Spatial externalities between Brazilian municipios and their neighbours

Philippe De Vreyer* Gilles Spielvogel†

April 2005

Prepared for the 45th Congress of the European Regional Science Association
Vrije Universiteit Amsterdam
23-27 August 2005

Abstract

Clustering of economic performance and growth in space has generated considerable research on the spillovers and linkages among geographical neighbours. In this paper, we study the growth process of a large sample of Brazilian municipalities for the period 1970-1996 and attempt to evaluate the spatial externalities at work among them. We estimate the the convergence speed of per capita income among municipios and test whether spatial externalities are linked to local income growth. Conditionally on structural characteristics, we find evidence of convergence between municipios and of positive spatial dependence in growth. These two facts could help explain the persistent inequalities between municipios in the Northeast region.

Keywords: Local growth, convergence, spatial externalities, spatial econometrics, Brazil.
JEL codes: O40, R11, R12.

*CADRE, University of Lille 2 and IRD DIAL, Paris. E-mail: devreyer@dial.prd.fr.
†Sciences Po Paris, University of Lille 2 and IRD DIAL, Paris. E-mail: spielvogel@dial.prd.fr.

1 Introduction

Economic growth is not a uniform process over space. Within countries, some regions grow more rapidly than others. These differences may result in poorer regions catching up with wealthier ones, or, on the contrary, increasing income gaps. Many developing and newly industrialised countries are characterized by important spatial inequalities between regions, which seem to be very persistent or even increasing over time (Ravallion, 1998). These poverty traps result from disparities in growth among neighbours and reducing them requires a better understanding of their formation.

Among the determinants of local growth, the role of externalities has been much discussed in the recent literature (Glaeser et al., 1992). These externalities not only matter for growth within a given city or region but also for growth between neighbouring regions (Lopez-Baso et al., 2004).

Growth at a given location may affect growth of neighbours through several channels. First, due to technological externalities, a locality may benefit from improved economic conditions in another. For instance, if some firms in a locality have developed innovative processes, knowledge spillovers may favor the diffusion of new technologies to firms at neighbouring locations. Linkages between input suppliers and final producers may also be critical: if a final consumption good produced at a particular location benefits from a booming demand, upstream firms in the same region will thrive. Finally, proximity of an important economic centre may improve matching on the labor market, thus reducing costs and increasing labour productivity.

Pecuniary externalities may also matter in spatial growth differentials. On the one hand, growth at a given location may create new market opportunities for firms in neighbouring localities, through the increased demand resulting from higher incomes. On the other hand, the same process may attract new firms and workers, thus increasing land rents. Transmission of this land market tension to nearby localities can reduce incentives for firms to locate there, and therefore attenuate growth prospects.

Finally, local economic growth may foster immigration from less dynamic places. The impact of this migration on both the departure and arrival locations depends on various factors, notably the migrants' education level, the substitutability between skilled and unskilled workers in production and the state of local labour markets.

Understanding how local growth may spread to neighbours or may hinder their economic performance is critical for policy design. Local policies aiming at fostering growth may have positive or adverse effects on nearby localities. Sorting between the "good" and the "bad" channels may help designing more efficient policies. Land and transportation policies are also a closely related issue: some spatial externalities are driven by the functioning of the land market. When rising rents in a growing locality are transmitted to adjacent locations, for instance, public policies may be needed to reduce market tensions, through the development of new land plots or the improvement of transportation networks. In this case again, evaluating the strength and spatial scope of pecuniary externalities can help improving these policies.

In this paper we examine the determinants of local growth in Brazil. In the recent years, several papers have analysed the dynamics of regional growth in this country. Azzoni (2001) investigates the evolution of regional inequality over the period 1939-1995, using standard statistical and regression methods for analysing σ and β-convergence between the Brazilian states. He finds signs of regional income convergence, but with important oscillations in the evolution of inequality over time as well as across regions within the country.[1] The methods used in this paper are "standard" in the sense that, as most surveys studying regional convergence at that time, it did not consider the issue in a spatial econometric perspective. In other words, regional economies are considered in isolation, independently of their spatial location and/or the spatial links with other economic units. However, as shown by Anselin and Bera (1998) the failure to hold account of spatial dependance in linear regression models may lead to biased and/or inefficient estimators. This obviously applies to growth regressions for which there are plenty of good theoretical arguments suggesting that spatial dependance is likely to occur, and has been confirmed, among others, by the works of Rey and Montouri (1999) for the United States, and of Magalhães, Hewings and Azzoni (2000) for Brazil. Recent papers on this topic are therefore using spatial econometric methods. Abreu, de Groot and Florax (2004) provide an extensive survey of the empirical literature on growth and convergence that has taken the role of space into account.

Another trend in the convergence literature, following Quah (1997), focuses on the dynamics of income distribution. Few works combine this approach with the possible role of space in the growth process (see Magrini, 2004). Bosch Mossi et al. (2003) use local indicators of spatial association (LISA, see Anselin 1995) together with Markov transition matrices and stochastic kernels to study the convergence of per capita income among Brazilian states over the 1939-1998 period and to what extent spatial spillovers are apparent. They find strong evidence of spatial clustering, with poor (rich) states tending to be close in proximity to other poor (rich) states. Their results also indicate that regions are becoming more homogeneous internally but that differences between regions are increasing. Moreover they find evidence of spatial spillovers among states. First, states with wealthier neighbours have a greater chance of moving upward on the income ladder. Second, the clustering between the rich, southeastern, states and the poor, northeastern, states tends to become stronger over time, to the extent that states that originally did not belong to a cluster ultimately ended up being part of one of the two distinct clusters. Intradistribution dynamics is investigated at a finer geographical level by Andrade et al. (2004), though without the spatial dimension: they test the convergence hypothesis among the Brazilian municipalities over the 1970-1996 period. They find no evidence of convergence. On the opposite, results suggest that municipalities form convergence clubs and that these clubs are persistent over time, so that poor and rich municipalities maintain their relative income status. However there is also some mobility within clubs, with some poor and rich municipalities becoming respectively relatively richer and poorer.

Using finely disaggregated spatial data in the analysis of the growth process clearly is a progress: first, it permits to take intraregional disparities into account and second, it makes it

[1] See also Ferreira (2000) for a closely related paper using the same methodsbut on a shorter period. As Azzoni (2001), Ferreira finds evidence of σ and β-convergence across regions.

easier to relate findings of spatial dependence to the potential role of local externalities. Focusing on the Brazilian Northeast, Lall and Shalizi (2003) test for β-convergence across municipalities using spatial econometrics methods. Using the growth in labour productivity, measured as earnings per worker, as the dependant variable in the econometric analysis, they find that conditionally on structural characteristics, earnings per workers exhibit signs of convergence. Surprisingly they also find that growth in municipalities is negatively influenced by growth in their neighbourhood. Lall and Shalizi offer two alternative explanations for this result. One is that productivity growth in one locality is likely to attract capital and labour from the neighbouring localities, thereby having a negative effect on growth in these areas. As the authors point out, this assumes that productive factors are mobile across regions and can be efficiently used in their new locations. These assumptions might be unrealistic in a low income country context where mobility is low. The second is that, due to the low level of opportunities for local producers of the Northeast to increase the scale of production, productivity enhancements in any location are likely to result in productivity or profitability reductions in neighbouring locations. Whatever the explanation, it would be interesting to determine whether this result is specific to the Northeast, in which case the second explanation would become the most likely, to the extent that producers in other regions are less limited in their opportunities to extend the markets for their goods.

In this paper, we study the growth process of a large sample of Brazilian municipalities for the period 1970-1996 and attempt to evaluate the spatial externalities at work among them. In the next section, we briefly present the data used and the adjustment needed to hold account of the growing number of municipios in the country. In Section 3, using a two-stage nested decomposition of the Theil inequality index, we provide an overview of the evolution of income inequalities between municipios during the period. Section 4 is devoted to exploratory spatial data analysis. In Section 5, we present theoretical developments that can help understand the origin of spatial dependance in growth rates among regions and the creation of poverty traps. Finally in section 6, several models of β-convergence are estimated, using spatial econometrics methods when needed to account for potential spillovers across municipalities.

2 Data

Today, Brazil is made up of 5561 municipios. In 1970, there were only 3951 municipios. The permanent creation of new municipalities through the redistricting of existing units has been particularly intense in the North region (the number of municipios in this region has more than doubled between 1980 and 2001), while it has been slower in the Southeast, already endowed with a greater number of municipios. When trying to study the growth process of local units, such a variation in their number over time is clearly a nuisance, since it makes it impossible to compare municipio-level variables over time. It is therefore necessary to work as if no new municipios were created after 1970. The same approach is followed by Andrade et al. (2004). This lead us to work with units defined by the Instituto de Pesquisa Econômica Aplicada (IPEA) as "Áreas

Mínimais Comparáveis" (minimum comparable areas, henceafter AMC, see IPEA's website[2] for details). AMC-level data were generally directly available from IPEA. When this was not the case, we reconstituted AMC data from available municipio-level data (for instance, this was the case for education variables). In what follows, we use indifferently the terms municipio and AMC. The description of the variables used in the analysis, their sources and summary statistics are presented in Table 1.

Since we want to examine the role of spatial externalities in the growth process of local units, heterogeneity in their geographical sizes may be a problem. Indeed, it seems difficult to assume that externalities between very large municipios may be similar in nature as those arising between small units. Size differences between AMC being huge, we chose to exclude the states made up of very large units and to restrict the analysis on the Eastern part of the country, where AMC are smaller and more homogenous in size. We also excluded the island of Fernado de Noronha, belonging to the state of Pernambuco, far away in the Atlantic Ocean. We therefore work with a sample of 3487 AMC over a total of 3659 (our sample represents more than 95% of the Brazilian AMC). As a result, the mean size of AMC in the sample is 1054 square kilometers, while it is 2332 square kilometers if all AMC are included (the mean size of out-of-sample AMC is 28398 square kilometers). Our sample of AMC comprises all of Northeast, Southeast and South regions, plus the states of Tocantins (region North) and Goias and the Federal District of Brasilia (region Centre West). Though it accounts for only 43% of the Brazilian territory, it represented more than 90% of the population and GDP over the period.

3 Inequalities between municipios

Over the last thirty years, global inequalities between Brazilian municipalities have decreased: the Theil ($GE(1)$) index of income inequalities between municipalities has decreased from 0.41 in 1970 to 0.3 in 1996.[3] However, Brazil is a vast country with huge differences between regions and it seems necessary to provide a more detailed account of this evolution. We compute the share of different components of global spatial inequalities using a two-stage nested decomposition of the Theil index. As is well known, general entropy indexes are additively decomposable, so that any index of this family can be written as the sum of exclusive and exhaustive sub-indexes (Shorrocks, 1984). If we use the AMC as the basic unit of observation, and since each AMC belongs to one of the 27 Brazilian states and each state belongs to one of the 5 regions (North, Northeast, Center-West, Southeast and South), the familiar Theil index ($GE(1)$) can be written as:

$$T = \sum_i \sum_j \sum_k \left(\frac{Y_{ijk}}{Y}\right) \log\left(\frac{Y_{ijk}/N_{ijk}}{Y/N}\right)$$

where Y_{ijk} and N_{ijk} are respectively the income and the population of the AMC k in state j and region i, and Y and N are the total income and population of the country (i.e. $Y =$

[2]http://www.ipeadata.gov.br

[3]In this section, we do not restrict the analysis to the sample described above and provide results for the whole country.

Table 1: Description of the variables used and summary statistics

Variable	Description (sources)	Brazil		Sample	
		Mean	St. Dev.	Mean	St. Dev.
Number of observations		3659		3487	
Initial income (y_0)	Per capita GDP, 1970, R$ of 2000 (IPEA, IBGE)	1399	1816	1412	1852
Income 1996	Per capita GDP, 1996, R$ of 2000 (IPEA, IBGE)	3221	3538	3242	3586
Growth 1970-1996 (g)	Growth of per capita GDP	0.812	0.572	0.815	0.563
Education	Mean number of years of education, people aged 25 +, 1970 (IPEA, IBGE)	1.37	0.82	1.37	0.82
Illiteracy	Illiteracy rate, people aged 15 +, 1970 (IPEA, IBGE)	0.44	0.18	0.44	0.18
Urbanization	Share of urban population, 1970 (IBGE)	0.33	0.21	0.33	0.21
Electricity	Share of households with electricity, 1970 (IBGE)	0.24	0.23	0.25	0.23
Agriculture		0.46	0.22	0.46	0.22
Industry	Share of sector in GDP, 1970 (IPEA)	0.16	0.17	0.16	0.17
Services		0.38	0.15	0.38	0.15
Labour force	Share of people aged 25 +, 1970 (IBGE)	0.36	0.04	0.36	0.04
Household size	Mean size of households, 1970 (IBGE)	5.48	0.46	5.45	0.44
Region dummies					
Center-West		0.061		0.046	
North		0.039		0.01	
Northeast		0.355		0.372	
South		0.162		0.17	
Southeast		0.383		0.402	

$\sum_i \sum_j \sum_k Y_{ijk}$ and $N = \sum_i \sum_j \sum_k N_{ijk}$). This can be rewritten as:

$$
\begin{aligned}
T &= \sum_i \sum_j \sum_k \left(\frac{Y_{ijk}}{Y}\right) \left\{ \log\left(\frac{Y_{ijk}/N_{ijk}}{Y_{ij}/N_{ij}}\right) + \log\left(\frac{Y_{ij}/N_{ij}}{Y_i/N_i}\right) + \log\left(\frac{Y_i/N_i}{Y/N}\right) \right\} \\
&= \frac{1}{Y} \left\{ \sum_i \sum_j \sum_k Y_{ijk} \log\left(\frac{Y_{ijk}/N_{ijk}}{Y_{ij}/N_{ij}}\right) + \sum_i \sum_j Y_{ij} \log\left(\frac{Y_{ij}/N_{ij}}{Y_i/N_i}\right) + \sum_i Y_i \log\left(\frac{Y_i/N_i}{Y/N}\right) \right\} \\
&= \sum_i \left(\frac{Y_i}{Y}\right) \sum_j \left(\frac{Y_{ij}}{Y_i}\right) T_{ij} + \sum_i \left(\frac{Y_i}{Y}\right) T_i + T_{BR}
\end{aligned}
$$

where T_{BR} is the Theil index of inequality between regions, T_i is the inequality index between states in region i and T_{ij} is the inequality index between AMC in state j in region i. The weighted sum of within-state indexes form the within-state component of the global index, and the weighted sum of between-state indexes form the intermediate between-state component.

In the top panel of Table 2, we present the evolution of the global index and of the share of each component: the between-regions and between-states components have decreased over the period, while the within-state component has increased, which is consistent with the formation of two convergence clubs, observed by Andrade et al. (2004). The bottom panel of Table 2 presents the evolution of the Theil index for each region: inequalities between municipios have increased in the Northeast and the North, while they have decreased in the South and the

Southeast. This evolution is compatible with the existence of persistent poverty traps in the North and Northeast. Those findings clearly require a more detailed investigation of the spatial pattern of income distribution and growth.

Table 2: Evolution of income inequalities between municipios – 1970-1996

	1970	1975	1980	1985	1996
Global Theil	0.415	0.411	0.382	0.314	0.297
Between regions	33 %	34 %	28 %	26 %	28 %
Between states	18 %	15 %	11 %	11 %	11 %
Within states	49 %	51 %	61 %	63 %	61 %
Intraregional between-municipios Theil indexes					
Center-West	0.26	0.17	0.17	0.22	0.22
North	0.19	0.18	0.2	0.23	0.28
Northeast	0.36	0.39	0.41	0.49	0.41
South	0.2	0.13	0.16	0.15	0.16
Southeast	0.29	0.3	0.29	0.2	0.18

4 Exploratory spatial data analysis

In this section we use statistical measures of global and local spatial association to investigate the dependance across the level and growth of the municipios average per capita incomes. The extent of spatial dependance of a given variable among a set of spatially distributed units can be assessed by computing a measure of global statistical dependence such as the Moran's I statistic:

$$I = \frac{n}{S_0} \sum_i \sum_j w_{ij} z_i z_j / \sum_i z_i^2 = \frac{n}{S_0} \frac{z'Wz}{z'z} \tag{1}$$

where n is the number of municipios, $(W)_{ij} = w_{ij}$ is a weight indicating how region i is spatially connected to region j, $S_0 = \sum_i \sum_j w_{ij}$ is a scaling factor and z_i and z_j are values of the log-average income per capita in municipios i and j (i.e. $z_i = \log(y_i/\bar{y})$ where y_i is the income per capita in municipio i). Various forms of the weight matrix W can be considered, ranging from a simple binary contiguity matrix to a distance-based matrix where regions that are faraway may be assumed to have a negligeable influence. In our case, the Moran's I statistic has been computed using first and second order contiguity matrices, where, in the first order case, $w_{ij} = 1$ if i and j share a common border and 0 otherwise and, in the second order case, $w_{ij} = 1$ if i and j share a common border or if j shares a border with an municipio bordering i and 0 otherwise ($w_{ii} = 0$ for all i). In order to normalize the outside influence upon each region, the weights are normalized, so that $\sum_j w_{ij} = 1$ for each i. In this case expression (1) simplifies since $S_0 = n$. Positive values of the Moran's I indicate positive spatial dependence, that is the clustering of similar attribute values, whereas negative values are associated with the clustering of dissimilar values.[4] The Moran's I statistic can be decomposed into a set of local indicators of spatial association (LISA), as developped by Anselin (1995). For municipio i the value of the LISA is given by:

$$I_i = \frac{n z_i \sum_j w_{ij} z_j}{\sum_i z_i^2} \tag{2}$$

and we have $I = \frac{1}{S_0} \sum_i I_i$. Using a method suggested by Anselin (1995), it is possible to generate an empirical distribution of the LISA index. This distribution can then be used to assess the statistical significance of the local statistics. The LISA for each municipio therefore gives a indication of significant spatial clustering of similar values around that observation. A positive value indicates spatial clustering of similar values (high or low) whereas a negative value indicates spatial clustering of dissimilar values between a region and its neighbours.

The Moran scatterplot is another tool for studying the local clustering of similar or dissimilar values. For each locality it plots the spatial lag, $\sum_j w_{ij} z_j$, against the original value z_i. The four different quadrants of the scatterplot correspond to the four possible types of local spatial association between a region and its neighbours. Regions with a high value (relative to the mean) surrounded by regions with high values are in the top right quadrant (HH). On the opposite, regions with low values, surrounded by regions with low values are found in the bottom left quadrant (LL). At the top left, one finds regions with low values, surrounded by regions with high values (LH) and at the bottom right, regions with high values, surrounded by regions with low values (HL). Quadrants HH and LL (respectively LH and HL) refer to positive (resp. negative) spatial autocorrelation indicating the spatial clustering of similar (resp. dissimilar) values.

We computed the values of the global Moran's I statistic for the log-average GDP per capita in years 1970 and 1996, as well as for the growth rate of the per capita GDP over the 1970-1996 period. Results are reported in table (3) for the first order and second order contiguity matrices.

Table 3: I-Moran Test

	First-Order Contiguity	Second-Order Contiguity
log GDP per capita 1970	0.702***	0.659***
log GDP per capita 1996	0.743***	0.711***
Growth GDP per capita 1970-96	0.259***	0.201***
*** : significant at 1%.		

In all cases we find highly (1%) significant and positive values of the Moran's global statistic, indicating clustering of similar values of the GDP per capita level in 1970 and 1996 and of the growth rate. In other words, municipios with relatively high (resp. low) values of per capita GDP are localized close to other municipios with relatively high (resp. low) per capita GDP

[4]The Moran's I statistic gives an indication of the degree of linear association between the vector z of observed values and the vector Wz of spatially weighted averages of neighbouring values. Values of I larger (resp. smaller) than the expected value under the hypothesis of no spatial autocorrelation, $E(I) = -1/(n-1)$, indicate positive (resp. negative) spatial autocorrelation, that is the clustering of similar (resp. dissimilar) attribute values.

more often than if their localization were purely random. This tendency appears to reinforce over time, since the Moran statistic is found higher in 1996 than in 1970. The same kind of evidence is found for the per capita GDP growth rate. One can notice that the value of the Moran's statistic decreases with the order of the contiguity matrix. This is not surprising if one expects the degree of spatial association to decrease with the distance between municipios.

The Moran scatterplot in Figures 1, 2 and 3 give a visual representation of this association. Figures 1 and 2 show the scatterplots obtained for GDP per capita in 1970 and 1996 respectively. We can see that most municipios are found in either quadrant HH or LL. Only a small proportion of municipios are found in quadrants LH or HL. Tables 4 and 5 show, for each quadrant, the percentage of municipios with a LISA statistic significant at the 5% level in a given state. It must be mentionned that due to global spatial autocorrelation, pseudo-significance levels of inference must be used (Anselin, 1995), that is if, in the absence of spatial autocorrelation, the significance level is set to α, in the present case, the significance level is set to α/k, where k is the number of municipios in the contiguity set.[5]

We can see that no municipio is found with a significant LISA statistic in the LH quadrant and only a very small proportion of the municipios is found in the HL quadrant. Moreover, the tables also show that the municipios in the HH quadrant mostly belong to the Southeast and South regions, whereas those in the LL quadrant mostly belong to the Northeast. This shows evidence of a spatial clustering between the Northeast on one side, and the Southeast and South on the other side, a result also found by Bosch Mossi et al. (2003) at the state level. Comparisons between Tables 4 and 5) show the changes in this clustering between 1970 and 1996. What we find is that the proportion of municipios in the LL quadrant tends to increase very much for states in the Northeast, whereas the proportion of municipios in the HH quadrant is rather stable. Thus over the period 1970-1996 the extent of spatial clustering seems to increase in Brazil as a whole, but this dynamics appears to be mainly due to the specific growth pattern of the Northeast.

The Moran scatterplot for growth is presented in Figure 3 together with the states percentage of significant LISA statistics in Table 6. The pattern is not as clear as with the GDP per capita levels. The percentage of municipios with a significant LISA statistic is much lower. However, the same opposition between the Northeast and the southern states appears. The states with a significant proportion of municipios in the LL quadrant (namely Alagoas (AL), Bahia (BA) and Maranhão (MA)) all belong to the Northeast region. In the HH quadrant, the only state with a sizeable proportion of municipios presenting a significant level of spatial association with neighbouring municipios is Paraná (PR), located in the South region.

Altogether, these results confirm the emergence of convergence clubs found by Bosch Mossi et al. (2003) at the state level, and by Andrade et al. (2003) at the municipios level, but with a different method of investigation. This pattern of spatial statistical association between GDP per capita levels and growth rates, does not tell us anything about causal relationships. In order

[5] This is the so-called Bonferroni pseudo-significance level. Another possible choice is the Sidàk pseudo-significance level that is equal to $1-(1-a)^{1/m}$. However, this requires the local statistics to be multivariate normal, which is unlikely to be the case with LISA. The chosen pseudo-significance level does not require this assumption.

to go beyond these results one needs to employ econometric methods of analysis, to which we now turn.

5 Explaining spatial dependence in growth and levels of income: some theoretical developments

In a recent paper López-Bazo, Vayá and Artís (2004) present a simple model of growth that allows for externalities across economies. This model provides the basis of our empirical investigations. Output, Y, is produced using labour, L, physical, K, and human, H, capital. The technology is assumed to be of the Cobb-Douglas type with constant returns to scale, so that output per capita in municipio i in period t, y_{it}, is a function of the levels of per capita physical and human capital, k_{it} and h_{it} and of the state of technology, A_{it}:

$$y_{it} = A_{it} k_{it}^{\tau_k} h_{it}^{\tau_h}$$

where τ_k and τ_h are internal returns to physical and human capital respectively. The assumption of constant returns to scale in labour and both types of capital implies that $\tau_k + \tau_h < 1$.

Technology in municipio i, A_{it}, is assumed to depend on the technological level of the neighbouring municipios, which is in turn related to their stocks of both types of capital:

$$A_{it} = A_t (k_{\rho it}^{\tau_k} h_{\rho it}^{\tau_h})^{\gamma}$$

where A_t is an exogenous component, common to all municipios and $k_{\rho it}$ (resp. $h_{\rho it}$) denotes the average physical (resp. human) capital ratio in the neighbouring municipios. The γ coefficient measures the externality across municipios. If γ is positive, a one percent increase in the level of the per capita average physical stock of neighbouring municipios increases technology in municipio i by $\gamma\tau_k$ percent. Thus under this assumption, a municipio benefits from investments made by its neighbours.

Given the assumptions of internal constant returns to scale and of technological externalities, the growth rates of physical and human capital in each municipio are assumed to be decreasing functions of their stocks, but are increasing functions of the stocks of these factors in the neighbouring municipios. As pointed out by López-Bazo et al., this means that investments in physical and human capital are going to be more profitable, and therefore larger, in municipios surrounded by other municipios with high stocks of these factors. In contrast, incentives to invest will be lower in municipios surrounded by others with low capital intensity. This could explain the emergence of convergence clubs.

6 Estimating β-convergence between municipios

Our assumptions on technological spillovers across municipios lead to the following empirical growth equation (see López-Bazo et alli. (2004) for details):

$$g = c - (1 - e^{-\beta T}) \ln y_0 + \frac{(1 - e^{-\beta T})\gamma}{1 - \tau_k - \tau_h} \ln y_{\rho 0} + \gamma g_\rho + \varepsilon \tag{3}$$

where g is the per capita GDP growth rate, y_0 is the per capita GDP at the beginning of the observation period, g_ρ and $y_{\rho 0}$ are the average values of g and y_0 over neighbouring municipios and ε is a random term that is assumed centered, normally distributed with variance σ^2. If the rate of convergence, β, is significantly positive, poorer areas tend to grow faster than wealthier ones. When γ is equal to zero, this model reduces to the standard neo-classical growth model of unconditional convergence. In the presence of positive technological externalities, γ is positive and both the average level of per capita GDP in neighbouring municipios at the beginning of the observation period and the average growth rate have a positive effect on the steady state growth rate. Growth will be higher in municipios surrounded by neighbours with high initial per capita GDP and high rates of growth.

We complete equation (3) by adding on the right-hand side a set, X, of control variables that could cause differences in the rate of technological progress and the steady state across municipios:

$$g = c - (1 - e^{-\beta T}) \ln y_0 + \frac{(1 - e^{-\beta T})\gamma}{1 - \tau_k - \tau_h} \ln y_{\rho 0} + \gamma g_\rho + X\delta + \varepsilon. \tag{4}$$

This inclusion is also necessary in order to control for similarities between neighbouring municipios, which, in the absence of these variables, could cause the coefficients of $\ln y_{\rho 0}$ and g_ρ to be found spuriously significant. In the set of control variables we include the shares of the primary and secondary sectors in GDP, to account for the heterogeneity in the industrial mix across municipios, the illiteracy rate among individuals aged 15 or over, the share of people aged 25 or over, the mean size of households, the share of urban population and the share of households with electricity. All variables are measured in 1970.

The spatial lags of GDP per capita in 1970 and growth rates are computed using the row-standardized spatial weight matrix, W. The econometric model is thus written:

$$g = c + \alpha \ln y_0 + \theta \ln(W y_0) + \gamma W g + X\delta + \varepsilon. \tag{5}$$

We estimate several versions of the model presented in equation (4), starting with the standard OLS specification ($\gamma = 0$) and then including spatial lag variables. Note that according to our structural model, $\gamma = 0$ implies $\theta = 0$ in the above equation. In what follows we shall not impose this restriction. We also contrast the results of the spatial lag model with those of the spatial error model, in which the model residuals follow a spatially auto-regressive process:

$$g = c + \alpha \ln y_0 + \theta \ln(W y_0) + X\delta + \varepsilon$$
$$\varepsilon = \lambda W \varepsilon + u$$

where u is an uncorrelated and homoskedastic error term.

The results are presented in Tables 7 to 10. Table 7 shows the results obtained when spillovers across municipios are neglected ($\gamma = \theta = 0$). First, we consider the absolute convergence model, where the only regressor is the log of initial per capita income. This model assumes that all municipios have the same steady-state. The fit we obtain being very poor, this assumption does not seem correct. As expected, conditional convergence estimations lead to higher rates of convergence across municipios. The variable proxying for human capital initial endowment has the expected sign and is strongly significant: municipios less richly endowed with human capital tend to grow at a slower pace. Infrastructures also play a key role in growth prospects: municipios where households had better access to electricity in 1970 have grown faster. The level of urbanization in 1970 has a negative impact on growth. The best fit is obtained when regional fixed effects are included (column 3) and we will use this model in the subsequent analysis. For this model, the point estimate of the yearly rate of convergence between municipios is 3.4%.[6]

Results obtained in section 4 clearly indicate that levels and growth of per capita GDP are spatially clustered. For this reason, we compute various tests of residual spatial autocorrelation using the weight matrices defined above. The Moran's I test is simply the application of the Moran's I to OLS residuals. A significant value indicates that the residuals are spatially correlated, which is the case for all the OLS models we have estimated. Note that tests results reported in Table 7 are obtained using the second order contiguity matrix; the first order matrix leads to similarly significant results. LM (Lagrange multiplier) tests are used to obtain a more precise idea of the kind of spatial dependence involved (Anselin and Bera, 1998). We first conducted the test while imposing $\theta = 0$. The Lagrange multiplier test for the spatial lag model (LM lag) tests the null hypothesis $\gamma = 0$. This test is significant for all the convergence models proposed, indicating that the null hypothesis $\gamma = 0$ must be rejected. Since the spatial lag model of equation (5) reduces to the simple convergence model (equation 4) when $\gamma = \theta = 0$, this latter model must be rejected. The LM test for the presence of spatial error autocorrelation (LM error) tests the null hypothesis $\lambda = 0$, where λ is the spatial autoregressive coefficient for the error lag $W\varepsilon$ in the following model:

$$g = c + \alpha \ln y_0 + X\delta + \varepsilon \tag{6}$$
$$\varepsilon = \lambda W \varepsilon + u.$$

The LM error test is significant in all cases, indicating that the hypothesis $\lambda = 0$ must be rejected. Both the spatial lag and the spatial error models are therefore preferable to our initial model. The robust LM lag test (test of $\gamma = 0$ in the presence of local misspecification involving

[6]The yearly rate of convergence is given by $\beta = -\ln(1+\alpha)/T$.

spatial-dependent error process), the robust LM error test (test of $\lambda = 0$ in the local presence of γ) and the SARMA test (joint test of $\gamma = \lambda = 0$ in the spatial autoregressive moving-average model, see Anselin and Bera, 1998) theoretically permit to determine which of these two models should be chosen. However in our case, all these tests are significant, and we are therefore not able to discriminate between the spatial lag and the spatial error models at this stage.

We estimated both models using different spatial weight matrix definitions (the first and second order contiguity matrices, as well as a spatial weight matrix based on 100 kilometers distance cut-off). In the spatial lag model, since the spatially lagged dependent variable Wg is correlated with the error term, OLS estimation will yield biased inconsistent estimates. In the spatial error model, OLS estimates are not biased but inefficient, due to the error covariance matrix being non spherical. As shown by Anselin and Bera (1998), both models can be consistently and efficiently estimated by maximum likelihood and this is the choice we made. Results reported in Table 8 are those obtained with the second order contiguity matrix since the Akaike Information Criterion was systematically lower for models estimated with this weight matrix. Interestingly, the application of this criterion would lead to prefer the spatial error model over the spatial lag model. For the spatial lag model, we find a positive spatial dependence between the growth rates of municipios belonging to the same neighbourhood and, for the spatial error model, we find a positive spatial autocorrelation in measurement errors or in possibly omitted variables. The estimated yearly rates of convergence are quite different: 3.04% for the spatial lag model and 4.32% for the spatial error model, a figure quite higher than what was estimated in the initial model.

We now relax the assumption that the initial level of income in neighbouring municipios does not affect the growth rate of GDP per capita (θ is no longer imposed to equal 0). The spatial cross-regressive model is obtained when $\theta \neq 0$ and $\gamma = 0$ (Anselin, 2003):

$$g = c + \alpha \ln y_0 + \theta \ln(Wy_0) + X\delta + \varepsilon. \tag{7}$$

This model can be safely estimated by means of OLS. The model was estimated using first and second order contiguity matrices, but the latter one provided the best fit. Results are presented in Table 9. Compared to the initial model (Table 7), the inclusion of the spatially lagged initial income improves the estimation. We find a significant positive impact of the initial income of the neighbourhood on growth. The tests of residual spatial autocorrelation indicate that this model does not capture the full extent of spatial effects. We therefore estimate the following two models:

$$\begin{aligned} g &= c + \alpha \ln y_0 + \theta \ln(Wy_0) + X\delta + \varepsilon \qquad (8)\\ \varepsilon &= \lambda W\varepsilon + u \end{aligned}$$

which we could call the spatial cross-regressive spatial error model, and:

$$g = c + \alpha \ln y_0 + \theta \ln(Wy_0) + \gamma Wg + X\delta + \varepsilon \tag{9}$$

which corresponds to our structural model (equation (5)) and which could be termed spatial cross-regressive spatial lag model.

Estimation results of these models using the second order contiguity matrix are presented in Table 10. Both models clearly outperform the spatial cross-regressive model of equation (7). We obtain a point estimate of 4.2% for the yearly rate of convergence between municipios in the spatial cross-regressive spatial error model and of 3.9% in the spatial cross-regressive spatial lag model. Both the coefficient of the spatially lagged initial income and the spatial autoregressive coefficient for the error lag or for the growth lag are positive and highly significant.

Based on the AIC criterion, our preferred specification is that of the spatial cross-regressive spatial error model, though the difference between the spatial lag and the spatial error models is small.

One can give a structural interpretation to the results of the spatial cross-regressive spatial error model. Computing the reduced form of ε in terms of u and replacing in equation (8), one gets:

$$g = (I-\lambda W)c+\alpha \ln y_0+\theta \ln(Wy_0)+\lambda Wg+X\delta-\lambda\alpha W \ln y_0-\lambda\theta W \ln(Wy_0)-\lambda WX\delta+u \quad (10)$$

This equation is directly comparable to equation (9). Several comments are in order. First, the values of γ and λ are found quite close to each other in Table 10, which indicates that both models predict a quite similar impact of the neighbours' growth on the growth of a given municipio. Second, provided that $\ln(Wy_0)$ and $W \ln y_0$ are close enough[7], the overall impact on growth of the initial income of neighbours in the spatial cross-regressive spatial error model can be approximated by $\theta - \lambda\alpha$, while it is directly measured by θ in the spatial cross-regressive spatial lag model. This impact is thus higher in the former ($\theta - \lambda\alpha = 0.728$) than in the latter ($\theta = 0.409$). Third, the spatial cross-regressive spatial error model implies that the initial income of more distant neighbours, through the variable $W \ln(Wy_0)$, has a modest negative impact on growth.

7 Conclusion

This paper has shown the extent of interdependencies between the Brazilian municipios, using a simple growth model in which externalities across economies positively influence the process of production. These externalities counterbalance the neoclassical law of diminishing returns to capital, thus making further investment more productive. This could explain the emergence of spatial poverty traps that has been found in Brazil in several previous surveys and that has been confirmed in this paper.

These results raise some important policy issues. Policies designed to reduce regional poverty should take account of the potential spillovers of geographically targeted investments in physical and human capital stocks.

(To be concluded).

[7] Which is the case, with a correlation of 0.9931.

References

[1] **Abreu, M., H.L.F. de Groot and R.J.G.M. Florax (2004):** "Space and Growth", *Tinbergen Institute Discussion Paper* TI 2004-129/3.

[2] **Andrade, E., M. Laurini, R. Madalozzo, P. L. Valls Pereira (2004):** "Convergence clubs among Brazilian municipalities", *Economics Letters*, **83**: 179-184.

[3] **Anselin, L. (1995):** "Local indicators of spatial association-LISA", *Geographical Analysis*, **27**: 93-115.

[4] **Anselin, L. (2003):** "Spatial externalities, spatial multipliers and spatial econometrics", *International Regional Science Review*, **26**: 153-166.

[5] **Anselin, L. and A.K. Bera (1998):** "Spatial Dependance in Linear Regression Models with an Introduction to Spatial Econometrics", in *Handbook of Applied Economic Statistics*, A. Ullah and D. E.A. Gilles eds., Marcel Dekker, Inc., New-York - Basel - Hong Kong.

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

[7] **Bosch Mossi, M., P. Aroca, I.J. Fernandez and C.R. Azzoni (2003):** "Growth Dynamics and Space in Brazil", *International Regional Science Review*, **26**, 3: 393-418.

[8] **Ferreira, A. (2000):** "Convergence in Brazil: recent trends and long-run prospects", *Applied Economics*, **32**: 479-489.

[9] **Glaeser, E., H. Kallal, J. Scheinkman and A. Shleifer (1992):** "Growth in cities", *Journal of Political Economy*, **100**: 1126-1152.

[10] **Lall, S.V. and Z. Shalizi (2003):** "Location and growth in the Brazilian Northeast", *Journal of Regional Science*, **43**, 4: 663-681.

[11] **Lopez-Bazo, E., E. Vaya and M. Artis (2004):** "Regional externalities and growth: Evidence from European regions", *Journal of Regional Science*, **44**: 43-73.

[12] **Magalhães A., G. J.D. Hewings and C.R. Azzoni (2000):** "Spatial Dependance and Regional Convergence in Brazil", *REAL Discussion Paper* 00-T-11.

[13] **Magrini, S. (2004):** "Regional (di)convergence", in *Handbook of Regional and Urban Economics, 4*, V. Henderson and J.-F. Thisse eds., North Holland, Amsterdam.

[14] **Mankiw, G., D. Romer and D. Weil (1992):** "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, **107**: 407-437.

[15] **Quah, D. (1997):** "Empirics for growth and distribution: stratification, polarization and convergence clubs", *Journal of Economic Growth*, **2**: 27-59.

[16] **Ravallion, M. (1998):** "Poor areas", in *Handbook of Applied Economic Statistics*, A. Ullah and D. E.A. Gilles eds., Marcel Dekker, Inc., New-York - Basel - Hong Kong.

[17] **Rey, S.J. and B.D. Montouri (1999):** "US regional income convergence: A spatial econometric perspective", *Regional Studies*, **33**, 2: 143-156.

[18] **Shorrocks, A.F. (1984):** "Inequality decomposition by population subgroups", *Econometrica*, **52**: 1369-1386.

Figure 1: Moran scatterplot – log GDP per capita 1970

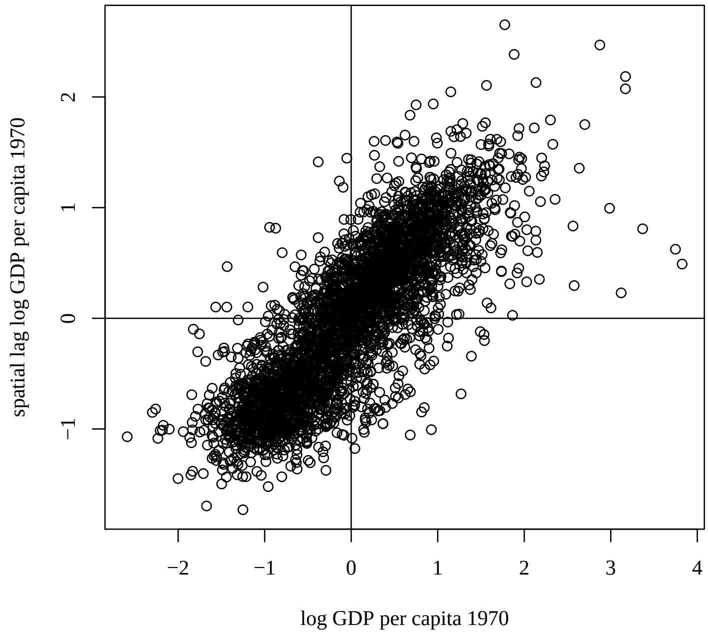

Table 4: Percentage of municipios in each state with LISA significant at 5% in each quadrant of the Moran scatterplot – log GDP per capita 1970

LH:	**HH:** BA (1.5) ES (1.9) GO (1.2) MG (2.5) PR (4.7) RJ (17.7) RS (51.8) SC (13.3) SP (47.6)
LL: AL (13.6) BA (15.6) CE (45.7) GO (1.9) MA (15.9) MG (4) PB (41.7) PE (17.3) PI (56.2) RN (44.2) SE (8.1) TO (8.8)	**HL:** BA (0.3) MG (0.1) RN (0.7)
Interpretation: 47.6% of the municipios of São Paulo are in the HH quadrant and have a LISA significant at 5%. 56.2% of the municipios of Piaui are in the LL quadrant and have a LISA significant at 5%.	

Figure 2: Moran scatterplot – log GDP per capita 1996

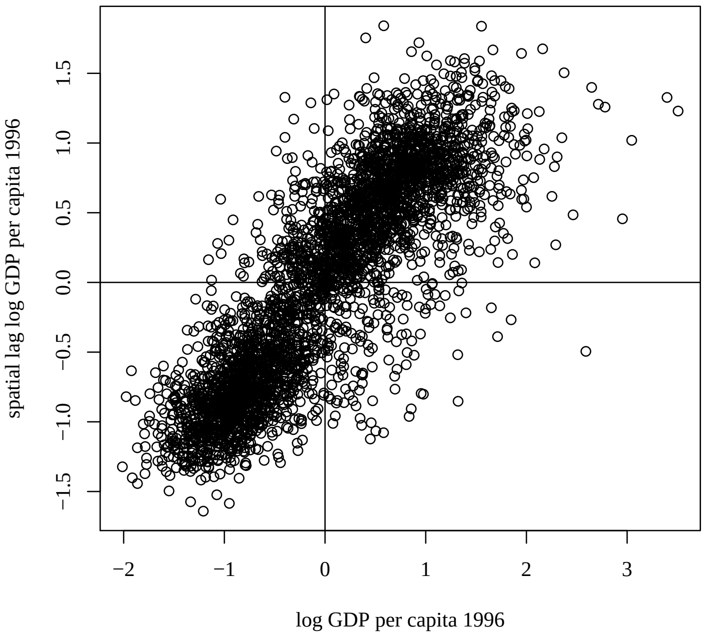

Table 5: Percentage of municipios in each state with LISA significant at 5% in each quadrant of the Moran scatterplot – log GDP per capita 1996

LH:	**HH:** BA (1.5) ES (1.9) GO (1.2) MG (4.4) PR (4.3) RJ (6.5) RS (43.1) SC (29.4) SP (49.7)
LL: AL (58) BA (23.5) CE (49.3) GO (1.2) MA (61.9) MG (2.4) PB (20.8) PE (6.8) PI (58.8) RN (64.6) SE (1.4) TO (20.6)	**HL:** BA (0.3) PI (1.2) SE (1.4)

Figure 3: Moran scatterplot – growth 1970-1996

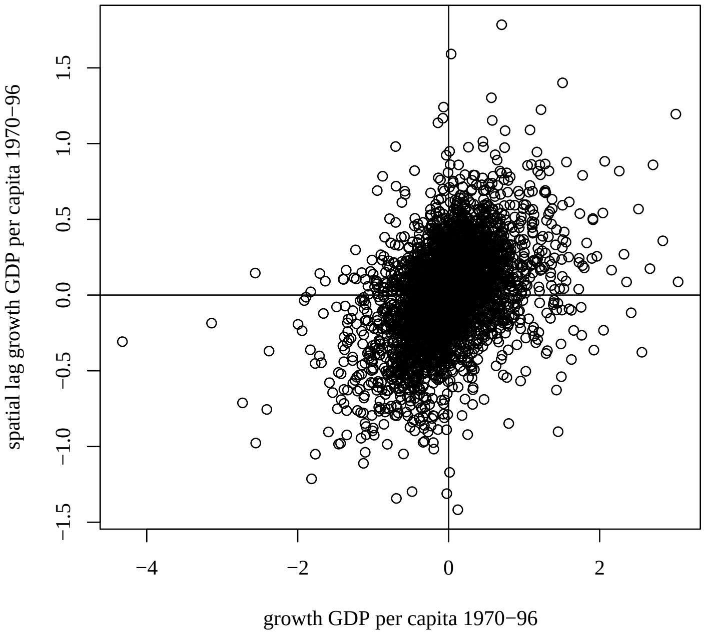

Table 6: Percentage of municipios in each state with LISA significant at 5% in each quadrant of the Moran scatterplot – growth 1970-1996

LH: AL (1.1) MG (0.6) PR (0.4) RJ (1.6)	**HH:** BA (3.1) CE (2.2) ES (1.9) GO (3.8) MA (2.7) MG (2.2) PB (1.8) PE (1.2) PI (1.2) PR (13.4) SC (1.1) SP (5.1)
LL: AL (34.1) BA (15.6) MA (14.2) MG (0.7) PE (1.2) PR (1.4) RJ (1.6) RN (2.7) SE (4.1) SP (2.3)	**HL:** AL (4.5) BA (0.9) ES (1.9) MG (0.3) RJ (1.6) RN (0.7) SE (4.1) SP (0.4)

Table 7: Estimation of the standard growth equation and tests of residual spatial dependence

	Absolute convergence	Conditional convergence (1)	Conditional convergence (2)
Constant	2.098***	5.705***	4.752***
	(26.75)	(20.83)	(16.72)
Log initial income	-0.186***	-0.536***	-0.588***
	(-16.46)	(-28.58)	(-31.61)
Illiteracy		-1.843***	-1.022***
		(-24.74)	(-11.02)
Urbanization		-0.524***	-0.248***
		(-7.67)	(-3.54)
Electricity		0.652***	0.558***
		(8.18)	(7.01)
Agriculture		-0.098	-0.018
		(-1.5)	(-0.28)
Industry		-0.117	0.006
		(1.44)	(0.08)
Labour force		-0.713**	0.21
		(2.31)	(0.68)
Household size		-0.01	0.013
		(-0.41)	(0.56)
Region fixed effects	No	No	Yes
Adj. R^2	0.072	0.286	0.329
I-Moran	0.281***	0.204***	0.193***
LM lag	1475.1***	637.2***	415.7***
Robust LM lag	669.6***	26.5***	85.3***
LM error	2741.6***	1435.4***	1282.4***
Robust LM error	1936.2***	824.7***	952***
SARMA	3411.3***	1461.9***	1367.7***

t-statistics in parentheses.

*, ** and *** : significant at 10%, 5% and 1%.

Table 8: Estimation of the growth equation with externalities across economies

	Spatial lag model	Spatial error model
Constant	4.382***	4.296***
	(45.43)	(37.59)
Log initial income	-0.547***	-0.675***
	(-28.99)	(-36.42)
Illiteracy	-0.932***	-0.791***
	(-11.56)	(-7.76)
Urbanization	-0.285***	-0.290***
	(-4.37)	(-4.22)
Electricity	0.615***	0.5***
	(8.08)	(5.4)
Agriculture	0.009	-0.162***
	(0.14)	(-2.7)
Industry	0.063	0.07
	(0.82)	(0.93)
Labour force	0.052	0.692***
	(0.27)	(3.09)
Household size	-0.036**	0.096***
	(-2.2)	(5.53)
γ (W*growth rate)	0.485***	
	(5.73)	
λ (W*error term)		0.751***
		(84.24)
Region fixed effect	Yes	Yes
Log likelihood	-906.069	-742.986
AIC	1838.1	1512

Asymptotic t-statistics in parentheses.
*, ** and *** : significant at 10%, 5% and 1%.

Table 9: Estimation of the growth equation (7) and tests of residual spatial dependence

	Spatial cross-regressive model
Constant	3.247***
	(10.03)
Log initial income	-0.647***
	(-33.3)
Illiteracy	-0.857***
	(-9.18)
Urbanization	-0.151**
	(-2.16)
Electricity	0.394***
	(4.9)
Agriculture	-0.095
	(-1.49)
Industry	-0.079
	(-0.99)
Labour force	0.242
	(0.79)
Household size	0.072***
	(3.06)
θ (W*log init.inc.)	0.23***
	(9.34)
Region fixed effects	Yes
Adj. R^2	0.345
I-Moran	0.186***
LM lag	898.5***
Robust LM lag	16.1***
LM error	404.2***
Robust LM error	142.6512***
SARMA	1302.7***

t-statistics in parentheses.

*, ** and *** : significant at 10%, 5% and 1%.

Table 10: Estimation of the growth equations (8) and (9)

	Spatial cross-regressive spatial error model	Spatial cross-regressive spatial lag model
Constant	2.484***	1.578***
	(8.14)	(5.2)
Log initial income	-0.667***	-0.636***
	(-40.7)	(-49.28)
Illiteracy	-0.718***	-0.61***
	(-6.95)	(-7.06)
Urbanization	-0.246***	-0.126**
	(-3.68)	(-2.04)
Electricity	0.417***	0.343***
	(4.49)	(4.71)
Agriculture	-0.197***	-0.119**
	(-3.36)	(-2.38)
Industry	0.031	-0.07
	(0.43)	(-1.07)
Labour force	0.66***	0.056
	(2.58)	(0.23)
Household size	0.114***	0.052***
	(5.81)	(2.62)
θ (W*log init.inc.)	0.252***	0.409***
	(7.09)	(17.44)
γ (W*growth rate)		0.648***
		(24.85)
λ (W*error term)	0.714***	
	(53.99)	
Region fixed effect	Yes	Yes
Log likelihood	-722.217	-768.219
AIC	1472.4	1564.4

Asymptotic t-statistics in parentheses.
*, ** and *** : significant at 10%, 5% and 1%.