

Cravo, Túlio

**Conference Paper**

## Regional Economic Growth and SMEs in Brazil: a Spatial Analysis (Submission for the Refereed Y-session Papers).

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Cravo, Túlio (2010) : Regional Economic Growth and SMEs in Brazil: a Spatial Analysis (Submission for the Refereed Y-session Papers)., 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/118904>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Regional Economic Growth and SMEs in Brazil: a Spatial Analysis<sup>\*</sup>

Túlio A. Cravo<sup>†</sup>

Department of Economics, Loughborough University  
t.antonio-cravo@lboro.ac.uk

June 13, 2010

## Abstract

This paper examines regional economic growth for a panel of 508 Brazilian micro-regions for the period 1980-2004 using spatial panel econometrics and paying particular attention to the importance of small and medium enterprises (SMEs). The empirical findings indicate the presence of significant spatial dependence in the process of economic growth. Human capital is an important growth determinant, but does not generate positive spillovers, reinforcing the need of localised investments in human capital. Furthermore, human capital embodied in SMEs seems to be more important than the size of this sector for regional growth but does not generate positive spatial spillovers either. Conversely, the interaction of the relative size of the SME sector across space is positively related to economic growth, suggesting that more SMEs in neighbouring regions incites more productive entrepreneurship and new economic possibilities that affect economic performance positively. This is in line with the idea that entrepreneurial activities have a more positive influence on neighbouring regions than on distant ones due to cultural aspects. Therefore, a SME policy that subsidizes the creation of SMEs does not generate economic growth directly but through spatial interactions. Moreover, the promotion of local human capital in SMEs seems to be important for regional growth.

**Keywords:** Firm size, market structure, economic growth, spatial econometrics.

**JEL Classification:** O10, L11, R11, C31.

---

<sup>\*</sup> I would like to thank Adrian Gourlay, Bettina Becker, Huw Edwards and Ahmad H. Ahmad for helpful comments, and André Rodrigues Nagy and Cristiano Aguiar for useful comments that helped in the construction of the data used in this paper. Finally, I would like to thank the Brazilian Ministry of Labour for providing access to RAIS database.

<sup>†</sup> Corresponding address: T. A. Cravo, Department of Economics, Loughborough University, Leicestershire, LE11 3TU, United Kingdom. E-mail: t.antonio-cravo@lboro.ac.uk

## 1 Introduction

This paper studies Brazilian regional economic growth considering the presence of spatial spillovers and paying particular attention to the importance of small and medium sized enterprises (SMEs) in this process. The importance of spatial dependence when analyzing regional economic growth is well documented in the literature and has been broadly considered by the use of spatial econometrics. Also, the interest in the study of SMEs and economic growth has been attracting increasing attention, largely because this sector is responsible for the majority of employment generation either in developed or in developing countries. This explains the large amount of financial resources allocated to the development of the SME sector. For instance, in 2003 the World Bank approved US\$ 1.3 billion in programs related to SMEs support (Beck et al., 2005) and the budget of the Brazilian small business support service (SEBRAE) amounted to US\$ 1.25 billion in 2009<sup>1</sup>.

There is a growing consensus that regional growth models exhibit spatial dependence and require spatial econometrics to be estimated correctly (Lesage and Fischer, 2008). The seminal work of Anselin (1988) demonstrated the need for considering spatial dependence in the form of spatial autocorrelation and spatial heterogeneity. The former can be defined as the coincidence of value similarity with location similarity, and the latter means that economic behaviour is not stable across geographic space, generating spatial regimes with clusters of high or low values of a variable. The omission of this spatial phenomenon in economic growth regressions can produce biased results as shown by Rey and Montouri (1999). After these seminal works, spatial econometrics has become an essential tool when regional economic growth is analysed. For instance, the survey by Abreu et al. (2005) reviews empirical and theoretical works and shows that spatial dependence is an important aspect of the economic growth process. The importance of spatial dependence when analysing economic growth was incorporated theoretically in growth models (e.g., López-Bazo et al., 2004; Ertur and Koch, 2007) and has been extensively documented in empirical papers (e.g., Badinger et al., 2004; Ertur et al., 2006; Ramajo et al., 2008; Arbia et al., 2008; Lesage and Fisher, 2008; Mohl and Hagen, 2010). Also, for the Brazilian case, recent studies using cross-sectional data recognize the importance of spatial spillovers to the process of economic

---

<sup>1</sup> This information is available at [www.sebrae.com.br](http://www.sebrae.com.br). Conversion is based on the official exchange rate of 31/12/2009 provided by the Brazilian Central Bank. Approximately 75% of SEBRAEs' budget comes directly from compulsory contribution collected from enterprises.

growth (e.g., Mossi et al., 2003; Silveira-Neto and Azzoni, 2006; Resende, 2009). Hence, testing for spatial dependence and estimating regional growth models using spatial econometrics are becoming the standard procedure in order to accommodate models that incorporate spatial spillovers and obtain non-biased estimations.

Recently, some authors have also been arguing that entrepreneurship and small firms are another crucial omitted aspect in the neoclassical growth framework (e.g., Audretsch and Keilbach, 2004; Audretsch, 2007). Solow (2007), for instance, recognizes entrepreneurship as an important force that drives a wedge between knowledge and total factor productivity. It can bridge the gap between specific pieces of technological knowledge and innovations through the creation of new firms. Therefore, explaining how SMEs affect growth could add to the explanatory power of growth theory. For instance, the interest in the study of small and medium sized enterprises (SMEs) and economic growth is illustrated by an important paper by Beck et al. (2005). They estimate the standard growth regression including the relative size of the SME sector (measured by the share of SME sector employment in the total manufacturing employment) into the growth framework and find an insignificant impact of this sector on economic growth for a cross-section of countries. A similar approach is used in Audretsch and Keilbach (2004), Mueller (2007), and Carree and Thurik (1998)<sup>2</sup>. Their findings suggest a positive impact of measures of entrepreneurship on economic growth in the context of developed countries. Although there is a paucity of studies for developing countries relative to developed ones, Cravo (2010) provides evidence of a negative impact of entrepreneurship on regional growth in Brazil. In common, all these studies consider the importance of entrepreneurship for economic growth without incorporating the spatial dimension in the analysis and their conclusions might be based on biased results.

Therefore, the available empirical evidence might be biased and suggests that whilst studies that focus on developing nations suggest a negative or neutral impact of SMEs and entrepreneurship on economic growth, studies examining developed countries suggest a positive impact. Acs et al. (2008) has attributed these differences in empirical results to different entrepreneurship responses to institutional arrangements. Heterogeneity in institutional arrangements and human capital levels across countries and regions are likely to provide different incentives to productive entrepreneurship. Baumol (1990) suggests that

---

<sup>2</sup> The start-up rates used in Audretsch and Keilbach (2004) and Mueller (2007) is a dynamic measures of entrepreneurship and the stock of SMEs' employment used in Beck et al. (2005) is a static measure. They represent different dimensions of entrepreneurship but are clearly related concepts. Higher small businesses start-up rates will lead to a higher share of workers in the SME sector, while a higher mortality rate will adjust the size of this sector in opposite direction.

while the total supply of entrepreneurs differs across economies, the productive contribution of the society's entrepreneurial activities varies much more due to their allocation between productive and unproductive activities. Hence, due to differences in institutions, human capital and rent-seeking levels, the presence of SMEs in a developing economy probably does not have the same meaning as in a knowledge based economy<sup>3</sup> and policy makers should aim at providing a good institutional arrangement that promotes productive entrepreneurship at the expense of rent seeking<sup>4</sup>. Institutions also incite human capital formation for productive entrepreneurs. Dias and McDermott (2006) propose a model where structural changes towards a modern economy depend on the role of entrepreneurs, human capital and institutions. In their model, entrepreneurs come from a pool of individuals that belong to the managerial class, which is specialized in two activities; rent-seeking and entrepreneurship. The important point is that more (productive) entrepreneurs lead to more human capital formation. Workers will look for education that suits a more productive job (that requires higher level of human capital) offered by the entrepreneurs. Therefore, barriers that prevent productive entrepreneurship to develop should be removed and better institutional policies are required to create more productive entrepreneurs and improve economic performance. If unproductive entrepreneurship dominates, educational improvement will be neutralized and will have little long run effect.

Nevertheless, the empirical literature on economic growth using spatial econometrics in Brazil is scant and based only on cross-section estimates and the papers that analyse the relationship between growth and measures related to entrepreneurship in a regional growth model framework have been ignoring the spatial dependence. Therefore, the aim of this paper is to address these gaps in the literature by providing an analysis of the importance of spatial spillovers for regional economic growth in Brazil for a panel of micro-regions from 1985 to 2004 paying particular attention to the importance of SMEs in this process. Understanding better the relationship between geography, economic growth and SMEs is important because the SME sector employs the majority of the labour force, and attracts considerable financial support from government, national and multilateral institutions. Hence, this paper contributes to the literature in various ways. Firstly, it provides comprehensive empirical evidence on the

---

<sup>3</sup> For example, Table 1 in Beck et al. (2005) shows that SMEs' employment share in Brazil is 58.8% while in Finland is 58.15%. However, the effect of SMEs on growth is likely to be different due to different characteristics of the SME sector.

<sup>4</sup> According to Baumol (1990), unproductive entrepreneurship takes many forms. Litigation and takeovers, tax evasion and avoidance efforts seem to constitute the prime threat to productive entrepreneurship.

importance of space for the regional economic growth process in Brazil using spatial panel econometrics. Secondly, to provide an extended analysis of the relationship between SMEs and economic growth by focusing on the existence of spatial spillover, this paper draws on Beck et al. (2005) and uses growth regressions to encompass the importance of the SME sector into the growth analysis. Besides, this paper investigates how different aspects of the SME sector (size and human capital) influence economic growth in different spatial regimes. More detailed empirical evidence on these matters is important to better understand how spatial aspects affect different growth determinants. Particularly, the inclusion of aspects of the SME sector (size and human capital) is very important to determine the appropriate SMEs' public policy focus across regions.

In this paper we address these issues considering the use of spatial panel econometrics. It finds that space cannot be ignored when studying regional economic growth in Brazil. Furthermore, the relative size of the SME sector affects regional economic growth only through spatial spillovers and the quality of the SME sector expressed by its level of human capital is important for growth but does not generate spillovers. The remainder of the paper is organized as follows. In the second section we present the baseline model and its extensions to consider the spatial structure. In section three we present the data and the spatial weight matrices and focus on the initial analysis of the spatial dependence. In section four we present the results, and the last section concludes.

## **2 The Baseline Model**

The baseline specification stems from the neoclassical growth model based on Solow (1956) and Mankiw et al. (1992) that consider physical and human capital as the main variables conditioning the convergence process. However, Temple (1999) argues that the most common approach is the use of a more *ad hoc* regression that encompasses other factors that influence growth. Regressions of this type are known as "Barro Regressions", after Barro (1991)'s seminal work. In this approach, variables are chosen based on previous results in the literature rather than on an explicit theoretical model. In this context, Levine and Renelt (1992) and Durlauf et al. (2005) listed an extensive number of variables that were used in growth regressions and Sala-i-Martin (2002) argues that one important contribution made by the growth literature that follows this tradition is that it has exerted influence on other economic literature such as development, economic geography, macroeconomics, econometrics and industrial organisation. Recently, this influence was also extended to study the importance of

SMEs and entrepreneurship for economic growth (e.g., Beck et al., 2005; Audretsch and Keilbach, 2004; Mueller, 2007). Hence, for a panel data framework that controls for individual and time fixed effects, the standard specification used to study economic growth in the Brazilian micro-regions considering the importance of SMEs takes the following form:

$$gr_{it} = +\beta \ln y_{i,t-1} + \phi \ln SMER_{it} + \gamma \ln SMEH_{it} + \psi \ln X_{it} + \alpha_i + \mu_t + v_{it} \quad (1)$$

where  $gr$  denotes the GDP annual per capita growth,  $\ln y_{i,t-1}$  is the initial GDP per capita,  $i$  denotes each individual economy,  $\beta$  the convergence coefficient,  $t$  represents each period of time considered and  $X$  represents a vector of control variables that can encompass growth determinants suggested by the original Solow model as well as other growth determinants that come from outside the model<sup>5</sup>. If the convergence coefficient is negative ( $\beta < 0$ ) and the coefficients of conditioning variables ( $\phi$ ,  $\gamma$  or  $\psi$ ) are different from zero, the data exhibits conditional convergence. The additional feature of our specification is the inclusion of the relative size of the SME sector ( $SMER$ ) as in Beck et al. (2005) and the human capital embodied in this sector ( $SMEH$ ) to extend their model as in Cravo (2010). However, Equation (1) ignores the existence of spatial dependence, it assumes that regional observations are independent and this can result in major model misspecification (Rey and Montouri, 1999).

## 2.1 The Model Specification with Spatial Dependencies

The underlying theoretical justification to the inclusion of spatial spillovers in growth regressions is provided in López-Bazo et al. (2004) and Ertur and Koch (2007). They explicitly take into account technological interdependencies that are assumed to operate through spatial externalities. Empirically, there are various ways of incorporating the spatial dependence in Equation (1) and we employ the most common model specifications in spatial econometrics according to Lesage and Pace (2009): the spatial error model (SEM), the spatial autoregressive model (SAR) and the spatial Durbin model (SDM).

The SEM specification reflects spatial dependence working through the random error terms and the specification of the baseline Equation (1) becomes:

---

<sup>5</sup> The variables included in vector  $X$  are the overall level of human capital and the population growth. Population growth is adjusted for depreciation ( $\delta$ ) and technological growth ( $g$ ), under the usual assumption that  $\delta + g$  equals 0.05. The terms  $\alpha_i$  and  $\mu_t$  control for the individual and time fixed effects, respectively.

$$gr_{it} = -\beta \ln(y_{i,t-1}) + \psi X_{it} + \alpha_i + \mu_t + \varepsilon_{it} \quad (2)$$

$$\varepsilon_{it} = \lambda \sum_{j=1}^N w_{ij} \varepsilon_{jt} + v_{it}$$

where variables are defined as in Equation (1), but hereinafter to simplify the notation, the relative size of the SME sector and the human capital of the SME sector are also included in vector  $X$ . The term  $\varepsilon$  is the error term, where  $w_{ij}$  contains information about the spatial structure and connectivity between regions  $i$  and  $j$ . The term  $\lambda$  is a scalar spatial error coefficient, and  $v \sim N(0, \sigma^2 I_n)$ . A random shock introduced into a specific region will not only affect the growth rate in that region but through the spatial transformation in the error term will impact the growth rates of other regions (Rey and Montouri, 1999; Ertur et al., 2006).

An alternative way of considering the spatial dependence is through the SAR specification, where the space lagged values of the dependent variable are included in the model as follows:

$$gr_{it} = \rho \sum_{j=1}^N w_{ij} gr_{jt} - \beta \ln(y_{i,t-1}) + \psi X_{it} + \alpha_i + \mu_t + v_{it} \quad (3)$$

where  $\rho$  is the scalar spatial autoregressive parameter and all other terms are defined as in Equation 2. As argued in Arbia et al. (2008), the spatially lagged dependent variable incorporates a sort of spatial conditional convergence, where convergence is conditioned and influenced by the spatial structure. Lesage and Pace (2009) argue that this specification is a hallmark of spatial econometrics, and can be used to provide useful extended versions of SAR models, such as the SDM specification that has the advantage of being sufficiently general to allow for three types of spatial interdependencies in the growth process. First, spatial effects working through the dependent variable, second, spatial effects working through the initial income variable, and third, spatial effects working through a set of conditioning variables (Lesage and Fisher 2007). The unconstrained SDM takes the following form:

$$gr_{it} = \rho \sum_{j=1}^N w_{ij} gr_{jt} - \beta_1 \ln(y_{i,t-1}) + \beta_2 \sum_{j=1}^N w_{ij} \ln(y_{j,t-1}) + \psi_1 X_{it} + \psi_2 \sum_{j=1}^N w_{ij} X_{jt} + \alpha_i + v_{it} \quad (4)$$

where everything is defined as in Equation (3), with the inclusion of the vectors  $WX$  and  $W\ln(y_{t-1})$  used to account for the spatially lagged values of all conditioning variables. This is



an attractive specification because when  $\psi_2 = 0$  and  $\beta_2 = 0$  we have the SAR model and if  $-\rho\psi_1 = \psi_2$  and  $-\rho\beta_1 = \beta_2$  the model is reduced to the SEM model (e.g., Anselin, 1988; López-Bazo et al., 2004; Bivand and Brunstad, 2006; Ertur and Koch, 2007).

Lesage and Fisher (2008) and Lesage and Pace (2009) provide a discussion about the motivations and advantages of the SDM specification for growth models. They show that the use of a SDM specification rests on the plausibility of two circumstances that are likely to arise in applied regional spatial growth regressions: the spatial dependence in the disturbances of an OLS regression and endogeneity in the form of an omitted explanatory variable (that follows a spatial autoregressive process) that exhibits non-zero covariance with the variables in the model<sup>6</sup>. Therefore, these plausible circumstances observed in applied spatial growth regressions make the SDM model a natural econometric choice over competing alternatives. Elhorst and Fréret (2009) also prefer an unconstrained SDM model to test whether there is evidence of political yardstick competition in France. The SDM model is also the natural empirical specification to capture externalities and spillovers considered theoretically in Ertur and Koch (2007) and López-Bazo et al. (2004).

Importantly, the spatial econometrics literature has shown that OLS estimation is inappropriate for models incorporating spatial effects. In the SEM specification, parameters' estimation will be unbiased, but inefficient due to the non-spherical structure of the disturbance variance matrix (Anselin 1988). In addition, the OLS estimator will be biased and inconsistent for the parameters of SAR model due to the simultaneity in the nature of the spatial autocorrelation process caused by the introduction of the spatial lag (Anselin 1988 and Lesage and Pace, 2009)<sup>7</sup>. Thus, inferences about the convergence process are based on the maximum likelihood estimator proposed by Anselin (1988) and Anselin and Hudak (1992) and on its extended version for panel data described in Elhorst (2009).

### **3 Data, Spatial Matrices and Exploratory Spatial Data Analysis**

#### **3.1 Data**

To analyse micro-regional economic growth and SMEs in Brazil, a data set combining micro-regional data with the aggregate results of individual establishment data to account for the SMEs measures at micro-regional level was constructed. This administrative geographic level

---

<sup>6</sup> For instance, data on physical capital is not available in Lesage and Fisher (2008) and they argue that this information is usually not available at regional level.

<sup>7</sup> The spatial lag is an average of neighbouring values and leads to a connectivity relation where the covariance of the error term between two regions is not zero.

consists of municipalities that share a common geographic area and characteristics. It provides a great number of observations and reduces the distortions observed at municipality level. What follows is a description of the variables and their sources.

Real GDP at factor price data for each micro-region (GDP) was collected from IBGE (*Instituto Brasileiro de Geografia e Estatística*)<sup>8</sup> and IPEADATA (*Instituto de Pesquisa Econômica Aplicada Database*)<sup>9</sup>. As a result, our data points for this series are: 1980, 1985, 1996, 2000 and 2004. This series constitutes the main constraint in terms of data availability for the construction of the panel data<sup>10</sup>. From these data points, the average annual growth rates for each time span were calculated for the GDP per capita. The data for population, used to calculate the GDP per capita and population growth ( $n$ ) were collected from IBGE. The data on average years of schooling of the population over 25 years old (*School*) were taken from IPEADATA<sup>11</sup>.

In addition, RAIS (*Relação Anual de Informações Sociais*)<sup>12</sup> provided the information necessary to construct the SMEs' measures. RAIS is a comprehensive administrative census dataset that starts in 1985 and is collected annually by the Brazilian Ministry of Labour. It covers the formal sector of the economy with demographic information for workers and establishments<sup>13</sup>. The relative size of the small and medium enterprise sector (*SMER*) is measured by the share of the SME sector employment in the total formal labour force in manufacturing, using the cut-off of 250 employees, as in Beck et al. (2005)<sup>14</sup>. Finally, we follow Muendler (2007) and Cravo (2010) and attribute a number of years of schooling to each level of educational attainment provided by RAIS to generate a continuous series<sup>15</sup> in

---

<sup>8</sup> Brazilian Institute of Geography and Statistics.

<sup>9</sup> Institute of Applied Economic Research Database.

<sup>10</sup> See Appendix A for details about this data. It is organised in intervals as close as possible to 5-years to avoid business cycle influences (e.g. Caselli et al. 1996), according to data the availability. The year of 2004 is the last of our panel because from this year onwards the GDP information is based on a new methodology.

<sup>11</sup> The data for the variable *School* is available only in a ten year time span interval from 1980 to 2000 (based on census data), and data points in between these years were constructed by interpolation. We used this variable lagged five years because 2000 was the most recent information available.

<sup>12</sup> Brazilian Annual Report of Social Information.

<sup>13</sup> Every year establishments are obliged to report all required information to the Ministry of Labour. In 2004, RAIS covers 31.5 million workers (97% of formal employment according to the Ministry of Labour). If an establishment fails to report the information required by RAIS, employers face fines that are proportional to their workforce size. However, because the payment of the annual wage supplement is based on RAIS, employers and workers have strong incentive to fulfil RAIS records.

<sup>14</sup> The classification of SMEs varies across and within countries. Ayyagari et al. (2007), for instance, provide a thorough discussion about the difficulties of collecting data and finding a common measure for SMEs. For instance, Beck et al. (2005) and the European Union adopt 250 employees as a cut-off to classify SMEs.

<sup>15</sup> See Appendix A (Table A.1) for further details.

order to construct the new variable for the average years of schooling in SMEs in manufacturing (*SMEH*) as follows:

$$SMEH = \frac{\sum_{i=0}^N (E^i \times S^i)}{E}$$

where  $E^i$  is the number of SME employees in each range of education,  $S^i$  the number of years of schooling for that particular range and  $E$  is the total number of SME employees.

The data concerning SMEs are not without caveats. First, data is only for formal sector. Second, as observed by Beck et al. (2005), SMEs' measures are static and do not account for the entry of new firms, graduation of successful SMEs into large and the exit of failing ones. In addition, RAIS is filled on a self-classification basis (according to the RAIS reference guide) and as a result, establishments might classify themselves incorrectly. In terms of measurement, accordingly to the Ministry of Labour, omissions of information are frequent in small municipalities and in construction, agriculture and public administration sectors<sup>16</sup>. However, we believe that most of these problems are offset by the census nature of RAIS and by its size.

The final panel is a balanced panel data from 1985 to 2004 with 508 (out of 558)<sup>17</sup> Brazilian micro-regions. A drawback in the dataset is the paucity of information about physical capital, however, we believe that a substantial part of the initial stock of physical capital is captured by the initial GDP per capita level. Unfortunately, this information is not available in Brazil, even at state level. Researches usually use industrial electricity consumption as proxy for physical capital at state level. However, this information is not available for a more disaggregated geographic level, such as micro-regions<sup>18</sup>.

### 3.2 Spatial Matrices and Exploratory Spatial Data Analysis

In this section, we want to check the existence of spatial dependence. Hence, the first task is to quantify the spatial structure for the Brazilian micro-regions in a form of a spatial matrix

---

<sup>16</sup> This is an additional reason for the use of micro-regions. This more aggregated territorial unit encompasses many municipalities and reduces individual bias from inaccurate data for small municipalities.

<sup>17</sup> The sample loses 50 micro-regions that did not present data at all data points. These micro-regions are a set of small municipalities usually located in poor and isolated areas, most of them in the Amazon Basin region. In 2004, this set of micro-regions encompassed only 2.3% of the Brazilian population and we believe that the omission of these regions due to lack of data does not generate a serious bias.

<sup>18</sup> Energy suppliers by law are not obliged to provide this information by municipality to the national Ministry of Energy and Mines or to any other regulator.

weight. As noted by Le Gallo and Ertur (2003) and Abreu et al. (2005), the choice of the spatial matrix should be made with caution once this choice can have a substantive impact on the results. The construction of the spatial structure should reflect the fundamental theorem of regional science, that is, distance matters. Therefore, the strength of spatial dependence should decline with the distance between observations<sup>19</sup>.

A simple way of modelling the spatial structure is based on the Queen contiguity binary spatial matrix, where regions that share a common border or vertices are assigned the value of 1. Conversely, regions that do not share borders or vertices are assigned the value of 0. We associate rows of the matrix  $w$  with the index  $i$ , and columns with the index  $j$  representing neighbouring regions to region  $i$ . For the purpose of forming a spatial lag of values from neighbours, we normalize the spatial matrix to have a row-standardized matrix  $W$ , where the elements of a row sum up to one. If we assume that regions that are further away also have influence in a given region, the same process of assigning ones to neighbours of neighbours is considered. We constructed four spatial weights based on the Queen contiguity structure from the first to the fourth spatial lag order.

Nevertheless, spatial weights constructed using this criterion can result in great variability in terms of number of neighbours for each region and create methodological problems (see Le Gallo and Ertur, 2003). Therefore, the concept of *k-nearest* neighbours calculated from the distance between regions centroids is used to construct the row-standardised spatial weights as in Le Gallo and Ertur (2003):

$$W = \begin{cases} w_{ij}(k) = 0 & \text{if } i = j \\ w_{ij}(k) = 1 & \text{if } d_{ij} \leq d_i(k) \text{ and } w_{ij}(k) = w_{ij}(k) / \sum_j w_{ij}(k) \\ w_{ij}(k) = 0 & \text{if } d_{ij} > d_i(k) \end{cases}$$

where  $d_i(k)$  is a critical cut-off distance defined for each region  $i$ . It is the  $k^{th}$  order smallest distance between regions  $i$  and  $j$  in a manner that each region  $i$  has exactly  $k$  neighbours. Matrices based on 20, 40 and 60 nearest neighbours were constructed<sup>20</sup>.

---

<sup>19</sup> This is the expression of the first law of geography (e.g. Tobler, 1970, p. 236), where “everything is related to everything else, but near things are more related than distant things”

<sup>20</sup> The spatial weight matrices for Queen contiguity and *k-nearest* neighbours were created using the software GeoDa 0.9.5-I and all versions of spatial weight matrices are row-standardised. The shapefile with all 558 micro-regions was edited to fit our data with 508 Brazilian micro-regions with data available. We obtained this shapefile from the Regional and Urban Economics Lab (NEREUS) of the University of São Paulo (USP).

However, the matrices above are a simpler representation of the spatial structure and do not represent the real dimension of distance. To overcome these limitations, we also use a weight matrix based on road distances from each pair of our 508 micro-regions<sup>21</sup>. This weight matrix is expressed as the inverse of the square distance to account for the intuition of the first law of geography that closer regions are more related than further regions and is defined as:

$$W = \begin{cases} w_{ij} = 0 & \text{if } i = j \\ w_{ij} = \frac{1}{d_{ij}^2} & \text{if otherwise} \end{cases}$$

where  $d_{ij}$  denotes the geographical distance between regions  $i$  and  $j$ <sup>22</sup>.

Following the construction of the spatial weights, we employ the Moran's  $I$  statistic to check for the global autocorrelation and the Local Indicator for Spatial Autocorrelation (LISA) based on the decomposition of the global Moran's  $I$  for each region to check for the local autocorrelation. The Moran's  $I$  statistic is given by the following expression:

$$I_G = \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2} = \frac{Z'WZ}{Z'Z} \quad (5)$$

where  $Z$  is the vector of a given variable in deviation from its mean and  $W$  is the spatial weight matrix. This index gives a formal indication of the association between the original vector of variables  $Z$  and its spatially lagged transformation  $WZ$ .

**TABLE 1: Global Spatial Autocorrelation (Moran's  $I$ )**

	$\ln GDP_{t-1}$	$\ln GDP_t$	$\ln(n+g+d)$	$\ln School$	$\ln SMEH$	$\ln SMER$
Queen 1 <sup>st</sup>	0.6773	0.6723	0.4862	0.6270	0.3398	0.2901
Queen 2 <sup>nd</sup>	0.6494	0.6143	0.3929	0.6066	0.3153	0.1925
Queen 3 <sup>rd</sup>	0.6182	0.5642	0.3113	0.5802	0.2883	0.1594
Queen 4 <sup>th</sup>	0.5930	0.5315	0.2324	0.5571	0.2582	0.1101
$K$ -nearest ( $k=20$ )	0.6458	0.6143	0.3813	0.6110	0.3464	0.1529
$K$ -nearest ( $k=40$ )	0.6231	0.5732	0.3021	0.5837	0.2991	0.1170
$K$ -nearest ( $k=60$ )	0.6111	0.5553	0.2345	0.5700	0.2798	0.0801
Inverse Distance-Row	0.5680	0.5237	0.2481	0.5258	0.2612	0.1564

Notes: Results are for the GDP per capita of 1980( $t-1$ ) and 2004 ( $t$ ). The results for the remaining variables are for their average between 1985 and 2004. The calculations were carried on using the software *R*.

<sup>21</sup> We would like to thank Prof. Eduardo Haddad from the University of Sao Paulo (USP) for providing the road distances for each pair of Brazilian micro-regions.

<sup>22</sup> The inverse squared distance weight matrix is also row-standardized. As an alternative, we also use the inverse squared distance standardized by its largest eigenvalue to comply with the stationary requirement without losing quantitative information on distances (see Anselin 1988, p. 23-24 and Elhorst 2009, p. 3-4). Estimation results based on this matrix are presented in the Appendix E and provide similar qualitative results.

Table 1 reports the results of Moran's  $I$  statistic for all variables included in Equation (1) using all spatial weight matrices described above. Regardless of the spatial structure imposed, all variables present a positive association between the original variable and its spatially lagged version, suggesting the existence of spatial autocorrelation. The global spatial autocorrelation may also be visualized graphically in the Moran scatterplot, since the Moran's  $I$  is formally equivalent to the slope coefficient of a linear regression of  $WZ$  on  $Z$  (Anselin, 1996; Anselin and Bao, 1997). Figure 1 presents the Moran scatterplot for the GDP per capita in 1980 using the inverse squared distance weight matrix. We observe that relatively high (low) income regions tend to be located nearby other high (low) income regions. Therefore, each region should not be viewed as an independent observation, as implicitly assumed in many studies of regional growth convergence (Rey and Montouri, 1999).

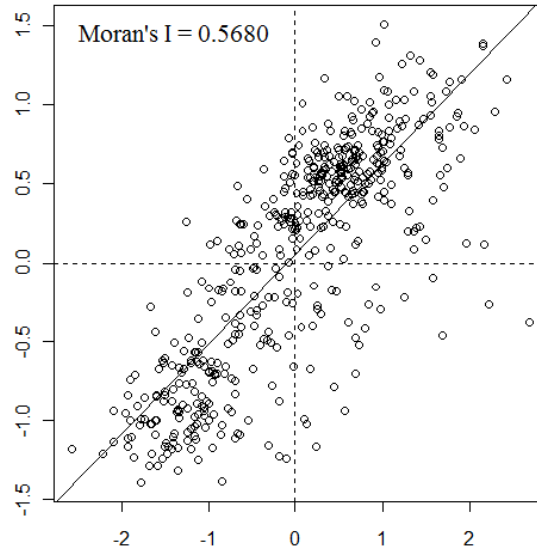


FIGURE 1: Moran Scatterplot of  $\ln GDP_{t-1}$ .

Nevertheless, Moran's  $I$  cannot assess the significance of the regional structure of spatial autocorrelation. On the other hand, LISA allows for the decomposition of Moran's  $I$  into the contribution of each observation and can be used to assess the significance of local spatial clustering around an individual location and for the identification of pockets of spatial nonstationarity (Anselin, 1995). In this index, only neighbouring values of the regions are included and LISA is expressed as follows:

$$I_i = \frac{z_i \sum_j w_{ij} z_j}{\sum_i z_i^2} \quad (6)$$

where variables are defined as in Equation (5). Local spatial clusters can be identified as those locations or set of locations where LISA is significant<sup>23</sup>. Combining the information of a Moran scatterplot with the significance of LISA yields the so called “Moran significance map”, showing the regions with significant LISA and indicating by a colour code the quadrants they belong to in the Moran scatterplot (Anselin, 1996; Anselin and Bao, 1997). The cluster is classified as such when the value at a location is more similar to its neighbours than would be the case under spatial randomness.

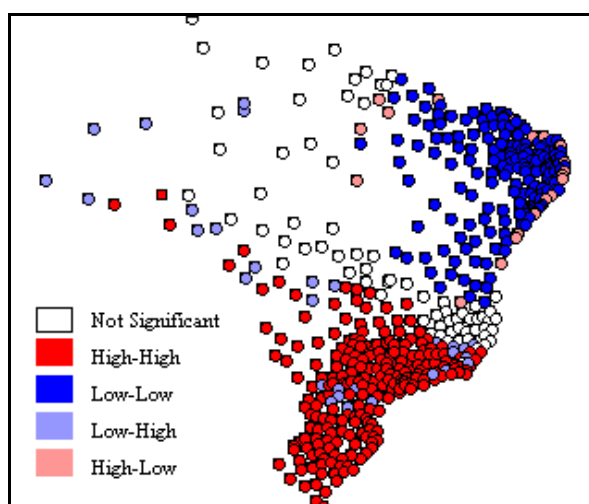


FIGURE 2: Local Moran's Significance Map for  $\ln GDP_{t-1}$ .

Figure 2 presents the significance map for the GDP per capita in 1980, and one important thing to notice is the presence of two strong regional regimes with the presence of a cluster of high-high values in the Southern part of Brazil and another of low-low values dominating the spatial pattern in the Northern part of the territory. There seems to be a clear Northern-Southern polarization that is in line with the existing literature on Brazil (e.g., Ferreira, 2000; Azzoni, 2001; Laurini et al., 2005; Coelho and Figueiredo, 2007)<sup>24</sup>. The presence of two spatial regimes is also in line with Silveira-Neto and Azzoni (2006) who use a local indicator of spatial association and suggest the existence of two geographical income clusters in Brazil. One cluster in the Southern part of the country encompassing South, Southeast and Centre-West states, and another one grouping states in the Northern part of the territory (North and Northeast states). Therefore, the exploratory spatial data analysis

<sup>23</sup> Inference of LISA is based on the permutation approach (Anselin, 1995). It works by allowing an approach in terms of conditional permutation where the value  $z_i$  of the region  $i$  is fixed and other values are permuted on all other locations of the sample.

<sup>24</sup> Alternative LISA maps for the GDP per capita in 1980 and 2004 using different spatial weights are reported in the Appendix B and confirm the regularity of the pattern observed in Figure 2 using the third-order Queen contiguity spatial weight. LISA maps were constructed using GeoDa 0.9.5-I.

confirms the presence of spatial autocorrelation and spatial heterogeneity that need to be considered when estimating growth models.

## **4 Empirical Results**

### **4.1 Results for Brazil**

The existing economic growth literature provides evidence of convergence for Brazilian regions and suggests that convergence is conditioned to structural factors such as population growth, physical capital and human capital using either standard (e.g., Lau et al., 1993; Ferreira, 2000; Azzoni, 2001; Nakabashi and Salvato, 2007) or spatial econometrics (e.g., Mossi et al., 2003; Silveira-Neto and Azzoni, 2006; Resende, 2009). However, to the best of our knowledge, there are no studies considering growth with spatial spillovers in a panel data setting for Brazil and the existing evidence that incorporates the SME sector into the growth framework is limited and also does not consider the spatial dimension into the analysis. Panel data models have advantages over the cross-section ones as they are generally more informative, control for individual and time fixed effects, and contain more variation and less collinearity among the variables in the model. The use of panel data also results in greater availability of degrees of freedom and therefore increases efficiency in the estimation (Elhorst 2009). For instance, Islam (1995) in the context of growth models uses panel estimations to allow for time specific and individual (country or region) specific levels of technology, the same argument is valid in the spatial context. The need to account for spatial heterogeneity stems from the fact that spatial units have individual characteristics that are space-specific and time invariant (Elhorst 2005), and cross-section estimates cannot control for spatial and time fixed effects and their results might be biased (Elhorst and Fréret 2009).

Spatial panel models only recently found application in regional convergence studies. The application of this setting is at a very early stage (Arbia et al., 2008). In this paper, the estimation of these models uses the developments made by Elhorst (2003, 2005 and 2009). He shows that estimations of spatial panel models with fixed effects can be carried out with standard techniques developed by Anselin (1988) and Anselin and Hudak (1992) after demeaning the variables in the model in order to control for the space specific and time fixed effects as in the within estimator (equivalent to the LSDV). Therefore, we draw on Elhorst



(2009) who presents the maximum likelihood functions used to estimate spatial panel models along with a link to Matlab codes to estimate these models<sup>25</sup>.

Regression results obtained from estimating the non-spatial Equation (1) and its spatial variants are presented in Table 2. Firstly, we present the non-spatial OLS within estimator controlling for space and time fixed effects together with the LM diagnostic tests used to formally detect spatial dependence in the regression residuals in column 1. The LM test statistics for spatial dependence refer to the spatial model as the alternative to the non-spatial OLS model. In relation to the spatial error model as the alternative, we report the  $LM_{ERR}$  and its robust version ( $LMR_{ERR}$ ), whereas for the spatial lag model the  $LM_{LAG}$  and its robust version ( $LMR_{LAG}$ ) are presented. If  $LM_{ERR}$  is more significant than  $LM_{LAG}$  or if  $LMR_{ERR}$  is more significant than the  $LMR_{LAG}$  in case both  $LM_{LAG}$  and  $LM_{ERR}$  are significant, then the SEM model is the most appropriate model<sup>26</sup>.

The results indicate that the model suffers from spatial dependence since the LM tests are statistically significant. Interestingly, all panel data estimates present similar qualitative results for the original variables (non-spatially lagged variables) that they have in common. Also, the scalar spatial error coefficient,  $\lambda$ , or the spatial autocorrelation parameters,  $\rho$ , are positive and significant, confirming the importance of spatial effects in all spatial panel estimates. Nevertheless, the SDM model is the only specification that sheds light on how spatial effects work through dependent and independent variables as discussed in Section 2. Besides, the spatial dependence in the disturbances of an OLS regression and endogeneity in the form of an omitted explanatory variable (in our study, for instance, regional information on physical capital is not available) make the SDM model a natural econometric choice over competing alternatives (Lesage and Fisher 2008). These arguments are particularly useful in supporting the choice of the SDM over the SEM model once these models are intrinsically linked as described in Section 2. The SDM specification is also the most natural way of estimating empirically neoclassical models with externalities (e.g., López-Bazo et al., 2004;

---

<sup>25</sup> The log-likelihood functions for panel data are presented in the Appendix C. The time and space (individual) specific effects are removed through the demeaning process that is formally presented in the Appendix D.

<sup>26</sup> The OLS model is tested against a specification with spatial dependence (SAR and SEM) using the inverse squared distance weight matrix. All spatial estimations presented in Table 2 use the inverse squared distance weight matrix but the use of alternative weight matrices present similar qualitative results (see Table A.2 in the Appendix). We performed the LM tests considering only the SAR and SEM models as alternatives to OLS because as noted by Elhorst (2009), the performance of the LM diagnostic tests presented in Table 2 (column 1) must still be investigated when one uses an extended panel data that includes spatially lagged independent variables (SDM). Refer to Burridge (1980) and Anselin et al (1996) for the non-robust and robust forms of these tests, respectively.

Ertur and Koch, 2007). Also, the fit of the SDM specification expressed by log-likelihood (LIK) is the best among the three spatial panel alternatives and supports our choice.

**TABLE 2: Regional Growth in Brazil (508 micro-regions): Non-Spatial *versus* Spatial Models**

	LSDV (1)	SAR (2)	SEM (3)	SDM (4)
$\ln GDP_{t-1}$	-0.1425*** (-31.7675)	-0.1343*** (-24.5779)	-0.1478*** (-33.2429)	-0.1499*** (-33.8944)
$\ln(n+d+g)$	-0.0082 (-1.0212)	-0.0050 (-0.6507)	-0.0063 (0.8137)	-0.0071 (-0.9070)
$\ln SCHOOL$	0.0534*** (4.5169)	0.0530*** (4.6929)	0.1067*** (6.6326)	0.1229*** (7.6101)
$\ln SMER$	-0.0090*** (-2.8321)	-0.0088*** (-2.8947)	-0.0109*** (-3.6697)	-0.0114*** (-3.8737)
$\ln SMEH$	0.0186*** (2.9929)	0.0183*** (3.0941)	0.0211*** (3.6716)	0.0213*** (3.7040)
$W * \ln GDP_{t-1}$				0.1545*** (9.5253)
$W * \ln(n+d+g)$				-0.0269 (-0.8211)
$W * \ln SCHOOL$				-0.1431*** (-5.1530)
$W * \ln SMER$				0.0276** (2.3739)
$W * \ln SMEH$				-0.0301 (-1.2695)
$\rho$ (SAR) / $\lambda$ (SEM)		0.5690** (2.3817)	0.7940*** (10.4037)	0.7740*** (19.1034)
Observations	2032	2032		2032
Log likelihood (LIK)	3387.60	3461.7323	3510.8821	3519.7317
$LM_{ERR}$	335.6792(0.0000)			
$LMR_{ERR}$	134.5770(0.0000)			
$LM_{LAG}$	206.3683(0.0000)			
$LMR_{LAG}$	5.2661(0.0220)			
$LR_{FE}$		1168.89(0.0000)	1152.90(0.0000)	1157.63(0.0000)

Note: \* p-value<0.10, \*\* p-value<0.05, \*\*\* p-value<0.01. The results for panel data are based on Elhorst's codes for Matlab available at <[www.regional.nl/elhorst/software.shtml](http://www.regional.nl/elhorst/software.shtml)>. The LM tests compare the classic OLS regression against the alternative SAR and SEM specifications under the null of no spatial dependence. The likelihood ratio for fixed effects ( $LR_{FE}$ ) tests for the existence of spatial specific effect against the specification with both, time and spatial fixed effects. The results throughout this section always suggest the consideration of both fixed effects. Numbers in brackets for the coefficients are the *t*-statistics and for the diagnostic tests are the levels of significance.

In the results for the SDM specification, the coefficients on the original independent variables represent the direct effect of those variables on growth and the coefficients of the spatially lagged independent variables capture the spillover effects of those variables on growth. They capture the interaction effects of a given independent variable with the values for this variable in the neighbouring regions. For instance, the spatial autoregressive parameter,  $\rho$ , indicates the magnitude of interregional spillovers (e.g., Ramajo et al 2008; Mohl and Hagen 2010). This implies that economic growth rates change 0.774% in reaction to a one percentage point increase in the growth rates of neighbouring regions. In the light of

López-Bazo et al. (2004)'s model, for instance, this coefficient (of the spatially lagged GDP per capita growth) formally measures the strength of externalities in the technology of production across Brazilian micro-regions. In addition, the positive and significant coefficient on the spatial lag of initial GDP per capita level supports the idea that higher levels of GDP in neighbouring regions impact positively on economic growth in a given region.<sup>27</sup>

The coefficients of population growth and its spatially lagged variable are negative and insignificant. The fact that the former is not negatively significant is a result contrary to the Solow growth model prediction, but in line with the literature about Brazil (e.g., Nakabashi and Salvato, 2007). They argue that this result might be related to the endogeneity between economic and population growth. The coefficient on human capital is positive and significant, indicating that localised investments in education at micro-regional level are important to promote economic growth. The importance of local investment in education is reinforced by the result that the interaction of human capital across space is negatively related to growth, indicating lack of positive human capital spillovers in Brazil at micro-regional level. This lack of human capital spillovers at regional level was also found for the overall level of human capital in Lesage and Fisher (2008) for the European NUTS 2 and in Resende (2009) for the case of Brazilian micro-regions. Qualified workers seem to be attracted to one location and the concentration of more qualified people in this region does not translate into human capital spillovers to the neighbouring regions<sup>28</sup>.

Turning to the SMEs' aspects, the estimates show that the size of the SME sector is negatively related to economic growth, a result that is in line with the idea that the size of the SME sector does not promote economic growth in the context of a developing country (e.g. Cravo, 2010). On the contrary, the interaction of this aspect of the SME sector across space is positive and statistically significant, suggesting that the interaction effects of the size of the SME sector with the size of this sector in neighbouring regions contribute to economic growth. The fact that there are more SMEs next door seems to incite a more productive entrepreneurship and new economic possibilities that affect economic performance positively. This is in line with Audretsch and Keilbach (2007) who argue that entrepreneurial activities

---

<sup>27</sup> This means that regions with the same parameters can have different growth rates if their neighbours have different level of productivity and GDP per capita, possibly due to different physical capital endowments. Thus, this result can also be interpreted as the spatial externalities stemming from physical capital. As noted in López-Bazo et al. (2004), an alternative interpretation to this lies in the fact that some of the spillovers could have been caused by pecuniary externalities present in models of new economic geography (e.g. Krugman 1991).

<sup>28</sup> This is also in line with the argument linked to New Economic Geography (NEG) models that more skilled labour is an important mobile factor that constitute a centripetal force towards geographical concentration (e.g. Krugman 1999).

have a more positive influence on neighbouring regions than on distanced regions due to cultural aspects, for instance. Using data for Germany and spatial econometrics, they provide evidence that an increase in entrepreneurial activity in adjacent regions increases entrepreneurship in a given region. Additionally, human capital of SMEs has a coefficient with the opposite sign from that associated with the spatial lags of this variable. Higher levels of human capital in SMEs are positively related with growth, indicating that this aspect of SMEs is more important than the size of this sector for promoting economic growth. However, the negative, but not significant, sign of the spatial lag of the SMEs' human capital indicates that regions do not benefit from a higher level of human capital in SMEs in neighbouring areas. This follows the same logic of the overall level of human capital and locations with SMEs that require higher levels of human capital only attract more skilled labour from neighbouring regions and do not generate any spillover for those regions<sup>29</sup>. Therefore, the presence of SMEs *per se* does not generate economic growth directly but through spatial interactions, small businesses in a given region seem to learn good practices from SMEs in neighbouring regions. Conversely, the level of human capital applied in the SME sector seems to be more relevant for regional economic growth in Brazil, although there is no human capital spillover in the SME sector. The small businesses in regions with higher levels of human capital seem to attract more skilled labour from neighbouring regions and do not generate spillovers for those regions.

## 4.2 Results for Regional Regimes

The economic growth literature on Brazil (e.g., Ferreira, 2000; Azzoni, 2001; Laurini et al., 2005; Coelho and Figueiredo, 2007; Silveira-Neto and Azzoni, 2006) and the results from LISA discussed in Section 3 suggest the existence of two different spatial regimes, one in the Northern Brazil (Northeast and North regions) and another in the Southern part of the territory (South, Southeast and Centre-west regions). This spatial heterogeneity might cover some distinctive patterns specific for those regimes that are not captured by the estimates for the country. Therefore, we consider two alternative samples for the regional analysis. The first regional sample comprises 207 contiguous micro-regions in the poorer Northern part of Brazil and the second sample comprises 301 contiguous micro-regions in the richer Southern part of

---

<sup>29</sup> This result can also support the idea that population growth in Brazil is endogenous in the sense that more qualified people endogenously migrate seeking better opportunities in richer regions.

the country<sup>30</sup>. This procedure is in line with Le Gallo and Dall’erba (2006) and Ramajo et al. (2008) that estimated spatial growth regressions for two spatial regimes in the European Union NUTS II in order to analyse the differences in the growth process in regions belonging to different regimes.

Table 3 reports the results for the two groups of regions and again we focus on the estimations of the SDM model that also provides the best fit. Column 3 reports the SDM results for the Northern regions that are in line with the results for the country as a whole. For instance, the autoregressive parameter suggests the existence of interregional spillovers and the spatial lag of initial GDP per capita indicates that richer neighbours impacts positively on the economic performance of a given region. Regarding the SME sector, we also observe that the size of the SME sector is negatively related with per capita GDP growth and the coefficient on *SMEH* is positive, indicating that this aspect of SMEs is more important than the size of this sector for economic performance. The proxies related to the SME sector have coefficients with the opposite sign from those associated with the spatial lags of these variables. The insignificant negative sign of the spatial lag of *SMEH* indicates that a given region does not benefit from higher level of human capital in SMEs in the neighbouring regions, while the positive and significant sign of the spatial lag of *SMER* indicates that a given region benefits from entrepreneurial activities in neighbouring regions. As in the results for the country as a whole, this indicates that SMEs in regions with higher levels of human capital only attract more skilled labour from neighbouring regions and do not generate any spillover for those regions.

The results for the Southern part of the country are presented in Table 3 from columns 4 to 6. The results for the SDM model (column 6) confirm the existence of spatial spillovers at micro-regional level in Brazil and that a given region achieves a better economic performance when its neighbours are richer. The coefficient on human capital is positive and significant and this aspect seems to be an important growth determinant for this set of regions. As in the case of the country as a whole, the coefficient of the spatial lag of human capital

---

<sup>30</sup> When we consider two groups of contiguous regions separately we are implicitly assuming that there are barriers preventing spillovers from one cluster to another in order to better observe structural differences within each cluster. For instance, Acemoglu et al. (2001) and Rodrik et al. (2004) provide evidence that GDP per capita and the quality of institutions are positively related and Menezes-Filho et al. (2006) and Naritomi et al. (2009) document that income per capita is positively related with institutional development in Brazil. Thus, spatial regimes might be the expression of different institutional regimes where the cluster of poorer regions has a worse institutional quality that might prevent spatial spillovers from the cluster of richer regions. A technological shock coming from one cluster to the other might eventually occur but does not need to be spatially correlated. It might have a direct impact on all economies of the poor cluster and then diffuse within that cluster.

presents a negative sign. Regarding the SME sector, we found the same indication that human capital in SMEs is more important than the size of the SME sector for regional economic performance. The coefficient on *SMER* is negatively related with economic growth and the *SMEH* is not negatively related with GDP per capita growth<sup>31</sup>.

**TABLE 3: Panel Results for the Two Regional Regimes**

	(1) SEM Northern	(2) SAR Northern	(3) SDM Northern	(4) SEM Southern	(5) SAR Southern	(6) SDM Southern
$\ln GDP_{t-1}$	-0.1443*** (-21.0625)	-0.1353*** (-19.7972)	-0.1485*** (-21.3746)	-0.1521*** (-26.7485)	-0.1383*** (-24.2044)	-0.1540*** (-26.9650)
$\ln SCHOOL$	0.0967*** (4.9143)	0.1004*** (5.1885)	0.0963*** (4.7740)	0.1636*** (6.4229)	0.1036*** (4.8533)	0.1804*** (6.5502)
$\ln(n+d+g)$	-0.0061 (-0.6463)	-0.0093 (-0.9990)	-0.0019 (-0.2037)	-0.0062 (-0.4559)	-0.0021 (-0.1546)	-0.0082 (-0.5792)
$\ln SMEH$	-0.0072* (-1.8327)	-0.0036 (-0.9002)	-0.0075* (-1.8848)	-0.0150*** (-3.3134)	-0.0148*** (-3.1839)	-0.0151*** (-3.3026)
$\ln SMEH$	0.0220*** (3.4082)	0.0209*** (3.1739)	0.0210*** (3.2643)	0.0167 (1.1525)	0.0192 (1.3314)	0.0160 (1.1041)
$W * \ln GDP_{t-1}$			0.1045*** (4.7162)			0.1865*** (8.4187)
$W * \ln SCHOOL$			-0.0007 (-0.0118)			-0.2396*** (-3.9347)
$W * \ln(n+d+g)$			-0.0569* (-1.7537)			0.0038 (0.0650)
$W * \ln SMEH$			0.0400*** (2.7966)			0.0226 (1.1208)
$W * \ln SMEH$			-0.0340 (-1.3342)			0.0310 (0.5277)
$\lambda(SEM) / \rho(SAR)$	0.5230*** (7.8880)	0.3320*** (5.1407)	0.4950*** (7.3868)	0.8579*** (23.6393)	0.6930*** (14.2021)	0.8480*** (23.7923)
$R^2$	0.6558	0.6692	0.6872	0.5501	0.6075	0.6384
log-likelihood (LIK)	1411.4709	1400.5040	1418.6311	2122.8757	2089.9441	2127.2736
Observations	828	828	828	1204	1204	1204
$LR_{FE}$	490.1073	481.3675	496.7470	686.2501	655.2885	691.1910
$p$ -value	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Note: \* p-value<0.10, \*\* p-value<0.05, \*\*\* p-value<0.01. The likelihood ratio for fixed effects ( $LR_{FE}$ ) tests for the existence of spatial specific effect against the specification with both, time and spatial fixed effects. Numbers in brackets for the coefficients are the  $t$ -statistics and for the diagnostic tests are the level of significance.

Hence, the regional panel data results are in line with the analysis for the country as a whole. An important point, for instance, is that economic growth is significantly affected by local human capital but there is no positive regional spillover of this capital. Also the size of the SME sector is negatively correlated with economic growth in all samples, and this sector seems to affect growth via the local human capital embodied in SMEs, evidenced by the

<sup>31</sup> The estimations presented in Table 3 also use the inverse squared distance weight matrix. As robustness checks, alternative regional results using different weight matrices are presented in the Appendix E (Tables A.3 and A.4) and provide the same qualitative results.

positive sign on the coefficient of this variable. In addition, results indicate that a given region does not benefit from higher levels of human capital embodied in SMEs in the neighbouring regions: *SMEH* has no positive spatial externality effect on economic growth, regardless of the group of regions analysed.

Nevertheless, although the general conclusion is similar, some differences emerged in the regional analysis in terms of the level of significance and magnitudes of the coefficients. For instance, in the Southern region, there is no significant indication that a given region benefits from a bigger SME sector in neighbouring regions. Besides, another significant difference is that the autoregressive parameter,  $\rho$ , and the coefficient of the spatially lagged initial GDP have a much greater magnitude in the case of Southern regions. The greater magnitude of the autoregressive parameter indicates that externalities in the technology of production are stronger in the Southern group while the greater coefficient of the spatially lagged initial GDP suggests greater physical capital spillovers in this set of regions.

In sum, the empirical results show that externalities detected in the case of the country as a whole were not a result of the omission of spatial regimes and is a phenomena that occur within each regime as well. In addition, the Southern regions seem to be more directly connected economically with neighbours as evidenced by the greater magnitude of the spatial autoregressive parameter and the coefficient of the spatial lag of the initial GDP per capita. This result might also be influenced by some externalities associated with the new economic geography models. For instance, Krugman (1999) argues that the Brazil's south is a more attractive place to produce due to better transport infrastructure and higher market potential. Thus, Southern Brazil attracts firms and investments that want to exploit economies of scale and access to consumers. Additionally, the trade liberalization that occurred during the 90's, due to the unilateral trade liberalization and the MERCOSUL regional trade agreement might have reinforced the advantages of Southern regions. For instance, Martincus (2010) provides empirical evidence of an accentuated tendency during that decade for industries with a high degree of sectoral openness to locate in (Southern) Brazilian states that are closer to Argentina, the biggest trading partner in MERCOSUL. In relation to the SME sector, as argued in Audretsch and Keilbach (2007), entrepreneurship might be linked to cultural phenomena that are clustered in space. Hence, differences in the regional empirical results are in line with the idea that entrepreneurship and SMEs are locally driven by cultural differences and impact on entrepreneurial activities and regional economic performance differently.

## 5 Conclusion

This paper has analysed the economic growth process in the Brazilian micro-regions using spatial panel econometrics and considering the importance of the SME sector in this process. The findings from the spatial panel estimates indicate the presence of a significant spatial dependence in the process of economic growth in the Brazilian micro-regions. The results show that growth rates in a given micro-region are affected positively by the level of the GDP per capita and its growth in neighbouring regions, a clear indication of the spatial connectiveness of the economic activity. Noticeably, human capital is an important growth determinant, but there is no indication of positive human capital spillovers and this finding reinforces the need of localised investments in human capital. Furthermore, the presence of SMEs *per se* does not generate economic growth directly but through spatial interactions, small businesses in a given region seem to learn good practices from SMEs in its surrounding area. Conversely, the level of human capital applied in the SME sector seems to be important for regional economic growth in Brazil, but there is no human capital spillover in the SME sector.

Measures of global and local spatial autocorrelation suggest the existence of spatial autocorrelation and heterogeneity expressed in the presence of two spatial regimes. In general, the analysis of the two different regimes is in line with the results for the country as a whole and growth rates of a region and initial GDP per capita also depend on the values of these variables in neighbouring regions. The overall level of human capital is an important factor for growth, but does not generate positive spillovers. In addition, human capital embodied in SMEs also seems to be more important than the size of this sector for regional economic growth but does not generate positive spatial spillovers either, regardless of the regime analysed. However, there are some differences between the two spatial regimes, which are consistent with the argument that the technological spillover is stronger in the Brazil's Southern regions and that entrepreneurship and SMEs are influenced by local characteristics.

Importantly, the empirical evidence suggests that localised investments in human capital are an important tool to promote regional economic growth. Regarding the SME sector, results do not support the public policy view that the size of the SME sector should be locally increased *per se* to improve economic performance. Instead, policy makers should better understand the spatial interactions of directly supporting SMEs in a given region to promote growth in its neighbourhood through the promotion of entrepreneurship. Thus, SMEs' policy should be coordinated with a broader regional focus in order to explore



entrepreneurship externalities. However, if the focus is on localised direct support to SMEs instead, educational policies should be the focus of the public policy by enabling more human capital formation for the economy and for a more productive SME sector that can contribute to the process of economic growth.

## References

Abreu M, de Groot H, Florax R (2005) Space and Growth: A Survey of Empirical Evidence and Methods. *Région et Développement* 21:13-44.

Acemoglu D, Johnson S, Robinson JA (2001) The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5): 1369-1401.

Acs Z, Desai S, Hessels J (2008) Entrepreneurship, economic development and institutions. *Small Business Economics* 31(3): 219-234.

Anselin L (1988) Spatial Econometrics: Methods and Models. Kluwer, Dordrecht.

Anselin L, Hudak S (1992) Spatial econometrics in practice : A review of software options. *Regional Science and Urban Economics* 22(3): 509-536.

Anselin, L (1995) Local Indicators of Spatial Association—LISA. *Geographical Analysis* 27: 93–115.

Anselin L (1996) The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association. In Fischer M, Scholten H, Unwin D (eds) *Spatial Analytical Perspectives on GIS*. London: Taylor and Francis.

Anselin L, Bao S (1997) Exploratory spatial data analysis linking SpaceStat and ArcView. In: Fisher M, Getis A (eds) *Recent developments in spatial analysis*. Springer, Berlin Heidelberg New York.

Anselin L, Bera AK, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26(1):77-104.

Arbia G, Le Gallo J, Piras G (2008) Does Evidence on Regional Economic Convergence Depend on the Estimation Strategy? Outcomes from Analysis of a Set of NUTS2 EU Regions. *Spatial Economic Analysis* 3(2):209-224.

Audretsch DB (2007) Entrepreneurship capital and economic growth. *Oxford Review of Economic Policy* 23(1): 63-78.

Audretsch D, Keilbach M (2004) Entrepreneurship capital and economic performance. *Regional Studies* 38(8): 949-959.

Audretsch D, Keilbach M (2007) The Localisation of Entrepreneurship Capital: Evidence from Germany. *Papers in Regional Science* 86(3): 351-365.

Ayyagari M, Beck T, Demirguc-Kunt A (2007) Small and medium enterprises across the globe. *Small Business Economics* 29(4): 415-434.

Azzoni, CR (2001) Economic growth and regional income inequality in Brazil. *Annals of Regional Science* 35(1): 133-152.

Badinger H, Müller WG, Tondl G (2004) Regional Convergence in the European Union, 1985-1999: A Spatial Dynamic Panel Analysis. *Regional Studies* 38(3): 241-253.

Barro R (1991) Economic growth in a cross section of countries. *Quarterly Journal of Economics* 106(2): 407-43.

Baumol W (1990) Entrepreneurship - Productive, unproductive, and destructive. *Journal of Political Economy* 98(5): 893-921.

Beck T, Demirguc-Kunt A, Levine R (2005) SMEs, growth, and poverty: Cross-Country evidence. *Journal of Economic Growth* 10(3): 199-229.

Bivand R, Brunstad R (2006) Regional Growth in Western Europe: Detecting Spatial Misspecification Using the R Environment. *Papers in Regional Science* 85(2): 277-297.

Bosker M (2007) Growth, agglomeration and convergence: a space-time analysis for European regions. *Spatial Economic Analysis* 2(1): 91-110.

Burridge P (1980) On the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society B*(42): 107-108.

Carree M, Thurik A (1998) Small Firms and Economic Growth in Europe. *Atlantic Economic Journal* 26(2):137-146.

Caselli F, Esquivel G, Lefort F (1996) Reopening the Convergence Debate. *Journal of Economic Growth* 1(3): 363-389.

Coelho R, Figueiredo L (2007) Uma Análise da Hipótese da Convergência para os Municípios Brasileiros. *Revista Brasileira de Economia* 61(3): 331-352.

Cravo T (2010) SMEs and Economic Growth in the Brazilian Micro-regions. *Papers in Regional Science*, in Press, Accepted Manuscript, doi 10.1111/j.1435-5957.2010.00301.x.

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.

Dias J, McDermott J (2006) Institutions, education, and development: The role of entrepreneurs. *Journal of Development Economics* 80(2): 299-328.

Durlauf S, Johnson P, Temple J (2005) Growth Econometrics. In: Aghion P, Durlauf S (eds) *Handbook of Economic Growth*. Chapter 8: Elsevier.

- Elhorst JP (2003) Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review* 26(3): 244-268.
- Elhorst JP (2005) Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels. *Geographical Analysis* 37(1): 85-106.
- Elhorst JP, Freret S (2009) Evidence of Political Yardstick Competition in France using a Two-Regime Spatial Durbin Model With Fixed Effects. *Journal of Regional Science* 49(5): 931-951.
- Elhorst JP (2009) Spatial Panel Data Models. In Fischer M, Getis A. (eds) *Handbook of Applied Spatial Analysis*. Springer, Berlin.
- 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(1): 3-34.
- Ertur C, Koch W (2007) Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics* 22(6):1033-1062.
- Ferreira A (2000) Convergence in Brazil: Recent trends and long-run prospects. *Applied Economics* 32(4): 479-489.
- Fingleton B (2000) Spatial econometrics, economic geography, dynamics and equilibrium: a 'third way'? *Environment and Planning A* 32: 1481-1498.
- Islam N (1995) Growth Empirics: A Panel Data Approach. *Quarterly Journal of Economics* 110(4): 1127-1170.
- Krugman P (1991) Increasing returns and economic geography. *Journal of Political Economy* 99(3): 483-499.
- Krugman P (1999) The Role of Geography in Development. *International Regional Science Review* 22(2): 142-161.
- Lau LJ, Jamison DT, Liu S, Rivkin S (1993) Education and economic growth: some cross-sectional evidence from Brazil. *Journal of Development Economics* 41(1): 45-70.
- Laurini M, Andrade E, Valls P (2005) Income convergence clubs for Brazilian municipalities: A non-parametric analysis. *Applied Economics* 37(18): 2099-2118.
- 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(2): 175-201
- Le Gallo J, Dall'erba S (2006) Evaluating the Temporal and Spatial Heterogeneity of the European Convergence Process, 1980–1999. *Journal of Regional Science* 46(2): 269-288.
- Levine R, Renelt D (1992) A sensitivity analysis of cross-country growth regressions. *The American Economic Review* 82(4): 942-963.

Lesage JP, Fischer MM (2007) Spatial Growth Regressions: Model Specification, Estimation and Interpretation, SSRN Working Paper. Available at: <http://ssrn.com/abstract=980965>.

Lesage JP, Fischer MM (2008) Spatial Growth Regressions: Model Specification, Estimation and Interpretation. *Spatial Economic Analysis* 3(3): 275-304.

Lesage JP, Pace RK (2009) *Introduction to Spatial Econometrics*. CRC Press/Taylor & Francis Group: London.

Lopez-Bazo E, Vaya E, Artis M (2004) Regional Externalities and Growth: Evidence from European Regions. *Journal of Regional Science* 44(1): 43-73.

Mankiw G, Romer D, Weil D (1992) A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107(2) 404-437.

Martincus C (2010) Spatial Effects of Trade Policy: Evidence from Brazil. *Journal of Regional Science*, 50(2):541-569

Menezes-Filho NA, Marcondes R, Pazello E, Scorzafave LG (2006). Instituições e diferenças de renda entre os estados brasileiros: Uma análise histórica. In: Anais do XXXIV Encontro Nacional de Economia. ANPEC, 2006. v. 1.

Mohl P, Hagen T (2010) Do Structural Funds Promote Regional Growth? New Evidence from Various Panel Data Approaches. *Regional Science and Urban Economics*, in Press, Accepted Manuscript, doi:10.1016/j.physletb.2003.10.071

Mossi MB, Aroca P, Fernández IJ, Azzoni CR (2003) Growth Dynamics and Space in Brazil. *International Regional Science Review* 26(3): 393-418.

Mueller P (2007) Exploiting entrepreneurial opportunities: The impact of entrepreneurship on growth. *Small Business Economics* 28(4): 355-362.

Muendler M (2007) Trade and workforce changeover in Brazil. National Bureau of Economic Research (NBER) Working Papers No. 12980, Cambridge, USA.

Nakabashi L, Salvato M (2007) Human capital quality in the Brazilian States. *Revista ANPEC* 8: 211-229.

Naritomi J, Soares R, Assunção J (2009) Institutional Development and Colonial Heritage within Brazil. Pontifical Catholic University of Rio de Janeiro (PUC-Rio). Discussion Paper No. 7579.

Ramajo J, Márquez MA, Hewings GJD, Salinas MM (2008) Spatial heterogeneity and interregional spillovers in the European Union: Do cohesion policies encourage convergence across regions? *European Economic Review* 52(3): 551-567.

Resende G (2009) Multiple dimensions of regional economic growth: The Brazilian case, 1991-2000. Proceedings of the 49<sup>th</sup> congress of the European Regional Science Association (ERSA). Lodz, Poland.

Rey SJ, Montouri BD (1999) US regional income convergence: A spatial econometric perspective. *Regional Studies* 33(2): 143-156.

Rodrik D, Subramanian A, Trebbi F (2004) Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth* 9(2):131-165

Sala-i-Martin X (2002) 15 years of new growth economics: What have we learnt? Universitat Pompeu Fabra Economics and Business Working Paper No. 620.

Silveira-Neto R, Azzoni CR (2006) Location and regional income disparity dynamics: The Brazilian Case. *Papers in Regional Science* 85(4): 599-613.

Solow R (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economic* 70(1): 65-94.

Solow R (2007) The last 50 years in growth theory and the next 10. *Oxford Review of Economic Policy* 23(1): 3-14.

Temple J (1999) The new growth evidence. *Journal of Economic Literature* 37(1): 112-156.

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

## **Appendix A: Data**

### **A.1. Description of Micro-regional GDP data**

The series of the real GDP at factor price (GDP) were constructed from the information available for Brazilian municipalities and aggregated into micro-regions. The available data at municipality level for 1980, 1985 and 1996 were constructed by IPEA based on the IBGE measurement and data for 2000 and 2004 were constructed by IBGE. The GDP at municipality level calculated by different institutes (IPEA and IBGE) are based on similar methodologies and always upon IBGE measurement. It was calculated by allocating the GDP of each state across its municipalities according to the observable production factors in each municipality. The major difference is that they use different proxies to allocate the state's GDP across municipalities because the information for some production factors is not available every year. The data processing described above may create some measurement errors but we believe this is the best possible series that could be constructed for the real GDP at factor price at micro-regional level. Problems with regional data are not restricted to Brazil. For instance, many regional studies for the European Union are based on the compilation of information from Cambridge Econometrics that put together data comprised from different

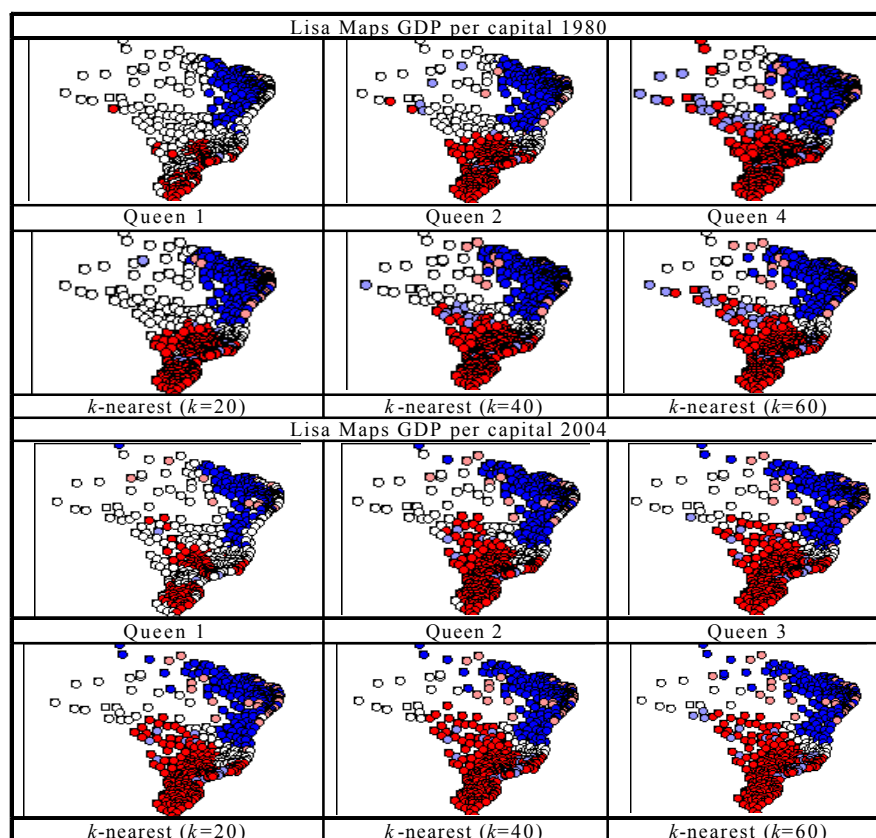
statistics offices that are not fully compatible, therefore, being subject to measurement errors (e.g., Fingleton, 2000; Bosker, 2007; Dall'erba and Le Gallo, 2008).

## A.2. Range of Education to Generate a Continuous Series for *SMEH*

**TABLE A.1: Educational Range and Years of Schooling**

RAIS Educational Range Level	Imputed years of Schooling
Illiterate	0
Primary School Dropout	1
Primary School Graduate	4
Middle School Dropout	5
Middle School Graduate	8
High School Dropout	9
High School Graduate	11
College Dropout	12
College Graduate	15

## Appendix B - Lisa Maps



**FIGURE A.1: LISA Maps for GDP per capita with Alternative Spatial Weights.**

## Appendix C. Log-likelihood functions for Panel Data Estimates

Elhorst (2003) and Elhorst (2009) have shown that we can estimate the spatial models using the Maximum Likelihood procedure proposed by Anselin (1988) and Anselin and Hudak (1992) using the demeaning process to remove the individual and time specific effects.

The log-likelihood function corresponding to the demeaned specification for panel data for the SAR (and SDM) specification is given by:

$$\ln L = -\frac{NT}{2} \ln(2\pi\sigma^2) + T \ln|I - \rho W| - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (y_{it}^* - \rho [\sum_{j=1}^N w_{ij} y_{jt}]^* - x_{it}^* \beta)^2$$

where the asterisk denotes the demeaning procedure (Appendix D) to extract the individual and time specific effects (see Elhorst, 2009). The variables  $y$  and  $x$  are the dependent and independent variables of the model.

The log-likelihood function for a panel SEM specification if the spatial effects are assumed to be fixed is:

$$\begin{aligned} \ln L = & -\frac{NT}{2} \ln(2\pi\sigma^2) + T \ln|I - \lambda W| \\ & - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T \{y_{it}^* - \lambda [\sum_{j=1}^N w_{ij} y_{jt}]^* - (x_{it}^* - \lambda [\sum_{j=1}^N w_{ij} x_{jt}]^*) \beta\}^2 \end{aligned}$$

## Appendix D. Demeaning Process

The individual fixed effect can be eliminated by demeaning the variables of the model, where variables are taken in deviation from their average over time. This transformation takes the following form for a given variable  $y$ :

$$y_{it}^* = y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}$$

A similar procedure should be taken to remove the time fixed effect through the following transformation:

$$y_{it}^* = y_{it} - \frac{1}{N} \sum_{j=1}^N y_{jt}$$

## Appendix E. Alternative Results for Panel Estimates

**TABLE A.2: Alternative Panel Results for Brazil (SDM Model)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln \text{GDP}_{t-1}$	-0.1454*** (-32.7812)	-0.1481*** (-32.9908)	-0.1451*** (-32.3995)	-0.1451*** (-32.3013)	-0.1585*** (-36.2192)	-0.1487*** (-33.3295)	-0.1441*** (-32.3378)	-0.1428*** (-31.9840)
$\ln \text{SCHOOL}$	0.0532*** (4.5355)	0.1076*** (6.4731)	0.1109*** (6.6702)	0.1068*** (6.5084)	0.0979*** (6.2166)	0.1021*** (6.1721)	0.1029*** (6.2367)	0.1017*** (6.1421)
$\ln(n+d+g)$	-0.0031 (-0.3890)	-0.0077 (-1.0052)	-0.0137* (-1.7383)	-0.0154* (-1.9366)	-0.0028 (-0.3815)	-0.0014 (-0.1781)	-0.0107 (-1.3503)	-0.0142* (-1.7854)
$\ln \text{SMER}$	-0.0097*** (-3.1098)	-0.0094*** (-3.1702)	-0.0083*** (-2.7421)	-0.0075** (-2.4439)	-0.0116*** (-4.1043)	-0.0101*** (-3.4264)	-0.0081*** (-2.6566)	-0.0089*** (-2.8782)
$\ln \text{SMEH}$	0.0203*** (3.3759)	0.0236*** (4.0983)	0.0230*** (3.8721)	0.0205*** (3.4280)	0.0224*** (4.0367)	0.0238*** (4.1530)	0.0236*** (4.0030)	0.0218*** (3.6293)
$W * \ln \text{GDP}_{t-1}$	0.1407*** (6.7983)	0.1204*** (9.9051)	0.1146*** (5.9932)	0.1191*** (4.9292)	0.1127*** (29.2122)	0.1171*** (8.0609)	0.1135*** (6.1602)	0.1101*** (5.359)
$W * \ln \text{SCHOOL}$	-0.0123 (0.3748)	-0.0789*** (-3.1976)	-0.0688** (-2.4345)	-0.0789** (-2.4170)	-0.0853*** (-4.2516)	-0.0990*** (-4.1114)	-0.0865*** (-3.1327)	-0.0886*** (-2.8011)
$W * \ln(n+d+g)$	-0.1387** (-2.0849)	-0.0224 (-1.0647)	0.0247 (0.8294)	0.0531 (1.3821)	-0.0240 (-1.6344)	-0.0440** (-2.1369)	0.0115 (0.4246)	0.0483 (1.3855)
$W * \ln \text{SMER}$	0.0388** (2.0453)	0.0137 (1.3378)	0.0191 (1.3779)	0.0030 (0.1825)	0.0167*** (2.8559)	0.0195** (2.0851)	-0.0074 (-0.5731)	-0.0006 (-0.0390)
$W * \ln \text{SMEH}$	-0.0679* (-1.7926)	-0.0466** (-2.3236)	-0.0792*** (-2.7687)	-0.0712* (-1.9035)	-0.0189 (-1.6166)	-0.0279 (-1.4870)	-0.0570** (-2.2034)	-0.0487 (-1.4329)
$\rho$ (SDM)	0.6900*** (10.4947)	0.6750*** (19.3923)	0.7090*** (9.6820)	0.7070*** (8.5502)	0.5280*** (10.9103)	0.6649*** (9.1968)	0.7050*** (9.2305)	0.6970*** (16.3592)
log-likelihood	3427.0403	3533.5497	3484.3651	3461.69	3577.6234	3539.9268	3487.6394	3451.5683
Observations	3032	2032	2032	2032	2032	2032	2032	2032
$\text{LR}_{\text{FE}}$	1329.20	1155.2962	1141.6921	1128.5750	1262.1200	1146.9836	1124.7989	1112.0610
p-value	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
<b>Spatial Weight</b>	<b>Inverse Distance Eigenvalue- standardized</b>	<b>K-20<sup>th</sup></b>	<b>K-40<sup>th</sup></b>	<b>K-60<sup>th</sup></b>	<b>Queen 1<sup>st</sup></b>	<b>Queen 2<sup>nd</sup></b>	<b>Queen 3<sup>th</sup></b>	<b>Queen 4<sup>th</sup></b>

Note: \* p-value<0.10, \*\* p-value<0.05, \*\*\* p-value<0.01. The likelihood ratio for fixed effects ( $\text{LR}_{\text{FE}}$ ) tests for the existence of spatial specific effect against the specification with both, time and spatial fixed effects. Numbers in brackets for the coefficients are the  $t$ -statistics and for the diagnostic tests are the level of significance.



**TABLE A.3: Alternative Panel Results for the SDM Model – Northern Regions**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln \text{GDP}_{t,l}$	-0.1451*** (-21.1606)	-0.1428*** (-20.1743)	-0.1419*** (-20.5015)	-0.1432*** (-20.9284)	-0.1502*** (-21.4159)	-0.1412*** (-20.0980)	-0.1399*** (-20.0287)	-0.1418*** (-20.6650)
$\ln \text{SCHOOL}$	0.1116*** (5.6577)	0.0747*** (3.6120)	0.0758*** (3.7021)	0.0770*** (3.7871)	0.0794*** (3.9574)	0.0717*** (3.4552)	0.0762*** (3.7021)	0.0794*** (3.8899)
$\ln(n+d+g)$	-0.0059 (-0.6125)	-0.0074 (-0.7993)	-0.0101 (-1.0864)	-0.0086 (-0.9281)	-0.0043 (-0.4605)	-0.0034 (-0.3703)	-0.0086 (-0.9089)	-0.0095 (-1.0233)
$\ln \text{SMER}$	-0.0065 (-1.6023)	-0.0061 (-1.5485)	-0.0048 (-1.2243)	-0.0042 (-1.0730)	-0.0067* (-1.7528)	-0.0057 (-1.4409)	-0.0050 (-1.2501)	-0.0051 (-1.2748)
$\ln \text{SMEH}$	0.0217*** (3.2961)	0.0222*** (3.4048)	0.0218*** (3.3403)	0.0193*** (2.9649)	0.0211*** (3.3349)	0.0221*** (3.4013)	0.0225*** (3.4242)	0.0193*** (2.9304)
$W^* \ln \text{GDP}_{t,l}$	0.0665*** (2.8187)	0.0791*** (3.6239)	0.0724** (2.3405)	0.0402 (1.0152)	0.0831*** (6.0223)	0.0650*** (3.1646)	0.0503* (1.8245)	0.0187 (0.5039)
$W^* \ln \text{SCHOOL}$	0.0161 (0.5065)	0.1061* (1.8510)	0.2643*** (3.2416)	0.5783*** (4.6859)	-0.0056 (-0.1525)	0.0521 (0.9774)	0.1752** (2.4155)	0.3236*** (3.3772)
$W^* \ln(n+d+g)$	-0.1774* (-1.9255)	-0.0315 (-1.1087)	-0.0243 (-0.5523)	-0.0164 (-0.2900)	-0.0153 (-0.7708)	-0.0336 (-1.1471)	0.0010 (0.0276)	0.0074 (0.1683)
$W^* \ln \text{SMER}$	0.0488** (2.5409)	0.0065 (0.3867)	0.0472* (1.8990)	0.0782** (2.1068)	0.0187** (2.3262)	0.0179 (1.2857)	0.0096 (0.4838)	0.0639** (2.4386)
$W^* \ln \text{SMEH}$	-0.0497 (-1.3510)	-0.0689*** (-2.6321)	-0.1654*** (-3.9808)	-0.2715*** (-4.2225)	-0.0182 (-1.3247)	-0.0385* (-1.6705)	-0.1313*** (-3.7748)	-0.2443*** (-4.7377)
$\rho$ (SDM)	0.3390*** (3.4078)	0.4830*** (6.6303)	0.4920*** (4.9810)	0.5000*** (4.6422)	0.3940*** (8.6167)	0.4480*** (6.4919)	0.4050*** (4.1843)	0.215991 (1.5212)
log-likelihood	1397.44	1415.99	1415.64	1414.86	1431.18	1416.71	1410.77	1408.61
Observations	828	828	828	828	828	828	828	828
$\text{LR}_{\text{FE}}$	620.3240	467.4763	491.1883	495.5561	523.1667	463.2095	471.6229	493.2492
$p$ -value	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
<b>Spatial Weight</b>	<b>Inverse Distance Eigenvalue- standardized</b>	<b>K-20th</b>	<b>K-40th</b>	<b>K-60th</b>	<b>Queen 1<sup>st</sup></b>	<b>Queen 2<sup>nd</sup></b>	<b>Queen 3<sup>th</sup></b>	<b>Queen 4<sup>th</sup></b>

Note: \* p-value<0.10, \*\* p-value<0.05, \*\*\* p-value<0.01. The likelihood ratio for fixed effects ( $\text{LR}_{\text{FE}}$ ) tests for the existence of spatial specific effect against the specification with both, time and spatial fixed effects. Numbers in brackets for the coefficients are the  $t$ -statistics and for the diagnostic tests are the level of significance.

**TABLE A.4: Alternative Panel Results for the SDM Model – Southern Regions**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln \text{GDP}_{t-1}$	-0.1502*** (-25.5798)	-0.1536*** (-26.8031)	-0.1498*** (-25.9636)	-0.14950*** (-25.8663)	-0.1628*** (-29.0339)	-0.1568*** (-27.8316)	-0.1497*** (-25.9980)	-0.1473*** (-25.4734)
$\ln \text{SCHOOL}$	0.0900*** (3.8005)	0.2047*** (7.0033)	0.2166*** (7.3110)	0.1990*** (6.9319)	0.1699*** (6.1282)	0.1939*** (6.7533)	0.1908*** (6.6280)	0.1952*** (6.8362)
$\ln(n+d+g)$	-0.0004 (-0.0276)	-0.0100 (-0.6949)	-0.0273* (-1.8212)	-0.0312** (-2.0878)	0.0056 (0.4086)	0.0030 (0.2138)	-0.0100 (-0.6597)	-0.0204 (-1.3386)
$\ln \text{SMER}$	-0.0147*** (-3.0643)	-0.0146*** (-3.1814)	-0.01465*** (-3.0430)	-0.0134*** (-2.7636)	-0.0173*** (-3.9633)	-0.0169*** (-3.7867)	-0.0125*** (-2.6332)	-0.0128*** (-2.6360)
$\ln \text{SMEH}$	0.0301** (1.9968)	0.0246* (1.7058)	0.0199 (1.3147)	0.0183 (1.2028)	0.0298** (2.1402)	0.0227 (1.6084)	0.0146 (0.9745)	0.0096 (0.6334)
$W^* \ln \text{GDP}_{t-1}$	0.1855*** (5.1843)	0.13532*** (7.8095)	0.0958*** (3.6319)	0.1151*** (3.4736)	0.1266*** (11.6638)	0.1563*** (9.7159)	0.1288*** (5.1551)	0.1157*** (3.1639)
$W^* \ln \text{SCHOOL}$	-0.0472 (-0.6448)	-0.1988*** (-4.2065)	-0.3128*** (-5.2356)	-0.3919*** (-5.2715)	-0.1555*** (-4.2548)	-0.2044*** (-4.6472)	-0.2476*** (-4.7531)	-0.3190*** (-4.9351)
$W^* \ln(n+d+g)$	-0.0540 (-0.5911)	-0.0269 (-0.6788)	0.1281** (2.0608)	0.4091*** (4.7752)	-0.0409 (-1.6103)	-0.0662** (-1.9745)	0.0094 (0.2122)	0.1026* (1.7320)
$W^* \ln \text{SMER}$	0.0553 (1.5696)	0.0127 (0.8242)	-0.0399* (-1.8238)	-0.1133*** (-3.989)	0.0153* (1.6558)	0.0369*** (2.5936)	-0.0214 (-1.0436)	-0.0742*** (-2.6960)
$W^* \ln \text{SMEH}$	-0.0568 (-0.6574)	-0.0544 (-1.1943)	-0.0900 (-1.4253)	-0.1394* (-1.7247)	-0.0162 (-0.6081)	-0.0187 (-0.4669)	-0.0333 (-0.5876)	-0.0369 (-0.4785)
$\rho$ (SDM)	0.8030*** (16.2386)	0.7320*** (18.8322)	0.6680*** (11.0510)	0.5300*** (5.9600)	0.5680*** (17.5819)	0.7510*** (20.8099)	0.7340*** (14.4344)	0.6350*** (8.1820)
log-likelihood	2067.98	2139.12	2097.66	2088.77	2165.19	2157.13	2106.19	2081.44
Observations	1204	1204	1204	1204	1204	1204	1204	1204
$\text{LR}_{\text{FE}}$	721.8830	714.5724	695.7395	704.5312	750.4950	724.3710	690.5047	680.9804
$p$ -value	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
<b>Spatial Weight</b>	<b>Inverse Distance Eigenvalue- standardized</b>	<b>K-20th</b>	<b>K-40th</b>	<b>K-60th</b>	<b>Q ueen 1<sup>st</sup></b>	<b>Q ueen 2<sup>nd</sup></b>	<b>Q ueen 3<sup>th</sup></b>	<b>Q ueen 4<sup>th</sup></b>

Note: \* p-value<0.10, \*\* p-value<0.05, \*\*\* p-value<0.01. The likelihood ratio for fixed effects ( $\text{LR}_{\text{FE}}$ ) tests for the existence of spatial specific effect against the specification with both, time and spatial fixed effects. Numbers in brackets for the coefficients are the  $t$ -statistics and for the diagnostic tests are the level of significance.