Union regions; and Sala-i-Martin (1996a, 1996b) present results for US states, Japanese prefectures, European regions and Canadian provinces. As noted by Islam (1995) most of these studies have been conducted in the framework of single cross-country regression in which it is econometrically difficult to allow for differences in the production function as are not easily measurable. Then, Islam (1995) proposed panel data framework that allows for differences of the aggregate production functions in the form of unobservable individual fixed effects. Moreover, in recent years, the spatially augmented Solow model has been examined using the appropriate spatial statistics and econometric methods due to the recognition that the traditional growth equation may suffer from a misspecification due to omitted spatial dependence (Rey and Montouri, 1999; Fingleton, 1999; López-Bazo et al., 2004; Elhorst et al., 2010). This debate focuses on identifying and testing for factors involved in growth processes and respective spatial interaction effects across regions.

Regarding the literature about Brazil, most papers on economic growth use state level data to run growth regressions (Ferreira and Diniz, 1994; Ferreira, 2000; Azzoni et al., 2000; Azzoni, 2001; Silvera Neto and Azzoni, 2006). Recently, growth regressions have been used to examine economic growth among Brazilian municipalities (Lall and Shalizi, 2003; Laurini et al., 2005; De Vreyer and Spielvogel, 2005; Coelho and Figueiredo, 2007). Cravo (2010) and Cravo and Resende (2012) are the few studies on Brazilian economic growth using micro-regional data. Most of these papers were revised in Section 3.2 (Chapter 3).

Despite much corroboration of conditional β-convergence in the empirical literature, the conditional β-convergence result remains controversial, suffering from substantial drawbacks as shown by, for instance, Friedman (1992) and Quah (1993) who stress that a negative coefficient for initial per capita income can be due to the more general phenomenon of mean reversion, and by reading convergence into this scenario, growth researchers are falling into Galton's fallacy (Islam, 2003). To sum up, the main problems with this growth literature include: (i) identification of β-convergence and economic divergence; (ii) endogeneity; (iii) outliers; (iv) missing data; (v) parameter heterogeneity; (vi) measurement error; (vii) robustness with respect to choice of control variables; (viii) spatial correlation in errors; and the (ix) MAUP. For a comprehensive discussion of topics (i) to (vii), see Durlauf et al. (2005). The problem of spatial correlation in errors (viii) is discussed in the next section. For this reason, herein, I will focus on MAUP and EF only. In addition, a more detail discussion on MAUP and EF is provided in Section 2.2.3 (Chapter 2).

According to Rey and Janikas (2005), while a number of studies have examined the robustness of growth regression with respect to choice of control variables (Levine and Renelt, 1992; Sala-i-Martin, 1997; Sala-i-Martin et al., 2004), changes in spatial scale have yet to be incorporated into this important line of research. Magrini (2004) points out that other authors call for greater attention to the issue of what spatial scale is most appropriate for regional analysis (Cheshire and Hay, 1989; Cheshire and Carbonaro, 1995; Cheshire and Magrini, 2000). Recently, Resende (2011) analyses Brazilian economic growth on four spatial scales (municipalities, micro-regions, spatial clusters and states) between the years of 1991 and 2000. Growth equations were systematically estimated – using the same time period and explanatory variables – across those spatial scales to demonstrate that the estimated coefficients change with the scale level. Moreover, robustness tests identified variables that are simultaneously significant on different spatial scales – higher educational and health capital and better local infrastructure were related to higher rates of economic growth – although their impacts on growth may differ across spatial scales (Resende, 2011).

This strategy returns to the exploration of the statistical literature, which proposes two approaches to analysing this measurement issue: the MAUP and the EF. MAUP refers to the variability in statistical results endemic to the selection of different area units (Openshaw and Taylor, 1979). EF appears when parameters estimated from macro-level data are used to make inferences about behavioral and socio-economic relations at a more disaggregate level (individual/micro-level). Basically, MAUP and EF can be related to the measurement issue because both indicate an aggregation bias or effect. Peeters and Chasco (2006) note that the term EF – typically used in social sciences (e.g., Hannan, 1971) – is similar to the MAUP in geography (e.g., Openshaw and Taylor, 1979, 1981; Openshaw, 1984). Anselin (2002) observes that even in very simple situations, the ecological approach creates problems of interpretation. Glaeser et al. (2003) associated this problem with the social multiplier argument, showing that aggregation may strongly influence coefficient sizes. Basically, the social multiplier is the ratio of aggregate coefficients (estimated at some macro-level) to individual coefficients. In one of the examples given by the authors, they found that the social multiplier of human capital is 2.172, by regressing wages on years of schooling at the state level and at the individual level and then taking the ratio of these two coefficients. This result supports the existence of a social multiplier that rises with the level of aggregation, and it corroborates the idea that there are human capital spillovers (Glaeser et al., 2003).

Finally, it is worth noting that recent empirical literature has been focused on the analysis of MAUP in several areas of urban and regional economics. For instance, Briant et al. (2010) evaluate, in the context of economic geography estimations, the magnitude of the distortions possibly induced by the choice of various French geographic stratifications. Fingleton (2011) also examines agglomeration processes operating at multiple levels of spatial aggregation using the UK and the EU regional data sets. Yamamoto (2008) investigates regional per capita income disparities in the USA on multiple spatial scales between 1955 and 2003, focusing on methods such as inequality indices, kernel density estimation and spatial autocorrelation statistics.

The next section describes the empirical strategy to evaluate how the results of growth equations change with various scale levels in Brazil. Specifically, it focuses on estimating non-spatial dynamic panel data growth models at different spatial scales and then evaluates the variability in the estimated coefficients and the extent of spatial correlation in errors that may be associated with the strength of the spatial interactions across regions.

5.3. Empirical Strategy

Panel data models have been widely used in empirical growth literature (Islam, 1995; Caselli et al., 1996; Lee et al. 1997, 1998)[91]. Indeed, Islam (2003: 324) points out that "*the convergence research gradually moved from the cross-section to the panel approach.*" Islam (1995), Temple (1999), Islam (2003) and Durlauf et al. (2005) present a detailed literature review of this line of inquiry. The main usefulness of using the panel data approach lies in its ability to address the omitted variable bias (OVB) problem often detected in the cross-sectional regressions by controlling for the omitted variables that are constant over time in the form of individual effects[92].

The specifications used in this chapter to study economic growth are the traditional panel data growth regressions, as presented in Durluaf et al. (2005), wherein income per capita growth rates are regressed on conditioning variables and income per capita levels. As discussed in the dataset section, the dependent variable comprises averaged income per capita growth rates over ten-year periods between 1970 and 2000; this implies that the panel data set contains three time-periods only (T=3).

[91] Spatial panel econometrics estimators suggested by Elhorst (2010a, 2010b, 2012) and Lee and Yu (2010a, 2010b, 2010c, 2010d) are not the focus of this chapter, but there has been a growing empirical literature on this topic (See Chapter 2 for a concise literature review and discussion of this topic).

[92] However, panel data models are not without problems, which include a small sample bias and a short frequency at which data are considered. See Islam (2003) for details.

Moreover, the explanatory variables are given in terms of initial values in each decade. As noted by Temple (1999), to mitigate endogeneity problems, researchers often make use of initial values. Four alternative methods to estimate panel data are used at four spatial scales. First, the pooled Ordinary Least Squares (OLS) model is an estimate that assumes that there is not any omitted variable correlated with the included variables. The following dynamic panel data growth model (Equation 5.1) is estimated via the pooled OLS specification:

$$g_{i,t} = \beta y_{i,t} + \psi X_{i,t} + \pi Z_{i,t} + \eta_t + \varepsilon_{i,t} \tag{5.1}$$

where $g_{i,t}$ is equivalent to $(\ln y_{i,t+1} - \ln y_{i,t})/10$ and represents a vector with observations for averaged per capita income growth rates for each spatial unit *i* at each decade *t* (*i* = 1, ..., N; *t* = 1, ..., T)[93]. Moreover, $y_{i,t}$, the initial income per capita, and $X_{i,t}$ represent those growth determinants that are suggested by the Solow growth model. $X_{i,t}$ contains a constant, human capital variable (proxied by averaged years of schooling) and the population growth ($n_{i,t}$+d+g) adjusted for depreciation (d) and technological growth (g), under the usual assumption that d + g equals 0.05. $Z_{i,t}$ represents other growth determinants not included in Solow's theory. η_t is a time-specific effect and $\varepsilon_{i,t}$ is the vector of error terms. As explained by Durlauf et al. (2005: 628) "*the inclusion of time-specific effects is important in the growth context, not least because the means of the log output series will typically increase over time, given productivity growth at the world level.*"

However, as previously stated, a major motivation for using the panel data approach has been the ability to allow for differences in the aggregate production function across countries or regions (Islam, 1995). With this aim, panel data with

[93] More precisely, the denominator of $(\ln y_{i,t+1} - \ln y_{i,t})/10$ is 10 only for the 1970-1980 period. Because per capita income is only available in 1991 (and not in 1990), the denominator for the 1980-1991 period is 11 and for the 1991-2000 period is 9.

individual fixed effects (also known as Least Squares with Dummy Variables, LSDV) is estimated by means the following regression (Equation 5.2):

$$g_{i,t} = \beta y_{i,t} + \psi X_{i,t} + \pi Z_{i,t} + \eta_t + \mu_i + \varepsilon_{i,t} \tag{5.2}$$

In this fixed-effect (FE) formulation, μ_i is included and represents individual-specific effects. This framework allows for unobservable differences in the production function, which is an improvement in relation to the single cross-section regressions. Indeed, Islam (1995) advocates that this panel data framework makes it possible to reconcile neoclassical empirics of growth and development economics because much of the discussion of development economics is thought to have been directed at ways to improve the country (or region) specific aspect of the aggregate production function, which focus attention on all the tangible and intangible factors (e.g., institutional characteristics) that may enter into its respective individual effect.

Another way to deal with unobserved fixed effects that are constant over time is to use the first-differencing (FD) method. Equation (5.3) takes first differences over time to get rid of unobserved fixed effects. The first-differenced equation has the following formulation:

$$\Delta g_{i,t} = \beta \Delta y_{i,t} + \Delta X_{i,t} \psi + \Delta Z_{i,t} \pi + \eta_t + \varepsilon_{i,t} - \varepsilon_{i,t-1} \tag{5.3}$$

where Δ is the first-difference operator. It is worth noting that in Equation (5.3) the $y_{i,t-1}$ component of $\beta \Delta y_{i,t}$ is correlated with the $\varepsilon_{i,t-1}$ component of the new composite error term, which means at least one of the explanatory variables in the first-differenced equation will be correlated with the residuals (Durlauf et al., 2005). In this case, instrumental variable procedures would be required to address the endogeneity problem that emerges in growth regressions with this formulation. This point also applies to dynamic panel data models with fixed effects (Equation 5.2) when the number of time periods (T) is small, as demonstrated by Nickell (1981). The strategy developed by Arellano and Bond (1991) uses the first-differences to eliminate the cross-sectional fixed effects, and then it applies GMM using lagged levels of the series as instruments for lagged first-differences (DIFF-GMM). However, Bond et al. (2001) point out that lagged levels can be weak instruments for first-differences because a variable such as educational attainment may influence income per capita growth with a considerable delay (i.e., the presence of highly persistent series) and the first-differenced GMM estimator is expected to be poorly behaved. To overcome the problems of the standard DIFF-GMM estimator, Blundell and Bond (1998) proposed the system GMM estimator (SYS-GMM), which can substantially reduce biases, and thus, more precise parameter estimates can be obtained. The SYS-GMM estimator uses not only lagged levels as instruments for first-differences, but also, lagged first-differences are used as instruments for levels. As noted by Durlauf et al. (2005: 634) "*[t]his builds in some insurance against weak identification, because if the series are persistent and lagged levels are weak instruments for first differences, it may still be the case that lagged first differences have some explanatory power for levels.*" However, the SYS-GMM approach is not without problems, as discussed in Roodman (2009)[94]. The results using the pooled OLS, FE, FD (without instruments) and SYS-GMM are presented in Section 5.5. The motivation to use four panel data methods to evaluate the extent of spatial autocorrelation across different spatial scales is derived from the fact that alternative methods may differently deal with spatial autocorrelation, and this fact might be of interest to place a set of results into a meaningful spatial perspective.

The final step of the empirical strategy is to measure the extent of spatial autocorrelation by means of the analysis of the variability of estimated coefficients and spatial correlation in errors across multiple spatial scales. As explained earlier, the estimated coefficients may carry information on the strength of the spatial interactions across individuals and regions, a phenomenon that was called by Glaeser et al. (2003)

[94] For instance, SYS-GMM easily generates instruments that are numerous. Roodman (2009: 139) points out that "*[s]imply by being numerous, instruments can overfit instrumented variables, failing to expunge their endogenous components and biasing coefficient estimates towards those from non-instrumenting estimators*."

the "social multiplier" effect. The drawback in the application of this approach to the empirical growth literature is the reliance of such literature on aggregate data to conduct the estimations. However, it is still possible to give some evidence of this social multiplier effect, or simply, aggregation effect, if we compare the minimum comparable area (MCA) level – which should be assumed to be the micro level of analysis (instead of the individual) – with the estimations using another aggregate level (e.g., micro-regional, meso-regional or state level). These results will be examined in Section 5.5.

Moreover, the primary focus of this chapter is to investigate the spatial autocorrelation in the residuals. It is expected that part of the spatial autocorrelation remains as a residual of the regressions in at least one or more spatial scales. The failure to account for spatial autocorrelation in the error term in economic growth regressions yields unbiased estimates for the estimated parameters but a biased estimate of the parameters' variance. Furthermore, ignoring substantive spatial dependence (e.g., a spatial lag model) will produce biased estimates of the coefficients. Indeed, this is the motivation of all spatial econometric literature, including the estimation of spatially augmented Solow models by means of spatial econometric techniques. To analyse the spatial autocorrelation in the residuals of the traditional panel data regressions, the global Moran's I statistic is calculated for the regression's residuals using different spatial weight (W) matrices (distance-based matrices), which can capture the degree to which spatial interdependencies become less important with geographical distance.

Global spatial autocorrelation is calculated based on Moran's I statistic (Cliff and Ord, 1981). To evaluate the spatial autocorrelation in each time span of the panel (1970s, 1980s and 1990s), Moran's I is calculated for cross-sectional errors generated by the panel data estimations. This approach provides information on the spatial autocorrelation effects across three decades. Furthermore, the residuals of the panel data estimations were time-averaged, and then, Moran's I statistics were also calculated. Moran's I statistic is written in the following form[95]:

$$I = \frac{\sum_i \sum_j w_{ij}(\varepsilon_i - \bar{\varepsilon})(\varepsilon_j - \bar{\varepsilon})}{\sum_i(\varepsilon_i - \bar{\varepsilon})^2}, \tag{4.4}$$

where w_{ij} are elements of a spatial weighting matrix that is row-standardised such that the elements w_{ij} in each row sum to 1. ε_i and ε_j are the values of the cross-sectional errors, $\bar{\varepsilon}$ is the mean of the errors and $\sum_i(\varepsilon_i - \bar{\varepsilon})^2$ is the variance normalisation factor. The spatial weighting matrices employed in this analysis are discussed in the next section. If I ≈ 0, then there is no evidence of spatial autocorrelation in the residuals, i.e., residuals tend to move independently. If Moran's I statistic is greater than zero, there is a positive spatial autocorrelation, i.e., areas with high residual values tend to be close to areas with high residual values (and vice-versa). Finally, if Moran's I statistic is smaller than zero, there is a negative autocorrelation, i.e., places with high residual values are close to neighboring places with low residual values, and vice versa. The statistical significance of Moran's I is calculated using the permutation approach (Anselin, 1995). The next section describes the data set including the spatial scales and the spatial weight matrices used in the analysis.

5.4. Dataset and Spatial Weight Matrices

To investigate the extent of spatial autocorrelation on different scale levels in the context of growth regression estimates, the chapter makes use of four Brazilian geographic stratifications: 27 states, 134 meso-regions, 522 micro-regions and 3,657 minimum comparable areas (MCAs)[96]. As explained earlier in this thesis, it was

[95] In the matrix form, Moran's I using a row-standardised W matrix is $I = e'We/e'e$, where e are cross-sectional residuals (Anselin and Rey, 1991).
[96] The total number of MCAs is 3,659, but this chapter uses 3,657. Fernando de Noronha (in the state of Pernambuco) and Ilhabela (in the state of São Paulo) were excluded because they are islands and do not adjust to the spatial weight matrices used in the analyses. These exclusions do not alter the results of the chapter.

necessary to make some adjustments in the data because the number of municipalities increased from 3,920 municipalities in 1970 to 5,507 municipalities in 2000. To address this problem, municipalities were merged into 3,657 MCAs – defined by Reis et al. (2005) as sets of municipalities whose borders were constant from 1970 to 2000. All data had then been aggregated to match these MCAs, which are the most disaggregated spatial units in this study. The data was drawn from the MCA level and then grouped to form other spatial scales.

The first part of Table 5.1 shows the four spatial scales and some statistics concerning their sizes (in square kilometres). Brazil is divided into 27 states[97] that are the main political-administrative units in the country. Municipalities (MCAs in the case of this chapter) represent the smallest administrative level, dealing with local policy implementation and management. Micro- and meso-regions are homogeneous regions defined by IBGE (Brazilian Institute of Geography and Statistics – Instituto Brasileiro de Geografia e Estatística) as being a group of contiguous municipalities within the same state. The micro-regions were grouped according to natural and production characteristics. Meso-regions are lager areas than the micro-regions and were also proposed by IBGE. This spatial scale is based on the following dimensions: the social aspects, the natural setting, and the communication and place network as an element of space articulation.

The information in the panel data was collected from IPEADATA (Institute of Applied Economic Research – Instituto de Pesquisa Econômica Aplicada), which has organized the population census information (from IBGE) of 1970, 1980, 1991 and 2000. From these data points, the dependent variable is the averaged annual income per capita growth rates for each time span, that is, the panel data set contains three time periods: 1970-1980, 1980-1991 and 1991-2000. Per capita income information is deflated to Real (R$, the Brazilian currency) in 2000. The income per capita growth rates are averaged over ten years because MCA data is only available from the

[97] More precisely, there are 26 states and one federal district.

Brazilian population censuses conducted every ten years. Furthermore, given the presence of business cycle effects, the choice of ten-year growth averages seems to be a reasonable approach to avoid those influences (Caselli et al.,1996). For instance, the 1973 and 1979 "oil price shocks" affected the Brazilian economy. In 1994, Brazil launched the 'Plano Real' (Real Plan), the stabilization program that ended a long period of high inflation rates that had started in the 1970s. However, the chosen periods may be influencing the results, and specific problems might emerge when business cycles are not symmetric across space. Explanatory variables are given in terms of initial values, that is, values in 1970, 1980 and 1991. The socioeconomic data are logged per capita income, logged average years of schooling, logged population density and population growth[98]. Logged transportation costs between MCAs and São Paulo city are from IPEADATA. These transportation cost data are for the years 1968, 1980 and 1995. I estimated this variable for 1970 and 1991 via interpolation. The cost of transportation to São Paulo is calculated through a linear program procedure as the minimum cost (given road and vehicle conditions) of traveling between a MCA's major headquarters and São Paulo. Finally, the econometric specifications include time dummies for the decades of 1980 and 1990 (the time dummy for the 1970 decade was excluded from the regressions to avoid perfect multicollinearity). Tables II.A.1 to II.A.4 (in Appendix II.A) present the correlation matrix of the explanatory variables. Table II.A.5 (in Appendix II.A) shows the summary statistics of the variables in the panel.

[98] Population growth is adjusted for depreciation (δ) and technological growth (g), under the usual assumption that $\delta + g$ equals 0.05 (e.g., Mankiw et al., 1992). I do not take the natural log of this variable because it has some negative values.

Table 5.1

Spatial Scales in Brazil (Areas and Distances to Ensure Connectivity)

Spatial Scales	States (n = 27)	Meso-regions (n = 134)	Micro-regions (n = 522)	MCAs* (n = 3,657)
				
	Area (in kilometres^2)			
Mean	312,994.0	63,065.9	16,189.3	2,310.9
Minimum	5,771.0	2,937.0	190.0	8.0
Maximum	1,558,987.0	650,338.0	439,498.0	367,284.0
Std. dev.	372,070.2	103,804.4	42,082.8	14,157.4
	Minimum distance between centroids to ensure connectivity for all regions (in kilometres)**			
	786.5	571.5	520.9	401.5
Cut-off	**Average number of neighbouring regions using alternative W matrices**			
Queen	3.8	5.1	5.6	5.9
400 km	-	-	-	520.4
500 km	-	-	84.3	707.4
600 km	-	21.7	101.8	894.5
700 km	-	26.8	123.6	1,080.9
800 km	4.4	31.9	145.6	1,267.3
900 km	5.0	37.3	167.1	1,444.2
1,000 km	6.2	42.6	188.5	1,611.6
1,500 km	10.4	67.1	283.7	2,274.7
2,000 km	15.7	90.9	377.6	2,916.5

Note: Own elaboration from data of IBGE. * Minimum Comparable Areas (MCAs). ** The centroids were calculated using the GeoDa software. In the menu option of GeoDa, centroids are central points. Central points are the average of the x and y coordinates of a polygon's vertices.

A spatial weight (W) matrix is used to model spatial relationships between regions to calculate the Moran's I statistic discussed in the previous section. I consider pure geographical weights, which are exogenous, to mitigate endogeneity problems. The W matrices used herein are based on geographical distance (distance between centroids) with the same fixed-distance critical cut-off for all regions. The $n \times n$ standardised W matrix provides the 'structure' of spatial relationships by defining neighbouring areas that should be connected. In this chapter, given the uneven size of the spatial units, I prefer to use distance-based W matrices. Furthermore, the results using distance-based W matrices are comparable across spatial scales, and these W matrices allow for measuring the extent of the spatial autocorrelation (in kilometres) across those spatial scales. Specifically, the element w_{ij} in the matrix is 1 if areas i and j are within "d" kilometres, and 0 otherwise. Moreover, by convention, the diagonal elements w_{ii} = 0. Table 5.1 shows the minimum distance "d" between centroids to ensure connectivity for all spatial units in each spatial scale. For instance, at the MCA level, the minimum distance between centroids to ensure connectivity for all MCAs is 400 kilometres. On the other hand, a cut-off distance of 786.5 kilometres is needed to ensure connectivity for all states. Given these minimum cut-off distances, I have chosen the following cut-offs (in kilometres) to conduct the analysis of the extent of spatial autocorrelation: 400 (only MCAs), 500 (only MCAs and micro-regions[99]), 600 (MCAs, micro- and meso-regions), 700 (MCAs, micro- and meso-regions), 800 (all scales), 900 (all), 1,000 (all), 1,500 (all) and 2,000 (all). This criterion avoids a situation in which rows and columns in W have only zero values[100]. The bottom part of Table 5.1 shows the average number of neighbouring regions using alternative distance-based matrices. In the next section, I show the results using these spatial weight matrices based on different cut-off distances. In addition, the standardised first-order contiguity matrix (also called the queen contiguity matrix) is used for comparative purposes, in which the element w_{ij} in the matrix is 1 if areas i and j share borders or vertices, and 0 otherwise.

5.5. Results

This section aims to report and to discuss the results of growth regressions at four spatial scales (MCAs, micro-regions, meso-regions and states), applying four alternative panel data methods [pooled OLS, fixed-effects (FE), first-difference (FD) and SYS-GMM]. The empirical strategy was to include all available data in the models

[99] Actually, for the micro-regional level, the initial cut-off is 520 kilometers, which is the minimum distance between centroids to ensure connectivity for all micro-regions.
[100] LeGallo and Ertur (2003) note that if unconnected observations are found, they are implicitly eliminated from the computed global Moran's I statistic, but this leads to a change in the sample size and thus must be explicitly accounted for in statistical inference.

to try to control for factors that drive economic growth and also may explain spatially autocorrelated economic growth. Therefore, the diagnostics for spatial autocorrelation in the residuals of these growth equations represents such spatial correlations that are left unexplained after controlling for some observable determinants of economic growth in Brazil. It is important to note that when conditioning variables were dropped from the models and only per capita income growth was regressed on initial per capita income (the absolute β-convergence case), the values for spatial autocorrelation in the residuals for all spatial scales and methods increased as expected[101]. For instance, Silvera Neto and Azzoni (2006) found that after conditioning their models on variables with strong geographic patterns across states in Brazil, spatial dependence disappeared. These authors suggest that statistically significant explanatory variables reveal the potential channels through which spatial dependence in the process of income convergence may occur.

The baseline specification (Equation 5.1) is estimated via pooled OLS for the four spatial scales under analysis. Spatial dependence was assessed by applying the Moran's I statistics in the error terms. Table 5.2 shows the results for these estimations in columns (1), (5), (9) and (13). This specification includes all the available explanatory variables and time dummies for the decades of 1980 and 1990 that control for time-specific effects. Of note, high R-squared values can be observed in all estimations. For instance, the R-squared in column 1 (pooled OLS at the AMC level) is 0.805. However, if the time dummies are dropped from the regression the R-squared goes to 0.4520 (not shown in Table 5.2). It means that time dynamics has a relevant explanatory power in the Brazilian case. This fact is observed for all estimation techniques and spatial scales. Moreover, unobserved heterogeneity between areas might be an important issue in the current analysis. To take into account this aspect, three panel

[101] For the absolute β-convergence case, estimations at all spatial scales suffer from higher spatial autocorrelation than the conditional β-convergence ones; because the Moran's I statistics in the residuals of former estimations present the highest values. For instance, Silvera Neto and Azzoni (2006) also demonstrate this evidence.

data methods are used to control for unobserved heterogeneity between areas that may be helpful in dealing – at least partially – with unobservable factors that can be correlated across neighbouring areas. The growth regression specifications represented by Equations (5.2) and (5.3) are estimated, which represent, respectively, the fixed-effect (FE) and first-difference (FD) methods. Table 5.2 presents the results for the FE estimations in columns (2), (6), (10) and (14) and for the FD estimations in columns (3), (7), (11) and (15). In addition, the SYS-GMM estimations aim to alleviate biases due to endogenous explanatory variables, and their results are in columns (4), (8), (12) and (16). However, the SYS-GMM results should be viewed with caution because the Sargan/Hansen tests are rejected under the null hypothesis that instruments are valid – for regressions at the four spatial scales – suggesting that the instruments of the GMM-SYS are not valid. This indicates that the use of the GMM-SYS might be adding more endogeneity to the system. Unfortunately, due to data unavailability, this empirical exercise uses only three time periods (T=3), a fact that does not allow for using lags of earlier periods as instruments that might be more exogenous. In this sense, Roodman (2009: 144) points out that "*[w]here system GMM offers the most hope, it may offer the least help*." Nevertheless, the SYS-GMM results are shown for comparative purposes.

First, it is important to examine the magnitude of the estimated coefficients at multiple levels of spatial aggregation. The coefficients of initial income per capita are negative and statistically significant in all estimations except for the SYS-GMM estimation at the state level (column 16 in Table 5.2), where the coefficient is not statistically significant. This negative correlation between the growth rate and the initial per capita income suggest mean reversion, conditional β-convergence, or both. Of note, given the estimation method, the coefficients of the initial income per capita are relatively similar at different spatial scales. For instance, using the pooled OLS method, the coefficient is -0.0290 (s.d. = 0.0007) at the MCA level and -0.0305 (s.d. = 0.0016) at the micro-regional level. On the other hand, the coefficient of the initial income per capita, using the FE method, is -0.1051 (s.d. = 0.0010) at the MCA level and -0.0851 (s.d. = 0.0028) at the micro-regional level. Therefore, it seems that distortions of conditional β-convergence coefficients due to panel data method choices are much larger than variations due to the spatial scale of analysis. Indeed, Islam (1995) provides a statistical explanation for the faster convergence rate in the FE framework compared to the pooled OLS approach[102]. He shows that in the framework of single cross-section regression (or even, pooled OLS regressions), the technology variable, A(0), being unobservable or unmeasurable, is left out of the equation (or, incorporated in the error term):

> "*[t]his actually creates an omitted variable problem. Since this omitted variable is correlated with the included explanatory variables, it causes the estimates of the coefficients of these variables to be biased. The direction of bias can be assessed from the standard formula for omitted variable bias. The partial correlation between A(0) and the initial value of y (income per capita) is likely to be positive, and the expected sign of the A(0) term in the full regression, (...), is also positive. Thus, the estimated coefficient of $y_{i,t-1}$, is biased upward. (...) This explains why we get lower convergence rates from single cross-section regressions and pooled regressions that ignore correlated individual country effects*" Islam (1995: 1147).

Here, we can use the same intuition provided by Islam (1995) to present the results, which indicate that persistent differences in technology level and, for instance, institutions are an important factor in understanding economic growth across regions; because when these variables are included in the regressions in the form of fixed effects, the convergence process occurs at a faster rate at all spatial scales. Then, he points out "*[c]ontrary to what may appear at first sight, the finding of a higher rate of conditional convergence actually calls for more policy activism*" (Islam, 1995: 1128). Improvements in these unobserved factors (e.g, technology levels and institutions) may have direct positive effects on the region's long-run income level, including a higher transitional growth rate. Abreu et al. (2005) also argue that as regional-fixed effects control unobserved heterogeneity, the results that include fixed effects lead to higher estimates of the rate of convergence. Of note, when the spatial distribution of these

[102] See also the seminal paper of Nickell (1981).

fixed effect terms is analysed, it is possible to observe a clustering of high values in the south, southeast and central-west of Brazil at the MCA spatial scale, for instance (see Figure II.A.1, in Appendix II.A). This fact suggests that fixed effects are really capturing a higher level of, for example, technology and institutions in the south, southeast and central-west which are the most developed areas in Brazil. However, this spatial distribution of the fixed effects is not able to mitigate the spatial autocorrelation that is presented in the errors of the regression as discussed below.

Furthermore, it is worth noting that the club convergence hypothesis cannot be ruled out given the growing evidence that this hypothesis is the correct one for the Brazilian case (Andrade et al., 2004; Laurini et al. 2005; Coelho and Figueirêdo, 2007). Interestingly, as noted by Islam (1995: 1149) "*instead of adopting the panel data approach, the other way to control for differences in technology and institutions is to classify the countries into similar groups*". Exactly, classifying regions into similar groups (or clubs) was the approach adopted in Chapter 4. It is worth noting the similarities of Figure I.D.1 in Appendix I.D - two spatial regimes at the municipal level in the initial income per capita identified by means of the Moran scatterplot – and Figure II.A.1, in Appendix II.A (the spatial distribution of estimated fixed-effects). In this sense, the panel data findings described above are consistent with the finding of faster convergence among groups of similar regions that have been reported earlier in the results section of Chapter 4, where I have conducted a club convergence analysis. The panel data approach with fixed effects allows for differences in the aggregate production function across individual regions (MCA, micro-region, meso-region or state), and the club convergence analysis allows for differences in the aggregate production function across groups of regions. Both approaches obtain higher rates of convergence when the results are compared to the estimations that do not control for such differences.

If we analyse the variability of other estimated coefficients at different spatial scales, it suggests that the choice of the spatial scale is more important than the panel data method choice. For instance, the average-year-of-schooling coefficient may carry information about the strength of the spatial interactions across individuals and regions, a phenomenon coined the "social multiplier" effect by Glaeser et al. (2003). The strategy followed by Glaeser et al. (2003) returns to the exploration of the ecological fallacy literature, which demonstrates that the estimated coefficients in aggregate models represent a blend of individual and contextual effects (Manski, 1993; Anselin, 2002). Herein, there are not estimations at the individual level, yet it is possible to give some evidence of this social multiplier effect if we compare the MCA level – which should be assumed to be the micro level of analysis (instead of the individual) – with the estimations using another aggregate level (e.g. micro-regional, meso-regional or state level). For instance, the years-of-schooling coefficient at the MCA level is 0.0091 (s.d. = 0.0004) and at the state level, 0.0457 (s.d. = 0.0095) using the pooled OLS method[103]. This result is in line with the social multiplier argument, suggesting that there are human capital spillovers.

The coefficients of population growth are statistically not significant for all estimations, excepting the SYS-GMM estimations at the MCA and micro-regional levels. As noted by Barro and Sala-i-Martin (2003: 26), "*the growth of population reflects the behavior of fertility, mortality, and migration*". Population growth effects on economic growth may present different results across different spatial scales, because migration pattern – which is one component of population growth – may vary across different scale levels (e.g., intra- versus inter-regional migration). For instance, the contrast in area sizes means that daytime commuting across municipalities can be more significant if compared to states. Moreover, if we are able to analyse only the migration effects, we need to bear in mind that, unlike newly born persons, migrants come with accumulated human capital; and for this reason, the results depend on

[103] Results based on standardised coefficients provide similar qualitative findings. To move from the metric to the standardised coefficients, the following formula should be applied: $\theta_k \text{standardized} = \theta_k \text{unstandardized} \square (Sx_k / Sg)$, where θ_k represents the coefficients of the explanatory variable k, Sx_k is the standard deviation of the explanatory variable K and Sg is the standard deviation of the dependent variable. Table II.A.5 (in Appendix II.A) provides the standard deviations of all variables.

whether immigrants have more or less human capital (i.e., are typically skilled or unskilled) than the residents of the receiving region (see Chapter 9 in Barro and Sala-i-Martin, 2003). It is probably because of a balance between countervailing effects, the population growth coefficient may become statistically insignificant. Other variables under analysis are population density and transportation cost to São Paulo. As suggested in Chapter 2, population density coefficients might present different results for the different spatial scales being analysed because the strength of agglomeration effects might vary with the size of the spatial unit (e.g., agglomeration-related centripetal forces may be much more relevant at the municipal than at the state level; in addition, analysing the influence of reductions in transportation costs on economic growth at multiple scale levels allows us to distinguish, for instance, this influence within the borders of a state from that occurring between states. Population density coefficients are negative and statistically significant at the 1% level on the MCA, micro- and meso-regional spatial scales; however, the coefficients are no longer significant at the 5% level on the state-level stratification. These results are contrary to the argument that agglomeration effects foster economic growth because the negative signs of the population density coefficients mean that higher populated areas are harmful to economic growth demonstrating somehow that congestion effects might be operating for the analysed period (1970-2000) For the pooled OLS results, the transportation cost coefficients are statistically significant (and negative) at all spatial scales. This finding indicates that reductions in transportation costs over the period of 1970 to 2000 have a positive impact on economic growth at all Brazilian scales. The analysis of FE and FD results reveal that reductions in transportation costs may have a negative impact on economic growth. However, these latter results might be imprecise because these estimates control for fixed effects when transportation costs already have a clear component that is fixed, the distance between each spatial unit and the spatial unit represented by São Paulo. Therefore, the FE and FD estimates for the transportation cost coefficients might only be picking up the variable part of transportation costs (for instance, roads conditions or quality). Furthermore, the SYS-GMM results show the significance of the coefficients for the transportation costs at the MCA and micro-regional levels, suggesting that reductions in transportation costs had a positive impact on growth only within the borders of the meso-regions and states between 1970 and 2000, a fact that has already been documented by Resende (2011) for the 1990s in Brazil.

Finally, the bottom part of Table 5.2 shows the spatial autocorrelation diagnostics using a first-order contiguity matrix (the queen contiguity matrix) for all spatial scales and methods. The test is based on Moran’s I statistic applied to cross-sectional residuals generated by the panel data estimations. This approach allows for measuring the strength of spatial autocorrelation in the residuals across three decades (1970s, 1980s and 1990s). Moreover, Moran's I statistics in the time-averaged residuals are calculated to provide an idea of the average pattern of spatial autocorrelation over the whole period of analysis. For fixed-effects (FE) estimations, the time-averaged residuals are zero by construction, and therefore, Moran’s I statistics are not calculated. A preliminary spatial analysis of the residuals using the most common contiguity matrix (i.e., the queen matrix) shows that spatial phenomena might be relevant for the study of Brazilian regional growth determinants, depending on the spatial scale of analysis. Indeed, growth estimations at the state level using four econometric methods show weak evidence of spatial autocorrelation across the error terms. In sum, Moran’s I statistics are not statistically significant in cross-sectional and time-averaged residuals at the state level. Only in the residuals from the period of the 1990s is Moran’s I statistically significant at the 5% level (in the pooled OLS, FE and SYS-GMM specifications). On the other hand, spatial autocorrelation in the residuals has been detected at three other spatial scales (MCAs, micro-regions and meso-regions) using alternative econometric methods. These results suggest that standard panel data methods (pooled OLS, FE, FD and SYS-GMM) do not have the ability to deal completely, with spatial phenomena in the Brazilian regional growth case.

However, this preliminary conclusion can be subjected to the choice of the queen W matrix. For this reason, subsequently, a sensitivity analysis is conducted using alternative distance-based W matrices. Among other advantages, this kind of W matrix can indicate the degree to which spatial autocorrelation behaves with geographical distance.

Of note, Baltagi and Pirotte (2010) pointed out that tests of hypothesis based on the standard panel data estimators that ignore spatial dependence, can lead to misleading inference. Specifically, these authors investigate the standard panel data estimators under spatial dependence using Monte Carlo experiments. This fact highlights that the coefficients analysed above should be interpreted with caution. The analyses of the residuals of the regressions conducted in this Chapter investigate the presence and extent of the spatial autocorrelation in errors of the standard panel data models across multiple spatial scales. This descriptive analysis is a first step in implementing a framework to measure and interpret the presence of spatial spillovers using spatial data models across multiple spatial scales. For instance, for some spatial scales there may be some spatial process that needs the use of spatial econometrics, for other scales the use of spatial econometrics might not be necessary because the spatial autocorrelation does not appear in the residuals.

Table 5.2 – Panel Data Model Results and Diagnostics for Spatial Autocorrelation

Spatial scale	Minimum Comparable Area (MCA)				Micro-region				Meso-region				State			
Method	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Explanatory variables																
ln (income per capita)	-0.0290***	-0.1051***	-0.1250***	-0.0681***	-0.0305***	-0.0851***	-0.1084***	-0.0348***	-0.0346***	-0.0779***	-0.1072***	-0.0327***	-0.0487***	-0.0931***	-0.1135***	-0.0480
	(0.0007)	(0.0010)	(0.0010)	(0.0101)	(0.0016)	(0.0028)	(0.0027)	(0.0078)	(0.0032)	(0.0057)	(0.0057)	(0.0126)	(0.0084)	(0.0120)	(0.0127)	(0.0300)
population growth (n+g+d)	-0.0164	0.0244	0.0282	0.4190***	-0.0095	-0.0516	0.0104	0.2770***	0.0203	-0.1475	-0.0203	0.2643	-0.0100	-0.1947	0.00746	0.4272
	(0.0112)	(0.0180)	(0.0180)	(0.1099)	(0.0211)	(0.0394)	(0.0366)	(0.0990)	(0.0432)	(0.0918)	(0.0897)	(0.1635)	(0.0893)	(0.1471)	(0.1548)	(0.2833)
ln (years of schooling)	0.0091***	0.0009**	0.0015***	0.0122***	0.0263***	0.0039	0.0050	-0.0020	0.0303***	0.0124**	0.0166***	0.0233*	0.0457***	0.0242*	0.0314**	0.0686**
	(0.0004)	(0.0004)	(0.0005)	(0.0040)	(0.0015)	(0.0028)	(0.0031)	(0.0070)	(0.0032)	(0.0056)	(0.0064)	(0.0119)	(0.0095)	(0.0134)	(0.0159)	(0.0322)
ln(population density)	-0.0006***	-0.0089***	-0.0098***	-0.0066***	-0.0013***	-0.0132***	-0.0121***	-0.0054	-0.0011**	-0.0183***	-0.0155***	0.0094	-0.0021*	-0.0141*	-0.0094	0.0106
	(0.0002)	(0.0013)	(0.0015)	(0.0015)	(0.0003)	(0.0024)	(0.0027)	(0.0040)	(0.0005)	(0.0047)	(0.0053)	(0.0063)	(0.0011)	(0.0084)	(0.0098)	(0.0089)
ln (transportation cost to SP)	-0.0129***	0.0056*	0.0031	-0.0307***	-0.0067***	0.0106**	0.0151***	-0.0281***	-0.0071***	0.0120	0.0173**	0.0015	-0.0101***	0.0468***	0.0476***	0.0096
	(0.0004)	(0.0028)	(0.0028)	(0.0059)	(0.0006)	(0.0050)	(0.0046)	(0.0068)	(0.0011)	(0.0084)	(0.0077)	(0.0109)	(0.0028)	(0.0161)	(0.0152)	(0.0167)
constant	0.3026***	0.4060***	0.0076***	0.5623***	0.2552***	0.3722***	-0.0022	0.4332***	0.2708***	0.3596***	-0.0043	0.1526	0.3463***	0.1260	0.0097	0.1064
	(0.0052)	(0.0226)	(0.0014)	(0.0750)	(0.0091)	(0.0418)	(0.0030)	(0.0778)	(0.0167)	(0.0729)	(0.0058)	(0.1304)	(0.0443)	(0.1375)	(0.0122)	(0.2439)
Time-dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	10,971	10,971	7,314	10,971	1,566	1,566	1,044	1,566	402	402	268	402	81	81	54	81
R-squared	0.805	0.923	0.954	-	0.910	0.960	0.978	-	0.935	0.969	0.983	-	0.940	0.977	0.987	-
Adjusted R-squared	0.805	0.884	0.954	-	0.909	0.940	0.978	-	0.934	0.953	0.983	-	0.934	0.961	0.985	-
Diagnostics for spatial autocorrelation:																
Moran's I in the residuals (70's)	0.2737***	0.3411***	-	0.4541***	0.3640***	0.3558***	-	0.4807***	0.2630***	01201**	-	0.4354***	0.1439	-0.1471	-	-0.0366
Moran's I in the residuals (80's)	0.2704***	0.1555***	0.2816***	0.4001***	0.3665***	0.3388***	0.3330***	0.3668***	0.1436***	0.2358***	0.0847*	0.5517***	0.0018	0.0570	-0.1174	0.1483
Moran's I in the residuals (90's)	0.2426***	0.3189***	0.2656***	0.3992***	0.3978***	0.4231***	0.4358***	0.4385***	0.4778***	0.4166***	0.4964***	0.5328***	0.3306**	0.2397**	0.2016*	0.3197**
Moran's I in the averaged residuals	0.4529***	++	0.4029***	0.5357***	0.4490***	++	0.4459***	0.4610***	0.3380***	++	0.2911***	0.5996***	0.0586	++	0.1000	0.1210

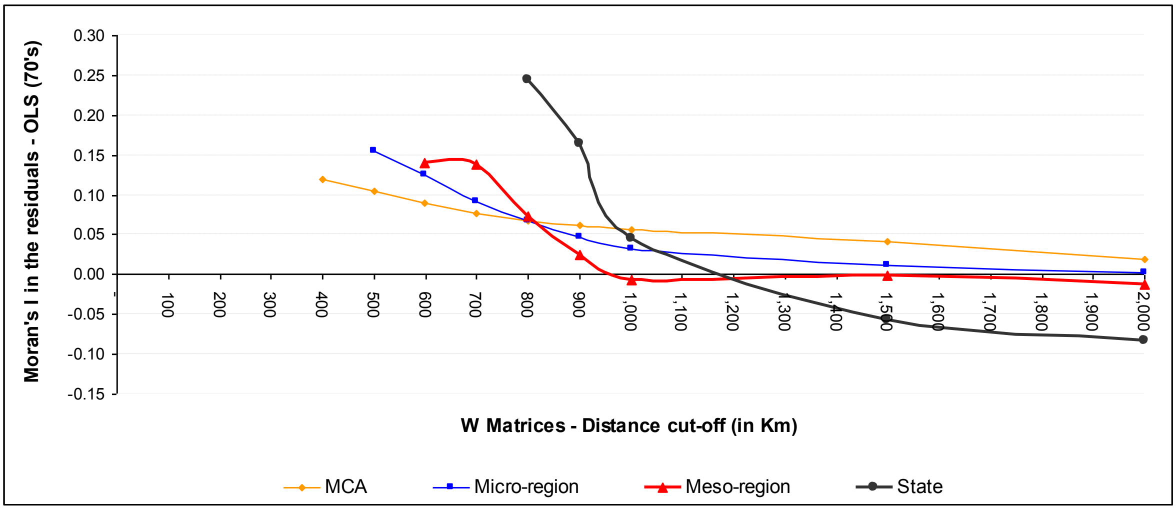
Note: Standard errors in parentheses; *** significant at 1%; ** significant at 5%; * significant at 10%. Moran's I significance test based on the permutation approach with ten thousand permutations and the spatial weight matrix is the queen contiguity matrix. ++ For fixed-effects (FE) estimations, the time-averaged residuals are zero by construction, therefore Moran's I statistics are not calculated. All SYS-GMM results are from the one-step estimates and include 11 instruments. In all SYS-GMM estimations, all growth determinants are treated as potentially endogenous.

Table II.B.1 (in Appendix II.B) reports the evolution of the spatial autocorrelation in the cross-sectional residuals derived from the panel data estimations at multiple levels of spatial aggregation over the period 1970–2000 using the Moran's I statistic. Three decades are analysed separately: the 1970s, 1980s and 1990s. Furthermore, Table II.B.2 (in Appendix II.B) documents the assessment of spatial autocorrelation in the time-averaged residuals. As explained earlier, Moran's I is a global measure of spatial correlation, which evaluates the degree of similarity or dissimilarity among values in spatially close areas. Along with the test statistics presented in columns (1) to (16), Tables II.B.1 and II.B.2 provide the associated significance level (*** significant at 1%; ** significant at 5%; * significant at 10%) based on the permutation approach with 10,000 permutations. The results for alternative weight matrices based on different cut-off distances of 400, 500, 600, 700, 800, 900, 1,000, 1,500 and 2,000 kilometres are listed.

For a better illustration of the results in Table II.B.1, Figures 5.1, 5.2 and 5.3 present Moran's I statistics in the residuals of pooled OLS estimations on the y-axis and the cut-off distances in the x-axis, respectively, for the 1970s, 1980s and 1990s. These figures have the following interesting features. First, irrespective of the spatial weight matrix used, there is evidence that the residuals of growth estimations have become more clustered over time, except at the MCA level. Indeed, spatial autocorrelation in the 1990s is at the highest level at the micro-regional, meso-regional and state levels. The positive signs of Moran's I indicate that the error terms are becoming more and more similar among neighbouring spatial units. In other words, a positive spatial autocorrelation demonstrates that areas with relatively high (low) residual values – which may be explained by unobservable variables – are located close to other areas with relatively high (low) residual values more often than it would be observed if their locations were purely random. This similar pattern can be observed when FE and FD methods are used (see Table II.B.1).

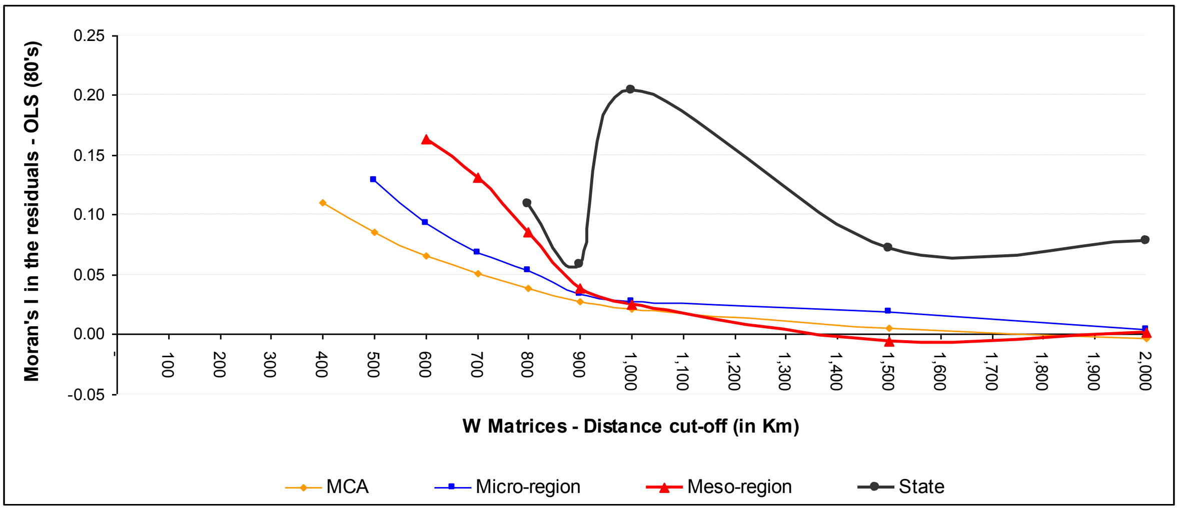
Secondly, focusing on the analysis of the spatial autocorrelation in each decade, the results show evidence that the magnitude of interactions varies at different spatial scales and declines with distance. Indeed, at the MCA and the micro-regional levels, the spatial autocorrelation in residuals shows a decay trend, and it is statistically significant across distances of more than 1,500 kilometres, but the spatial autocorrelation is largely reduced when the distance is more than 900 kilometres, in particular for the 1970s and 1980s. At the meso-regional level, the values of the Moran's I statistics suggests that the spatial linkages decrease steadily over distances up to 900-1,000 kilometres, beyond which the null hypothesis of no spatial correlation cannot be rejected (at least for the 1970s and 1980s). Interestingly, at the state level, where the minimum distance between centroids to ensure connectivity for all states is 800 kilometres, there is evidence of no statistical significance of Moran's I for any choice of the spatial weight matrix. The diagnostics for spatial autocorrelation by means of Moran's I in the time-averaged residuals presented in Table II.B.2 (in Appendix II.B) confirms this observation. Therefore, the analysis of the residuals at multiple scale levels suggests that spatial autocorrelation may be bounded within each state.

Figure 5.1 – Moran's I in the Residuals of the OLS Regression (1970s Period)



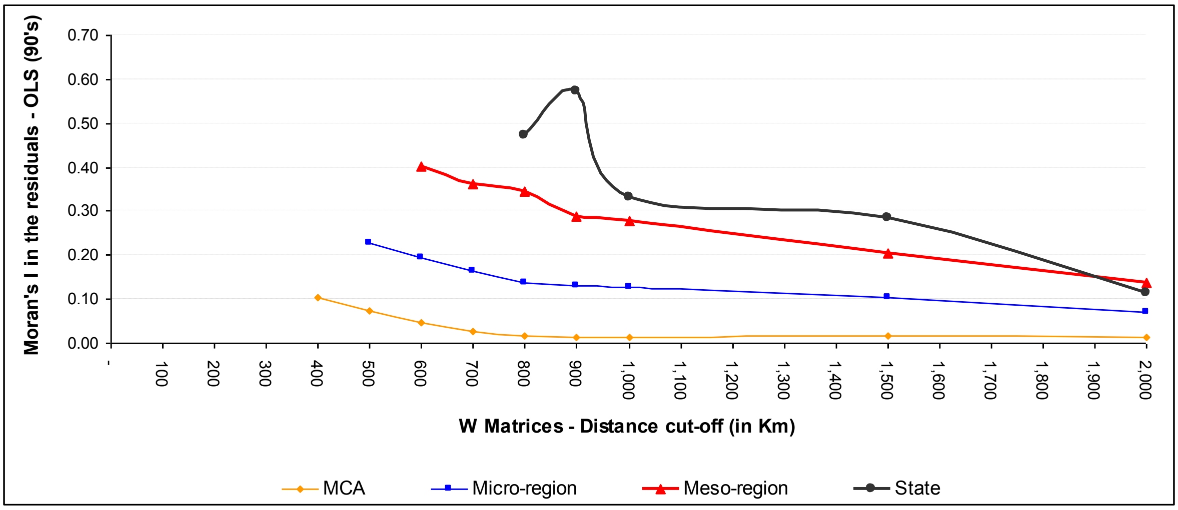
Note: All Moran's I statistics at the state level are not statistically significant at the 1% level. Moran's I statistics using distance cut-offs of 900, 1,000, 1,500 and 2,000 kilometres at the meso-regional level are not statistically significant at the 1% level.

Figure 5.2 – Moran's I in the Residuals of the OLS Regression (1980s Period)



Note: All Moran's I statistics at the state level are not statistically significant at the 1% level. Moran's I statistics using distance cut-offs of 900, 1,000, 1,500 and 2,000 kilometres at the meso-regional level are not statistically significant at the 1% level.

Figure 5.3 – Moran's I in the Residuals of the OLS Regression (1990s Period)

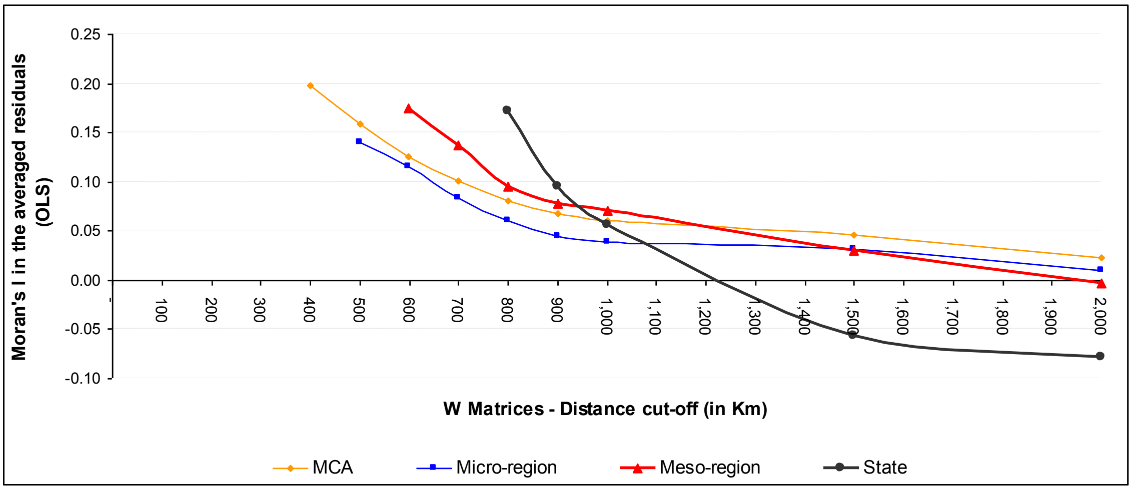


Note: All Moran's I statistic using a distance cut-off of 2,000 kilometres at the meso-regional and state levels are not statistically significant at the 1% level.

Finally, it appears from Tables II.B.1 and II.B.2 that the Moran's I statistic provides similar values of residual spatial correlation for each panel data method. In other words, irrespective of the panel data strategy choice (i.e. using pooled OLS, FD, FE or SYS-GMM), the results of the Moran's I statistic are very similar. Note that for FE estimations, the time-averaged residuals are zero by construction, and therefore, Moran's I is not calculated. Figures 5.4, 5.5 and 5.6 illustrate this empirical evidence for the time-averaged residuals of pooled OLS, FD and SYS-GMM regressions. These results indicate that traditional dynamic panel data models were not able to address spatially autocorrelated residuals in this empirical exercise. It was expected that FE, FD and SYS-GMM methods could, at least partially, address the problem of spatially correlated omitted variables. However, the results show that the Moran's I statistics for FD and SYS-GMM estimations are very similar (or even higher) to those of pooled OLS estimations. Of note, the FE estimations were able to partially deal with spatially autocorrelated residuals because as can be observed in Table II.B.1, the Moran's I statistics for the FE estimations – at least for the MCA scale – are lower than those of pooled OLS estimations. For the case of SYS-GMM, this technique introduced more spatial autocorrelation in the models. Moreover, it is important to note that the state-level spatial scale was able to address the spatial autocorrelation irrespective of the panel data method choice. Therefore, this empirical exercise indicates that traditional panel data models (and the included explanatory variables) do not incorporate all of the channels of interdependence between spatial units within states.

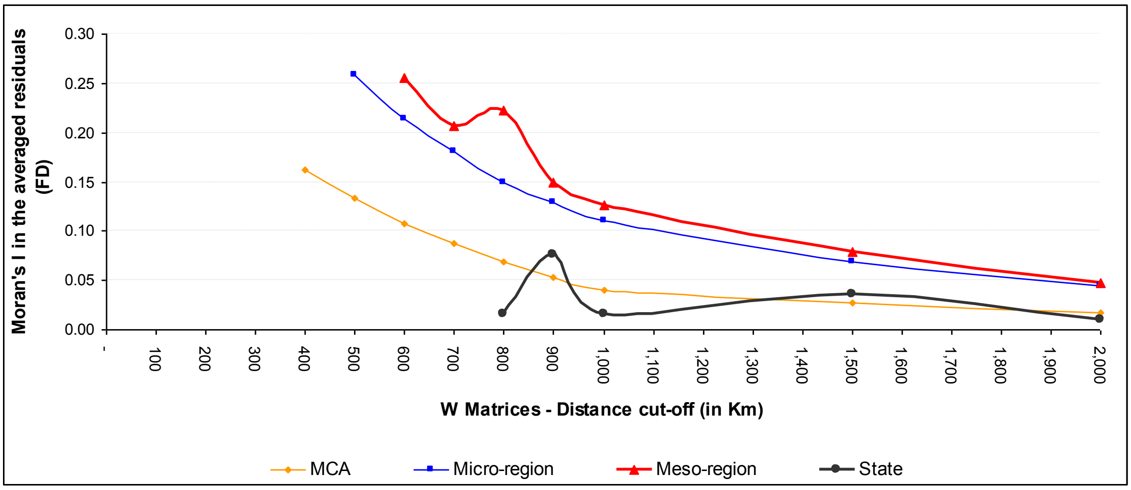
To sum up, the evidence collected thus far demonstrates that estimated coefficients are dependent on the spatial scale of analysis and that spatial autocorrelation in the residuals of panel data growth models varies at different spatial scales, increases over time and declines with distance. The mechanisms that can explain this phenomenon may be related to nuisance or substantive arguments, as discussed earlier, and the exact identification of the origin of spatial linkages observed in this empirical exercise is still a challenging issue to be addressed by the spatial econometric literature [see, for instance, McMillen (2010) and Gibbons and Overman (2012)].

Figure 5.4 - Moran's I in the Time-Averaged Residuals of Pooled OLS Regressions



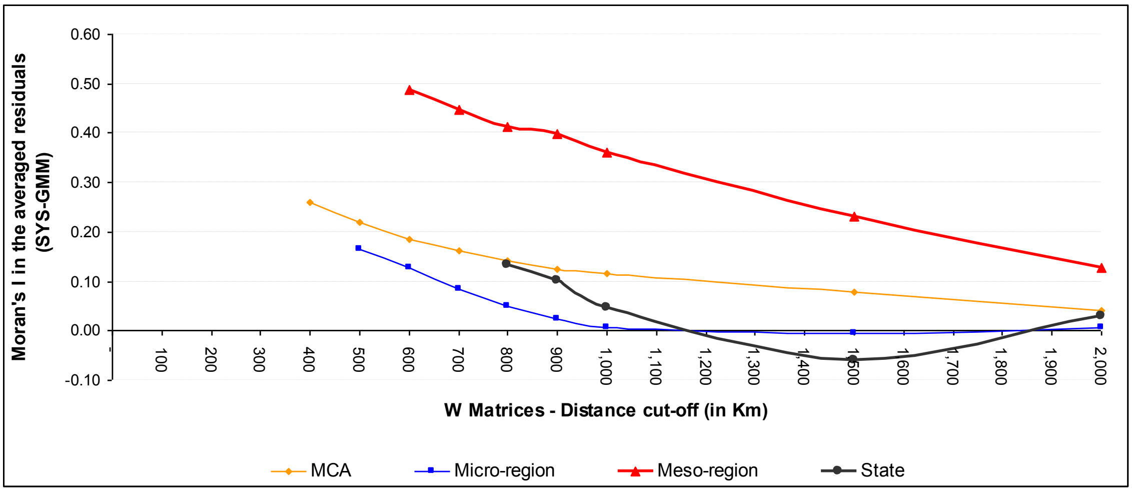
Note: Moran's I statistics at the state level are not statistically significant at the 1% level. Moran's I statistic using distance cut-off of 2,000 kilometers at the meso-regional level is not statistically significant at the 1% level.

Figure 5.5 - Moran's I in the Time-Averaged Residuals of FD Regressions



Note: Moran's I statistics at the state level are not statistically significant at the 1% level.

Figure 5.6-Moran's I in the Time-Averaged Residuals of SYS-GMM Regressions



Note: Moran's I statistics at the state level are not statistically significant at the 1% level. Moran's I statistics using distance cut-offs of 1,000, 1,500 and 2,000 kilometres at the micro-regional level are not statistically significant at the 1% level.

5.6. Conclusions

This study provides empirical evidence that the Brazilian economic growth process between 1970 and 2000 varied according to the spatial scale under analysis. Four Brazilian spatial scales are used in the current analysis: MCAs, micro-regions, meso-regions and states. Alternative panel data models were systematically estimated across those spatial scales to evaluate the extent of spatial autocorrelation in the residuals of growth equations as well as to demonstrate that the estimated coefficients change with the scale level. First, it was shown that data aggregation may strongly influence coefficient sizes because they may carry information on the strength of the spatial interactions across individuals and regions, as suggested by Glaeser et al. (2003). For instance, the current analysis corroborates the notion that that there are human capital spillovers by showing evidence that the magnitude of the average-years-of-schooling coefficient is much larger at the state level than the one estimated at the MCA level, which might be assumed to be the micro level of analysis (instead of the individual level). Moreover, it has been observed that differences of conditional β-convergence coefficients due to panel data method choices are much larger than variations due to the spatial scale of analysis. For instance, faster convergence rate in the FE framework compared to the pooled OLS approach is verified. This fact indicates that persistent differences in technology levels and institutions (represented by fixed effects) are important factors in understanding economic growth across regions (Islam, 1995).

Regarding the analysis of the extent of residual spatial autocorrelation often detected in traditional regional growth regressions, it was shown that such spatial autocorrelation seems to vary across spatial scales. Indeed, spatial autocorrelation seems to be bounded at the state level because it appears that Moran’s I statistic is not statistically significant for any choice of the spatial weight matrix, beginning from an 800-kilometer cut-off. On the other hand, although the spatial autocorrelation in residuals of the other three spatial scales shows positive and statistically significant values across distances of more than 1,500 kilometres, their levels are largely reduced when the distance is more than 900 kilometres, in particular for the 1970s and 1980s. Interestingly, an increasing clustering of the regressions' residuals over time was demonstrated, in particular over the 1990s period.

Of note, the causes of spatial correlation in the residuals are related to nuisance and substantive factors. For instance, augmented Solow growth models demonstrate that regional spillovers of the diffusion of technology across regions are caused by the spatial dimension of investments in physical and human capital (López-Bazo et al., 2004). Among other arguments, a plausible one is related to pecuniary externalities, such as those created by a specialised labour market or forward and backward linkages that might be the cause of externalities across regions. The traditional panel data models used herein were not able to distinguish between alternative explanations of spatial interactions across regions. Furthermore, in this empirical exercise, the FE, FD and SYS-GMM approaches do not completely eliminate the problem of spatially correlated omitted variables. However, it is important to note that at the state level, there is not spatial autocorrelation in the residuals of non-spatial panel data models (including the pooled OLS model) demonstrating that these models are appropriate to investigate growth determinants and convergence process in the Brazilian states case. However, for the MCAs, micro-regions, meso-regions, the results show that non-spatial panel data techniques are not able to deal completely with spatially autocorrelated residuals across those spatial units. Indeed, addressing this identification problem is still a challenging issue for the (spatial) econometric literature. One contribution of this empirical exercise is to highlight all of these features using alternative panel data strategies, demonstrating that a multiple spatial dimensional analysis may be useful to investigate regional economic growth determinants and convergence process. Future work might employ panel data spatial econometric techniques to further investigate the spatial scope of economic growth in Brazil.

6. Measuring Micro- and Macro-Impacts of Regional Development Policies: The Case of the FNE Industrial Loans in Brazil, 2000-2006

6.1. Introduction

It is well known that regional inequalities have persisted in Brazil for decades. For example, the Gross Domestic Product (GDP) per capita of the poorest region, the Northeast, was only 43% of the national average in 1989 and 47.5% in 2006. On the other hand, the per capita GDP in the Southeast region, the wealthiest region, was 139% of the national average in 1989 and 133% in 2006 (see Figure III.A.1 in Appendix III.A). The main Brazilian regional development policy has been enacted since 1989 by the Constitutional Financing Funds for the Northeast (FNE), the North (FNO), and the Central-West (FCO) (henceforth referred to as regional funds). This policy seeks the economic and social development of lagging regions through loans at below-market interest rates that are given mainly to small-scale farmers and small industrial firms (see section 3.5 for detailed information on these regional funds). However, assessment of the impacts of regional policies in Brazil has rarely occurred over the years. Thus, the goal of this chapter is to contribute to filling – at least partially – this gap in the literature by measuring the micro- and macro-impacts of the FNE loans. As analysed in Chapter 3 the Northeast region is the poorest macro-region – in terms of income per capita, for instance – in Brazil. It also receives the most part of the resources allocated to these regional funds each year, namely, 60% goes to the FNE; 20%, to the FNO; and 20%, to the FCO. For this reason the Northeast region and the FNE deserve greater attention and is the focus of this chapter. Given the availability of data, the analysis includes only those FNE loans for firms in the industrial, commerce and services sectors, which represent roughly 40% of the FNE loans granted during the analysis period.

The inquiry involves two levels of analysis that are often implemented separately in the impact-evaluation literature[104]. First, it measures the effectiveness of the FNE: does the FNE create jobs and/or increase labour productivity (proxied by wage growth) at the firm level? Second, it assesses the impact of the FNE on regional inequalities: does the FNE reduce the regional GDP per capita gap? This combined approach is useful because evaluations can suggest, for instance, that the regional funds create jobs and/or increase labour productivity at the micro level. However, it is still necessary to demonstrate that the program has solved, or at least reduced, regional inequalities. These micro- and macro-effects have been overlooked in the literature that deals with the impacts of regional funds. In the international context (for example, in the context of the European Union (EU) regional policy), most studies focus on the macro-impacts of EU regional funds on regional inequalities. These studies include Rodríguez-Pose and Fratesi (2004), Dall'erba (2005), Leonardi (2006), Esposti and Bussoletti (2008) and Dall'erba and Le Gallo (2008). Other studies are focused on the micro-impacts of specific policies; an example is Romero and Noble (2008), which evaluates England's 'New Deal for Communities' programme.

With regard to the Brazilian literature, there are few papers related to FNE impact evaluation at the firm level (but see, Silva et al., 2009; Soares et al., 2009) and none at the macro level[105]. For instance, Silva et al. (2009) measured the effectiveness of regional fund (FNE, FNO and FCO) loans using propensity score estimates for firms that received loans (the treatment group) and others that did not receive them (the

[104] Evaluation can be defined in several ways: in terms of time (e.g., ex ante, mid term or ex post), levels of complexity (e.g., monitoring daily tasks or assessing impact on the problem) or as an internal or external evaluation. The definition of Bartik and Bingham (1995) is employed here, which looks at evaluation as a continuum moving from the simplest form of evaluation (monitoring daily tasks) to the more complex (assessing impact on the problem). To demonstrate that a program (or policy) accomplishes its targets, the evaluation must be at the highest level: measuring effectiveness (for instance, it actually does create jobs) or assessing impact (e.g., there has been an improvement in the situation). Herein, I use the terms micro-impact evaluation for measuring effectiveness and macro-impact evaluation for assessing impact on the problem. Also, a cost-benefit analysis needs to be carried out to demonstrate that the program benefits outweigh the costs. However, because the data necessary to carry out this cost-benefit analysis are not available, this type of evaluation is left for future research. See Khandker et al. (2010) for further discussion on impact-evaluation strategies.

[105] Oliveira and Domingues (2005) employ a municipal dataset to examine whether the Brazilian regional funds (only FNO and FCO are analysed) have a positive impact on regional inequality. The results show that regional growth in Brazil between 1991 and 2000 was not affected by these funds.

control group) over 2000-2003. The results show that FNE has a positive impact on the employment growth rate but no impact on the growth rate for wages. The study found that employment growth was approximately 60 percentage points higher for those firms that received loans than for those that did not receive them over the period 2000-2003. With regard to FNO and FCO, there is no observed impact of the regional funds. It is possible to improve the work done by Silva et al. (2009) that assessed the effectiveness of the Brazilian regional development funds, in several ways. For instance, with firm-level data or aggregate data at different spatial scales, other methods of evaluation can be used such as fixed effects or first-difference designs, which deal with unobserved fixed factors that are constant over time (Angrist and Pischke, 2009).

The contribution of the current chapter based on those previous studies is threefold. First, it brings together the measurement of the micro- and macro-impacts of regional funds, which are often implemented separately in the evaluation literature. Second, the chapter employs the first-difference estimation technique to eliminate unobservable factors that are constant over time and may be biasing the previous results. This method has not been used in previous studies on FNE impact evaluation. Third, this is the first study to evaluate the macro-effects of FNE loans awarded to firms in the industrial and commerce/services sectors on GDP per capita growth at the municipal level. With regard to this macro-analysis, other levels of spatial aggregation of the observational units are employed because an aggregation problem [the Modifiable Areal Unit Problem (MAUP)] might prevent us from identifying the effects of FNE loans on GDP per capita growth at the municipal level. In this sense, this chapter seeks to provide a more complete picture of FNE performance during 2000-2006.

It is important to note that the goal of the FNE is to reduce regional inequalities through the financing of productive sectors in the Northeast. This imprecisely defined objective is the major obstacle to impact evaluations. As pointed out by Jann and Wegrich (2007: 54) "*given the strong incentive of blame-avoidance, governments are* *encouraged to avoid the precise definition of goals because otherwise politicians would risk taking the blame for obvious failure*". In the Brazilian case, the combined assessment of micro- and macro-impacts of the FNE is relevant because the FNE goal is broadly defined at the macro level (reducing regional inequalities), but the FNE contains a policy tool that operates at the micro level by means of subsidised loans to producers in the Northeast region. Whilst the objectives at the firm level are not defined by federal law, some official documents[106] have highlighted the fact that the FNE loans seek more efficient resource allocations to increase the productivity of firms and generate new jobs. In this sense, and given the availability of the data, this study defined the reduction of the GDP-per-capita gap as the policy objective at the macro-scale level and job creation and increases in labour productivity (proxied by wage growth)[107] as the objectives at the firm (micro) level.

Note that the rate of wage growth is the proxy for the rate of labour-productivity growth because, with the Cobb-Douglas production function [$Y = K^{\alpha} L^{1-\alpha}$] in mind (assuming that capital (K) and labour (L) factors are the major inputs into production), the derivative of the production function with respect to labour (dY / dL) equals real wages. Thus, as firms maximise profits, the result should be equal to marginal labour productivity (Solow, 1956). Finally, the connection between labour productivity growth and the macro variable, GDP per capita growth, stems from the fact that marginal productivity is supposed to be proportional to the average productivity [$dY / dL = (1-\alpha) Y / L$] that is proxied by GDP per capita[108]. It is worth noting that rising employment (with all else constant) might be negatively related to labour productivity.

[106] For example, see the webpage of the Ministry for National Integration <http://www.integracao.gov.br/fundos/fundos_constitucionais/diretrizes.asp?id=diretrizes> and Banco do Nordeste (2001, 2009).

[107] Another proxy for labour productivity at the firm level could be value added per worker; however this information is not available for the current study.

[108] In fact, GDP per worker should be used to indicate average productivity; however, given the lack of accurate data for small municipalities, the total number of workers at the municipal level is not available for most years under analysis. On the other hand, there is a high correlation (the correlation coefficient is +0.997) between total population and total workers at the municipal level in 2000, which is a census year for which reliable information is available. For this reason, the results are not affected by the use of GDP per capita.

This is because the more employees there are, the lower is the labour productivity, and thus rising employment also has negative effects on average productivity. Alternatively, employment growth can be viewed as a contribution to the increase in total income, which fosters consumption and positively impacts GDP per capita growth (at least in the short term) in an economy that is below full employment. From this perspective, Pichelmann and Roeger (2008) point out that newly employed people contribute more to GDP than they did previously even if their productivity is below average. The joint consideration of these three variables (employment, wages and GDP per capita growth) allows for a more consistent evaluation of FNE[109].

Another issue that requires explanation is that the micro-impact evaluation focuses only on those firms that can be traced in RAIS[110] during the period under analysis. Most FNE loans (approximately 60%) are granted to individuals who have small farming businesses in the informal sector (i.e., they do not have a CNPJ identifier[111]), and for this reason, they are not covered by RAIS (which is the source of information for the micro-analysis). The few formal rural firms found in RAIS are not statistically representative of the FNE rural population. For this reason, the agricultural sector was excluded from the micro-evaluation in this chapter. The government still needs to formulate a specific survey to cover individuals and small rural businesses in the Northeast to evaluate this important, targeted FNE population. For the sake of comparability between the micro- and macro-impact evaluations, the macro-impact evaluation at the municipal level was restricted to assessing only those FNE loans granted to the industrial and commerce/services sectors.

[109] As suggested by one anonymous referee, a computable general equilibrium (CGE) model would have given a wider perspective on the connections between the variables under investigation. The partial equilibrium approach that was applied herein may have limited the current empirical analysis.
[110] Annual Social Information Report (Relação Anual de Informações Sociais) of the Ministry of Labor. The RAIS data were used under a cooperation agreement between the Labor Ministry and the "Instituto de Pesquisa Econômica Aplicada" (IPEA). More details can be found in Appendix B.
[111] "Cadastro Nacional da Pessoa Jurídica" – literally, national juridical person registration – as opposed to the CPF number for persons. CNPJ is an identification number for Brazilian companies assigned by the Brazilian Ministry of Revenue ("Ministério da Fazenda"). The CNPJ number is comprised of a base of 8 digits, a 4-digit radical and 2 check digits, such as 22.222.222/0001-05.

The chapter is organised as follows. Section 6.2 describes the framework employed to measure the micro- and macro-impacts of the FNE loans. Section 6.3 carries out the micro-impact effectiveness estimates for employment and labour-productivity growth and the assessment of the macro-impact estimates, both of which use the first-difference method. Section 6.4 presents the conclusions.

6.2. Micro- and Macro-Impact-Measurement Framework

This section explains the empirical method, first-differencing (FD), that is used to measure the micro-impacts of FNE loans on employment and labour productivity growth and the macro-impacts on GDP per capita growth. Initially, it is important to note that the main challenge of any policy evaluation is to deal with the selection bias introduced when random assignments are not possible. To describe this problem more precisely, it is possible to think about the FNE loans granted to firms in the Brazilian Northeast region by means of the binary variable that represents participation, $D_i = \{0,1\}$, which is 1 if the firm is treated (i.e., received the loan) and 0 otherwise[112]. The observed outcome of variable Z for firm *i* would be as follows:

$$Z_i = DZ_{1i} + (1-D)Z_{0i} \tag{6.1}$$

where Z_{0i} is the employment growth, for instance, of a firm had it not received the FNE loan and irrespective of whether it was actually received, whereas Z_{1i} is the firm's employment growth if it had received the loan. The result of interest is the difference between Z_{1i} and Z_{0i} , which is the causal effect of the FNE loan for firm *i*. This analysis would be feasible if it were possible to go back in time and change a firm's treatment status. However, Angrist and Pischke (2009) point out that, because it is never possible to observe both potential outcomes for any one firm, the effects of this policy must be

[112] Note that information on FNE participation at the firm level is a binary variable given availability of data. At the macro level, the available data is a non-binary treatment variable that represents the amount of FNE loans as a ratio of the total GDP at the municipal level. See Wooldridge (2002: 638-642) for further discussion on the similarities between binary and non-binary treatments.

elucidated by comparing the average employment growth of those firms that were and were not granted the FNE loans.

The comparison of the averages of those firms granted and not granted FNE loans is formalised in the following equation:

$$\underbrace{E[Z_i \mid D_i = 1] - E[Z_i \mid D_i = 0]}_{\substack{Observed\_differences\_in\_average\_\\ employment\_growth}} = \underbrace{E[Z_{1i} \mid D_i = 1] - E[Z_{0i} \mid D_i = 1]}_{Average\_treatment\_effect\_on\_the\_treated} + \underbrace{E[Z_{0i} \mid D_i = 1] - E[Z_{0i} \mid D_i = 0]}_{Selection\_bias}$$

(6.2)

The term "average treatment effect on the treated" (ATT) represents the average differences between the employment growth of firms that received the FNE loan, $E[Z_{1i} \mid D_i = 1]$, and what would have happened to these firms had the loans not been granted, $E[Z_{0i} \mid D_i = 1]$. Nevertheless, the "selection bias" must be considered in addition to this causal effect. This last term is the difference in the average for Z_{0i} between those firms that were $(D_i = 1)$ and those that were not $(D_i = 0)$ granted the FNE loans. The point here is that it is not possible to know what would have happened if the granted firms had not received the loans, $E[Z_{0i} \mid D_i = 1]$. The selection bias may be negative or positive and may cause us to underestimate or overestimate the treatment effect. For instance, one might argue that those firms seeking the loans are more motivated and entrepreneurial, and that even without the FNE loans, they would perform better than the others.

The random assignment of FNE loans (D_i) solves the selection problem. Formally, in the absence of the selection bias, it would be expressed by the following:

$$E[Z_{0i} \mid D_i = 1] - E[Z_{0i} \mid D_i = 0] = 0, \qquad (6.3)$$

that is, on average, there would be no differences between the potential outcomes for the untreated and treated firms if the loans had not been granted. This does not mean that the FNE loans should be randomly granted, but it demonstrates that the goal of empirical economic research is to overcome this selection bias by using the appropriate policy-evaluation techniques.

The benchmark estimation carried out in the next section is the difference between the means of employment growth (without controls) for those firms that are treated and those that are not. The regression of $Z_{i,growth}$ on $D_{i,2000}$ can be used to test the significance of the treatment effect:

$$Z_{i,growth} = \alpha + \rho D_{i,2000} + \eta_i \, , \tag{6.4}$$

where α is the constant term, ρ is the coefficient of the treatment effect and η_i is the random term. The subscript 'growth' is included in *Z* to highlight the fact that the dependent variables are expressed in terms of average annual growth rates covering two periods (2000-2003 and 2000-2006). With regard to the micro-analysis, $Z_{i,growth}$ represents the employment growth and labour-productivity growth of firm *i,* and $D_{i,2000}$ is a binary variable that represents participation in the FNE program in the year 2000. In the case of the macro-analysis, $Z_{i,growth}$ is the GDP per capita growth of municipality *i,* and $D_{i,2000}$ represents the amount of FNE loans to the industrial sector in 2000 as a ratio of the total GDP in 2000 at the municipal level. In Equation (5.4), the treatment-effect estimation relies on the weaker assumption that D_i is independent of Z_{0i} without placing any restriction on the relationship between D_i and Z_{1i} (Wooldridge, 2002: 606). It is worth noting that, under this assumption of absence of selection bias, it is unlikely that these estimates provide a reliable value for the impact of the policy because the firms were not selected randomly. However, these estimates are shown in the results section for the sake of comparison with the other empirical strategies that are discussed next.

One way to improve the estimation of Equation (6.4) is to add some controls (X_i) that tend to affect the variables of interest ($Z_{i,growth}$) by both directly and indirectly motivating participation, D_i. See Appendix III.B for the description and data sources for all covariates (X_i) that are included in the micro- and macro-analyses and for the dependent variables[113].

$$Z_{i,growth} = \alpha + \rho D_{i,2000} + X_{i,2000}'\beta + \upsilon_i \tag{6.5}$$

In this case, the assumption is that there are differences among firms in the treatment and control groups in terms of their observable characteristics. For this reason, it is necessary to explicitly include all covariates (X_i) that are important for the determination of $Z_{i,growth}$ and participation, D_i. In this situation, as explained by Angrist and Pischke (2009: 59), the residual υ_i is uncorrelated with the regressors D_i and X_i, and the regression coefficient ρ is the causal effect of interest. This is the selection-on-observables assumption for regression models (Barnow et al., 1981), which assumes that the observable characteristics X_i are the only reason why η_i and D_i are correlated in Equation (6.4)[114]. It is worth noting that the estimates might be biased if Equation (6.5) does not consider all of the variables that are important in determining participation and that also affect the variable of interest, $Z_{i,growth}$.

An issue that has not been examined in the FNE evaluation literature is the likely bias due to unobservable characteristics. For instance, this is the case for some dimensions of motivation/ability/entrepreneurship as related to applying for or receiving the FNE loan. Furthermore, at the macro level, it is likely that only those firms located in municipalities with good access to information and a banking infrastructure, or to other unobservable institutional advantages, have access to these funds. Here, the FD method is employed to eliminate those unobservable effects that are constant over

[113] At the macro level, Equation (6.5) can be motivated by the so-called Barro-regression. Indeed, many papers that examine the impacts of regional funds on GDP per capita growth – for instance, the studies about the EU structural funds, such as Rodriguez-Pose and Fratesi (2004), Dall'erba and Le Gallo (2008), and Esposti and Bussoletti (2008) – are based on the neoclassical growth model described in Barro and Sala-i-Martin (1991, 1992). Interestingly, Armstrong (2002) discusses some practical steps to reconcile the evaluation evidence on regional policy with the evidence from the growth literature.

[114] In other words, Wooldridge (2002: 607) highlights that "*when D_i and (Z_{0i} , Z_{1i}) are allowed to be correlated, we need an assumption in order to identify treatment effects. Rosenbaum and Rubin (1983) introduced the following assumption, which they call ignorability of treatment (given observed covariates X_i): Conditional on X_i, D_i and (Z_{0i} , Z_{1i}) are independent*".

time. At least two time periods are needed to carry out this strategy. Now, all of the subscripts are included in the equation to indicate the time periods, as shown below:

$$Z_{i,growth2003-2000} = \alpha + \rho D_{i,2000} + X_{i,2000}'\beta + A_i'\gamma + \upsilon_{i,2000}, \tag{6.6}$$

More precisely, $Z_{i,growth2003-2000}$ is, for instance, the average annual employment growth of firm *i* between 2000 and 2003. In the results section, I also show the results using employment, labour productivity and GDP per capita growth between 2000 and 2006[115]. More important is the vector of unobserved but fixed covariates, A_i, which will be ruled out with the first-difference strategy. As described in the Appendix B, information was collected at the firm and municipal levels for a previous period represented by the subscript 1999:

$$Z_{i,growth2000-1999} = \alpha + \rho D_{i,1999} + X_{i,1999}'\beta + A_i'\gamma + \upsilon_{i,1999} \tag{6.7}$$

Subtracting (6.7) from (6.6) yields the following:

$$Z_{i,growth2003-2000} - Z_{i,growth2000-1999} = \alpha + \rho(D_{i,2000} - D_{i,1999}) + (X_{i,2000} - X_{i,1999})'\beta + (\upsilon_{i,2000} - \upsilon_{i,1999})$$

(6.8)

Note that, in the FD regression model, the unobserved fixed effect, A_i, is eliminated by subtracting the observation for the previous time period from the observation for the current time period[116]. Subsequently, equation (6.8) can be estimated by OLS, and the coefficient of the FNE variable, ρ, indicates the average impact on the differences for the variable of interest (employment, labour productivity and GDP per capita growth).

Finally, it is worth noting that there are alternative impact-evaluation strategies, such as instrumental variables, matching and propensity score techniques (as discussed in detail by Khandker et al. 2010). The instrumental-variable approach tries

[115] The average annual growth rates are calculated as follows: $Z_{i,growth}$ = $((y_{i,t} / y_{i,0})$^$(1/T))-1$, where $y_{i,t}$ and $y_{i,0}$, are, respectively, the final period and the initial period of dependent variable for firm i, and T is the time period in years.
[116] In the micro analysis, $D_{i,1999}$ is a dummy variable that is now zero for all firms in 1999.

to solve the endogeneity problem by finding a variable (the instrument) that is correlated with the causal variable of interest (in this case, the FNE variable) but that is uncorrelated with any other determinants of the dependent variable; equivalently, the instrumental variable is uncorrelated with the error term (Angrist and Pischke, 2009). However, there is a lack of good instruments because there are so many variables that can be used to explain the dependent variables (in particular, GDP per capita growth) that it is difficult to find variables that are highly correlated with the FNE variable that can be excluded from the regression (Temple, 1999). Matching based on observable characteristics assumes that the selection of the FNE loans occurs only based on observable characteristics, so that firms with such characteristics have the same probability of participation. Then, the average effects of FNE loans can be obtained by averaging weighted effects for subgroups of firms with similar characteristics. The drawback of this estimator is its implementation when there are a large number of variables or when these variables are continuous. The propensity score estimates overcome this problem by summarising the similar characteristics of firms via the estimation of a logit or probit model that indicates the probability of receiving the FNE loan. Although these matching approaches are appealing, they are accompanied by the same explicit statement of the conditional independence assumption that is required to provide a causal interpretation of regression coefficients and, for this reason, we can say that matching and regression are both control strategies (Angrist and Pischke, 2009: 69)[117]. In the current work, I focus on the FD regression approach because it has the ability to control for observable and time-invariant unobservable characteristics.

6.3. Results

6.3.1. Micro-results

[117] Angrist and Pischke (2009: 69) argue that "*since the core assumption underlying causal inference is the same for the two strategies, it's worth asking whether or to what extent matching really differs from regression. Our view is that regression can be motivated as a particular sort of weighted matching estimator, and therefore the differences between regression and matching estimates are unlikely to be of major empirical importance*".

The micro-analysis measures the effects of FNE loans to firms in the industrial and commerce/services sectors and seeks to answer the question of whether the regional fund creates jobs and/or increases labour productivity (proxied by wage growth). Table 6.1 shows the results of the FNE impact evaluation using four different control samples and covering two periods (2000-2003 and 2000-2006). Heteroskedasticity-robust standard errors are indicated in parenthesis for all estimates because the diagnostics for this problem are statistically significant.

The most important results of this table are shown in the second and the third parts, which employ, respectively, a sample matched to the treated group and a sample 'perfectly' matched to the treated group (see Appendix III.B for details). These are better samples than the others for carrying out the micro-impact evaluation estimates because they ensure that the treatment and control groups have similar characteristics and thus make the two groups more comparable (as shown in the summary statistics in Table III.B.1 in Appendix III.B). For the sample matched to the treated group, the first results (ii. OLS without covariates in Table 6.1) are obtained using Eq. (6.4), which does not control for observable and unobservable characteristics. With this assumption in mind, the results show a statistically significant positive impact of FNE loans on employment growth between 2000 and 2003. If controls for observable characteristics are added – which means estimating the FNE impact using Eq. (6.5) – the results in Table 6.1 (ii. OLS with covariates) also show that there is only a positive impact of FNE on employment growth over the 2000-2003 period. More precisely, average annual employment growth is 5.14 percentage points higher for the financed firms than for the non-financed firms between 2000 and 2003. On the other hand, it is not possible to verify any impact of FNE on average annual wage growth for the two periods (2000-2003 and 2000-2006).

However, the most important estimation is the one that controls for observable and unobservable (constant over time) characteristics using Eq. (6.8) (ii. first difference with covariates, FD). When the first differences are estimated, a positive and statistically significant impact of FNE on employment growth is observed for the period between 2000 and 2003. This result suggests that the first difference of average annual employment growth among those firms that received the FNE loans was about 18.96 percentage points higher than the first difference of average annual employment growth for those firms that were not granted loans. Furthermore, the results are robust to alterations in the conditioning set of variables. Appendix III.C, Table III.C.1 shows the robustness checks for these FD estimations. Moreover, a statistically significant positive impact of 16.11 percentage points on the differences in the employment growth variable is found for the 2000-2006 period using the FD method. If the 'perfectly' matched sample is used (part iii in Table 6.1), the impact of FNE loans on the differences in employment growth between 2000 and 2003 is even larger (20.7 percentage points), but it disappears for the longer period (2000-2006).

The results in (ii) and (iii) using matched samples are in line with previous studies such as Silva et al. (2009) and Soares et al. (2009). For instance, Soares et al. (2009) employ the propensity score method and expand the evaluation of FNE that was conducted by Silva et al. (2009) by enlarging the time horizon under analysis. The results show significant impacts of FNE loans on employment growth for all periods between 1999 and 2005; however, no impact on the growth rate for wages was found. Comparing the impact of FNE on employment growth over the three-year period, it is 33 percentage points higher for financed firms in the Soares et al. (2009) analysis. In this current study, for example, it is 10.83 percentage points per annum higher for financed firms (which means 36 percentage points over 2000-2003 period) using the 'perfectly' matched sample and the OLS-with-covariates method (as shown in Table 6.1). Moreover, this outcome seems to be in line with the related literature. Sousa and Ottaviano (2009) measured the effects of BNDES loans on the productivity of Brazilian manufacturing firms. Their results suggest no impact of BNDES loans on firms' productivity, even though they found effects on employment. Pereira (2007) also finds positive impacts of BNDES loans on employment for granted firms.

Table 6.1

The Micro-Approach for Evaluating FNE Impact on Employment and Average Annual Wage Growth Using Four Control Samples

	Employment growth between 2000-2003	Wage growth between 2000-2003	Employment growth between 2000-2006	Wage growth between 2000-2006
i. All firms				
OLS without covariates	0.0052	-0.0022	-0.0113	-0.0057
	(0.0258)	(0.0064)	(0.0167)	(0.0045)
OLS with covariates	0.0089	-0.0051	-0.0056	-0.0079**
	(0.0260)	(0.0058)	(0.0167)	(0.0038)
First Difference with covariates (FD)	0.0247	0.0102	0.0079	0.0067
	(0.0721)	(0.0129)	(0.0696)	(0.0123)
Observations	111,960	111,960	111,960	111,960
ii. Matched sample to the treated group				
OLS without covariates	0.0482*	-0.0003	0.0198	-0.0029
	(0.0259)	(0.0065)	(0.0168)	(0.0045)
OLS with covariates	0.0514*	-0.0028	0.0229	-0.0048
	(0.0267)	(0.0058)	(0.0174)	(0.0036)
First Difference with covariates (FD)	0.1896***	-0.0017	0.1611**	-0.0041
	(0.0708)	(0.0126)	(0.0685)	(0.0120)
Observations	10,081	10,081	10,081	10,081
iii. 'Perfectly' matched sample to the treated group				
OLS without covariates	0.1079***	0.0012	0.0339	-0.00003
	(0.0311)	(0.0105)	(0.0229)	(0.0068)
OLS with covariates	0.1083***	0.0010	0.0331	-0.0004
	(0.0302)	(0.0098)	(0.0231)	(0.0057)
First Difference with covariates (FD)	0.2070*	-0.0049	0.1381	-0.0068
	(0.1251)	(0.0171)	(0.1228)	(0.0146)
Observations	182	182	182	182
iv. Random sample				
OLS without covariates	0.0293	0.0017	0.0046	-0.0022
	(0.0268)	(0.0071)	(0.0174)	(0.0048)
OLS with covariates	0.0281	-0.0007	0.0065	-0.0042
	(0.0290)	(0.0077)	(0.0186)	(0.0047)
First Difference with covariates (FD)	0.0974	0.0067	0.0720	0.0030
	(0.0866)	(0.0138)	(0.0844)	(0.0130)
Observations	905	905	905	905

Note: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. OLS Dependent variable: Employment growth (or wage growth)= (((yi,t / yi,0)^(1/T))-1) , where yi,t and yi,0 are, respectively, the final period and the initial period of employment stock (average wage) for firm i and T is the time period in years. OLS estimations with covariates include: average age of the employees in 2000, average years of schooling of the employees in 2000, dummy for the commerce/services sector, average wage in 2000, number of employees in 2000, dummies for Northeast states. FD Dependent variable= [(Employment growth between 2000-2003) - (Employment growth between 2000-1999)]. FD estimations with covariates include: diff. average age of the employees (2000-1999), diff. average years of schooling of the employees (2000-1999), diff. average wage (2000-1999), diff. number of employees (2000-1999).

The first (i) and the fourth (iv) part of Table 6.1 show, respectively, the results using the control sample with all firms identified in RAIS during the period under analysis and the control sample used in the paper of Silva et al. (2009). Actually, this last dataset is different from that used in Silva et al. (2009); the few firms in the agricultural sector and those not present in the RAIS dataset during the years 1999 and 2006 were excluded. The results in parts (i) and (iv) of Table 6.1 demonstrate that the evaluation of FNE impact depends on the choice of control sample. As can be observed, there are no statistically significant effects when the FD estimates are carried out using samples (i) and (iv). The results in Table 6.1 suggest that the control sample should be analysed with caution.

Altogether, the FD results for the micro-impact evaluation show that the FNE loans granted in 2000 played a role in attracting and stimulating employment growth during the following years. However, even if the FNEs were effective in creating jobs, this does not mean that the FNE loans have been able to eliminate (or even to reduce) Brazilian regional inequalities. Indeed, the observed employment growth may have a negative impact on GDP per capita via lowering labour productivity as discussed earlier. The macro-impact evaluation presented next aims to investigate this issue.

6.3.2. Macro Results

The goal of the macro-impact evaluation is to test whether the FNE loans fosters GDP per capita growth at the municipal level in the Northeast region to reduce the regional inequalities in Brazil. Table 6.2 shows the results regarding FNE impact on GDP per capita growth during the 2000-2003 and 2000-2006 periods. The FNE industrial variable is the amount of FNE loans to the industrial sector in 2000 as a proportion of total GDP in 2000. A robustness check using the FNE industrial ratio from between 2000 and 2003 is also provided. The results using OLS (both with and without covariates) and FD at the municipal level are shown in Table 6.2. As was done in the micro-analysis, the first step is to estimate Eq. (6.4) via OLS, paying attention to the

FNE industrial coefficient. Heteroskedasticity-robust standard errors are provided in parenthesis for all estimates because the diagnostics for this problem are statistically significant.

With regard to the impact of the FNE industrial ratio in 2000 on the GDP per capita growth for the periods 2000-2003 and 2000-2006, the OLS estimates without covariates are not statistically significant. When the OLS estimates with covariates are estimated, a negative and statistically significant impact of the FNE industrial ratio in 2000 on GDP per capita growth between 2000 and 2006 is observed. This indicates that municipalities with high FNE industrial ratios in 2000 experienced slower growth between 2000 and 2006. However, this result does not hold when the FD strategy is used to control for time-invariant unobservable factors that may be biasing the previous FNE macro-impact OLS estimates. In other words, the FD method suggests no significant impact of the FNE ratio on the differences for the GDP per capita growth rates. This result is robust to alterations in the conditioning set of the controlling variables and is reported in Table III.C.2 in Appendix III.C for the 2000-2003 period. The most important element of this result is the fact that the FNE industrial ratio does not have any negative effect on the GDP per capita growth. In fact, the positive and statistically significant impact of the FNE industrial loans on job creation that was verified in the micro-impact evaluation has no significant impact on the GDP per capita growth at the municipal level in Brazil.

Another possible explanation for this lack of impact on the GDP per capita growth may be the relatively small magnitudes of the FNE industrial loans to the industrial sector at the municipal level. For instance, in 2000, there were 129 “treated” municipalities that had firms that received the FNE loans; of these, only 13 municipalities have a FNE ratio greater than 0.01. The mean FNE ratio in 2000 is 0.0005 for all of the 1,731 municipalities in the Northeast. To overcome this issue (at least partially), the same models using the amount of FNE loans to the industrial sector between 2000 and 2003 as a proportion of total GDP in 2000 are estimated in an effort to identify any different significant impact. Using the FNE loans from between 2000 and 2003, there are 522 “treated” municipalities with firms that received the FNE loans; of these, 45 municipalities have a FNE ratio greater than 0.01. The mean FNE ratio for between 2000 and 2003 is 0.0016 for all 1,731 municipalities in the Northeast (see Table III.B.4 in Appendix III.B for details). Figure III.A.2 in Appendix III.A shows the spatial distribution of the FNE industrial ratios at the municipal level. Again, the estimates obtained with these regressions (part b in Table 6.2) do not indicate any significant impact of the FNE industrial ratio from between 2000 and 2003 on GDP per capita growth. One can note that the FNE ratio from between 2000 and 2003 might still represent investments that are too limited to have any significant impact on the GDP per capita growth at the municipal level. From Table III.A.1 in Appendix III.A, it is possible to note that from 2004 onwards, the total amount of FNE loans has substantially increased, and the impact of these loans should be assessed in future work when data become available. However, before advocating for more resources, it is still necessary to demonstrate that the FNE program is cost effective; that is, a cost-benefit analysis should be carried out to show that the FNE benefits outweigh its costs.

Table 6.2

The Macro-Approach for Evaluating FNE Impact on Average Annual Growth of the GDP per capita at the Municipal Level

	GDP per capita growth between 2000-2003	GDP per capita growth between 2000-2006
a. FNE industrial ratio in 2000		
OLS without covariates	0.0615	-0.0312
	(0.0878)	(0.0302)
OLS with covariates	0.0013	-0.0540*
	(0.0990)	(0.0304)
First Difference with covariates (FD)	0.2384	0.0917
	(0.2076)	(0.1446)
Observations	1,731	1,731
b. FNE industrial ratio between 2000 and 2003		
OLS without covariates	0.1467	-0.0294
	(0.1726)	(0.0664)
OLS with covariates	0.1742	-0.0123
	(0.1983)	(0.0637)
First Difference with covariates (FD)	0.1833	-0.0456
	(0.2497)	(0.2132)
Observations	1,731	1,731

Note: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. Dependent variable: GDP per capita growth = (((yi,t / yi,0)^(1/T))-1) , where yi,t and yi,0, are, respectively, the final period and the initial period of GDP per capita for firm i and T is the time period in years. OLS estimations with covariates include: average years of schooling of the working force population in the formal sector in 2000, population density in 2000, GDP per capita in 2000 and dummies for Northeast states. FD Dependent variable= [(GDP per capita growth between 2000-2003) - (GDP per capita growth between 2000-1999)]. FD estimations with covariates include: diff. years of schooling of the working force population in the formal sector (2000-1999), diff. population density (2000-1999), diff. GDP per capita (2000-1999).

Finally, the results may change if another spatial scale is used. This fact is linked to a measurement issue that can cause variability in the estimated coefficients due to the use of different levels of spatial aggregation of the observational units. This variability could occur because of the existence of the MAUP (Gehlke and Biehl, 1934; Robinson, 1950; Openshaw and Taylor, 1979). For this reason, a cautious analysis of FNE impact on GDP per capita growth at different spatial scales may allow one to identify the appropriate spatial scale for evaluating this regional policy. The macro-dataset is available at the municipal level and is merged to form other two spatial scales: 189 micro-regions and 22 spatial clusters in the Northeast region. The micro-regions were defined by IBGE in 1990 as being a group of contiguous municipalities in the same state. They were grouped according to natural and production characteristics.

The spatial-cluster level proposed by Carvalho et al. (2007) employs a cluster methodology (algorithmic) that groups contiguous municipalities that share similar characteristics using 46 variables reported in the Brazilian Census of 2000. In Table 6.3, the results for both spatial levels (micro-regions and spatial clusters) show conclusions that are similar to those at the municipal level, that is, the FD method suggests no significant impact of the FNE ratio on the differences for the GDP per capita growth rates. Altogether, the results at various spatial scales suggest that there are no statistically significant impacts of FNE industrial loans on GDP per capita growth.

Table 6.3

The Macro-Approach for Evaluating FNE Impact on Average Annual Growth of the GDP per capita at the Micro-regional and the Spatial Cluster Levels

	Micro-regions		**Spatial Clusters**	
	GDP per capita growth between 2000-2003	**GDP per capita growth between 2000-2006**	**GDP per capita growth between 2000-2003**	**GDP per capita growth between 2000-2006**
a. FNE industrial ratio in 2000				
OLS without covariates	2.7590	0.0205	-12.8747**	-9.5211***
	(3.2890)	(1.2208)	(4.8284)	(2.8275)
OLS with covariates	0.9851	-0.3001	-19.1873	-10.4059
	(3.3979)	(1.2764)	(10.9137)	(6.9806)
First Difference with covariates (FD)	0.7426	0.3455	-3.1388	-2.3692
	(0.7307)	(0.3827)	(9.5546)	(7.6720)
Observations	189	189	22	22
b. FNE industrial ratio between 2000 and 2003				
OLS without covariates	-0.9557**	-0.8828***	-1.9378	-0.7125
	(0.4676)	(0.2836)	(1.2745)	(0.9060)
OLS with covariates	-0.9332**	-0.6428**	-2.2072	0.2956
	(0.4391)	(0.2945)	(1.3419)	(1.0176)
First Difference with covariates (FD)	-0.6502	-0.6771	-0.1717	0.9095
	(0.5239)	(0.4753)	(1.9636)	(1.3953)
Observations	189	189	22	22

Note: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. Dependent variable: GDP per capita growth = (((yi,t / yi,0)^(1/T))-1) , where yi,t and yi,0, are, respectively, the final period and the initial period of GDP per capita for firm i and T is the time period in years. OLS estimations with covariates include: average years of schooling of the working force population in the formal sector in 2000, population density in 2000, GDP per capita in 2000 and dummies for Northeast states. FD Dependent variable= [(GDP per capita growth between 2000-2003) - (GDP per capita growth between 2000-1999)]. FD estimations with covariates include: diff. years of schooling of the working force population in the formal sector (2000-1999), diff. population density (2000-1999), diff. GDP per capita (2000-1999).

6.4. Conclusions

One contribution of this chapter is that it brings together two levels of analysis that are often implemented separately in the impact-evaluation literature; thus, it provides a more complete assessment of the FNE loans directed toward firms in the industrial/commerce/services sectors. These micro- and macro-effects have been overlooked in the literature that addresses the impacts of regional development funds. In this sense, the chapter presents a general framework for micro- and macro-impacts measurement that can be applied in regional development policy analyses.

The micro-impact evaluation seeks to answer the following question: did the subsidised FNE loans cause an increase in employment or in labour productivity in the targeted firms? The results, based on control samples matched to the treated group (using the FD method that controls for observable characteristics and unobserved fixed effects), suggest that the FNE industrial loans played a role in attracting and stimulating employment growth in the Northeast region between 2000 and 2003 and over the longer period of 2000-2006. However, the results show no impact of FNE industrial loans on firm labour productivity (proxied by wage growth) in the estimates that control for observable and unobservable characteristics.

As discussed in the chapter, these positive effects on job creation at the firm level do not mean that the FNE loans have been able to eliminate (or even to reduce) Brazilian regional inequalities. For this reason, a second impact evaluation was carried out to attempt to answer the following question: have regional inequalities been reduced as a result of the FNE loans? An empirical strategy was performed to test whether FNE loans foster GDP per capita growth at the municipal level in the Northeast region. The same FD strategy used at the firm level was implemented at the macro level. The FD results do not indicate any positive impact of FNE industrial ratios on GDP per capita growth during the 2000-2003 or 2000-2006 periods. This conclusion is robust to combinations of the set of controlling variables and to the use of different levels of spatial aggregation of the observational units. Altogether, the micro- and macro-impact evaluation results suggest that the effect of FNE industrial loans on employment growth at the firm (micro) level has not had any significant impact on GDP per capita growth at the macro level in the Northeast region. At least in the short term, this is probably because of a balance between countervailing effects: the negative effect of employment growth on average productivity and the positive impact of employment growth on GDP per capita growth via increases in total income. The lack of impact at the macro level is worrying because this is exactly the main objective of the FNE. The positive effects of the FNE industrial loans at the firm level seem to be very limited and localised. The positive effects verified only at the firm level suggest that the FNE industrial loans are not able to create backward and forward linkages that would produce positive effects at macro (municipalities, micro-regions or spatial clusters) levels. The joint analysis of micro and macro evidence simultaneously is important because it can show more precisely the spatial scale(s) where the impacts are observed and then the micro-macro paradox can be better understood. Of note, these localised positive impacts and absence of impacts at the macro level have not been examined at the European Union (EU) context, but the results discussed herein resemble the paradox that Dall'erba and Hewings (2003) found for the EU case using regional and country data. Dall'erba and Hewings (2003) show that there is convergence of Cohesion countries characterized by a catching-up of their income on the EU average; however it is also observed increasing regional disparities within each country. One conclusion is that the regional development policies in the EU that aims at reducing disparities at the EU aggregate level, has not impeded the process of increasing within-Cohesion country regional income inequalities. Complementarily to this evidence, Ramajo et al. (2008) results support the idea of a positive effect of the EU cohesion policies in fostering economic growth and convergence in the poorest EU members, but they do not investigate the trend of inequalities within these poorest EU countries. In the same direction, Mohl and Hagen (2010) suggest that Objective 1 payments in particular promote regional economic growth, whereas the total amount of

Objectives 1, 2, and 3 do not have a positive and significant impact on the EU regions' growth rates. In the Brazilian case analysed herein, the micro level positive effects may not be impeding increasing regional disparities within the Northeast region. As discussed in Chapter 3, some authors, such as Almeida Junior et al. (2007) and Oliveira and Domingues (2005), suggest that the resource allocation of the Brazilian regional development funds within each macro-region is guided by the demand side. In this sense, only entrepreneurs within the prosperous areas have contracted these loans which may be generating an increase of intra-regional inequalities (in terms of GDP per capita), i.e., the inequalities within the lagging macro-regions might be growing, or at least, are not being affected.

It is worth noting that the Brazilian regional funds have the broad objective of reducing regional inequalities as defined by federal law, but that no variable or measure of inequality was well defined when the policy was implemented. This study assumed that the reduction of the GDP per capita gap is the policy objective at the macro-scale level and that job creation and increasing labour productivity (proxied by wage growth) are the objectives at the firm (micro) level. The impact evaluation may be hampered by this lack of a precisely defined objective; it is always hard to define a measure for policy evaluation if it does not actually exist. Furthermore, it is important to highlight the fact that the results presented in this study cannot be extended to the FNE loans for the agricultural sector. The assessment of these loans may show completely different results given the features of the agricultural sector in the Northeast region, where FNE loans may have a substantial role in financing productive activities of individual and small farmers (especially those located in the “semi-árido” region). The results at the micro level discussed here focus exclusively on formal firms in the industrial and commerce/services sectors that could be traced in RAIS during the analysis period. In the same way, the macro-analysis of the FNE loans only includes loans to the industrial and commerce/services sectors, which represent approximately 40% of total loans. When more data (and with more quality) become available, the results can be reassessed to verify the robustness of the conclusions presented herein. Although the issues indicated above are somewhat problematic, they do not disqualify the impact evaluation carried out in this study. Instead, this evaluation should be seen as an important practical step towards formulating a framework to measure the micro- and macro-impacts of regional development policies in Brazil.

Finally, despite some recent changes in Brazilian regional policy (namely, the adoption of the National Regional Development Policy, PNDR, that was implemented by the Ministry for National Integration through Decree n. 6047 of 2007[118]), it is still necessary to define a relevant system for appraisal, monitoring and impact evaluation that covers all designed interventions at both the firm/individual and macro levels. Furthermore, it is important to demonstrate to public administrators and legislators the benefits and costs of more rigorous evaluations. As noted by Bartik and Bingham (1995), it is difficult to convince someone to do something (in this case, impact evaluation) that has not been done before. In addition, those authors argue that once policymakers have seen that high-quality evaluations of the regional development funds can help improve policy performance and political viability, the interest in impact evaluations should increase.

[118] Available at <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Decreto/D6047.htm>.

7. Concluding remarks

The contribution of this thesis was to shed light on the variability present in statistical results of regional economic growth estimates and of regional economic development policy evaluations when different spatial scales are used. Although there are serious problems associated with choosing only one spatial scale at which to conduct any regional analysis, the empirical literature tends to draw policy implications from results based on a single spatial scale. This thesis showed that this standard approach might be problematic and attempted to address this important research topic, thereby providing some guidance for future research on regional economic growth and regional development policy evaluations. Beyond the discussion of the statistical variations of the results when different levels of aggregation of data are used, the thesis also explores the substantive interpretation of such variations to provide some potential theoretical reasons for different results found across models estimated at different spatial scales.

First, Chapter 2 provides background discussion on theories and methods which form the basis for the analyses conducted in the following chapters. Regional economic development policies and growth literature are reviewed and the spatial scope of the economic growth theories is discussed including the potential theoretical reasons for different results found in the regional economic growth literature across models estimated at different spatial scales. Moreover, recent econometric issues related to regional economic growth is examined, which include the spatial econometrics literature and MAUP/EF.

Chapter 3 presents the Brazilian spatial scales for studying economic growth as well as reviewing the Brazilian empirical growth literature using these spatial scales. Furthermore, Brazilian socioeconomic information at different spatial scales is described. Finally, Chapter 3 reviews the justifications for regional economic development policies in Brazil and the evaluation literature addressing the Brazilian regional development funds. Indeed, Chapter 3 showed that these funds are rarely evaluated and discussed the policy process, objectives and the types of evaluation as well as reviewing the strategy of the Brazilian regional development funds since 1989.

Chapters 4 and 5 focused on the analysis of Brazilian economic growth at different spatial scales ranging from municipalities to state regions. Chapter 4 employed a cross-section from between the years 1991 and 2000 to study the determinants of regional economic growth in Brazil. It suggests a general framework for dealing with multiple spatial scales, spatial autocorrelation, spatial heterogeneity and model uncertainty. The spatial heterogeneity issue was addressed by grouping the Brazilian regions into different spatial regimes: a cluster of rich regions and a cluster of poor regions. This chapter showed that if a single regression is estimated on different scale levels, then the results will change with the scale level. The robustness tests were important in identifying variables that were simultaneously significant on different spatial scales: higher educational and health capital and better local infrastructure were related to higher rates of economic growth. This result suggests that because public policies (education, health and local infrastructure) are operating across all scale levels, these factors will be significant across spatial scales; although their impact on growth may differ across spatial scales as well as spatial regimes. Potential theoretical reasons were discussed to justify different results found across models estimated at different spatial scales. Moreover, the significance of the coefficients for the transportation costs of the spatial regime of poor regions at the municipal and micro-regional levels suggests that reductions in transportation costs had an impact only within the borders of the poor spatial clusters and the poor states in Brazil over the 1990s. Furthermore, the hypothesis of spatial club-convergence cannot be rejected, indicating different processes of convergence in Brazil. With regard to σ-convergence analysis, the results showed that per capita income distribution decreased within the spatial club of rich regions in Brazil between 1991 and 2000. On the other hand, this dispersion increased for the spatial club of poor regions. Finally, spatial autocorrelation only appeared at finer scales (the municipal and micro-regional levels). Of note, the results presented in this analysis are specific to the study period (1991-2000).

Chapter 5 attempted to expand the time horizon of the regional growth analysis presented in the previous chapter and investigate whether alternative non-spatial panel data models (that control for time invariant fixed effects) eliminate or, at least, mitigate the spatial autocorrelation in the growth estimates. In this sense, it addresses the omitted variable bias (OVB) problem often detected in the cross-sectional regressions carried out in Chapter 4 by controlling omitted variables that are constant over time in the form of individual effects. Of note, when the panel data framework is adopted in economic growth analyses, it creates a bridge between development economics and the neoclassical empirics of growth because this framework allows for unobservable differences in the production function which focus attention on all the tangible and intangible factors (e.g., institutional characteristics) that may enter into its respective individual effect (Islam, 1995). Then, the contribution of Chapter 5 was to explore the space and time dimensions of economic growth in Brazil using alternative panel-data techniques to provide a measure of the extent of spatial autocorrelation (in kilometres) over three decades (1970-2000). The analysis of the extent of residual spatial autocorrelation showed that it seems to vary across spatial scales. Indeed, spatial autocorrelation seems to be bounded at the state level because it appears that Moran's I statistic is not statistically significant for any choice of the spatial weight matrix, beginning from an 800-kilometer cut-off. On the other hand, although the spatial autocorrelation in residuals of the other three spatial scales shows positive and statistically significant values across distances of more than 1,500 kilometres, their levels are largely reduced when the distance is more than 900 kilometres, in particular for the 1970s and 1980s. Interestingly, an increasing clustering of the regressions' residuals over time was demonstrated, in particular over the 1990s period. The study suggests that the non-spatial panel-data techniques are not able to deal with spatially correlated omitted variables across different spatial scales, except for the state level where non-spatial panel data models seem to be appropriate to investigate growth determinants and convergence process in the Brazilian states case. To my knowledge, this is the first study of regional economic growth that explores both time and spatial scale dimensions. Both chapters shed new light on the MAUP (a measurement issue) when studying regional economic growth. However, in the context of the regional-growth literature, it is still necessary to develop a cross-level theory (i.e., a theory linking scale levels) to provide a better understanding of such variability in empirical results. The discussion provided in this thesis on potential explanations for the origin of this variability might be a first step in the conception of such a theoretical model. Moreover, future work should investigate economic growth in Brazil using these spatial scales and applying the recent spatial models of space-time econometrics discussed in Chapter 2. This investigation may shed light on the interpretation of the spatial autocorrelation found in the regressions of Chapter 5. Indeed, Resende et al. (2012) have a work in progress called “Evaluating Multiple Spatial Dimensions of Economic Growth in Brazil using Spatial Dynamic Panel Data Models, 1970-2000” seeking to investigate this issue.

Next, the thesis turned to the discussion of regional economic development policies, a topic that is intrinsically related to regional economic growth. Interestingly, Armstrong (2002) points out some obstacles that prevent the synthesis of evidence of economic growth in the European Union (EU) with evaluation research on the EU-Structural Funds. Such obstacles include the variety of economic growth theories that can be used to evaluate regional policy, data limitation for the conditioning variables, data limitation for policy variables and lack of policy maturity. Chapter 5 was a first step in constructing a framework to measure the micro- and macro-impacts of regional development funds in Brazil. Of note, this approach brings together two levels of analysis that are often implemented separately in the impact-evaluation literature and thus allows a more complete assessment of the FNE loans. These micro- and macro-effects have been overlooked in the literature that addresses the impacts of regional development funds. Again, the variability of the results across different levels of analysis was investigated. Specifically, the chapter evaluated the impact of the FNE industrial loans on employment and labour productivity growth at the micro (firm) level and on GDP per capita growth at macro (municipalities, micro-regions and spatial clusters) levels for the 2000-2003 and 2000-2006 periods. The empirical investigation focuses on the FD regression approach because it has the ability to control for observable and time-invariant unobservable characteristics. The results show a positive and statistically significant impact of the FNE industrial loans on job creation at the micro level. However, no significant impact on the GDP per capita growth at the macro level was found. At least in the short term, this is probably because of a balance between countervailing effects: the negative effect of employment growth on average productivity and the positive impact of employment growth on GDP per capita growth via increases in total income. Future research should update the micro- and macro-analysis focusing in a more recent period, 2004-2011 when the amount of resources allocated to all regional development funds (FNE, FNO and FCO) grew. This fact may have positive impacts on regional development at the different spatial levels. Moreover, loans to the rural sector should be assessed in future evaluations if data is available. Other studies using firm level data can be conducted to investigate funds' performance on other variables and on different groups of workers. Constant evaluations of the impacts of these regional development funds can inform how this policy can be improved to deliver better results both at micro and macro levels. Finally, as previously discussed in detail throughout this thesis, the Brazilian regional development funds have the broad objective of reducing regional inequalities (as defined by federal law), but no variable or measure of inequality was well defined when the policy was implemented. Therefore, it is still necessary that the Brazilian government define more precise targets for the regional development funds. Such a definition would be a fundamental step towards the implementation of a formal system of appraising, monitoring and evaluating the outcomes of all designed interventions of the regional development funds. I would suggest some steps to improve and overcome the lack of studies on regional development policy evaluation in Brazil. In the short-term, the government should make disaggregate data of resource allocation of the regional development funds available to the public. In the mid-term, it is important to demonstrate to the public administrators and legislators the benefits and costs of more rigorous outcome evaluations. As noted by Bartik and Bingham (1995), it is difficult to get people to do something that has not been done before. In addition, they argue that once "*policy makers have seen that a high quality evaluation of economic development programs can help improve the programs's performance and political viability, the interest in economic development evaluations should increase*" (Bartik and Bingham, 1995: 26). In the long-term, it is necessary to begin a wide debate about the actual causes of regional inequalities in Brazil and discuss the formulation of instruments to best deal with them. The discussion of the balance between policies place- and people-based should be at the centre of this debate for the Brazilian context. Concerning this debate, the Brazilian regional development policy has to be informed by the rethinking of the space-neutral versus the place-based approaches (Barca et al., 2012). Much empirical work should be done for the Brazilian case and then pros and cons of such policies need to be clear.

Appendices

APPENDIX I

APPENDIX I.A

Table I.A.1
Correlation matrix of the explanatory variables (states)

	Explanatory variables	1	2	3	4	5	6	7
1	local infrastructure in 1991	1	0.83	0.88	-0.14	-0.75	-0.78	0.62
2	ln(income per capita in 1991)	0.83	1	0.97	0.22	-0.87	-0.59	0.20
3	ln(average years of schooling in 1991)	0.88	0.97	1	0.16	-0.86	-0.59	0.30
4	ln(Gini index in 1991)	-0.14	0.22	0.16	1	0.01	0.38	-0.33
5	ln(infant mortality rate in 1991)	-0.75	-0.87	-0.86	0.01	1	0.62	-0.12
6	ln(transportation cost to SP in 1991)	-0.78	-0.59	-0.59	0.38	0.62	1	-0.65
7	ln(population density in 1991)	0.62	0.20	0.30	-0.33	-0.12	-0.65	1

Own elaboration.

Table I.A.2
Correlation matrix of the explanatory variables (municipalities)

	Explanatory variables	1	2	3	4	5	6	7
1	local infrastructure in 1991	1	0.83	0.82	0.08	-0.69	-0.76	0.48
2	ln(income per capita in 1991)	0.83	1	0.86	0.24	-0.78	-0.68	0.23
3	ln(average years of schooling in 1991)	0.82	0.86	1	0.27	-0.75	-0.62	0.31
4	ln(Gini index in 1991)	0.08	0.24	0.27	1	-0.08	-0.01	-0.12
5	ln(infant mortality rate in 1991)	-0.69	-0.78	-0.75	-0.08	1	0.64	-0.12
6	ln(transportation cost to SP in 1991)	-0.76	-0.68	-0.62	-0.01	0.64	1	-0.37
7	ln(population density in 1991)	0.48	0.23	0.31	-0.12	-0.12	-0.37	1

Own elaboration.

Table I.A.3
Correlation matrix of the explanatory variables (micro-regions)

	Explanatory variables	1	2	3	4	5	6	7
1	local infrastructure in 1991	1	0.86	0.87	-0.02	-0.77	-0.78	0.55
2	ln(income per capita in 1991)	0.86	1	0.92	0.15	-0.86	-0.67	0.26
3	ln(average years of schooling in 1991)	0.87	0.92	1	0.20	-0.82	-0.63	0.35
4	ln(Gini index in 1991)	-0.02	0.15	0.20	1	-0.08	0.10	-0.28
5	ln(infant mortality rate in 1991)	-0.77	-0.86	-0.82	-0.08	1	0.67	-0.15
6	ln(transportation cost to SP in 1991)	-0.78	-0.67	-0.63	0.10	0.67	1	-0.49
7	ln(population density in 1991)	0.55	0.26	0.35	-0.28	-0.15	-0.49	1

Own elaboration.

Table I.A.4
Correlation matrix of the explanatory variables (spatial clusters)

	Explanatory variables	1	2	3	4	5	6	7
1	ln(local infrastructure in 1991)	1	0.88	0.90	0.13	-0.74	-0.68	0.73
2	ln(income per capita in 1991)	0.88	1	0.95	0.33	-0.80	-0.55	0.54
3	ln(average years of schooling in 1991)	0.90	0.95	1	0.36	-0.77	-0.50	0.61
4	ln(Gini index in 1991)	0.13	0.33	0.36	1	-0.10	0.24	-0.10
5	ln(infant mortality rate in 1991)	-0.74	-0.80	-0.77	-0.10	1	0.62	-0.34
6	ln(transportation cost to SP in 1991)	-0.68	-0.55	-0.50	0.24	0.62	1	-0.58
7	ln(population density in 1991)	0.73	0.54	0.61	-0.10	-0.34	-0.58	1

Own elaboration.

APPENDIX I.B

Table I.B.1

Robustness test at state level

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: OLS								
Explanatory variables	**(Scale level: States)**							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(income per capita in 1991)	-0.0672***	-0.0595***	-0.0697***	-0.0509**	-0.0684***	-0.0647***	-0.0612***	-0.0386***
	(0.0177)	(0.0188)	(0.0210)	(0.0205)	(0.0179)	(0.0191)	(0.0090)	(0.0091)
ln(average years of schooling in 1991)					0.0144	0.0243		
					(0.0305)	(0.0323)		
ln(Gini index in 1991)	0.1349**	0.1131*	0.1508**	0.0864	0.1270**	0.1060*	0.1237**	0.0999
	(0.0596)	(0.0635)	(0.0707)	(0.0687)	(0.0530)	(0.0557)	(0.0513)	(0.0661)
ln(infant mortality rate in 1991)	-0.0243*				-0.0236*		-0.0246*	
	(0.0122)				(0.0123)		(0.0119)	
ln(transportation cost to SP in 1991)	-0.0054	-0.0044	-0.0123*		-0.0062	-0.0063	-0.0044	-0.0093
	(0.0056)	(0.0061)	(0.0060)		(0.0063)	(0.0068)	(0.0048)	(0.0059)
ln(population density in 1991)	0.0062**	0.0061**			0.0060**	0.0057**	0.0066***	
	(0.0022)	(0.0024)			(0.0022)	(0.0024)	(0.0019)	
local infrastructure in 1991	0.0024	0.0030	0.0107	0.0064				
	(0.0060)	(0.0065)	(0.0066)	(0.0068)				
constant	0.5721***	0.4178**	0.5654***	0.3284**	0.5569***	0.4221***	0.5257***	0.3477***
	(0.1638)	(0.1566)	(0.1666)	(0.1305)	(0.1330)	(0.1213)	(0.1127)	(0.1037)
regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	27	27	27	27	27	27	27	27
Adjusted R-squared	0.7367	0.6906	0.5948	0.5254	0.7377	0.6968	0.7497	0.5595
Diagnostic for spatial dependence								
Lagrange Multiplier-Lag	0.5477	1.6069	1.5058	0.2660	0.6031	1.6987	0.5137	1.2771
Robust Lagrange Multiplier-Lag	0.0211	0.4986	0.1561	0.5000	0.0004	0.7203	0.0129	0.1800
Lagrange Multiplier-Error	0.9958	1.1082	1.3978	1.0228	0.9124	0.9979	0.9073	1.1203
Robust Lagrange Multiplier-Error	0.4691	0.00002	0.0481	1.2568	0.3097	0.0195	0.4065	0.0232

Note: Standard errors in parentheses; *** significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991].

Table I.B.2

Robustness test at municipal level

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: ML								
Explanatory variables	**(Scale level: Municipalities)**							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(income per capita in 1991)	-0.0557***	-0.0526***	-0.0511***	-0.0506***	-0.0612***	-0.0590***	-0.0402***	-0.0295***
	(0.0012)	(0.0013)	(0.0013)	(0.0013)	-0.0012	(0.0012)	(0.0013)	(0.0010)
ln(average years of schooling in 1991)					0.0419***	0.0435***		
					(0.0013)	(0.0013)		
ln(Gini index in 1991)	0.0005	-0.0040	-0.0066**	-0.0080***	-0.0171***	-0.0211***	-0.0034	-0.0155***
	(0.0011)	(0.0030)	(0.0030)	(0.0030)	(0.0028)	(0.0028)	(0.0044)	(0.0031)
ln(infant mortality rate in 1991)	-0.0122***				-0.0098***		-0.0146***	
	(0.0011)				(0.0011)		(0.0012)	
ln(transportation cost to SP in 1991)	-0.0033***	-0.0042***	-0.0050***		-0.0079***	-0.0087***	-0.0079***	-0.0119***
	(0.0010)	(0.0010)	(0.0010)		(0.0009)	(0.0009)	(0.0010)	(0.0010)
ln(population density in 1991)	0.0025***	0.0024***			0.0016***	0.0015***	0.0046***	
	(0.0003)	(0.0004)			(0.0003)	(0.0003)	(0.0004)	
local infrastructure in 1991	0.0090***	0.0095***	0.0103***	0.0108***				
	(0.0004)	(0.0004)	(0.0004)	(0.0004)				
lambda (λ)	0.6529***	0.6692***	0.6562***	0.6558***	0.6494***	0.6631***	0.6380***	0.6406***
	(0.0160)	(0.0154)	(0.0154)	(0.0152)	(0.0158)	(0.0153)	(0.0164)	(0.0155)
constant	0.3601***	0.3040***	0.3069***	0.2695***	0.3576***	0.3124***	0.3208***	0.2482***
	(0.0100)	(0.0104)	(0.0103)	(0.0072)	(0.0107)	(0.0097)	(0.0133)	(0.0104)
regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	5,507	5,507	5,507	5,507	5,507	5,507	5,507	5,507
Log-likelihood	14335.8	14278.7	14255.0	14241.7	14586.8	14546.5	14106.4	13942.8
Likelihood Ratio test (LR)	1157.1***	1310.1***	1272.2***	1304.2***	1172.4***	1289.3***	1072.1***	1208.4***

Note: Asymptotic standard errors in parentheses; *** significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991]. The spatial weight matrix W is based on the 10-nearest neighbours.

Table I.B.3

Robustness test at micro-regional level

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: ML								
Explanatory variables	**(Scale level: Micro-regions)**							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(income per capita in 1991)	-0.0245***	-0.0207***	-0.0206***	-0.0206***	-0.0356***	-0.0328***	-0.0160***	-0.0080***
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0030)	(0.0030)	(0.0024)	(0.0017)
ln(average years of schooling in 1991)					0.0332***	0.0357***		
					(0.0035)	(0.0035)		
ln(Gini index in 1991)	0.0031	-0.0021	-0.0024	-0.0024	-0.0131	-0.0184	0.0030	-0.0050
	(0.0086)	(0.0087)	(0.0086)	0.0086	(0.0083)	(0.0083)	(0.0087)	(0.0088)
ln(infant mortality rate in 1991)	-0.0156***				-0.0121***		-0.0175***	
	(0.0033)				(0.0031)		(0.0033)	
ln(transportation cost to SP in 1991)	0.0003	0.00004	-0.00005		-0.0027	-0.0032	-0.0009	-0.0021
	(0.0026)	(0.0027)	(0.0027)		(0.0023)	(0.0024)	(0.0026)	(0.0026)
ln(population density in 1991)	0.0004	0.0002			-0.0008	-0.0011	0.0011	
	(0.0007)	(0.0007)			(0.0007)	(0.0007)	(0.0007)	
local infrastructure in 1991	0.0046***	0.0055***	0.0056***	0.0056***				
	(0.0012)	(0.0012)	(0.0012)	(0.0012)				
lambda (λ)	0.7695***	0.7787***	0.7783***	0.7783***	0.7479***	0.7542***	0.7601***	0.7686***
	(0.0387)	(0.0376)	(0.0376)	(0.0376)	(0.0412)	(0.0404)	(0.0398)	(0.0388)
constant	0.2046***	0.1257***	0.1259***	0.1255***	0.2206***	0.1583***	0.1762***	0.0761***
	(0.0317)	(0.0278)	(0.0277)	(0.0181)	(0.0290)	(0.0248)	(0.0310)	(0.0259)
regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	559	559	559	559	559	559	559	559
Log-likelihood	1697.9	1686.7	1686.6	1686.6	1732.7	1725.0	1690.8	1675.7
Likelihood Ratio test (LR)	167.9***	176.6***	177.5***	179.1***	143.9***	147.2***	160.6***	169.2***

Note: Asymptotic standard errors in parentheses; *** significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991]. The spatial weight matrix W is based on the 10-nearest neighbours.

Table I.B.4

Robustness test at spatial cluster level

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: OLS								
Explanatory variables	**(Scale level: Spatial Clusters)**							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(income per capita in 1991)	-0.0544***	-0.0508***	-0.0472***	-0.0469***	-0.0653***	-0.0626***	-0.0305***	-0.0114**
	(0.0102)	(0.0100)	(0.0098)	(0.0095)	(0.0091)	(0.0088)	(0.0077)	(0.0049)
ln(average years of schooling in 1991)					0.0694***	0.0729***		
					(0.0126)	(0.0122)		
ln(Gini index in 1991)	0.0300	0.0236	0.0118	0.0095	-0.0415	-0.0500	0.0194	-0.0125
	(0.0333)	(0.0332)	(0.0328)	(0.0296)	(0.0321)	(0.0312)	(0.0352)	(0.0352)
ln(infant mortality rate in 1991)	-0.0170				-0.0123		-0.0270**	
	(0.0120)				(0.0109)		(0.0123)	
ln(transportation cost to SP in 1991)	0.00001	0.0006	-0.0007		-0.0036	-0.0034	-0.0003	-0.0012
	(0.0042)	(0.0042)	(0.0042)		(0.0039)	(0.0039)	(0.0045)	(0.0046)
ln(population density in 1991)	0.0031*	0.0028			-0.0001	-0.0005	0.0042**	
	(0.0017)	(0.0017)			(0.0017)	(0.0017)	(0.0017)	
local infrastructure in 1991	0.0115***	0.0127***	0.0137***	0.0137***				
	(0.0035)	(0.0034)	(0.0034)	(0.0033)				
constant	0.3759***	0.2845***	0.2758***	0.2670***	0.3120***	0.2404***	0.2817***	0.0778
	(0.1007)	(0.0780)	(0.0786)	(0.0580)	(0.0879)	(0.0608)	(0.1024)	(0.0675)
regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	91	91	91	91	91	91	91	91
Adjusted R-squared	0.3174	0.3088	0.2941	0.3024	0.4373	0.4355	0.2333	0.1608
<u>Diagnostic for spatial dependence</u>								
Lagrange Multiplier-Lag	2.1173	1.7848	2.6090	2.5382	1.8981	1.5697	1.0161	1.2343
Robust Lagrange Multiplier-Lag	0.4585	0.0726	0.7341	0.6980	0.0743	0.0713	0.5727	2.6631
Lagrange Multiplier-Error	3.5713*	2.4710	1.9252	1.8830	2.9404*	2.4790	1.9338	0.3723
Robust Lagrange Multiplier-Error	1.9125	0.7588	0.0503	0.0428	1.1166	0.9806	1.4904	1.8011

Note: Standard errors in parentheses; *** significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991].

APPENDIX I.C

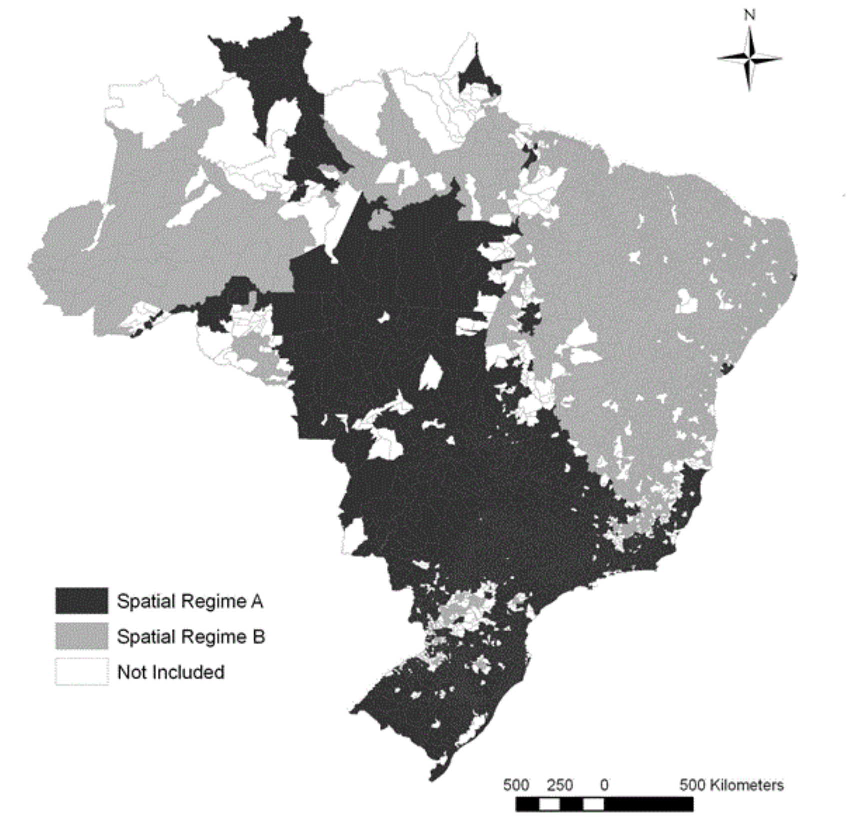
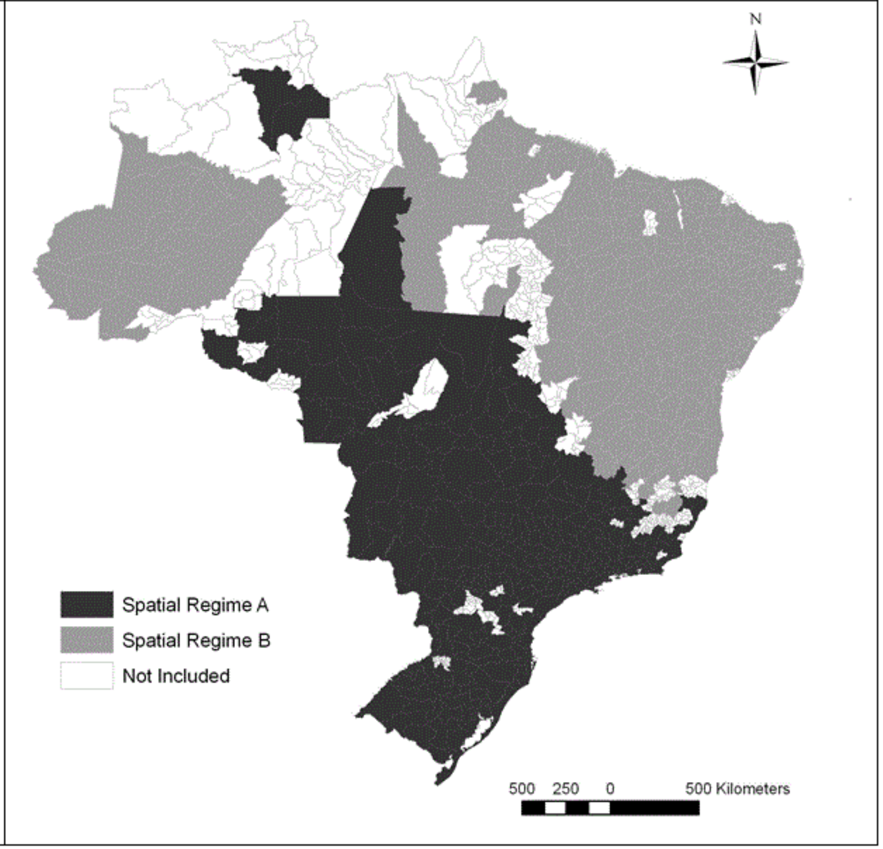
Table I.C.1

Moran's *I* for the residuals of municipal and micro-regional specifications

Scale level	k-nearest neighbours	k=5	k=10	k=15	k=30	k=60	k=120
Micro-regions	Moran's I	0.3205	0.2795	0.2186	0.1243	0.0346	0.0062
(559 units)	p-value*	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0004)	(0.0558)
Municipalities	Moran's I	0.3015	0.2787	0.2601	0.2309	0.1876	0.1268
(5,507 units)	p-value*	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)

Note: Based on the estimation of OLS Baseline Estimation for Diagnostics of Spatial Dependence (Table 2.1). Explanatory variables included in the model: ln(income per capita in 1991), ln(average years of schooling in 1991), ln(Gini index in 1991), ln(infant mortality rate in 1991), ln(transport cost to SP in 1991), ln(population density in 1991), local infrastructure in 1991 and regional dummies.
*P-values are based on the permutation approach with ten thousand permutations.

APPENDIX I.D

Figure I.D.1 - Two spatial regimes at the municipal level in the initial income per capita identified by means of the Moran scatterplot (with t=1991, k=10-nearest)	**Figure I.D.2 - Two spatial regimes at the micro-regional level in the initial income per capita identified by means of the Moran scatterplot (with t=1991, k=10-nearest)**
	

Note: Own elaboration.

Table I.D.1

Robustness test at municipal level for the two regimes spatial error model

Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: ML									
Spatial Regime	**Explanatory variables**	**(Scale level: Municipalities)**							
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spatial Regime A (High-High)	ln(income per capita in 1991)	-0.0468***	-0.0434***	-0.0408***	-0.0403***	-0.0650***	-0.0628***	-0.0383***	-0.0284***
		(0.0019)	(0.0019)	(0.0019)	(0.0018)	(0.0020)	(0.0020)	(0.0018)	(0.0016)
	ln(average years of schooling in 1991)					0.0662***	0.0677***		
						(0.0029)	(0.0029)		
	ln(Gini index in 1991)	-0.0095**	-0.0158***	-0.0198***	-0.0206***	-0.0168***	-0.0215***	-0.0157***	-0.0305***
		(0.0043)	(0.0044)	(0.0044)	(0.0044)	(0.0041)	(0.0041)	(0.0045)	(0.0044)
	ln(infant mortality rate in 1991)	-0.0100***				-0.0077***		-0.0110***	
		(0.0015)				(0.0014)		(0.0016)	
	ln(transport cost to SP in 1991)	-0.0007	-0.0005	-0.0013		-0.0025**	-0.0022*	-0.0017	-0.0030**
		(0.0015)	(0.0015)	(0.0015)		(0.0012)	(0.0012)	(0.0013)	(0.0013)
	ln(population density in 1991)	0.0024***	0.0022***			-0.0005	-0.0008*	0.0034***	
		(0.0004)	(0.0005)			(0.0004)	(0.0004)	(0.0005)	
	local infra-structure in 1991	0.0061***	0.0063***	0.0071***	0.0072***				
		(0.0006)	(0.0006)	(0.0006)	(0.0006)				
	Constant	0.2768***	0.2158***	0.2104***	0.1977***	0.2939***	0.2465***	0.2362***	0.1503***
		(0.0168)	(0.0166)	(0.0173)	(0.0113)	(0.0178)	(0.0149)	(0.0184)	(0.0158)
Spatial Regime B (Low-Low)	ln(income per capita in 1991)	-0.0680***	-0.0656***	-0.0656***	-0.0649***	-0.0690***	-0.0677***	-0.0483***	-0.0406***
		(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0020)	(0.0020)	(0.0020)	(0.0020)
	ln(average years of schooling in 1991)					0.0352***	0.0363***		
						(0.0015)	(0.0015)		
	ln(Gini index in 1991)	0.0090**	0.0088**	0.0084*	0.0090**	-0.0052	-0.0059	0.0141***	0.0117**
		(0.0044)	(0.0045)	(0.0046)	(0.0045)	(0.0041)	(0.0042)	(0.0046)	(0.0048)
	ln(infant mortality rate in 1991)	-0.0123***				-0.0083***		-0.0148***	
		(0.0019)				(0.0018)		(0.0020)	
	ln(transport cost to SP in 1991)	-0.0146***	-0.0150***	-0.0150***		-0.0193***	-0.0197***	-0.0197***	-0.0218***
		(0.0026)	(0.0027)	(0.0027)		(0.0025)	(0.0024)	(0.0027)	(0.0026)
	ln(population density in 1991)	0.0010*	0.0009			0.0016***	0.0015***	0.0049***	
		(0.0006)	(0.0006)			(0.0005)	(0.0005)	(0.0006)	

local infra-structure in 1991	0.0122***	0.0125***	0.0128***	0.0133***				
	(0.0007)	(0.0007)	(0.0006)	(0.0006)				
Constant	0.5166***	0.4604***	0.4613***	0.3380***	0.4839***	0.4466***	0.4562***	0.3833***
	(0.0261)	(0.0251)	(0.0253)	(0.0119)	(0.0249)	(0.0225)	(0.0268)	(0.0254)
lambda (λ)	0.6077***	0.6273***	0.6229***	0.6343***	0.5900***	0.5999***	0.5749***	0.5772***
	(0.0178)	(0.0180)	(0.0177)	(0.0170)	(0.0187)	(0.0184)	(0.0197)	(0.0189)
Regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	4,896	4,896	4,896	4,896	4,896	4896	4,896	4,896
Log-likelihood	12848.24	12807.27	12795.14	12779.33	13107.70	13083.57	12639.47	12527.95
Likelihood Ratio test (LR)	768.30***	877.44***	888.16***	975.60***	741.57***	789.52***	649.58***	709.23***

Note: Asymptotic standard errors in parentheses; ***significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991]. The spatial weight matrix W is based on the 10-nearest neighbours.

Table I.D.2

Robustness test at micro-regional level for the two regimes spatial error model

Spatial Regime	Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: income per capita growth between 1991 and 2000 - Estimation method: ML								
		(Scale level: Micro-regions)							
Spatial Regime A (High-High)	ln(income per capita in 1991)	-0.0361***	-0.0333***	-0.0306***	-0.0293***	-0.0497***	-0.0467***	-0.0300***	-0.0196***
		(0.0044)	(0.0043)	(0.0041)	(0.0041)	(0.0045)	(0.0044)	(0.0036)	(0.0027)
	ln(average years of schooling in 1991)					0.0447***	0.0483***		
						(0.0073)	(0.0073)		
	ln(Gini index in 1991)	0.0011	-0.0073	-0.0141	-0.0114	-0.0106	-0.0201	-0.0004	-0.0207
		(0.0138)	(0.0137)	(0.0133)	(0.0132)	(0.0129)	(0.0126)	(0.0140)	(0.0135)
	ln(infant mortality rate in 1991)	-0.0124***				-0.0112***		-0.0149***	
		(0.0041)				(0.0037)		(0.0040)	
	ln(transport cost to SP in 1991)	0.0032	0.0035	0.0023		0.0010	0.0013	0.0018	
		(0.0025)	(0.0026)	(0.0026)		(0.0021)	(0.0021)	(0.0024)	0.0003
	ln(population density in 1991)	0.0022**	0.0018*			0.0008	0.0004	0.0026***	(0.0025)
		(0.0009)	(0.0009)			(0.0009)	(0.0009)	(0.0009)	
	local infra-structure in 1991	0.0044**	0.0056***	0.0061***	0.0055***				
		(0.0018)	(0.0018)	(0.0018)	(0.0017)				
	Constant	0.2276***	0.1592***	0.1535***	0.1678***	0.2479***	0.1789***	0.2127***	0.1076***
		(0.0378)	(0.0307)	(0.0308)	(0.0241)	(0.0351)	(0.0268)	(0.0378)	(0.0285)
Spatial Regime B (Low-Low)	ln(income per capita in 1991)	-0.0283***	-0.0259***	-0.0258***	-0.0260***	-0.0349***	-0.0340***	-0.0159***	-0.0101***
		(0.0051)	(0.0050)	(0.0050)	(0.0050)	(0.0044)	(0.0044)	(0.0038)	(0.0034)
	ln(average years of schooling in 1991)					0.0274***	0.0287***		
						(0.0040)	(0.0039)		
	ln(Gini index in 1991)	0.0261**	0.0256**	0.0253**	0.0287**	0.0141	0.0137	0.0295**	0.0257**
		(0.0127)	(0.0129)	(0.0128)	(0.0128)	(0.0123)	(0.0124)	(0.0130)	(0.0132)
	ln(infant mortality rate in 1991)	-0.0114**				-0.0062		-0.0133***	
		(0.0047)				(0.0045)		(0.0048)	
	ln(transport cost to SP in 1991)	-0.0136***	-0.0136***	-0.0129***		-0.0169***	-0.0170***	-0.0133***	-0.0143***
		(0.0046)	(0.0047)	(0.0046)		(0.0042)	(0.0043)	(0.0047)	(0.0047)
	ln(population density in 1991)	0.00002	-0.0002			-0.0006	-0.0008	0.0018*	
		(0.0012)	(0.0012)			(0.0010)	(0.0010)	(0.0010)	

local infra-structure in 1991	0.0061***	0.0067***	0.0066***	0.0072***				
	(0.0017)	(0.0017)	(0.0016)	(0.0016)				
Constant	0.3308***	0.2740***	0.2675***	0.1661***	0.3272***	0.2983***	0.2716***	0.2007***
	(0.0525)	(0.0475)	(0.0457)	(0.0274)	(0.0465)	(0.0412)	(0.0507)	(0.0443)
lambda (λ)	0.6434***	0.6602***	0.6748***	0.6941***	0.5513***	0.5524***	0.6260***	0.6525***
	(0.0533)	(0.0515)	(0.0500)	(0.0479)	(0.0622)	(0.0621)	(0.0550)	(0.0523)
Regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	495	495	495	495	495	495	495	495
Log-likelihood	1568.66	1561.43	1559.79	1555.85	1597.48	1592.06	1559.34	1545.14
Likelihood Ratio test (LR)	85.61***	92.73***	103.01***	114.33***	51.36***	50.07***	76.43***	89.46***

Note: Asymptotic standard errors in parentheses; ***significant at 1%; ** significant at 5%; * significant at 10%. Dependent variable = (1/9)*ln[incomepercapita_in_2000/incomepercapita_in_1991]. The spatial weight matrix W is based on the 10-nearest neighbours.

APPENDIX II

APPENDIX II.A

Table II.A.1
Correlation matrix of the explanatory variables (Minimum Comparable Areas, MCAs)

Explanatory variables		1	2	3	4	5
1	ln (income per capita)	1	0.18	0.68	0.22	-0.69
2	(n + g + d)	0.18	1	0.08	0.07	-0.05
3	ln (average years of schooling)	0.68	0.08	1	0.21	-0.54
4	ln (population density)	0.22	0.07	0.21	1	-0.30
5	ln (transportation cost to SP)	-0.69	-0.05	-0.54	-0.30	1

Note: Own elaboration. Number of observations = 10,971.

Table II.A.2
Correlation matrix of the explanatory variables (Micro-regions)

Explanatory variables		1	2	3	4	5
1	ln (income per capita)	1	0.16	0.85	0.26	-0.67
2	(n + g + d)	0.16	1	0.06	-0.21	0.07
3	ln (average years of schooling)	0.85	0.06	1	0.34	-0.63
4	ln (population density)	0.26	-0.21	0.34	1	-0.47
5	ln (transportation cost to SP)	-0.67	0.07	-0.63	-0.47	1

Note: Own elaboration. Number of observations = 1,566.

Table II.A.3
Correlation matrix of the explanatory variables (Meso-regions)

Explanatory variables		1	2	3	4	5
1	ln (income per capita)	1	0.13	0.87	0.29	-0.62
2	(n + g + d)	0.13	1	0.02	-0.37	0.24
3	ln (average years of schooling)	0.87	0.02	1	0.38	-0.59
4	ln (population density)	0.29	-0.37	0.38	1	-0.54
5	ln (transportation cost to SP)	-0.62	0.24	-0.59	-0.54	1

Note: Own elaboration. Number of observations = 402.

Table II.A.4
Correlation matrix of the explanatory variables (States)

Explanatory variables		1	2	3	4	5
1	ln (income per capita)	1	0.03	0.88	0.21	-0.55
2	(n + g + d)	0.03	1	-0.10	-0.60	0.44
3	ln (average years of schooling)	0.88	-0.10	1	0.27	-0.52
4	ln (population density)	0.21	-0.60	0.27	1	-0.69
5	ln (transportation cost to SP)	-0.55	0.44	-0.52	-0.69	1

Note: Own elaboration. Number of observations = 81.

Table II.A.5

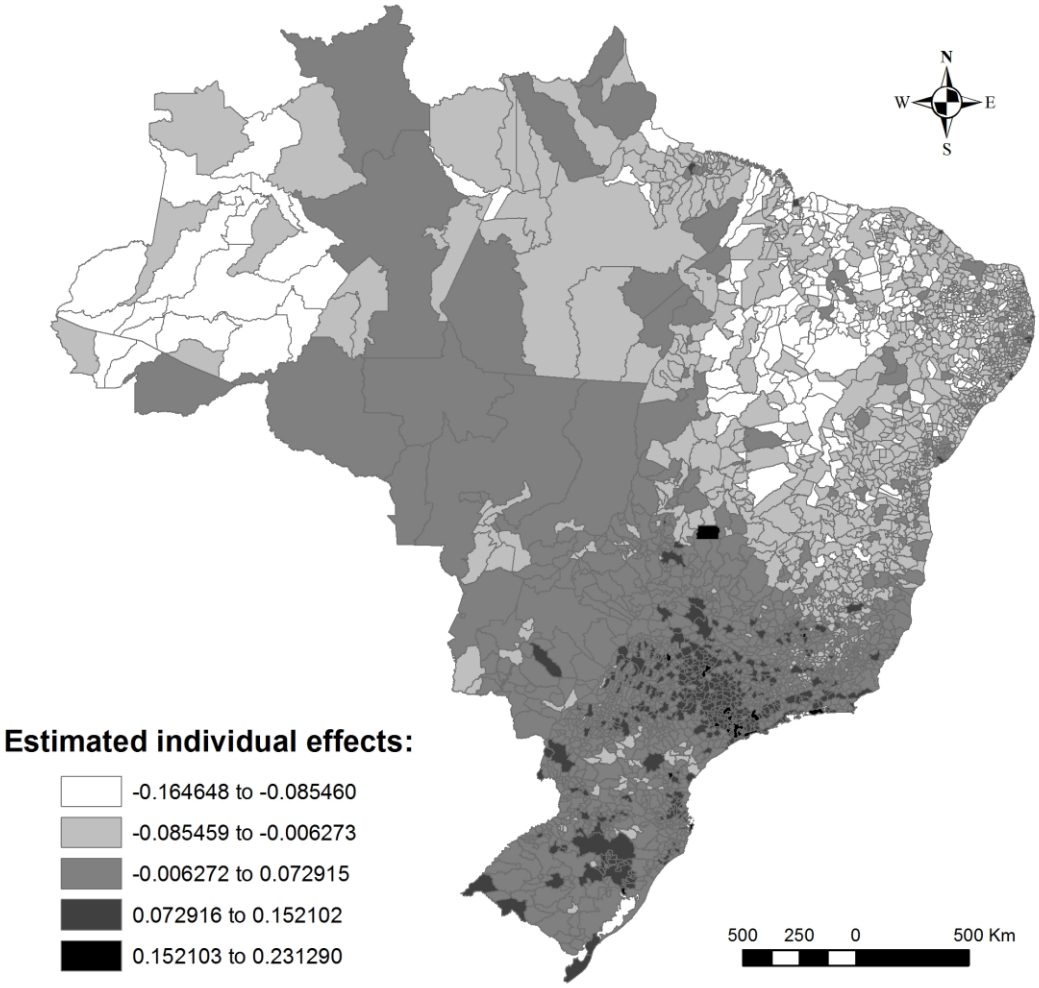
Summary statistics of the variables in the panel (1970-2000)

Variables		**States**	**Meso-regions**	**Micro-regions**	**MCAs***
Average annual income per capita growth rates between 1970-1980, 1980-1991 and 1991-2000.	Mean	0.0473	0.0475	0.0473	0.0483
	Minimum	-0.0266	-0.0756	-0.0786	-0.1934
	Maximum	0.1088	0.1329	0.1558	0.3709
	Std. dev.	0.0422	0.0467	0.0485	0.0541
ln (income per capita), 1970, 1980 and 1991.	Mean	4.6197	4.4842	4.3288	4.1966
	Minimum	3.2472	2.9627	2.5869	0.2718
	Maximum	5.8605	5.9564	6.0514	6.2442
	Std. dev.	0.5981	0.6674	0.6955	0.7041
(n+g+d), 1970, 1980 and 1991.	Mean	0.0778	0.0701	0.0656	0.0594
	Minimum	0.0581	0.0238	-0.0221	-0.1197
	Maximum	0.1986	0.1986	0.3220	0.3220
	Std. dev.	0.0210	0.0180	0.0197	0.0212
ln (years of schooling), 1970, 1980 and 1991.	Mean	0.9977	0.8516	0.7102	0.5803
	Minimum	-0.1039	-0.7391	-1.3856	-46.0517
	Maximum	2.0015	2.0015	2.0015	2.1748
	Std. dev.	0.4890	0.6064	0.6533	0.8238
ln(population density), 1970, 1980 and 1991.	Mean	2.5834	2.9085	3.0423	3.2179
	Minimum	-1.7278	-2.2713	-3.9118	-3.9118
	Maximum	5.6895	7.4888	8.5627	9.4510
	Std. dev.	1.7605	1.6641	1.4901	1.2664
ln (transportation cost to SP), 1970, 1980 and 1991.	Mean	7.7425	7.4226	7.3370	7.2085
	Minimum	5.3315	4.6706	3.7987	2.3026
	Maximum	9.2782	9.5446	9.6053	9.6175
	Std. dev.	0.8386	0.9197	0.9345	0.8734
Observations		81	402	1,566	10,971

Note: Own elaboration. * Minimum Comparable Areas (MCAs).

Figure II.A.1

Spatial distribution of estimated fixed-effects at MCA level



Note: In the map, the ranges were defined using equal intervals. Estimated individual effects come from the FE estimation at Minimum Comparable Area (MCA) in Table 5.2 (Panel Data Model Results and Diagnostics for Spatial Autocorrelation), column (2).

APPENDIX II.B

Table II.B.1
Diagnostics for spatial autocorrelation by means of Moran's I in the cross-sectional residuals of the estimations of Table 5.2

Spatial scale	Minimum Comparable Area (MCA)				Micro-region				Meso-region				State			
Method	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Spatial Weight Matrix: 400km cut-off																
Moran's I in the residuals (70's)	0.1204***	0.1306***	-	0.2238***	-	-	-	-	-	-	-	-	-	-	-	-
Moran's I in the residuals (80's)	0.1106***	0.0677***	0.1122***	0.1729***	-	-	-	-	-	-	-	-	-	-	-	-
Moran's I in the residuals (90's)	0.1033***	0.1286***	0.1104***	0.1963***	-	-	-	-	-	-	-	-	-	-	-	-
Spatial Weight Matrix: 500km cut-off																
Moran's I in the residuals (70's)	0.1046***	0.1069***	-	0.1897***	0.1546***	0.1951***	-	0.2065***	-	-	-	-	-	-	-	-
Moran's I in the residuals (80's)	0.0855***	0.0471***	0.0863***	0.1431***	0.1296***	0.1458***	0.1399***	0.1757***	-	-	-	-	-	-	-	-
Moran's I in the residuals (90's)	0.0741***	0.1039***	0.0867***	0.1608***	0.2291***	0.2572***	0.2463***	0.1861***	-	-	-	-	-	-	-	-
Spatial Weight Matrix: 600km cut-off																
Moran's I in the residuals (70's)	0.0890***	0.0857***	-	0.1605***	0.1252***	0.1511***	-	0.1623***	0.1407***	0.1388***	-	0.3561***	-	-	-	-
Moran's I in the residuals (80's)	0.0661***	0.0300***	0.0668***	0.1195***	0.0938***	0.1116***	0.1075***	0.1498***	0.1634***	0.2643***	0.1394***	0.4788**	-	-	-	-
Moran's I in the residuals (90's)	0.0482***	0.0800***	0.0641***	0.1295***	0.1955***	0.2117***	0.2033***	0.1409***	0.4009***	0.3118***	0.3740***	0.4324***	-	-	-	-
Spatial Weight Matrix: 700km cut-off																
Moran's I in the residuals (70's)	0.0771***	0.0687***	-	0.1418***	0.0915***	0.1212***	-	0.1180***	0.1374***	0.1006***	-	0.3416***	-	-	-	-
Moran's I in the residuals (80's)	0.0513***	0.0174***	0.0516***	0.1022***	0.0690***	0.0847***	0.0816***	0.1243***	0.1319***	0.2297***	0.1049***	0.4151***	-	-	-	-
Moran's I in the residuals (90's)	0.0283***	0.0617***	0.0472***	0.1053***	0.1641***	0.1813***	0.1722***	0.0942***	0.3633***	0.2768***	0.3314***	0.4125***	-	-	-	-
Spatial Weight Matrix: 800km cut-off																
Moran's I in the residuals (70's)	0.0676***	0.0518***	-	0.1270***	0.0668***	0.0891***	-	0.0807***	0.0739***	0.0990***	-	0.3077***	0.2449**	-0.2716*	-	-0.0141
Moran's I in the residuals (80's)	0.0388***	0.0100***	0.0379***	0.0869***	0.0532***	0.0631***	0.0558***	0.0990***	0.0857***	0.1460***	0.0665***	0.3691***	0.1096	0.1805*	-0.1362	0.1021
Moran's I in the residuals (90's)	0.0159***	0.0478***	0.0362***	0.0873***	0.1381***	0.1539***	0.1468***	0.0574***	0.3456***	0.2976***	0.3196***	0.3718***	0.4721***	0.3318**	0.2930**	0.4296***
Spatial Weight Matrix: 900km cut-off																
Moran's I in the residuals (70's)	0.0611***	0.0364***	-	0.1157***	0.0463***	0.0630***	-	0.0534***	0.0247	0.0084	-	0.2823***	0.1646	-0.2583*	-	-0.0463
Moran's I in the residuals (80's)	0.0282***	0.0077***	0.0258***	0.0727***	0.0342***	0.0505***	0.0318***	0.0775***	0.0385**	0.1088***	0.0091	0.3263***	0.0588	0.1788*	-0.1779	0.1086
Moran's I in the residuals (90's)	0.0124***	0.0389***	0.0311***	0.0758***	0.1314***	0.1418***	0.1381***	0.0373***	0.2867***	0.2192***	0.2595***	0.3350***	0.5725***	0.3763***	0.3631**	0.4639***
Spatial Weight Matrix: 1000km cut-off																
Moran's I in the residuals (70's)	0.0573***	0.0249***	-	0.1082***	0.0322***	0.0410***	-	0.0381***	-0.0059	-0.0280	-	0.2380***	0.0452	-0.0971	-	-0.1078
Moran's I in the residuals (80's)	0.0215***	0.0080***	0.0168***	0.0625***	0.0273***	0.0461***	0.0152***	0.0626***	0.0254*	0.0946	-0.0163	0.3020***	0.2046**	0.2634**	0.0596	0.0953
Moran's I in the residuals (90's)	0.0131***	0.0336***	0.0293***	0.0699***	0.1258***	0.1277***	0.1291***	0.0210***	0.2774***	0.2018***	0.2529***	0.3090***	0.3320***	0.0920	0.1167*	0.3531***
Spatial Weight Matrix: 1500km cut-off																
Moran's I in the residuals (70's)	0.0417***	0.0111***	-	0.0778***	0.0125***	0.0092**	-	0.0176***	-0.0009	-0.0367**	-	0.1562***	-0.0569	-0.1001	-	-0.1803**
Moran's I in the residuals (80's)	0.0055***	0.0018***	0.0030***	0.0276***	0.0190***	0.0375***	-0.0040	0.0296***	-0.0054	0.0518***	-0.0354***	0.1892***	0.0718*	0.1255**	0.0067	-0.0171
Moran's I in the residuals (90's)	0.0154***	0.0253***	0.0233***	0.0537***	0.1034***	0.0827***	0.0920***	0.0114***	0.2035***	0.1319***	0.1689***	0.1993***	0.2859***	0.0770	0.0765*	0.2449***
Spatial Weight Matrix: 2000km cut-off																
Moran's I in the residuals (70's)	0.0189***	0.0042***	-	0.0399***	0.0030**	0.0049**	-	0.0155***	-0.0119	-0.0171	-	0.0749***	-0.0839	-0.0643	-	-0.0618
Moran's I in the residuals (80's)	-0.0026***	0.0005**	-0.0018***	0.0099***	0.0047**	0.0267***	-0.0078***	0.0168***	0.0014	0.0521***	-0.0273***	0.1238***	0.0780**	-0.0146	-0.0801	0.0635**
Moran's I in the residuals (90's)	0.0120***	0.0189***	0.0184***	0.0317***	0.0710***	0.0597***	0.0666***	0.0116***	0.1370	0.0922***	0.1209***	0.1180***	0.1145**	0.0247	-0.0096	0.1253***

Note: *** Significant at 1%, ** significant at 5%, * significant at 10% based on the permutation approach with 10,000 permutations.

Table II.B.2

Diagnostics for spatial autocorrelation by means of Moran's I in the time-averaged residuals of the estimations of Table 5.2

Spatial scale	Minimum Comparable Area (MCA)				Micro-region				Meso-region				State			
Method	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM	OLS	FE	FD	SYS-GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Diagnostics for spatial autocorrelation - Spatial Weight Matrix:																
400 km cut-off	0.1983***	++	0.1621***	0.2589***	-	++	-	-	-	++	-	-	-	++	-	-
500 km cut-off	0.1586***	++	0.1331***	0.2184***	0.1401***	++	0.2577***	0.1653***	-	++	-	-	-	++	-	-
600 km cut-off	0.1262***	++	0.1076***	0.1852***	0.1156***	++	0.2134***	0.1274***	0.1742***	++	0.2561***	0.4874***	-	++	-	-
700 km cut-off	0.1010***	++	0.0874***	0.1614***	0.0842***	++	0.1812***	0.0835***	0.1373***	++	0.2070***	0.4484***	-	++	-	-
800 km cut-off	0.0810***	++	0.0685***	0.1416***	0.0601***	++	0.1488***	0.0487***	0.0950***	++	0.2224***	0.4130***	0.1718*	++	0.0158	0.1334
900 km cut-off	0.0675***	++	0.0524***	0.1253***	0.0448***	++	0.1292***	0.0239***	0.0785***	++	0.1493***	0.3990***	0.0947	++	0.0765	0.1026
1,000 km cut-off	0.0600***	++	0.0409***	0.1148***	0.0387***	++	0.1104***	0.0080**	0.0707***	++	0.1258***	0.3623***	0.0561	++	0.0160	0.0481
1,500 km cut-off	0.0459***	++	0.0274***	0.0784***	0.0318***	++	0.0687***	-0.0053	0.0304***	++	0.0789***	0.2300***	-0.0573	++	0.0358	-0.0594
2,000 km cut-off	0.0236***	++	0.0169***	0.0417***	0.0106***	++	0.0441***	0.0053*	-0.0029	++	0.0480***	0.1267***	-0.0783	++	0.0098	0.0286

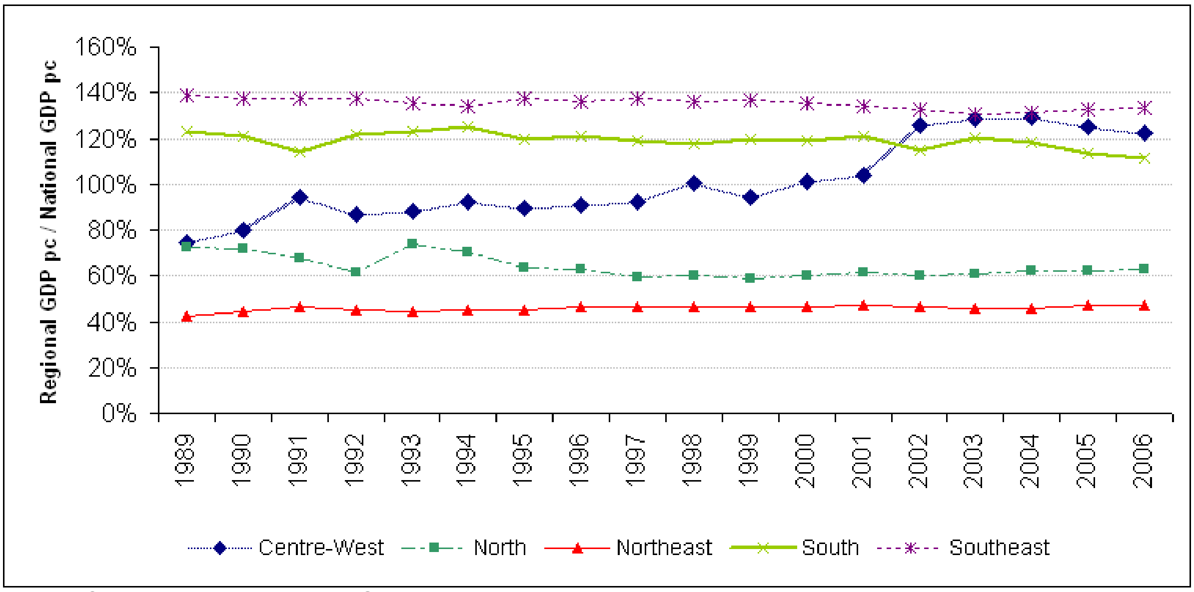
Note: *** Significant at 1%, ** significant at 5%, * significant at 10% based on the permutation approach with ten thousand permutations. The residuals of the panel data estimations were time-averaged and then Moran's I statistics were calculated. ++ For fixed-effects (FE) estimations, the time-averaged residuals are zero by construction; therefore, Moran's I is not calculated.

APPENDIX III

Appendix III.A

Figure III.A.1

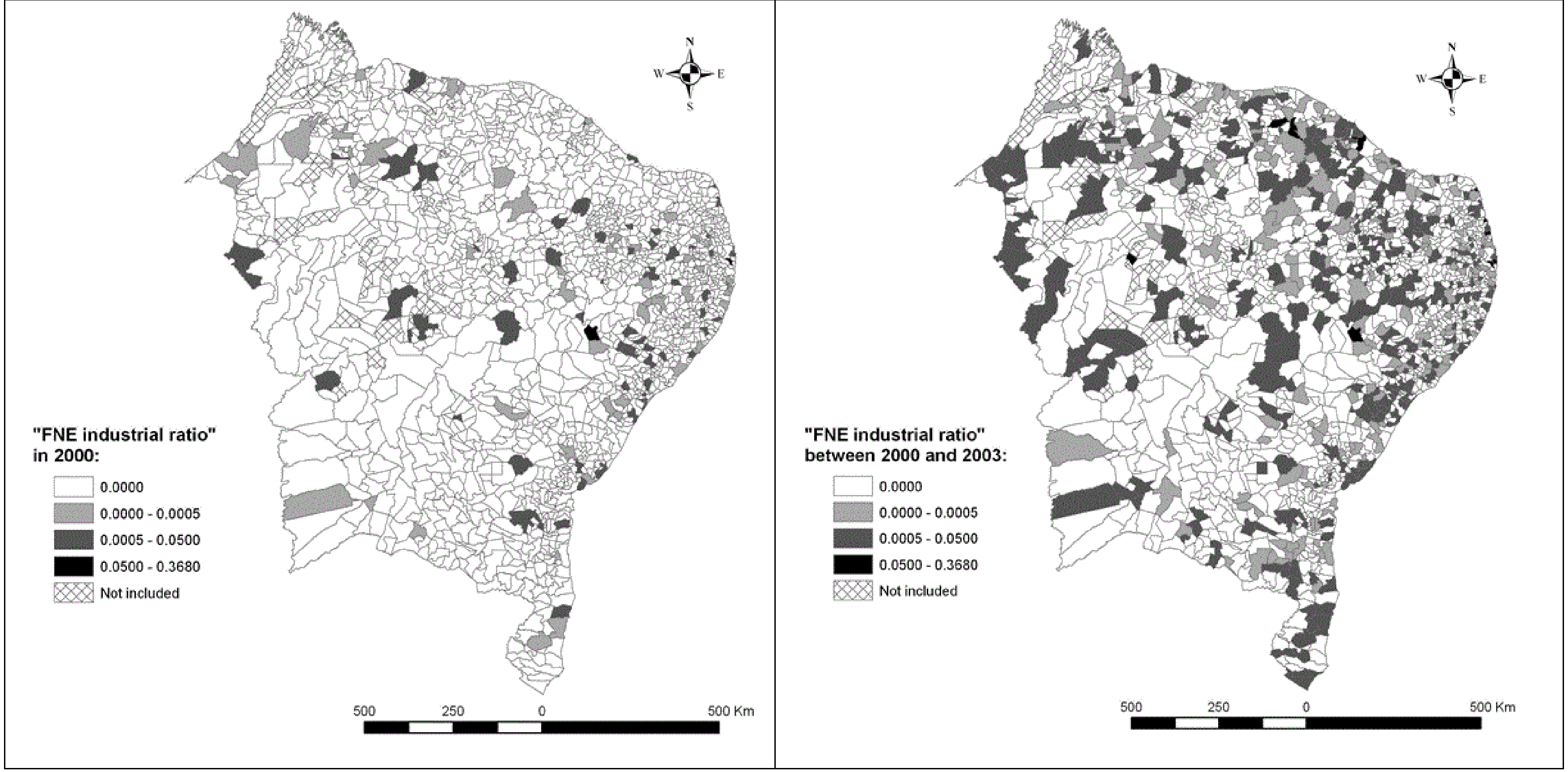
Regional GDP per capita as a proportion of national GDP per capita in Brazil between 1989 and 2006



Note: Own elaboration based on IBGE data.

Figure III.A.2 – Spatial distribution of "FNE industrial ratios" at municipal level

a. Total amount of FNE loans to the industrial sector in 2000 as a ratio of total GDP in 2000	b. Total amount of FNE loans to the industrial sector between 2000 and 2003 as a ratio of total GDP in 2000



Note: Own elaboration based on IBGE and BNB data. The "FNE industrial ratios" also include the loans to commerce/services sectors.

Appendix III.B

Micro-data Description

The micro-evaluation approach relies only on information about firms that are found in the annual censuses for firms in the formal sector (those that have the CNPJ identifier), namely, RAIS. Based on the information provided by the Bank of the Northeast (BNB), it was possible to identify those firms included in the RAIS dataset from 2000 that had received the FNE loans (the treatment group) and those that had not received the FNE loans (the control group). This matching between the BNB information and the RAIS dataset was made possible by using the CNPJ identifiers. It is worth noting that the construction of the dataset followed procedures necessary to guarantee the confidentiality of information. Table III.B.1 lists the data sources for the variables at the firm level.

Table III.B.1

Variables and data sources at the firm level

Dependent variables:	**Source**
Average annual employment growth rates for 1999-2000, 2000-2003 and 2000-2006	RAIS1999, RAIS2000, RAIS2003 and RAIS2006
Average annual wage growth rates for 1999-2000, 2000-2003 and 2000-2006*	RAIS1999, RAIS2000, RAIS2003 and RAIS2006
FNE variable:	
Dummy variable for participation in FNE	BNB
Covariates:**	
Average age of employees	RAIS1999 and RAIS2000
Average years of schooling of employees	RAIS1999 and RAIS2000
Number of employees	RAIS1999 and RAIS2001
Average wage*	RAIS1999 and RAIS2000

Note: * The average wages (in R$) for all years were converted to constant 2000 prices using a CPI index, namely IGP-M from Fundação Getúlio Vargas (FGV). **In addition, as covariates, a dummy for the commerce/services sectors, as defined by the CNAE/IBGE [National Economic Activity Classification (CNAE) from the Brazilian Institute of Geography and Statistics (IBGE)], was included, as well dummies for states where the firms are located (the dummy for the Ceará state was excluded).

The dataset is comprised of two different groups. The first group (the treatment group) consists of the 91 firms that received FNE loans in 2000 only and that could be traced to RAIS 2000 and the datasets for the years 1999, 2003 and 2006. This treatment group is almost the same as that used in the study of SILVA et al. (2009). The current study uses fewer firms because 17 firms in the agricultural sector were excluded, and only the firms that could be traced in the years 1999, 2000, 2003 and 2006 were used; in contrast, SILVA et al. (2009) considered the 2000-2003 period alone. The second group is the control group, which is comprised of firms that did not receive FNE loans in any year analysed. This group was selected from industrial and commercial/services firms in the Northeast region that were in the RAIS in 2000 and the progress of which could be traced through time. Estimates are carried out using four different control samples: (b.i) a sample using all firms identified in RAIS during 1999-2006 (111,869 firms); (b.ii) a sample matched to the treatment group (9,990 firms); (b.iii) a sample 'perfectly' matched to the treatment group (91 firms); and (b.iv) the control sample used in the SILVA et al. (2009) paper, which includes 914 firms drawn from a random and representative sample of the population of the Northeast with records in RAIS. Matching the control samples to the treated group (b.ii and b.iii) aims to ensure that the treatment and control groups have similar characteristics and thus makes the two groups more comparable. The sample selection strategy for the matched samples used value ranges for the control-group variables based on those values observed for the treatment group in order to find a "representative" or a "matched" control sample. The variables used in this strategy are schooling, wage, activity sector and number of employees. For the 'perfectly' matched sample, the ranges were shrunk to find the most similar firm. Furthermore, it is worth noting that in this 'perfect' matching, treated and matched firms are located in the same municipality. Table III.B.2 shows the summary statistics of the RAIS dataset for 2000.

Table III.B.2

Summary statistics for the dataset at the firm level in the RAIS 2000

Variables	(a) FNE sample* (treatment group) (91 firms)	(b.i) Control sample - All firms in RAIS2000 (excluding the FNE sample)* (111,869 firms)	(b.ii) Control sample - Matched sample to the treated group* (9,990 firms)	(b.iii) Control sample - 'Perfectly' matched sample to the treated group* (91 firms)	(b.iv) Control sample - Random sample* (914 firms)
Average age of employees	31.5	32.5	32.2	31.6	33.2
Average years of schooling	4.9	5.5	4.9	4.9	5.4
Average wage (R$ in 2000)	288.0	336.2	308.5	290.8	382.2
Activity sector (%)					
Agriculture	0.0%	0.0%	0.0%	0.0%	0.0%
Industry	63.7%	13.3%	65.0%	63.7%	16.8%
Commerce/Services	36.3%	86.7%	35.0%	36.3%	83.2%
Size (%)					
Small firms (1-49 employee)	76.9%	93.6%	74.7%	76.9%	84.2%
Medium firms (50-99 employees)	6.6%	2.7%	14.0%	6.6%	6.9%
Large firms (>99 employees)	16.5%	3.7%	11.3%	16.5%	9.0%
Northeast states (%)					
Maranhão	8.8%	5.7%	5.3%	8.8%	6.0%
Piauí	7.7%	5.1%	5.1%	7.7%	5.0%
Ceará	16.5%	15.9%	20.6%	16.5%	17.6%
Rio Grande do Norte	7.7%	7.1%	8.2%	7.7%	7.4%
Paraíba	20.9%	7.8%	7.9%	20.9%	6.5%
Pernambuco	14.3%	19.4%	19.9%	14.3%	19.8%
Alagoas	3.3%	4.8%	3.9%	3.3%	4.2%
Sergipe	7.7%	4.8%	4.2%	7.7%	4.5%
Bahia	13.2%	29.6%	25.0%	13.2%	29.0%

Note:* Only firms that can be linked through time in the RAIS in 1999, 2000, 2003 and 2006 are selected. Own elaboration based on RAIS data in 2000.

Macro-data Description

The macro-analysis employs a municipal dataset from the Brazilian Regional Accounts (IBGE, 2009), the Bank of the Northeast (BNB), IPEADATA and RAIS. Table III.B.3 lists the data sources for the variables at the municipal level:

Table III.B.3

Variables and data sources at the municipal level

Dependent variables:	**Source**
Average annual per capita gross domestic product (GDP) growth rates for 1999-2000, 2000-2003 and 2000-2006***	Brazilian Regional Accounts (IBGE, 2009), IPEADATA
FNE variable:	
FNE industrial loans as a proportion of total GDP*	BNB
Covariates:**	
Average years of schooling of the working force population in the formal sector	RAIS1999 and RAIS2000
Population density (inhab./km2)	IPEADATA
GDP per capita***	Brazilian Regional Accounts (IBGE, 2009), IPEADATA

Note: * The FNE industrial variable also includes the loans to commerce/services sectors. The FNE loans were converted to R$ at the 2000 level using a CPI index (namely, the IGP-M index from Fundação Getúlio Vargas-FGV). **In addition, dummies for states are included (the dummy for the Ceará state was excluded); ***GDP per capita variables were obtained directly at constant 2000 prices in R$ from IPEADATA.

The dependent variable is the average annual growth of the per capita gross domestic product (GDP) at the municipal level over the 2000-2003 and 2000-2006 periods. The GDP per capita of the municipalities has been computed annually by the IBGE since 1999. This thesis makes use of updated estimates of municipal GDP released by IBGE (2009), which provides comparable GDP at municipal level from 1999 onwards. This is an improvement of earlier estimates of municipal GDP produced by IBGE; these estimates only provided comparable GDPs from 2002 onwards due to an alteration in the IBGE methodology in 2002. At any rate, if there is concern about the consistency of the methodology across the analysed period, the first-difference approach allows for controlling for such global changes in the methodology, as their effects vanish when the first-difference method is conducted using the 1999-2000 period [see POSTALI (2009) for a similar approach to addressing global changes in GDP methodology at the municipal level in Brazil]. Furthermore, due to the permanent creation of new municipalities during 2000-2006, there are few municipalities that did not exist in 2000. To construct a consistent dataset, the boundaries of the 1,787 Northeastern municipalities in 2000 were used instead of those for the existing 1,793 municipalities in 2006. The solution was to assign the new municipalities to the existing municipalities in 2000. The final dataset has 1,731 (out of 1,787) Brazilian municipalities. The data lost 56 municipalities that did not present data for all variables of interest at all data points. The FNE variables come from the Bank of the Northeast (BNB), which provided information about the amount of loans for individuals and firms in the aggregate by programme at the municipal level over the 2000-2003 period. Using this information, the FNE industrial ratio was constructed; this ratio is the amount of FNE loans to the industrial sector in 2000 as a proportion of total GDP in 2000. A robustness check was also provided using the amount of FNE loans to the industrial sector between 2000 and 2003 as a proportion of the total GDP in 2000. It is worth noting that the FNE industrial variables also include the loans to commerce/services sectors (as in the micro-analysis). Table III.B.4 provides the summary statistics for variables used in the macro-analysis.

Table III.B.4

Summary statistics for the dataset at the municipal level in 2000

Variables	obs.	mean**	minimum	maximum	st. dev.
FNE industrial loans in 2000 as a proportion of total GDP in 2000*	1731	0.0005	0.0000	0.3680	0.0092
FNE industrial loans between 2000 and 2003 as a proportion of total GDP in 2000*	1731	0.0016	0.0000	0.3680	0.0121
Years of schooling of the work-force population in the formal sector	1731	4.7	1.3	9.0	0.9
Population density (inhab./km2)	1731	84.5	0.9	9404.2	416.7
GDP per capita (in thousands, R$)	1731	2.0	0.6	114.9	3.2

Note: Own elaboration based on IBGE, IPEADATA and BNB datasets.* The FNE industrial variable also includes the loans to commerce/services sectors; ** Arithmetic mean.

Appendix III.C

Table III.C.1

The micro-approach for evaluating FNE impact on average annual employment growth using the matched sample to the treatment group (Method: First-Difference)

Dependent variable= [(Employment growth between 2000 and 2003) - (Employment growth between 2000 and 1999)]					
	(1)	(2)	(3)	(4)	(5)
Diff. FNE industrial dummy (2000-1999)	**0.1624****	**0.1728****	**0.1706****	**0.1709****	**0.1896*****
	(0.0787)	**(0.0748)**	**(0.0734)**	**(0.0731)**	**(0.0708)**
Diff. average age of employees (2000-1999)		0.0487***	0.0508***	0.0493***	0.0428***
		(0.0137)	(0.0136)	(0.0135)	(0.0132)
Diff. average years of schooling of employees (2000-1999)			0.1380**	0.1353**	0.1124*
			(0.0697)	(0.0682)	(0.0676)
Diff. average wage (2000-1999)				0.0004	0.0000
				(0.0004)	(0.0003)
Diff. number of employees (2000-1999)					-0.0090***
					(0.0025)
Constant	-0.3368***	-0.3527***	-0.3644***	-0.3600***	-0.3318***
	(0.0173)	(0.0177)	(0.0189)	(0.0183)	(0.0170)
Observations (firms)	10,081	10,081	10,081	10,081	10,081
R-squared	0.00	0.01	0.01	0.01	0.05

Note: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The FNE industrial dummy also includes firms in the commerce/services sectors based on CNAE/IBGE. The "Diff. FNE industrial dummy (2000-1999)" is equal to the "FNE industrial dummy in 2000" because no firm in this sample had received an FNE industrial loan in 1999.

Table III.C.2

The macro-approach for evaluating FNE impact on average annual growth of the per capita GDP at the municipal level (Method: First-Difference)

Dependent variable= [(GDP per capita growth between 2000 and 2003) - (GDP per capita growth between 2000 and 1999)]				
	(1)	(2)	(3)	(4)
Diff. FNE industrial ratio (2000-1999)	**0.2142**	**0.2187**	**0.2168**	**0.2384**
	(0.2323)	**(0.2345)**	**(0.2377)**	**(0.2076)**
Diff. average years of schooling of working force population (2000-1999)		-0.0108	-0.0107	-0.0075
		(0.0095)	(0.0095)	(0.0080)
Diff. population density (2000-1999)			0.0015***	0.0014**
			(0.0005)	(0.0005)
Diff. GDP per capita (2000-1999)				-0.1544**
				(0.0627)
Constant	-0.0251***	-0.0245***	-0.0258***	-0.0108*
	(0.0041)	(0.0041)	(0.0041)	(0.0061)
Observations (municipalities)	1,731	1,731	1,731	1,731
R-squared	0.00	0.00	0.00	0.23

Note: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.. The FNE industrial ratio variable also includes the loans to commerce/services sectors.

Bibliography

ABRAMOVITZ, M. (1986) Catching up, forging ahead and falling behind. *Journal of Economic History*, 46: 385-406.

ABREU, M., DE GROOT, H. L. F., FLORAX, R. J. (2005) Space and growth: A survey of empirical evidence and methods. *Région et développement*, 21: 13-44.

AGHION, P., CAROLI, E., GARCIA-PENALOSA, C. (1999). Inequality and economic growth: the perspective of the new growth theories. *Journal of Economic Literature*, 37 (4): 1615–60.

ALESINA, A., RODRIK, D. (1994) Distributive politics and economic growth. *Quarterly Journal of Economics*, 109: 465-490.

ALMEIDA JUNIOR M., RESENDE G. M., SILVA A. M. (2007) Distribuição espacial dos fundos constitucionais de financiamento do Nordeste, Norte e Centro-Oeste. *Revista de Economia* (UFPR), 33 (2): 115-137.

ANDERSSON, R. (2005) The efficiency of Swedish regional policy. *The Annals of Regional Science*, 39(4): 811–832.

ANDRADE, E., LAURINI, M., MADALOZZO, R., PEREIRA, P. (2002) *Testing convergence across municipalities in Brazil using quantile regression*. Ibmec Working Papers n. 14.

ANDRADE, E., LAURINI, M., MADALOZZO, R., PEREIRA, P. (2004). Convergence clubs among Brazilian municipalities. *Economics Letters*, 83: 179–184.

ANDREWS, D. W. K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59: 817-59.

ANGRIST J., PISCHKE J. (2009) *Mostly harmless econometrics: an empiricist's companion.* Princeton University Press, London, UK.

ANSELIN, L. (1988) *Spatial econometrics: methods and models*. Dorddrecht: Kluwer Academic.

ANSELIN, L. (1990) Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science*, 30: 185–207.

ANSELIN, L. (1995) Local indicators of spatial association — LISA. *Geographical Analysis*, 27: 93–115.

ANSELIN, L. (2001) *Spatial econometrics*. In: Baltagi, B. (ed.) A companion to theoretical econometrics, Basil Blackwell, Oxford, UK: 310-330.

ANSELIN, L. (2002) Under the hood: issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, 27: 247-267.

ANSELIN, L. (2003) Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26: 156-166.

ANSELIN, L., FLORAX, R. (1995) Small sample properties of tests for spatial dependence in regression models. In: Anselin, L., Florax, R. (eds.) *New directions in spatial econometrics*, Berlin: Springer-Verlag.

ANSELIN, L., REY, S. (1991) Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, 23: 110-31.

ANSELIN L, LE GALLO J, JAYET H (2008) Spatial panel econometrics. In: Mátyás L, Sevestre, P. (eds.) *The econometrics of panel data, fundamentals and recent developments in theory and practice*, Kluwer, Dordrecht: 627–662

ANSELIN L, REY S (1991) Properties of tests for spatial dependence in linear-regression models. *Geographical Analysis*, 23: 112–131.

ARAÚJO, T. B. (1999) Por Uma Política Nacional de Desenvolvimento Regional. *Revista Econômica do Nordeste*, 30: 144-161.

ARELLANO, M., BOND, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58 (2): 277–297.

ARBIA, G., PETRARCA, F. (2011) Effects of MAUP on spatial econometric models. *Letters in Spatial and Resource Sciences*, 4: 173–185.

ARMSTRONG, H. (1995) Convergence among the regions of the European Union. *Papers in Regional Science*, 74: 143–152.

ARMSTRONG, H. W. (2002) European Union regional policy: Reconciling the evaluation and convergence evidence. In: Cuadrado-Roura, J., Parellada, M. (eds.) *Regional convergence in the European Union: Facts, prospects and policies*, Springer-Verlag (Advances in Spatial Sciences Series), Amsterdam: 231-272.

ARMSTRONG, H.W., KEHRER, B.,WELLS, P. (2001) Initial impacts of community economic development initiatives in the Yorkshire and Humber structural funds programme. *Regional Studies*, 35(8): 673-688.

ASCHAUER, D. (1989) Is public expenditure productive? *Journal of Monetary Economics*, 23: 177-200.

AUTANT-BERNARD, C., LESAGE J. (2011) Quantifying knowledge spillovers using spatial econometric models. *Journal of Regional Science*, 51 (3): 471-496.

AZARIADIS, C., DRAZEN, A. (1990). Threshold externalities in economic development. *The Quarterly Journal of Economics*, 105: 501–26.

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

AZZONI, C. R., MENEZES-FILHO, N., MENEZES, T., SILVEIRA-NETO, R. (2000) *Geography and income convergence among Brazilian states*. Research Network Working Paper R-395, Inter-American Development Bank, Washington, DC.

BACHTLER, J., WREN, C. (2006) Evaluation of European Union Cohesion policy: Research questions and policy challenges. *Regional Studies*, 40: 143–153.

BAER, W. (2003) *A Economia Brasileira*. São Paulo: Nobel Editora.

BALDWIN, R., FORSLID, R. (2000) The core-periphery model and endogenous growth: Stabilizing and destabilizing integration. *Economica*, 67: 307-324.

BALDWIN, R., FORSLID, R., MARTIN, P., OTTAVIANO, G., ROBERT-NICOUD, F. (2003) *Economic Geography and Public Policy*. Princeton University Press: Princeton, N.J.

BALTAGI, B. H., PIROTTE, A. (2010) Panel data inference under spatial dependence. *Economic Modelling*, 27(6): 1368-1381.

BANCO CENTRAL (2000). *Juros e Spread Bancário no Brasil*. Departamento de Estudos e Pesquisas do Banco Central (available at: http://www.bancocentral.gov.br/ftp/jurospread112000.pdf).

BANCO DO NORDESTE (2001) *Programação do Fundo Constitucional de Financiamento do Nordeste – FNE para o ano de 2001*. Banco do Nordeste, Fortaleza: 1-136.

BANCO DO NORDESTE (2009) *Programação do Fundo Constitucional de Financiamento do Nordeste – FNE para o ano de 2009*. Banco do Nordeste, Fortaleza: 1-159.

BANERJEE, A., NEWMAN, A. (1993) Occupational choice and the process of development. *Journal of Political Economy*, 101 (2): 274–298.

BARCA, F. (2009) *An agenda for a reformed Cohesion Policy: A place-based approach to meeting European Union challenges and expectations*. Independent report, prepared at the request of the European commissioner for regional policy, Danuta Hubner, European Commission, Brussels.

BARCA, F., MCCANN, P., RODRÍGUEZ-POSE, A. (2012) The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science*, 52(1): 134–152.

BARNOW, B., CAIN, G., GOLDBERGE, A. (1981) Selection on observables. *Evaluation Studies Review Annual*, 5: 43-59.

BARRO, R. J. (1990) Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98: 103-125.

BARRO, R. J. (1991) Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106 (2): 407-443.

BARRO,R., SALA-I MARTIN,X. (1991) Convergence across states and regions. *Brooking Papers on Economic Activity*, 1: 107-182.

BARRO,R., SALA-I MARTIN,X. (1992) Convergence. *Journal of Political Economy*, 100: 223–51.

BARRO, R., SALA-I MARTIN, X. (2003) *Economic growth*. Cambridge, Mass.: MIT, 2ed.

BARRO, R., MANKIW, G., SALA-I-MARTIN, X. (1995) Capital mobility in neoclassical models of growth. *American Economic Review*, 85 (1): 103-115.

BARROS, A. R. (1993) Some implications of new growth theory for economic development. *Journal of International Development*, 5 (5): 531 -558.

BARROS, A. R. (2011) *Desigualdades regionais no Brasil: Natureza, causas, origens e soluções*. Editora Campus / Elsevier.

BARTIK, T., BINGHAM, R. (1995) *Can economic development programs be evaluated?* Staff Working Papers 95-29, W.E. Upjohn Institute for Employment Research.

BAUMOL, W. (1986) Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review*, 76 (5): 1072–1085.

BEHRENS, K., THISSE, J. F. (2007) Regional economics: A new economic geography perspective. *Regional Science and Urban Economics*, 37: 457–465.

BENABOU, R. (1996). *Inequality and growth*. NBER Working Paper no. 5658.

BERNARD, A. B., JONES, C. I. (1996) Comparing apples to oranges: Productivity convergence and measurement across industries and countries. *American Economic Review*, 86: 1216–1238.

BERNARD, A. B., JONES, C. I. (2001) Comparing apples to oranges: Productivity convergence and measurement across industries and countries: Reply. *American Economic Review*, 91: 1168–1169.

BIRKLAND, T. (2007) Agenda setting in public policy. In: Fischer, F, Miller, G., Sidney, M. (eds). Handbook of public policy analysis. New York: CRC Press.

BLOOM, D., CANNING, D., SEVILLA, J. (2004) The effect of health on economic growth: A production function approach. *World Development*, 32 (1): 1–13.

BLUNDELL, R., BOND, S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87: 115-143.

BOARNET, M. (1998) Spillovers and the location effects of public infrastructure. *Journal of Regional Science*, 38: 381–400.

BODE, E., REY, S. J. (2006) The spatial dimension of economic growth and convergence. *Papers in Regional Science*, 85: 171-176.

BOND, S., HOEFFLER, A., TEMPLE, J. (2001) *GMM estimation of empirical growth models*. Centre for Economic Policy Research Discussion Paper No. 3048.

BRADLEY, J. (2006) Evaluating the impact of European Union Cohesion policy in less-developed countries and regions. *Regional Studies*, 40 (2):189-200.

BRASIL. (2008) *Constituição da República Federativa do Brasil* (available at: http://www.planalto.gov.br/ ccivil_03/Constituicao/Constituiçao.htm).

BRIANT, A., COMBES, P. P., LAFOURCADE, M. (2010) Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, 67: 287-302.

BROCK, W., DURLAUF, S. (2001) Growth empirics and reality. *World Bank Economic Review*, 15: 229–272.

BROCK, W., DURLAUF, S., WEST, K. (2003) Policy evaluation in uncertain economic environments. *Brookings Papers of Economic Activity*, 1: 235–322.

BRUECKNER, J. K. (1998) Testing for strategic interaction among local governments: The case of growth controls. *Journal of Urban Economics*, 44: 438–467.

BRUNDSON, C., FOTHERINGHAM, A. S., CHARLTON, M. (1999) Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, 39: 497-524.

CAF (2010) *Desarrollo local: Hacia un nuevo protagonismo de las ciudades y regiones*. Caracas: Corporación Andina de Fomento.

CANO, W. (1976) *Raízes da Concentração Industrial no Brasil*. São Paulo: Difel.

CARVALHO, A., LALL, S. V., TIMMINS, C. (2006) *Regional subsidies and industrial prospects of lagging regions*. World Bank Policy Research Working Paper n. 3843.

CARVALHO, A. Y., DA MATA, D., RESENDE, G. M. (2007) Clusterização dos municípios brasileiros. In: Carvalho, A., Alburqueque, C. W., Mota, J. A., Piancastelli, M. (eds) *Dinâmica dos Municípios*, Brasília: IPEA.

CASE, A. (1992) Neighborhood influence and technological change. *Regional Science and Urban Economics*, 22: 491-508.

CASELLI, F., ESQUIVEL, G., LEFORT, F. (1996). Reopening the convergence debate: A new look at cross country growth empirics. *Journal of Economic Growth*, 1 (3): 363–389.

CASTRO, A. B. (1971) *Sete ensaios sobre a economia brasileira*. São Paulo: Forense.

CHESHIRE, P., CARBONARO, G. (1995) Convergence-divergence in regional growth rates: an empty black box?, In: Armstrong, H.W., Vickerman, R.W. (eds.) *Convergence and divergence among European regions*, European Research in Regional Science, London: Pion.

CHESHIRE, P., HAY, D. G. (1989) *Urban problems in western Europe: An economic analysis*. London: Unwin Hyman.

CHESHIRE, P. C., MAGRINI, S. (2000) Endogenous processes in European regional growth: Convergence and policy. *Growth and Change*, 31 (4): 455-479.

CLIFF A. D., ORD, J.K. (1981) *Spatial processes: Models and applications*. Pion, London.

COELHO, R. L. P., FIGUEIREDO, L. (2007) Uma análise da hipótese da convergência para os municípios Brasileiros. *Revista Brasileira de Economia*, 61: 331–352.

CONLEY, T. (1999) GMM Estimation with cross-sectional dependence. *Journal of Econometrics*, 92 (1): 1–45.

CORRADO, L., FINGLETON, B. (2012) Where is the Economics in Spatial Econometrics? *Journal of Regional Science*, 52 (2): 210-239.

CRAVO, T. A. (2010) SMEs and economic growth in the Brazilian micro-regions. *Papers in Regional Science*, 89: 711–734.

CRAVO, T., RESENDE, G. M. (2012) Economic growth in Brazil: A spatial filtering approach. *The Annals of Regional Science*, DOI: 10.1007/s00168-012-0504-6.

DA MATA D., DEICHMANN U., HENDERSON J.V., LALL S.V., WANG H.G. (2007a) Determinants of city growth in Brazil. *Journal of Urban Economics*, 62: 252–272.

DA MATA, D., PIN, C, RESENDE, G. M. (2007b) Composição e Consolidação da Infra-estrutura Domiciliar nos Municípios Brasileiros. In: Carvalho, A., Alburqueque, C. W., Mota, J. A., Piancastelli, M. (eds) *Dinâmica dos Municípios*, Brasília: IPEA.

DALL'ERBA, S. (1995) Distribution of regional income and regional funds in Europe 1989–1999: An exploratory spatial data analysis. *The Annals of Regional Science*, 39: 121–148.

DALL'ERBA, S., LE GALLO, J. (2008) Regional convergence and the impact of European structural funds over 1989–1999: A spatial econometric analysis. *Papers in Regional Science*, 87 (2): 219-244.

DE GRAAFF, T., FLORAX, R. J.C.M., NIJKAMP, P., REGGIANI, A. (2001) A general misspecification test for spatial regression models: Dependence, heterogeneity, and nonlinearity. *Journal of Regional Science*, 41: 255–276.

DE GROOT, M. A. H. L. F., FLORAX, R. J. G. M. (2005), A meta-analysis of β-convergence: The legendary 2%. *Journal of Economic Surveys*, 19: 389–420.

DE LA FUENTE, A. (2004) Second-best redistribution through public investment: A characterization, an empirical test and an application to the case of Spain. *Regional Science and Urban Economics*, 34: 489– 503.

DE VREYER, P., SPIELVOGEL, G. (2005) *Spatial externalities between Brazilian municipios and their neighbours*. Ibero America Institute for Econ. Research ,IAI Discussion Papers 123, Ibero-America Institute for Economic Research.

DECREE nº 6047 (2007) *Política Nacional de Desenvolvimento Regional – PNDR* (available at: http:// www.planalto.gov.br/ccivil_03/Ato2007-2010/2007/Decreto/D6047.htm).

DESDOIGTS, A., (1999) Patterns of economic development and the formation of clubs. *Journal of Economic Growth*, 4 (3): 305-330.

DOLLAR, D., WOLFF, E. N. (1988) Convergence of industry labour productivity among advanced economies, 1963–1982. *Review of Economics and Statistics*, 70, 549–558.

DOLLAR, D., WOLFF, E. N. (1993) *Competitiveness, convergence, and international specialization*. Cambridge, MA: MIT Press.

DOMAR, E. (1946) Capital expansion, rate of growth, and employment. *Econometrica*, 14: 137-147.

DOMINICIS, L., FLORAX, R. J. G. M., DE GROOT, H. L. F. (2008) A meta-analysis on the relationship between income inequality and economic growth. *Scottish Journal of Political Economy*, 55 (5): 654-682.

DURANTON, G., MONASTIRIOTIS, V. (2002) Mind the gaps: The evolution of regional earnings inequalities in the U.K., 1982-1997. *Journal of Regional Science*, 42(2): 219-256.

DURLAUF, S. N., JOHNSON, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, 10: 365–84.

DURLAUF, S. N., KOURTELLOS, A., MINKIN, A. (2001) The local Solow growth model. *European Economic Review*, 45(4-6): 928-940.

DURLAUF, S. N., JOHNSON, P. A., TEMPLE, J. R. W. (2005) Growth econometrics. In: Aghion, P., Durlauf, S. (Eds) *Handbook of Economic Growth*, Elservier.

EASTERLY, W., LEVINE, R. (1998) Troubles with the neighbours: Africa's problems, Africa's opportunity. *Journal of African Economies*, 7: 120–142.

EASTERLY, W., REBELO, S. (1993) Fiscal policy and economic growth. *Journal of Monetary Economics*, 32: 417-58.

EICHER, T. S., PAPAGEORGIOU, C., ROEHN, O. (2007) Unraveling the fortunes of the fortunate: an Iterative Bayesian Model Averaging (IBMA) approach. *Journal of Macroeconomics*, 29: 494-514.

EGGER, P., PFAFFERMAYR, M. (2006) Spatial convergence. *Papers in Regional Science*, 85: 199-215.

ELHORST, J. P. (2001) Dynamic models in space and time. *Geographical Analysis*, 33(2):119–140.

ELHORST, J. P. (2010a) Spatial panel data models. In: Fischer, M., Getis, A. (eds) *Handbook of applied spatial analysis*. Springer, Berlin, Heidelberg and New York: 377–407.

ELHORST, J. P. (2010b) Dynamic panels with endogenous interaction effects when T is small. *Regional Science and Urban Economics*, 40: 272–282.

ELHORST, J. P. (2012) Dynamic spatial panels: models, methods, and inferences. *Journal of Geographical System*, 14: 5–28.

ERTUR C., KOCH, W. (2007) Growth, Technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22: 1033–1062.

ERTUR C., LE GALLO J., BAUMONT C. (2006) The European regional convergence process, 1980-1995: Do spatial regimes and spatial dependence matter? *International Regional Science Review*, 29: 3-34.

ERTUR, C., LE GALLO, J. (2008) *Regional growth and convergence: Heterogenous reaction versus interaction in spatial econometric approaches*, Working Papers hal-00463274, HAL.

ESPOSTI, R., BUSSOLETTI, S. (2008) Impact of objective 1 funds on regional growth convergence in the European Union: A panel-data approach. *Regional Studies*, 42 (2): 159-173.

FERNANDEZ, C., LEY, E., STEEL, M. F. (2001) Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics*, 16: 563–576.

FERREIRA, A. H. B. (1999) Concentração regional e dispersão das rendas per capita estaduais : um comentário. *Estudos Econômicos*, 29: 47-63.

FERREIRA, A., (2000). Convergence in Brazil: Recent trends and long-run prospects. *Applied Economics*, 32: 479-489.

FERREIRA, P. C. (2004) *Regional policy in Brazil: A review*. Mimeograph, Rio de Janeiro.

FERREIRA, A. H. B., DINIZ, C. C. (1994) *Convergência entre as rendas per capita estaduais no Brasil*. Belo Horizonte: UFMG/CEDEPLAR,Texto para Discussão, n.79.

FINGLETON, B. (1999) Estimates of time to economic convergence: An analysis of regions of the European Union. *International Regional Science Review*, 22: 3-34.

FINGLETON, B. (2010) The empirical performance of the NEG with reference to small areas. *Journal of Economic Geography*, 11: 267–279.

FINGLETON, B., LÓPEZ-BAZO, E. (2006) Empirical growth models with spatial effects. *Papers in Regional Science*, 85: 177-198.

FISCHER, M. (2011) A spatial Mankiw–Romer–Weil model: theory and evidence. Annals of Regional Science, 47(2): 419-436.

FISCHER, M., SCHERNGELL, T., REISMANN, M. (2009) Knowledge spillovers and total factor productivity: Evidence using a spatial panel data model. *Geographical Analysis*, 41(2): 204–220.

FISCHER, M. M., STIRBÖCK, C. (2006) Pan-European regional income growth and club-convergence. *Annals of Regional Science*, 40: 693-721.

FLEMING, J. M. (1955) External Economies and the Doctrine of Balanced Growth. *Economic Journal*, 65: 241-56.

FLORAX, R., FOLMER, H. (1992) Specification and estimation of spatial linear regression models: Monte-Carlo evaluation of pre-test estimator. *Regional Science and Urban Economics*, 22: 405-32.

FLORAX, R. J., FOLMER, H., REY, R. J. (2003) Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics*, 33: 557-579.

FLORIO, M. (2006) Cost-benefit analysis and the European Union cohesion fund: On the social cost of capital and labour. *Regional Studies*, 40 (2): 211 - 224.

FOLMER H., OUD J. (2008) How to get rid of W? A latent variable approach to modelling spatially lagged variables. *Environment and Planning A*, 40: 2526–2538.

FOOTE, C. L. (2007) *Space and time in macroeconomic panel data: Young workers and state-level unemployment revisited*. Working Paper No. 07-10, Federal Reserve Bank of Boston.

FOTHERINGHAM A.; BUNSDEN, C.; CHARLTON M. (2000). *Quantitative Geography*. London: Sage.

FRIEDMAN, M. (1992) Do old fallacies ever die? *Journal of Economic Literature*, 30: 2129–2132.

FUJITA, M., KRUGMAN, P., VENABLES, A. (1999) *The spatial economy*. Cambridge, Mass.: MIT.

FUJITA, M., THISSE, J. F. (2002) *Economics of agglomeration: cities, industrial location and regional growth*. Cambridge, UK.: Cambridge University.

FURTADO, C. (1997) Relatório do grupo de trabalho para o desenvolvimento do nordeste. *Revista Econômica do Nordeste*, 28 (4): 375-386.

FRENKEL, A., SHEFER, D., ROPER, S. (2003) Public policy, locational choice and the innovation capability of high tech firms: A comparison between Israel and Ireland. *Papers in Regional Science*, 82: 203-221.

GALOR, O., MOAV, O. (2004). From physical to human capital accumulation: inequality and the process of development. *Review of Economic Studies*, 71: 1001–26.

GALOR, O., ZEIRA, J. (1993) Income Distribution and Macroeconomics. *Review of Economic Studies*, 60: 35-52.

GAMBOA, O. R. E. (2010) The (un)lucky neighbour: Differences in export performance across Mexico's states. *Papers in Regional Science*, 89 (4):777-799.

GEHLKE,C., BIEHL,K. (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of American Statistical Association*, 29: 169-170.

GÉRARD, M., JAYET, H., PATY, S. (2010) Tax interactions among Belgian municipalities: Do interregional differences matter? *Regional Science and Urban Economics*, 40: 336–342.

GETIS, A., GRIFFITH, D. A. (2002) Comparative spatial filtering in regression analysis. *Geographical Analysis*, 34: 130-40.

GIBBONS, S., OVERMAN, H. (2012) Mostly pointless spatial econometrics? *Journal of Regional Science*, 52: 172–191.

GIBBONS, S., OVERMAN, H., PELKONEN, P. (2010) *Wage disparities in Britain: People or Place?* London School of Economics Spatial Economics Research Centre Discussion Paper 0060.

GIBBONS, S., OVERMAN, H., RESENDE, G. M. (2011) *Real earnings disparities in Britain*, London School of Economics Spatial Economics Research Centre Discussion Paper 0065.

GLAESER, E. L., SACERDOTE, B. I., SCHEINKMAN, J. A. (2003) The social multiplier. *Journal of the European Economic Association*, 1 (2-3): 345-353.

GREENLAND, S. (2002). A review of multilevel theory for ecologic analyses. *Statistics in Medicine*, 21: 389–395.

GRIER, K., TULLOCK, G. (1989) An empirical analysis of cross national economic growth, 1951–80. *Journal of Monetary Economics*, 24 (2): 259–276.

GROVES, T., LEDYARD, J. (1977) Optimal allocation of public goods: A solution to the free rider problem. *Econometrica*, 45 (4): 783-809.

HANNAN, M. T. (1971) *Aggregation and disaggregation in sociology*. Lexington, Mass.: Lexington Books.

HARRIS, R., KRAVTSOVA, V. (2009) *In search of 'W'*. SERC Discussion paper n.17, p.1-20.

HARRIS, R., TRAINOR, M. (2005) Capital subsidies and their impact on total factor productivity: Firm-level evidence from Northern Ireland. *Journal of Regional Science*, 45: 49–74.

HARROD, R. F. (1948) *Toward a dynamic economics: Some recent developments of economic theory and their application to policy*. London: Macmillan.

HAUGHWOUT, A. (1998) Aggregate production functions, interregional equilibrium, and the measurement of infrastructure productivity. *Journal of Urban Economics*, 44: 216–227.

HENDERSON, J. V. (2010) Cities and development. *Journal of Regional Science*, 50(1): 515–540.

HENDRY, D. F. (1979) Predictive failure and econometric modelling in macroeconomics: The transactions demand for money. In: Ormerod, P. (ed) *Economic modelling*. London: Heinemann.

HIRSCHMAN, A. (1958) *The strategy of economic development*. New Haven: Yale University Press.

IBGE (2009) *Sistema de contas regionais referência 2009 – Produto interno bruto dos municípios 1999-2008* (available at: http://www.sidra.ibge.gov.br/bda/pesquisas/PIBMun/default.asp).

IBGE (2011) *Division into regions*. Instituto Brasileiro de Geografia e Estatística. (available at: http:// www.ibge.gov.br/english/geociencias/geografia/default_div_int.shtm, accessed on 10/01/2011).

IPEADATA. *Dados macroeconômicos e regionais* (available at: http://www.ipeadata.gov.br).

IPEA, PNUD, FJP (2003) *Atlas do desenvolvimento humano no Brasil.* Brasília.

ISLAM, N. (1995) Growth empirics: A Panel data approach. *Quarterly Journal of Economics*, 110 (4): 1127-1170.

ISLAM, N. (2003) What have we learnt from the convergence debate? *Journal of Economic Surveys*, 17: 309–362.

JANN, W., WEGRICH, K. (2007) Theories of the policy cycle. In: Fischer, F., Miller, G., Sidney, M. (eds). *Handbook of public policy analysis*, New York: CRC Press.

JEANTY, P., PARTRIDGE, M., IRWIN, E. (2010) Estimation of a spatial simultaneous equation model of population migration and housing price dynamics. *Regional Science and Urban Economics*, 40: 343–352.

KELLER, W., SHIUE, C. H. (2007) The origin of spatial interaction. *Journal of Econometrics*, 140: 304–332.

KHANDKER, S. R., KOOLWAL, G. B., SAMAD, H. A. (2010) *Handbook on impact evaluation: Quantitative methods and practices*. The World Bank: Washington D.C.

KIRMAN, A. (1992) Whom and what does the representative individual represent? *Journal of Economic Perspectives*, 6: 117–136.

KORMENDI, R., MEGUIRE, P. (1985) Macroeconomic determinants of growth: Cross country evidence. *Journal of Monetary Economics*, 16 (2): 141–163.

KRUEGER, A. O. (1990) Government failures in development. *Journal of Economic Perspectives*, 4 (3): 9-23.

KRUGMAN, P. (1987) The Narrow moving band, the Dutch disease, and the competitive consequences of Mrs Thatcher. *Journal of Development Economics*, 27 (1-2): 41-55.

KRUGMAN, P. (1991) Increasing returns and economic geography. *Journal of Political Economy*, 99: 483-499.

KRUGMAN, P. (1995) *Development, geography, and economic theory*. Cambridge: MIT Press.

LAFOURCADE, M., THISSE, J. (2008) *New economic geography: A guide to transport analysis*. PSE Working Papers 2008-02, PSE ,Ecole normale supérieure.

LALL, S.V., SHALIZI, Z. (2003) Location and growth in the Brazilian Northeast. *Journal of Regional Science*, 43: 663-681.

LAMBRINIDIS, M., PSYCHARIS, Y., ROVOLIS, A. (2005) Regional allocation of public infrastructure investment: The case of Greece. *Regional Studies*, 39(9): 1231 - 1244.

LASSWELL, H. D. (1956) *The decision process: Seven categories of functional analysis*. College Park, Maryland: University of Maryland Press.

LAURINI, M., ANDRADE, E., VALLS, P., (2005). Income convergence clubs for Brazilian Municipalities: a non-parametric analysis. *Applied Economics*, 37 (18): 2099-2118.

LAW Nº 7827 (1989) (available at: http://www.planalto.gov.br/ccivil_03/Leis/L7827.htm).

LAW Nº 9394 (1996) (available at: http://www.planalto.gov.br/ccivil_03/LEIS/l9394.htm).

LAW Nº 11107 (2005) (available at: http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/ Lei/L11107.htm).

LE GALLO, J., ERTUR, C. (2003) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. *Papers in Regional Science*, 82: 175–201.

LEAMER, E. E. (1978) *Specification searches: ad hoc inference with nonexperimental data*. NewYork: Wiley & Sons.

LEAMER, E.E. (1983) Let's take the con out of econometrics. *American Economic Review*, 73: 31–43.

LEAMER, E. E. (1985) Sensitive analyses would help. *American Economic Review*, 75: 308-313.

LEE,L. (2004) Asymptotic distributions of quasi maximum likelihood estimators for spatial econometric models. *Econometrica*, 72: 1899-1926.

LEE, K., PESARAN, M., SMITH, R. (1997) Growth and convergence in multi country empirical stochastic Solow model. *Journal of Applied Econometrics*, 12 (4): 357–392.

LEE, K., PESARAN, M., SMITH, R. (1998) Growth empirics: a panel data approach: A comment. *Quarterly Journal of Economics*, 113 (1): 319–323.

LEE, L., YU, J. (2010a) Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40: 255–271.

LEE, L., YU, J. (2010b) Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2):165–185.

LEE, L., YU, J. (2010c). *Efficient GMM estimation of spatial dynamic panel data models with fixed effects* (available at: http://www.economics.smu.edu.sg/events/Paper/LungfeiLee.pdf).

LEE, L., YU, J. (2010d) A spatial dynamic panel data model with both time and individual fixed effects. *Economic Theory*, 26(2):564–597.

LEONARDI, R. (2006) Cohesion in the European Union. *Regional Studies*, 40 (2): 155 - 166.

LESAGE, J. P. (1999) *Spatial econometrics*. The Web Book of Regional Science, Regional Research Institute, West Virginia University, Morgantown, WV (available at: http://www.rri.wvu.edu/ WebBook/LeSage/spatial/spatial.html).

LESAGE, J. P., FISCHER, M. M. (2008) Spatial growth regressions: Model specification, estimation and interpretation. *Spatial Economic Analysis*, 3(3):275.

LESAGE, J. P., PACE, R. K. (2008) Spatial econometric modeling of origin-destination flows. *Journal of Regional Science*, 48: 941-967.

LESAGE, J. P., PACE, R. K. (2009) *Introduction to spatial econometrics*. CRC Press/Taylor and Francis Group, London.

LEVINE, R., RENELT, D. (1992) A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82: 942-963.

LEY, E., STEEL, M. (2007) Jointness in variable selection with applications to growth regression. *Journal of Macroeconomics*, 29: 476-493.

LÓPEZ-BAZO, E., VAYÁ, E., MORA, A., SURINACH, J. (1999) Regional economic dynamics and convergence in the European Union. *Annals of Regional Science*, 33: 343–70.

LÓPEZ-BAZO, E., ARTÍS, M., VAYÁ, E. (2004) Regional externalities and growth: evidence from European regions. *Journal of Regional Science*, 44: 43–73.

LUCAS, R. (1988) On the mechanics of economic development. *Journal of Monetary Economics*, 22 (1): 3-42.

MAGALHÃES, A., HEWINGS, G., AZZONI, C. (2000) *Spatial dependence and regional convergence in Brazi*l. Chicago: University of Illinois, Working Papers REAL.

MAGRINI, S. (2004) Regional (di)convergence. In: Henderson, J. V., Thisse, J. F. (eds) *Handbook of Regional and Urban Economics*, ed. 1, vol. 4, chapter 62: 2741-2796.

MANKIW, N. G., ROMER, D., WEIL, D. N. (1992) A Contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107: 407-37.

MANSKI, C. F. (1993) Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60: 531-42.

MANSKI, C. F. (2000) Economic analysis of social interactions. *The Journal of Economic Perspectives*, 14 (3): 115-136

MARTIN, R., TYLER, P. (2006) Evaluating the impact of the structural funds on objective 1 regions: An exploratory discussion. *Regional Studies*, 40 (2): 201 — 210.

MCDONALD, S., ROBERTS, J. (2002) Growth and multiple forms of human capital in an augmented Solow model: A panel data investigation. *Economics Letters*, 74: 271-276.

MCMILLEN, D.P. (2010) Issues in spatial data analysis. *Journal of Regional Science*, 50(1): 119-141.

MENEZES-FILHO, N.A (2001) Educação e desigualdade. In: Lisboa, M., Menezes Filho, N.A. (eds) *Microeconomia e Sociedade no Brasil*, Contra Capa Livraria, Rio de Janeiro.

MICHELIS, N., MONFORT, P. (2008) Some reflections concerning GDP, regional convergence and European cohesion policy. *Regional Science Policy and Practice*, 1 (1): 15-22.

MULDER, P., DE GROOT, H. L. F. (2007) Sectoral energy- and labour-productivity convergence. *Environmental & Resource Economics*, 36 (1): 85-112.

MURPHY, K. M., SHLEIFER, A., VISHNY, R. W. (1989) Industrialization and the big push. *The Journal of Political Economy*, 97 (5): 1003-1026.

NEWEY, W. K., WEST, J. D. (1987) A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55: 703-8.

NICKELL, S. (1981) Biases in dynamic models with fixed effects. *Econometrica*, 49: 1417–1426.

North, D. (1990) *Institutions, institutional change and economic performance*. Cambridge: Cambridge University Press.

NURSKE, R. (1953) *Problems of capital formation in underdeveloped countries*. New York: Oxford Press.

OATES, W. E. (1999) An essay on fiscal federalism. *Journal of Economic Literature*, 37: 1120-1149.

OECD (2009a) *How regions grow*. Paris: Organisation for Economic Growth and Development.

OECD (2009b) *Regions matter: Economic recovery, innovation and sustainable growth*. Paris: Organisation for Economic Growth and Development.

OLIVEIRA, H., DOMINGUES, E. P. (2005) *Considerações sobre o impacto dos FNO e FCO na redução da desigualdade regional no Brasil*.In: XXXIII Encontro Nacional de Economia, Natal.

OPENSHAW, S. (1984) Ecological fallacies and the analysis of areal census data. *Environment and Planning A*, 16: 17-31.

OPENSHAW, S., TAYLOR, P.J. (1979) A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In: Wrigley, N. (ed) *Statistical Applications in the Spatial Sciences*. London: Pion.

OPENSHAW, S., TAYLOR, P.J. (1981) The Modifiable Areal Unit Problem. In: Wrigley, N., Bennett, R. (ed) *Quantitative Geography, a British View*. London: Routledge and Kegan.

OTTAVIANO, G., PINELLI, D. (2006) Market potential and productivity: Evidence from Finnish regions. *Regional Science and Urban Economics,* 36: 636–657.

OTTAVIANO, G., PUGA, D. (1998) Agglomeration in the global economy: A survey of the new economic geography. *World Economy*, 21: 707-731.

PAGE, S. E., TASSIER, T. (2010) A characterization of equilibria in the Groves Ledyard mechanism. *Journal of Mathematical Economics*, 46 (6): 1229-1242.

PARTRIDGE, M. D., BOARNET, M., BRAKMAN, S., OTTAVIANO, G. (2012) Introduction: Whither spatial econometrics? Journal of Regional Science, 52 (2): 167-171.

PEETERS, L., CHASCO, C. (2006) Ecological inference and spatial heterogeneity: an entropy-based distributionally weighted regression approach. *Papers in Regional Science*, 85: 257- 276.

PEREIRA, R. O. (2007). Ação do BNDES sobre o emprego formal: Efeito nas empresas financiadas. *Revista do BNDES*, 27: 27-42.

PEREIRA, A., ANDRAZ, J. (2006) Public investment in transportation infrastructures and regional asymmetries in Portugal. *Annals of Regional Science*, 40: 803–817.

PÉREZ, J., DONES, M., LLANO, C. (2009) An interregional impact analysis of the EU Spanish structural funds in Spain (1995-1999). *Papers in Regional Science*, 88 (3): 509-529.

PEROTTI, R. (1996) Growth, income distribution, and democracy: What data say. *Journal of Economic Growth*, 1: 149–87.

PESSÔA, S. (2001) Existe um problema de desigualdade regional no Brasil? In: *Anais do XXIX Encontro Nacional de Economia – ANPEC, 2001*, Salvador/BA, Anais. Salvador.

PICHELMANN, K., ROEGER, W. (2008) Employment and labour productivity in the EU: Reconsidering a potential trade-off in the Lisbon strategy. In: Eichengreen, B., Landesmann, M., Stiefel, D. (eds) *The European economy in an American mirror*, Routledge: London and New York, p. 128-142.

PIKE, A., RODRIGUEZ-POSE, A., TOMANEY, J. (2007) What kind of local and regional development and for whom? *Regional Studies*, 41(9): 1253–1269.

POSTALI (2009) Petroleum royalties and regional development in Brazil: The economic growth of recipient towns. *Resources Policy*, 34: 205–213.

PRITCHETT, L. (1996) *Where has all the education gone?* World Bank working paper n. 1581.

PUGA, D., VENABLES, A. (1996) The spread of industry: Spatial agglomeration in economic development. *Journal of the Japanese and International Economics,* 10: 440–464.

PÜLZL, H., TREIB, O. (2007) Implementing public policy. In: Fischer, F., Miller, G., Sidney, M. (eds) *Handbook of public policy analysis*. New York: CRC Press.

QUAH, D. (1993) Galton's fallacy and tests of the convergence hypothesis. Scandinavian Journal of Economics 95: 427-443.

QUAH, D. (1996) Empirics for economic growth and convergence. *European Economic Review*, 40: 1353–1375.

QUAH, D. (1997) Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2: 27-59.

RAIS (1999, 2000, 2003, 2006) *Ministério do Trabalho. Relatório Anual de Informações Sociais*. Brasilia.

RAVALLION, M., JALAN, J. (1996) Growth divergence due spatial externalities. *Economics Letters*, 53: 227-232.

RAY, D. (1998). *Development economics*. Princeton: Princeton University Press.

REIS E. J., TAFNER P., PIMENTEL M., MEDINA M., MAGALHAES K., REIFF L. O., SERRA R.V. (2005) *O PIB dos municípios brasileiros: metodologia e estimativas, 1970-96*. IPEA Discussion Paper, n. 1064, Brasília.

RESENDE, G. M. (2011) Multiple dimensions of regional economic growth: The Brazilian case, 1991-2000. *Papers in Regional Science*, 90 (3): 629-662.

RESENDE, G. M., FIGUEIRÊDO, L. (2010) Testes de robustez: Uma aplicação para os determinantes do crescimento econômico estadual brasileiro entre 1960 e 2000. *Revista Econômica do Nordeste*, 41: 9-39.

RESENDE, G. M., CARVALHO, A. X. Y., SAKOWSKI, P. A. M. (2012) *Evaluating multiple spatial dimensions of economic growth in Brazil using spatial dynamic panel data models, 1970-2000*. Manuscript, Brasilia.

REY, S., ANSELIN, L. (2000) Regional science publication patterns in the 1990s. *International Regional Science Review*, 23 (4): 323–344

REY S. J., DEV, B. (2006) σ-convergence in the presence of spatial effects. *Papers in Regional Science*, 85: 217-234.

REY, S. J., JANIKAS, M. V. (2005) Regional convergence, inequality, and space. *Journal of Economic Geography*, 5: 155-176.

REY, S. J., MONTOURY, B. D. (1999) U.S. regional income convergence: A spatial econometric perspective. *Regional Studies*, 33: 143-156.

ROBINSON, W. S. (1950) Ecological correlations and the behaviour of individuals. *American Sociological Review*, 15: 351-357.

ROBINSON, S. (1971) Sources of growth in less developed countries: A cross-section study. *Quarterly Journal of Economics*, 85 (3): 391-408.

RODRIGUEZ-POSE, A., FRATESI, U. (2004) Between development and social policies: The impact of European structural funds in objective 1 regions. *Regional Studies*, 38 (1): 97-113.

ROMER, P. M. (1986) Increasing returns and long-run growth. *Journal of Political Economy*, 94 (5): 1002-1037.

ROMER, P. (1990) Endogenous technological change. *Journal of Political Economy*, 98 (5): 71-102.

ROMERO, R. (2009) Estimating the impact of England's area-based intervention 'New Deal for Communities' on employment. *Regional Science and Urban Economics*, 39: 323–331.

ROMERO, R., NOBLE, M. (2008) Evaluating England's 'New Deal for Communities' programme using the difference-in-difference method. *Journal of Economic Geography*, 8: 759-778.

ROODMAN, D. (2009) A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71: 135–158.

ROSENSTEIN-RODAN, P. (1943) Problems of industrialization of Eastern and Southeastern Europe. *Economic Journal*, 53: 202-211.

SALA-I-MARTIN, X. (1996a) The classical approach to convergence analysis. *Economic Journal*, 106: 1019–1036.

SALA-I-MARTIN, X. (1996b) Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review*, 40: 1325–1352.

SALA-I-MARTIN, X. (1997) I just ran two million regressions. *American Economic Review*, 87: 178–83.

SALA-I-MARTIN, X., DOPPELHOFER, G., MILLER, R. (2004) Determinants of long-term growth: a Bayesian averaging of classical estimates, BACE approach. *American Economic Review*, 94: 813–835.

SAPIR, A., AGHION, P., BERTOLA, G., HELLWIG, M., PISANI-FERRY, J., ROSATI, D., VINALS, J., WALLACE, H. (2004) *An agenda for a growing Europe: The Sapir report*. Oxford: Oxford University Press.

SCITOVSKY, T. (1954) Two concepts of external economies. *Journal of Political Economy*, 62:143-151.

SEN, A. (2000) *Development as freedom*. New York: Anchor Books.

SIDNEY, M. (2007) Policy formulation: Design and tools. In: Fischer, F., Miller, G., Sidney, M. (eds) *Handbook of public policy analysis*. New York: CRC Press.

SILVA, A. M., RESENDE, G. M., SILVEIRA NETO, R.. (2007) Uma avaliação da eficácia do FNE, no período 1995-2000. *Análise Econômica* (UFRGS), 25: 233-261.

SILVA, A. M., RESENDE, G. M., SILVEIRA NETO, R. (2009) Eficácia do gasto público: uma avaliação do FNE, FNO e FCO. *Estudos Econômicos*, 39 (1): 89-125.

SILVEIRA NETO, R. M. (2001) Crescimento e spillovers: a localização importa? Evidências para os estados brasileiros. *Revista Econômica do Nordeste*, 32: 524-545.

SILVEIRA NETO, R. M., AZZONI C. (2006) Location and regional income disparity dynamics: The Brazilian case. *Papers in Regional Science*, 85: 599-613.

SKURAS, D., TSEKOURAS, K., DIMARA, E., TZELEPIS, D. (2006) The effects of regional capital subsidies on productivity growth: A case study of the Greek food and beverage manufacturing industry. *Journal of Regional Science*, 46: 355–381.

SOARES R., SOUSA J., PEREIRA NETO, A. (2009) Avaliação de impactos do FNE no emprego, na massa salarial e no salário médio em empreendimentos financiados. *Revista Econômica do Nordeste*, 40 (1): 217-234.

SOLOW, R. M. (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70: 65-94.

SØRENSEN, A. (2001) Comparing apples to oranges: Productivity convergence and measurement across industries and countries: Comment. *American Economic Review*, 91, 1160–1167.

SOUSA, F., L., OTTAVIANO, G. I. P. (2009) *The effects of BNDES loans on the productivity of Brazilian manufacturing firms*. Mimeograph. Rio de Janeiro, Brazil.

STAKHOVYCH, S., BIJMOLT, T. H. A. (2008) Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, 88(2): 389-408.

TEMPLE, J. (1999) The new growth evidence. *Journal of Economic Literature*, 37: 112-156.

TEMPLE, J. (2000) Growth regressions and what the textbooks do not tell you. *Bulletin of Economic Research*, 52: 181-205.

TOBLER, W. R. (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46: 234-240.

TRIFFIN, R. (1940) *Monopolistic competition and general equilibrium theory*. Harvard University Press: Cambridge, MA.

VENABLES, A. (1998) Localization of industry and trade performance. *Oxford Review of Economic Policy*, 12: 52–60.

VERGOLINO, J., NETO, A. P., BARROS, M. (2004) Crescimento econômico regional no Brasil: Educação como fator de convergência,1970/96. *Economia (Curitiba)*, 30: 79-103.

WORLD BANK (2009) *World development report 2009: Reshaping economic geography*. Washington, DC: World Bank.

WOOLDRIDGE, J. (2002) *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

YAMAMOTO, D. (2008) Scales of regional income disparities in the USA, 1955–2003. *Journal of Economic Geography*, 8: 79-103.