TEXTO PARA DISCUSSÃO N° 324

REGIONAL IMBALANCES AND MARKET POTENTIAL IN BRAZIL

Pedro Vasconcelos Amaral
Mauro Borges Lemos
Rodrigo Ferreira Simões
Flávia Chein Feres

Dezembro de 2007

Ficha catalográfica

333.73981 A485r 2007	Amaral, Pedro Vasconcelos. Regional imbalances and market potential in Brazil / Pedro Vasconcelos Amaral et al. - Belo Horizonte: UFMG/Cedeplar, 2007. 27p. (Texto para discussão ; 324) 1. Economia regional. 2. Geografia econômica. 3. Brasil – Condições econômicas. I. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. II. Título. III. Série. CDD

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL

REGIONAL IMBALANCES AND MARKET POTENTIAL IN BRAZIL[*]

Pedro Vasconcelos Amaral
Cedeplar/UFMG.

Mauro Borges Lemos
Professor at Cedeplar/UFMG.

Rodrigo Ferreira Simões
Professor at Cedeplar/UFMG.

Flávia Chein Feres
Researcher at Cedeplar/UFMG.

CEDEPLAR/FACE/UFMG
BELO HORIZONTE
2007

[*] Primeira versão: Dezembro de 2007.
Versão revista: Junho de 2009.

SUMÁRIO

RESUMO

O objetivo deste trabalho é analisar empiricamente a estrutura regional de produção e salários no Brasil, tendo como base o arcabouço teórico da Nova Geografia Econômica (NGE). Como usual na literatura, uma maneira de fazê-lo é através da estrutura espacial salarial e os diferenciais de mercado potencial, que no presente trabalho têm como referência empírica os municípios brasileiros no período 1980-2000.

Um avanço deste trabalho é a utilização de um modelo de dados em painel com componentes do erro espacialmente e temporalmente correlacionados, modelagem esta que mais adequada para a estimação, que nos forneceu resultados robustos e ainda não havia sido aplicada à realidade brasileira. Os resultados apontam para uma forte relação entre o mercado potencial de determinada localidade e seu nível salarial, indicando que atributos regionais, e não apenas individuais, também estão correlacionados com o nível regional de salários e suas disparidades.

Palavras-chave: diferenciais salariais, Nova Geografia Econômica, Brasil.

Classificação JEL: J31, R12.

ABSTRACT

This paper estimates the effects of market potential on the regional wage imbalances among Brazilian municipalities over recent decades. More specifically, with the 1980, 1991 and 2000 Brazilian Census data (at 3,630 comparable municipality areas), we estimate the NEG's wage equation using a spatial panel data model with the presence of endogeneity. Our results bring out new evidence of a strong relationship between market potential and wages, indicating that regional, not only individual, attributes, are also correlated with wages and their regional imbalances.

Keywords: regional imbalances, market potential, New Economic Geography, Brazil.

*JEL Classificati*on: J31, R12

INTRODUCTION

Identification of the factors that affect the production level in different regions is a topic of major economic interest. Increases in production levels in depressed areas are directly correlated to gains in quality of life and reduction in regional disparities. The main objective of this work is to identify the role of market potential in regional wage imbalances. To do so, we start from the New Economic Geography's theoretical background to understand and explain the wage spatial structure among the Brazilian municipalities in the last few decades, and its differentials. As usual, this is carried out based on an extension of the NEG's wage equation.

For a long time, regional scientists and location theoreticians have dealt with the problem of the distribution and spatial concentration of economic activities. However, such authors have been kept out of the economics mainstream, mainly because of the difficulty in understanding their ideas in a competitive context and for their few mathematical and empirical formalizations (Krugman, 1991). Following Dixit-Stiglitz's monopolist competition model, New Economic Geography presents a theoretical and empirical approach in order to analyze the causes of production concentration. (Fujita et al., 1999).

The spatial concentration of economic activities would be a consequence of increasing returns due to agglomeration economies and transportation costs, in a context of monopolistic competition, according to the modeling of Dixit and Stiglitz (1977). Proximity to the consuming market or to the suppliers of inputs would imply lower transportation costs, propitiating pecuniary externalities. Such factors would lead to the agglomeration of productive activities in the centers of higher potential markets. (Lösch, 1954). Besides, production concentration in an only space would reduce the fixed costs and allow greater exploitation of scale profits (Marshall, 1890).

Market potential is one of the principal factors that weigh in a firm's choice of localization, together with outside factors: climate and initial endowment of resources being two such. Thus, as considered by Harris (1954), the potential demand of a determined place would be related to the total purchasing capability of other places, weighted by the transportation cost. In this way, nearby consumer markets would have greater impact on the local economy (demand linkages) than more distant consumer markets, if they possessed the same total income.

However, agglomeration does not possess only positive effects on productive activity. The spatial concentration of productive activity also generates negative externalities or congestion costs, related mainly to the demand for fixed factors, such as land, as well as factors like pollution, violence and urban congestion. These factors raise production costs. If we consider the labor supply as not perfectly elastic, the productive concentration and consequent rise of the demand for labor would also cause an increasing price of this factor, raising the nominal wages.

The relationship between potential market and nominal wages in a determined locality is one of the main points of New Economic Geography (NEG) and one of its main differences in relation to the traditional neoclassic theory. While the first states that nominal wages differentiation is possible, the latter, in its models of exogenous growth, contemplates the factors' price equalization in a competitive market, but restricted to convergence clubs. However, this theoretical conclusion of the NEG has not been the object of many empirical tests to identify its adequacy with the reality.

The purpose of this paper is to contribute to this research field and test empirically the relationship between market potential and wage differentials among Brazilian municipalities. The results point to a strong relationship between the market accessibility of a determined region and its wage level, indicating that not only individual characteristics are relevant to define wages, but also regional attributes.

THE NEW ECONOMIC GEOGRAPHY

The NEG theory starts from the assumption of two distinct economic sectors: the first, whose market structure is one of perfect competition, and another based on monopolistic competition, as considered by Dixit and Stiglitz (1977), with a great variety of differentiated goods and increasing returns of scale at the level of the individual company. The consumer utility function is a Cobb-Douglas:

$$U = M^{\mu} A^{1-\mu} \tag{1}$$

in which *M* represents the consumption of goods in the monopolistic competition sector (named industrial sector), *A* the consumption of the perfect competition sector goods (named agricultural or residual sector) and μ is a constant that represents the fraction of expense in industrialized goods.

The index of consumption *M* is a sub-utility function where *m* (i) denotes the consumption of each available variety and *n* the number of produced varieties, assuming that *M* is defined by a function of constant elasticity of substitution (CES):

$$M = \left[\int_0^n m(i)^{\rho} \, di \right]^{1/\rho}, \ 0 < \rho < 1 \tag{2}$$

where ρ represents the intensity of the preference for variety of industrialized goods.

The more ρ approaches 1, the more differentiated goods are to be considered perfect substitutes. As ρ approaches zero, greater importance is given to the variety of the goods. Making $\sigma \equiv 1/(1-\rho)$, σ represents the elasticity of substitution between any of two varieties of good. The same is valid for the index *A*, which represents consumption of the goods produced in perfect competition, whose elasticity of substitution between the varieties will be represented here by η .

From the modeling presented in Fujita et al. (1999), there are four groups of equations whose simultaneous solution determines the income, the products' index of price, the nominal and the real wage of each region.

The third group of equations, the most important for this work, is the one that determines the steady-state nominal wages:

$$
\begin{aligned}
w_r^M &= [\sum_s Y_s (G_s^M)^{\sigma-1} {T_{Mrs}}^{1-\sigma}]^{\frac{1}{\sigma}} \\
w_r^A &= [\sum_s Y_s (G_s^A)^{\eta-1} {T_{Ars}}^{1-\eta}]^{\frac{1}{\eta}}
\end{aligned}
\qquad (3)
$$

By the equations of nominal wages determination we can observe the backward demand effect, that is, the effect of potential market on wages in a determined region. As can be seen in equation (3), the higher the income in nearby producing regions, or the higher the market potential, the higher will be the nominal wage in this region, considering similar indices of prices in all regions.

In short, the equilibrium of the NEG model is achieved with the simultaneous solution of the income, price index, nominal and real wages equations. The dynamics is given by the forward and backward linkages. If we consider a balanced distribution of production activities between two regions and if, in any way, this concentration is modified, the possibility to save in transportation costs would lead to a reduced price index in the region with higher production. This price index reduction would raise the real wage, increasing the attractiveness of the region for workers – this is the forward linkage. Analogously, a higher concentration of consumers, i.e. a higher market potential, attracts more firms willing to save on transportation costs, which allows them to pay higher wages, which again induce migration – backward linkage. Therefore, there is a centripetal force caused by the attractiveness of agglomeration and lower transportation costs. On the other hand, the exogenous location of agricultural workers and competitiveness imply a centrifugal force.

Whether by means of the forward or backward linkages, there is a cumulative causation as in Myrdal (1957) inducing the concentration of workers and production. The close relationship between the market potential and the nominal wages implied by this cumulative causation is what we examine in this paper.

As demonstrated by Fujita and Mori (2005), there is a wide range of works related to the NEG, however only a few deal with the NEG specifically. Hanson (1998) was the first to estimate the structural parameters of Krugman (1991)'s model. When analyzing the economy of United States' counties, the author identified strong, highly concentrated, demand connections among the regions. Its results emphasize a reduction of the nominal wage proportional to the distance of the main consumer centers. From 1000 km or more, there would no longer be any impact on the local wages.

Following the same approach based on the wage equations, Brakman et al. (2004) find indications that confirm the presence of a spatial structure of wages in Germany, also highly concentrated, which indicates a great relevance of the distance in the wage determination. The authors also identify a strong effect of the former border between East and West Germany, such that demand connections are stronger internally within the old borders, than between regions situated on opposite sides of the old wall.

Mion (2004) uses the analysis of panel data in accordance with the methodology proposed by Arellano and Bond (1991) and Anselin and Kelejian (1997) in order to estimate a version of Helpman (1998)'s model for Italian regions. His model is an improved version of Krugman (1991) for multiple localities, considering the demand of land for housing. The author not only finds results that corroborate the theory of the demand connections but also suggests that the effect of the spatial externalities reaches larger areas than those found by Hanson (1998). However, the same author stresses that this difference can be partially credited to the adoption of a different matrix of distances.

Niebuhr (2006) finds evidence of rise of regional wages in accordance with potential markets in 158 regions of Europe. Her findings suggest a reduction of the effect of demand connections throughout the 90s. The work presents two empirical models to test the theory of a spatial wages structure. In the first, the regional prices index is calculated in accordance with the equilibrium condition of real wages equalization. In the second, the author assumes that there is no differentiation in the level of prices between the regions. From her estimations, Niebuhr (2006) concludes that assuming the same price level between the regions is the best choice in the absence of regional prices data. This hypothesis is also adopted in our work.

In his turn, Fingleton (2008a) collates the hypotheses of the neoclassic theory of conditional convergence and that of the NEG, searching for an explanation of the differentials in per capita production from 77 countries throughout years 70, 80, 90 and 2000. Using an artificially nestled model, auto-regressive and moving average spatial errors, the author finds indications that, although both backgrounds explain the differentials, NEG does so more efficiently.

Lin (2003) analyses the relationship between wages and access to international markets and input suppliers in Chinese provinces. The author estimates a gravitational equation to construct variables that measure the access to markets and input suppliers. Then she estimates the relationship among these variables and the nominal wage, and concludes that about 25% of the wage disparities among the provinces are related to the access to international markets and input suppliers.

An example of the scarcity of empirical works based on New Economic Geography is the rarity of works that deal with potential market and its relation to the spatial structure of nominal wages in Brazil using the theoretical background of the NEG. Figueiredo (2002) analyzes the effect of the scale economies on the distribution of productive activities between Brazilian states in the light of the NEG, in a context of decreasing transport costs. Batista da Silva and Silveira Neto (2005) also use the NEG background attempting to identify the role that the different types of agglomeration economies play in the growth of Brazilian industrial employment. Souza (2007) aims to identify the factors that induce production agglomeration. Among the factors, the author includes the forward and backward linkages of the NEG, the workforce density and some other factors that would affect the location of firms and workers. However, none of these aims to identify the relation between the potential market and the nominal wage among the Brazilian municipalities.

Apart from the NEG background, several works investigate the interaction between spatial localization and wages. Savedoff (1990) analyses the wage disparities between the Brazilian metropolitan areas in the 80s. According to the author, the classical factors such as cost of living disparities or workforce productivity were responsible for only a portion of the wage differentials. The imbalances in the demand of the labor market also played an important role. Following the same line, Servo (1999) finds that there is a permanent residual wage disparity even when controlling for individual attributes, employment characteristics and cost of living between the Brazilian metropolitan areas.

In Brazilian literature, remaining wage disparities after controlling for individual attributes in Brazil have also been emphasized by Reis and Barros (1990), Silveira-Neto and Campelo (2003), Fontes et al. (2009), Galinari et al. (2006), Paillacar (2007), among others.

Based on the development of the studies that empirically test the NEG equations, we decided to adopt a methodological approach inspired by Fingleton (2008a). Fingleton presents an approach theoretically simplified and methodologically complex, hoping to mitigate the usual econometrical issues in regional data estimations without abandoning the direct manner of interpreting the results.

Following Fingleton (2008a), we assume that the wages also depend on the efficiency level of the workforce. Representing the potential market by $P_r^M \equiv [\sum_s Y_s (G_s^M)^{\sigma-1} T_{Mrs}{}^{1-\sigma}]$ and inserting the effect caused by the workforce efficiency (X_r) on nominal wage, we can rewrite equation (3) as:

$$w_{rt} = P_{rt}{}^{1/\sigma} X_{rt} \tag{4}$$

We assume that the regional heterogeneity of labor efficiency derives from different levels of instruction (S_1), literacy ratio of adults older than 25 (S_2) and other residual factors (u)

$$\ln X_t = b_0 + b_1 \ln S_{1t} + b_2 \ln S_{2t} + u_t \tag{5}$$

where X_t, S_{1t}, S_{2t} and u_t are vectors and b_0, b_1 are b_2 are scalar parameters.

In such a way, substituting (5) in (4), we have the extended equation of the nominal wages:

$$\ln w_t = \frac{1}{\sigma} \ln P_t + b_0 + b_1 \ln S_{1t} + b_2 \ln S_{2t} + u_t \tag{6}$$

By the theory of New Economic Geography, equilibrium is reached when the migration of labour ceases, which would be the consequence of the equalization of real wages among the regions. In this situation, the economic activity of the monopolistic competition sector can be found equally divided between the diverse regions or concentrated in one or more regions, following a dynamic core-periphery, in accordance with the parameters of elasticity of substitution between the varieties, transportation costs and ratio of the expense in manufacturing goods.

Therefore, for the NEG there is the possibility of equalization of *per capita* production level in equilibrium in accordance with the sustentation point of the core-periphery structure or the rupture point of the productive activity's symmetrical distribution. Whether the equilibrium is reached with symmetry or a core-periphery structure, the NEG foresees the equalization of the real wages in the long run in order for labour migration to be interrupted. Regardless of real wage homogeneity, the nominal wages may vary according to transportation costs and price indexes.

In order to avoid the requirement of regional price indexes in equation (6), we focus on the short run, assuming that all Brazilian municipalities have the same price index. We also assume the equality of transport costs, price indices and elasticity of substitution between the competitive and the monopolistic competition sectors. The main advantage of our specification is that it is focused on the central point of the wage determination according to the NEG: the market potential. As far as we know, this is the first time this model has been applied to Brazilian data at the municipality level. As we show further, the Brazilian data on wages, education and market potential is highly spatially correlated, as would be expected in a municipality context. Therefore it requires a spatial modeling, since ignoring such a spatial pattern raises the variance of the estimated parameters and harms its efficiency. To address that, we estimate the NEG's wage equation adopting the method of panel data with spatially correlated error components proposed by Kapoor et al. (2007). As we show in the next section, the Kapoor et al. (2007)'s methodology considers only exogenous regressors. However, the nominal wage and market potential are endogenous variables. Hence, some modifications have been carried out in the methodology following Fingleton (2008b), to make it suitable for endogenous regressors.

ESTIMATION STRATEGY

Kapoor et al. (2007) present a model of panel data with components of the error correlated in space as well as in time, which we will call henceforth as the KKP model. According to the authors, spatial models that consider some measure of distance between the analyzed individuals use in general some modeling similar to those presented by Cliff and Ord (1973, 1981).

The analysis of panel data allows us to control the time-invariant effects specific to each region on the nominal wage, mainly those that we omit in our model. The regional heterogeneity is modeled by the KKP model as random effects. Besides, with the analysis of the error we try to identify the effect of the possible spillover that can happen between the regions throughout the period analyzed.

Based on the generalizations of the Generalized Moments Method (GMM) proposed by Kelejian and Prucha (1999), the KKP model is a panel data model involving a spatially first order auto-regressive disturbance term, whose innovations have an error component structure. The model's specification allows potentially auto-correlated disturbances in space and time, as well as heteroscedasticity.

The modeling proposed by Kapoor et al. (2007) considers a linear regression of panel data that allows for disturbance correlation throughout space and time. The authors assume that in each period of time *t* the data is generated in accordance with the following model:

$$y_N(t) = X_N(t)\beta + u_N(t) \tag{7}$$

where *N* indicates the locality, $y_N(t)$ is a N x 1 vector of observations of the dependent variable in time *t*, $X_N(t)$ is a N x K matrix of regressors that can contain the constant term, β is the K x 1 vector correspondent to the parameters of the regression and $u_N(t)$ denotes the N x 1 vector of the disturbances generated by a random error process.

To model the spatial dependence of the disturbances, we consider the spatial first order auto-regressive process for each period of time:

$$u_N(t) = \rho W_N u_N(t) + \varepsilon_N(t) \tag{8}$$

where W_N is a N x N matrix of constant weight independent of *t*, ρ is a scalar auto-regressive parameter and $\varepsilon_N(t)$ is a N x 1 vector of innovation in the period *t*.

Stacking the observations in (7) and (8), we have:

$$\begin{aligned} y_N &= X_N\beta + u_N \\ u_N &= \rho(I_T \otimes W_N)u_N + \varepsilon_N \end{aligned} \tag{9}$$

where I_T is a T x T identity matrix.

To allow for the innovations to be correlated over time, we assume the following error component structure for the innovation vector ε_N:

$$\varepsilon_N = (e_T \otimes I_N)\mu_N + v_N \tag{10}$$

where μ_N represents the vector of unit specific error components of each locality and v_N contains the error components that vary in space and time.

In this way, the innovations are correlated in time, but not in space. However, as presented in (9), the disturbance of any locality is affect by the weighted disturbances of its neighbors. For that reason, even the innovations, i.e. the spatial heterogeneities, spillover. We consider that this approach is more suitable to our analysis of the Brazilian municipalities because the interactions at this level are very high.

Solving the disturbance vector in terms of the innovation vector, results in:

$$\begin{aligned} u_N &= [I_T \otimes (I_N - \rho W_N)^{-1}]\varepsilon_N \\ y_N &= X_N \beta + [I_T \otimes (I_N - \rho W_N)^{-1}]\varepsilon_N \end{aligned} \tag{11}$$

Then, in the case of $\rho > 0$, the disturbances are correlated in space and time:

$$\mathrm{E}[u_N(t)u'_N(t)] = (\sigma_\mu^2 + \sigma_v^2)(I_N - \rho W_N)^{-1}(I_N - \rho W'_N)^{-1} \tag{12}$$

Kapoor et al. (2007) emphasize that, if the data are restricted to only one period of time ($T = 1$), the specification of its model is reduced to the traditional first order spatial auto-regressive Cliff-Ord model (AR(1)).

The estimation of the NEG wage equation requires some special econometrical considerations. The simplified hypotheses that we adopt for the market potential calculation, such as equivalent price indices among the cities and constant transport costs, induce to errors of measure in this variable, as in reality the price index and transportation fees can have regional variations. Besides, the market potential dependence on wages implies endogeneity. This requires a compatible methodology, as the method of Kapoor et al. (2007) considers exogenous regressors. Therefore, following Fingleton (2008b), some alterations have been carried out in the methodology.

The process involves three stages. In the first, considered here as Estimation 1, we used the instrumental variables model to estimate the residuals from equation (6). In the second, those residuals were used to estimate, through a non-linear optimization routine, a moments equation that gave us estimates for the parameters ρ, σ_v^2 e σ_1^2, and hence to the covariance matrix Ω_ξ:

$$\hat{\Omega}_\xi = \mathrm{E}(\xi\xi') = \hat{\sigma}_\mu^2 (J_T \otimes I_N) + \hat{\sigma}_v^2 I_{TN} = \hat{\sigma}_v^2 Q_0 + \hat{\sigma}_1^2 Q_1 \tag{13}$$

where $\sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2$, J_T is a T x T unity matrix and Q_0 and Q_1 are standard transformation matrices, symmetrical, idempotent and orthogonal between themselves.

The third stage uses the estimated values of ρ , σ_v^2 and σ_1^2 , calculated here by the estimate with complete weights, not only partial weights as Kapoor et al. (2007) alternatively present. With another instrumental variables estimation we can finally reach the estimated values of the parameters and their standard deviations. In this stage, the data is transformed via a Cochrane-Orcutt type of procedure in order to consider the spatial dependence of the residuals.

As our model presents heteroscedasticity and correlated errors, we cannot follow the standard assumption of a spherical errors structure. Therefore, we adopted the estimation of an instrumental variables model with non-spherical disturbances (Bowden and Turkington, 1990). In both the first and third stages, a set of linearly independent exogenous variables were used as instruments. Considering Z as the matrix of instruments, we have:

$$P_z = Z(Z'\hat{\Omega}Z)^{-1}Z'$$

Thus:

$$\hat{b}^* = (X^{*\prime}P_z X^*)^{-1} X^{*\prime}P_z Y^* \tag{14}$$

The estimated variance-covariance matrix of the parameter is given by:

$$\hat{C} = (X^{*\prime}P_z X^*)^{-1} \tag{15}$$

In this way, the square root of the constant values in the main diagonal line of the variance-covariance matrix is equivalent to the standard errors of the estimated parameters. However, this methodology does not provide the standard error of $\hat{\rho}$, the statistical significance of which can be tested by Bootstrap methods.

It is important to emphasize that, as in stage 1 we assume that $\rho = 0$, in this case, we have $Y^* = Y$ and $X^* = X$. Besides, we also assume that $\sigma_v^2 = 1$ and $\sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2 = 1$, then, in stage 1, the estimation with non-spherical disturbances corresponds to the estimation by standard instrumental variables.

Following (Fingleton, 2008a), we use the absolute latitude of each municipality's centroid (*L*) and its square (*L²*) as instruments for the endogenous variable *ln P*.

In order to estimate the model we used data from the three last Demographic Censuses in Brazil collected by the Brazilian Census Bureau (IBGE), every ten years. The use of the census data allows us to construct a database at the municipal level, the municipal meshes being made compatible according to Chein et al. (2007). By matching, 3,951 territorial units are reached by the analysis. Therefore, where one reads the term municipality in this work, it must be understand as a geographic area resulting from the matching of the municipal areas for the period 1970-2000, whose denomination was given according to cities with the largest population in 2000 pertaining to the resultant geographic area.

The variables used are the average years of schooling of the population of 25 or more in the city (S_1), the literacy ratio of the population 25 or more (S_2), the average income from the main occupation of the workers with at least the age of 12 (w), and the sum of the income from all sources of the population of 12 years of age or older (Y), all the values being priced in accordance with the deflators proposed by Corseuil and Foguel (2002) specifically for the Brazilian census data.

The market potential (P) is equivalent to the consideration of the external and internal markets conjointly. In order to calculate the external market, we used the total of the income from all the sources of the population of 12 or older divided by the distance to the reference locality, added for all the 3951 municipalities. The distance (d_{ij}) was measured in accordance with the great circle distance between the municipalities' centers. The domestic market is the addition of the incomes from all the sources of the population of 12 or older in the reference locality. Therefore, the potential market is defined as:

$$P_i = Y_i + \sum \frac{Y_j}{d_{ij}}$$

where i and j represent different municipalities.

The definition of the weight matrix W was carried out following Fingleton (2008a), substituting the distance limit of 1000 miles used by the author in his international analysis by 100 miles, which is more suited to our regional analysis and more related to daily commuting. Then, we have:

$$W_{ij} = \left(1 - \frac{d_{ij}}{100}\right)^2 \text{ if } d_{ij} \leq 100$$

$$W_{ij} = 0 \text{ if } d_{ij} > 100 \text{ or } i = j$$

where i and j represent municipalities.

The limitation of neighborhood to the maximum of 100 miles resulted in 21 municipalities without neighbors. Those municipalities had been excluded from the sample[1], resulting in 3,930 analyzed municipalities.

[1] The municipalities that have been excluded from the sample are: Altamira, Aripuanã, Atalaia do Norte, Barcelos, Barra do Garças, Boa Vista, Caracaraí, Carauari, Chapada dos Guimarães, Fernando de Noronha, Itaituba, Japurá, Lábrea, Luciara, Nobres, Santa Isabel do Rio Negro (Ilha Grande), Santo Antônio do Içá, São Félix do Xingu, São Gabriel da Cachoeira, Tapauá e Vila Bela da Santíssima Trindade. Those are municipalities that, with the exception of the islands, had had their areas roughly changed by the geographical compatibility. The resultant area would have great heterogeneity, which justifies its elimination from the sample. Besides, most of those municipalities are part of the Amazon Rainforest, where the main transportation routes are rivers and the average municipalities' area is huge. An example is Altamira, the world's biggest municipality, with an area of 161,445.9 Km², which is equivalent to the sum of the areas of Holland, Belgium and Portugal. Since they are concentrated in the Northwest region of Brazil, their removal has minor impact on the spatial structure of the other municipalities.

EXPLORATORY ANALYSIS

The evolution of the social and economic conditions of the municipalities analyzed in the period 1980-2000 was sufficiently significant. The wage value and the schooling level had risen by around 100%. The potential market and the literate population ratio also had increased significantly in the period, although on a lesser scale.

TABLE 1
Municipal averages per year

	1980	*1991*	*2000*
Wage	202.17	316.50	416.00
Market potential*	96.96	107.42	160.67
Years of schooling	2.11	3.17	4.17
% of literate population	56.07	64.33	73.76

* in millions of reais based on prices of 2002.
Source: Prepared by the authors based on data of the Demographic Censuses, IBGE.

TABLE 2
Municipal averages by region, year 2000

	Centre-West	*Northeast*	*North*	*Southeast*	*South*
N. of municipalities	248	1374	181	1410	717
Wage	499.51	253.48	382.16	499.83	542.22
Market potential*	128.75	87.83	68.46	249.40	160.10
Years of schooling	4.56	3.03	3.66	4.82	5.05
% of literate population	79.46	57.45	70.07	82.42	86.92

* in millions of reais based on prices of 2002.
Source: Prepared by the authors based on data of the Demographic Censuses, IBGE.

As shown in Table 2, the regional differentials in all of the variables were large: the average wage in the year 2000 in the Southern region surpassed the Northeast region by 114% and Southeastern market potential was 3.64 times larger than the North of Brazil.

In order to further investigate the spatial distribution of the considered variables we present in Illustrations 1 to 4 their Local Indicator of Spatial Association i.e. the local Moran's I at 5% of significance level (Anselin, 1995). Illustration 1 shows a clear imbalance in the spatial distribution of the average income in Brazil. While the Centre-South region presents a spatial pattern of high wages, the Northeastern municipalities represent a cluster of low wages. In relation to the spatial distribution of market potential (Illustration 2), São Paulo and Rio de Janeiro concentrated 46% of all income in Brazil in the year 2000. Besides the Rio-São Paulo region, there are some High-High clusters in the main metropolitan areas. The capitals of North and Northeast regions appear as High-Low outliers. All 16 states from the North and Northeast regions sum only 19.3% of the income in the country.

Illustrations 3 and 4 show the spatial association of the average years of schooling and the literacy ratio, respectively. As was to be expected, the spatial pattern of the education measures is much the same as the wage. Once again the Northeast region presents a spatial cluster of municipalities with low index values. On the other hand, the South and Southeast regions are clusters of high schooling indicators. As stated in Table 2, the Northeastern municipalities present an average schooling of 3.03 years and only 57.45% of the population is literate, while those in the South have on average 5.05 years of schooling and 86.92% of the population is literate.

ILLUSTRATION 1

Local Moran's I of the average income from the main occupation, year 2000

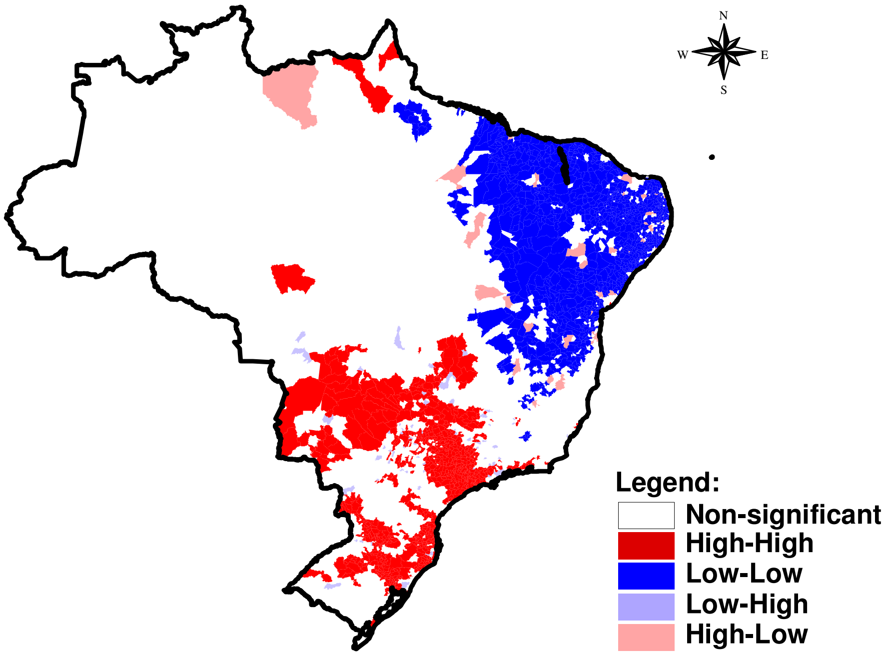


ILLUSTRATION 2

Local Moran's I of the market potential, year 2000

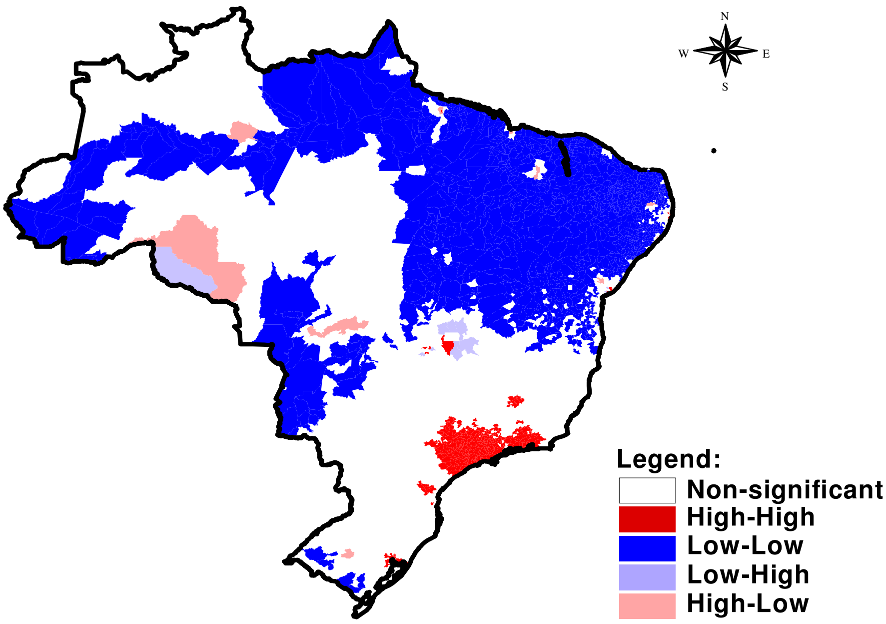


ILLUSTRATION 3

Local Moran's I of the average schooling years, year 2000

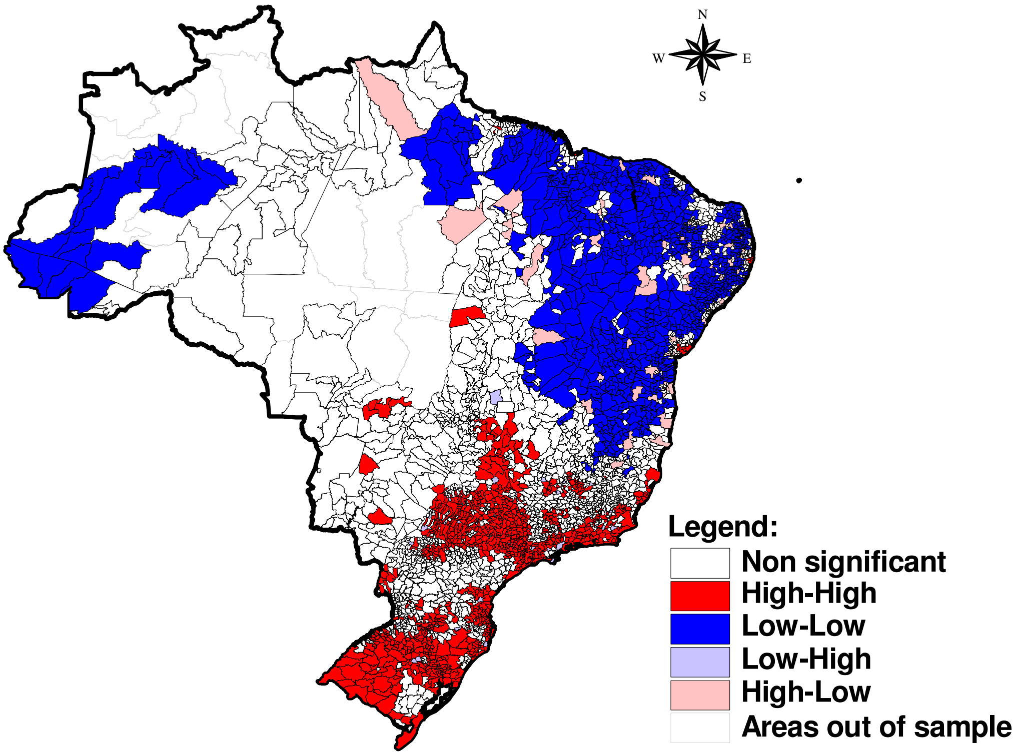
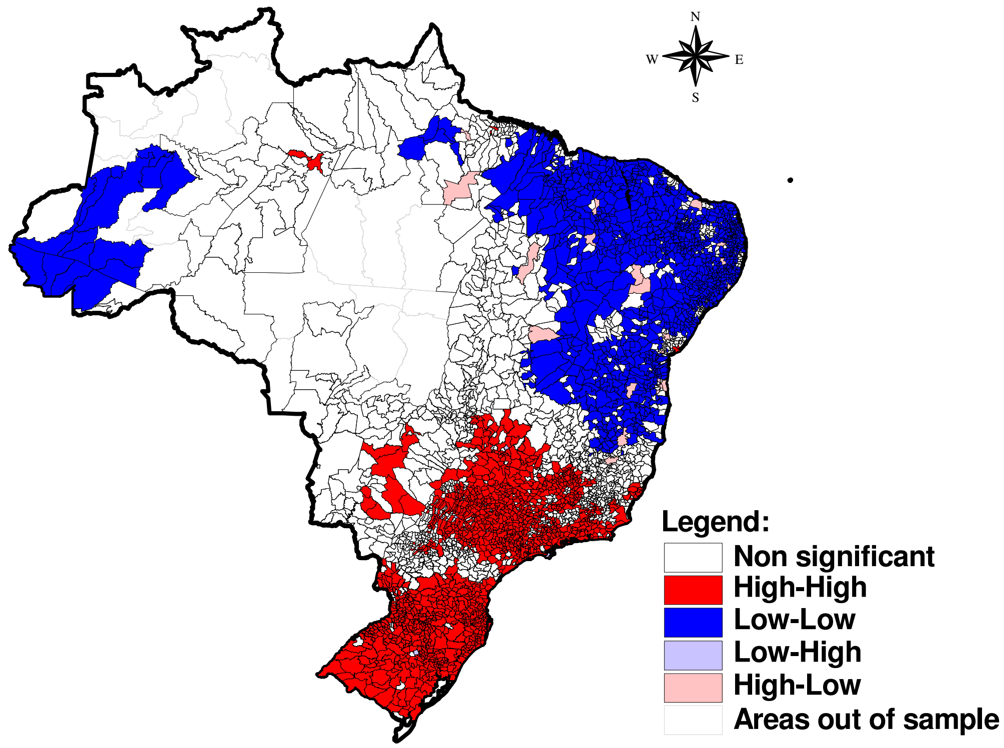


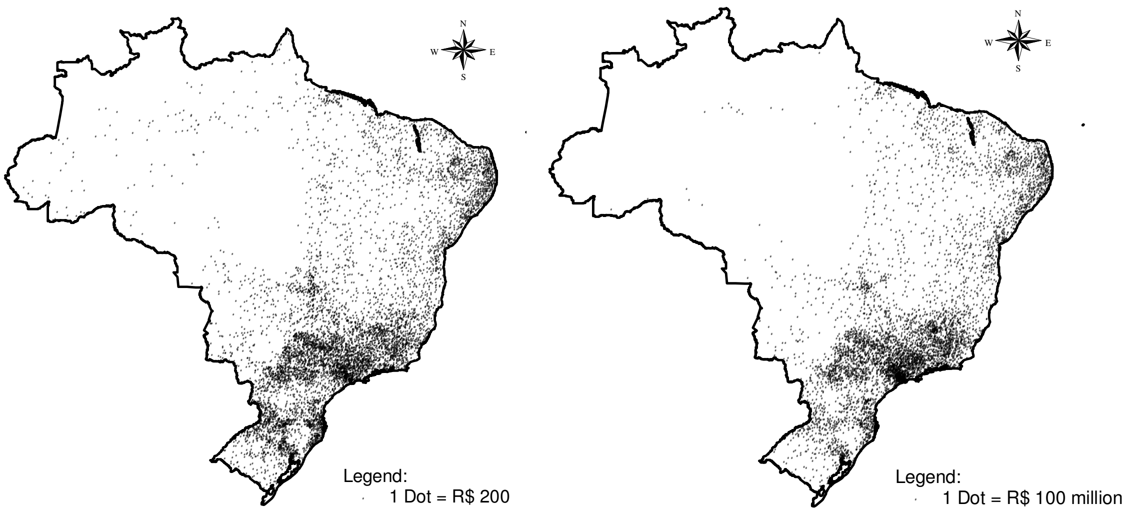
ILLUSTRATION 4

Local Moran's I of the literacy ratio, year 2000



To further explore the relation between the market potential and wages, Illustration 5 presents the density of the wage spatial distribution and the market potential in Brazil in the year 2000. As we can observe, there is a big concentration of income and market in the South and Southeast regions of Brazil. However, the concentration of market potential is still bigger, as São Paulo and Rio de Janeiro have the highest mass of incomes and, in consequence, of consumption. The North of the country seems to have relatively higher wages in relation to its market potential, such evidence we will confirm by the spatial analysis of the residues of Estimation 1.

ILLUSTRATION 5
Income density of the occupations and potential market, year 2000



The spatial feature of the variables included in our estimation model indicates the necessity of spatial modeling. Therefore, the motivation to adopt a model of panel data with spatially correlated error components arose from the spatial pattern presented by our variables. Such spatial dependence violates hypotheses of the ordinary least squares method, raising the variance of the estimated parameters and harming its efficiency.

ESTIMATES AND INFERENCES

The results presented in Table 3 show the parameters estimation in each model. Estimation 1 is based on 2-SLS and suggests that the reduced model of the New Economic Geography's wage equation provides a good explanation for the regional differentials within wages in the Brazilian municipalities in the period 1980-2000. The signs of the parameters are all positive, as expected, so that the higher the market potential of a given locality, the higher is the wage. According to NEG this is a consequence of the lower costs provided by the spatial concentration of the productive activity favoured by the proximity to the consuming market. The controlling variables, average schooling and literacy ratio, also presented appropriate signs:

the higher the schooling level and literacy ratio in the city, the higher its average wage. All but literacy ratio have high statistical significance.

The results of Estimation 2, considering the random effects and the spatial and time dependence of the error term, indicate that the relevance of the market potential in explaining the average nominal wage is even higher. Except for the literacy ratio, all estimated parameters have high statistical significance. However, the parameter's signal of the literacy ratio was inverted. This fact is further analyzed in Estimation 3 – presented in Table 4 – and can be credited to the high correlation between the literacy ratio and the average years of schooling, whose value is 0.91. This multicollinearity influences the estimation of these parameters.

The estimated market potential parameter is smaller than the one found by Fingleton (2008a) – 0.4546. This fact suggests that the relation between market potential and wages is higher when international terms are considered rather than regional ones. This can be explained by the fact that wages are more rigid on sub-national scale. According to Estimation 2, an increase of 1% in the market potential would be related to a wage improvement of 0.35%.

TABLE 3
Results of the estimations over the dependent variable *ln w*

	Estimation 1	**Estimation 2**
Intercept	-0.6556 (0.2191)	-1.4793 (1.0551)
Ln P	0.3037 (0.0119)	0.3498 (0.0586)
Ln S1	0.6595 (0.0129)	0.6162 (0.0291)
Ln S2	0.0483 (0.0214)	-0.0711 (0.0327)
ρ	-	0.8570
σv	-	0.0329
σ1	-	0.0766
R^2	0.7562	0.7180
$\hat{\sigma}$ *	3.2927	2.8588
Instruments	LnS1, LnS2, L, L^2.	LnS1, LnS2, L, L^2.

Source: Prepared by the authors based on data of the Demographic Censuses, IBGE.

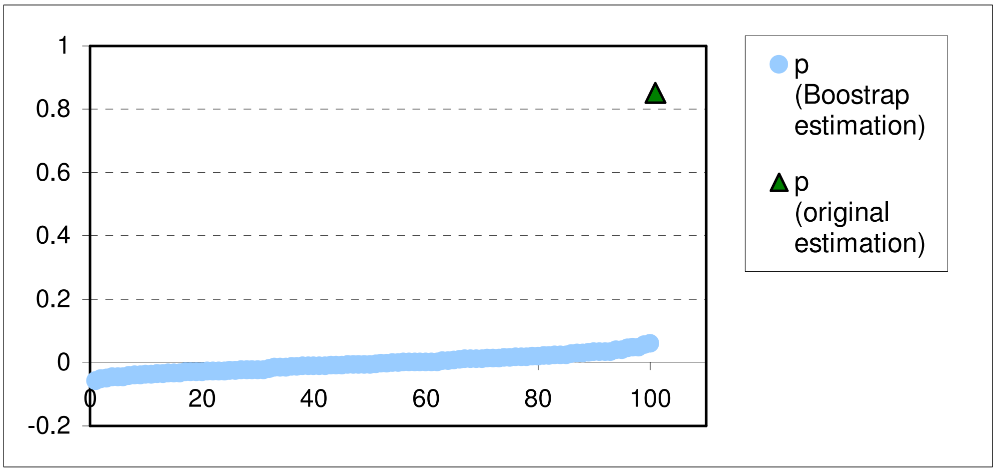
The estimated parameter for the market potential is nearly half that for average schooling. This result suggests that the market potential, a regional attribute, has a powerful influence over the nominal wage although the average schooling, an individual attribute, remains as the most important factor in the wage determination. The estimated value for the ratio $\sigma/(\sigma-1)$ is 1.53. According to Krugman (1991), this value is equal to the ratio of the marginal product of labor to its average product, i.e., the degree of economies of scale. Therefore, an estimated value above unity denotes the presence of economies of scale in the Brazilian economy. The relatively low value of the parameter of preference ($\hat{\sigma} = 2.86$) indicates that the varieties of the manufactured goods are considered as differentiated goods. Therefore, the estimated parameter satisfies to conditions for an acceptable estimation. The parameter is higher than 1, consistent with the implied elasticity of substitution, and is close to 2, since we have assumed $\hat{\sigma} = 2$ in our definition of the market potential transportation costs.

Besides, the estimated value of ρ found by us is much higher to that found by Fingleton (2008a): 0.0772. This difference can be attributed to the higher spatial influence among municipalities than among countries, as geographical, cultural, institutional and economic particularities are hardly kept contained inside such restricted borders as those of a municipality. As stated, the statistical significance of ρ can be checked by Bootstrap. As shown in Chart 1, bootstrapping the NEG wage equation residuals 100 times, keeping all other variables and the weight matrix constant, results in ρ values that go from -0.058 to 0.060. To randomly extract sub-samples from the wage equation residuals means that the residuals' spatial correlation should be broken. Therefore, the Bootstrap estimated ρ was expected to be close to zero, as it was. As our original estimation of ρ is 0.8570, we can safely reject the null hypothesis of $\rho = 0$. Such a high spillover parameter indicates that the regional context in which a locality is inserted has major importance on its wage level determination.

To test the robustness of our estimations and further investigate the relationship between market potential and nominal wage in Brazil we present in Table 3 the estimation of two additional models. The first, Estimation 3, is related to the correlation between the years of schooling and the literacy ratio. As can be seen, if the literacy ratio of the population over 25 years old is removed from the model, the other coefficients estimates do not suffer much change. The difference of Estimation 2 and 3 from the estimated coefficient for the market potential, which is an inverted measure of the elasticity of substitution, is no bigger than 0.0166.

CHART 1

Original p estimation from Regression 3 and p Bootstrap distribution as a result of re-sampling the residuals u



Estimation 4 checks the pertinence of one of Fingleton's conclusions to the Brazilian municipalities Fingleton (2008a) states that "basing market potential purely on foreign income levels and trade costs does not seem to be a viable option at the international level". To check if it is a viable option at the regional level, we consider only the effects of the external market, disregarding the internal purchase power. The result indicates that the external market gives reasonable estimations for elasticity of substitution: $\hat{\sigma} = 3.55$. The fact that the estimated elasticity of substitution is higher than 2 suggests that, when considering only the external market, we should have considered more distance friction in our definition of the market potential.

Within the framework of a municipality, it is reasonable to assume that interregional commerce is facilitated by low transportation costs, the absence of exchange rate and tariffs. Therefore, the interregional economic interaction has relevance on its own, mainly when considering the closest neighbors. These findings diverge from Fingleton (2008a). Of course, as Rosenstein-Rodan (1943), Nurkse (1953) and Furtado (1976) show, the internal market is decisive for the economic growth of a country. But that does not necessarily apply to municipalities. At the municipality level, the internal market plays a role less important than the external.

TABLE 4

Results of the estimations over the dependent variable *ln w*

	Estimation 3	**Estimation 4**
Intercept	-1.1126 (1.0665)	-0.2288 (0.8986)
Ln P	0.3332 (0.0593)	-
Ln S1	0.5902 (0.0267)	0.6667 (0.0156)
External Market	-	0.2820 (0.0497)
ρ	0.8590	0.8550
σv	0.0329	0.0338
$\sigma 1$	0.0763	0.0811
R^2	0.7189	0.7207
$\hat{\sigma}$ *	3.0012	3.5461
Instruments	LnS1, L, L^2.	LnS1, L, L^2.

Source: Prepared by the authors based on data of the Demographic Censuses, IBGE.

Moreover, Estimations 2, 3 and 4 indicate the robustness of our spillover parameter. In the three models it was kept between 0.8550 and 0.8590. Hence, not only the spatial friction represented by transportation costs is important when estimating regional wage models, but also the spatial localization and the neighbourhood effects.

CONCLUSION

The theoretical and empirical approach developed by New Economic Geography, considering economic concentration as a consequence of the increasing returns propitiated by the agglomeration economies and transport costs in a context of monopolistic competition, as modeled by Dixit and Stiglitz (1977), reinforced the analysis of the regional inequalities. Our results suggest that the market potential of one determined locality, and, in consequence, productive concentration has a strong relationship with nominal wage, confirming the results found by works such as Harris (1954), Hanson (1998), Brakman et al. (2004), Mion (2004), Niebuhr (2006) and Fingleton (2008a).

The choice of the methodology used here in order to make the work comparable to Fingleton (2008a) allowed us to conclude that the results found for the regional economy of Brazil does not only suggest a higher spatial dependence but also a higher relationship between the market potential and the nominal wages among the Brazilian regions.

The contribution of this work is mainly the use of the technique of panel data models with spatially correlated error components in the estimation of a reduced version of the New Economic Geography's wage equation.

Therefore this work opens for us a field of deepened research on the Brazilian situation in the context of the NEG, such as:

- Inquiries concerning the relationship between the external and internal markets in the determination of the market potential and its effects on the spatial distribution of the productive activities and on the wages;
- The search for consistent estimates using Brazilian data for other parameters adopted by the NEG;
- The comparison of the efficiency of the NEG and other theories in regional understanding and explanation of the current Brazilian situation and its disparities

All these are gaps that still need to be opened up for research and that can help to better understand, and even to participate actively in the regional dynamics of Brazil.

ACKNOWLEDGEMENTS

The authors are grateful to Danilo Igliori and Sueli Moro for their contributions to the essay of graduate conclusion from which this paper was derived and to Nancy Lozano and the other participants at the II World Conference of the Spatial Econometrics Association, New York, 2008. In addition, the authors acknowledge and gratefully thank the grant of a Travel Award by the Regional Studies Association for the presentation of a previous version of this paper at the mentioned conference.

REFERENCES

Anselin, L., 1995. Local Indicators of Spatial Association - LISA. Geographical Analysis 27(2), 93-115.

Anselin, L., Kelejian, H., 1997. Testing for spatial error autocorrelation in the presence of endogenous regressors. International Regional Science Review 20, 153-182.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277-297.

Batista da Silva, M., Silveira Neto, R., 2005. Determinantes da Localização Industrial no Brasil e Geografia Econômica: evidências para o período pós-Real. Proceedings of XXXIII Encontro Nacional de Economia. Natal, RN: ANPEC.

Bowden, R., Turkington, D., 1990. Instrumental Variables. Cambridge: Cambridge University Press.

Brakman, S., Garretsen, H., Schramm, M., 2004. The Spatial Distribution of Wages and Employment: estimating the Helpman-Hanson model for Germany. Journal of Regional Science 44(3), pp.437-466.

Chein, F., Lemos, M., Assunção, J., 2007. Desenvolvimento desigual: evidências para o Brasil. Revista Brasileira de Economia 61(3), 301-330.

Cliff, A., Ord, J., 1973. Spatial autocorrelation. London: Pion.

Cliff, A., Ord, J., 1981. Spatial processes: models and applications. London: Pion.

Corseuil, C., Foguel, M., 2002. Uma sugestão de deflatores para rendas obtidas a partir de algumas pesquisas domiciliares do IBGE. Texto para discussão do IPEA 897.

Dixit, A., Stiglitz, J., 1977. Monopolistic Competition and Optimum Product Diversity. American Economic Review 67(3) 297-308.

Figueiredo, L., 2002. The New Economic Geography and Regional Growth in Brazil and India. Ph.D. Thesis. Nottingham: University of Nottingham.

Fingleton, B., 2008a. Competing models of global dynamics: evidence from panel models with spatially correlated error components. Economic Modelling 25(3), pages 542-558.

Fingleton, B., 2008b. A Generalized Method of Moments Estimator for a Spatial Panel Model with an Endogenous Spatial Lag and Spatial Moving Average Errors. Spatial Economic Analysis 3(1) 27-44.

Fontes, G., Simões, R., Hermeto, A., 2009. Urban attributes and wage disparities in Brazil: a multilevel hierarchical model. Regional Studies (forthcoming).

Fujita, M., Krugman, P., Venables A., 1999. The spatial economy: cities, regions, and international trade. Massachusetts: MIT Press.

Fujita, M., Mori, T., 2005. Frontiers of the New Economic Geography. Institute of Developing Economies Discussion Paper 27.

Furtado, C., 1976. A economia latino-americana: formação histórica e problemas contemporâneos. São Paulo: Cia. Editora Nacional.

Galinari, R., Lemos, M., Amaral, P., 2006. Retornos crescentes urbanos: a influência do espaço na diferenciação da taxa salarial no Brasil. In: De Negri, J., De Negri, F., Coelho, D. (Eds.) Tecnologia, exportações e emprego. Brasília: IPEA.

Hanson, G., 1998. Market Potential, Increasing Returns, and Geographic Concentration. Journal of International Economics 67(1), pages 1-24.

Harris, C., 1954. The Market as a Factor in the Localization of Industry in the United States. Annals of the Association of American Geographers 44 315-348.

Helpman, E., 1998. The Size of Regions. In: Pines, D., Sadka, E., Zilcha, I. (Eds.) Topics in Public Economics. Cambridge: Cambridge University Press.

Kapoor, M., Kelejian, H., Prucha, I., 2007. Panel data models with spatially correlated error components. Journal of Econometrics 140(1) 97-130.

Kelejian, H. Prucha, I., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review 40 509-533.

Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99(3) 483-499.

Lin, S., 2003. International Trade, Location and Wage Inequality in China. UNU/WIDER Discussion Paper 2003/61.

Lösch, A., 1954. The economics of location. New Haven: Yale University Press.

Marshall, A., 1890. Principles of economics. London: Macmillan.

Mion, G., 2004. Spatial Externalities and Empirical Analysis: the case of Italy. Journal of Urban Economics 56(1), pages 97-118.

Myrdal, G., 1957. Economic Theory and Underdeveloped Regions. London: Duckworth.

Niebuhr, A., 2006. Market Access and Regional Disparities. The Annals of Regional Science 40(2), pages 313-334.

Nurkse, R., 1953. Problems of capital formation in underdeveloped countries. Oxford: Basil Blackwell.

Paillacar, R., 2007. Market Potential and Worker Heterogeneity as Determinants of Brazilian Wages (mimeo). Available at: http://team.univ-paris1.fr/teamperso/paillacar/

Reis, J., Barros, R., 1990. Desigualdade salarial e distribuição de educação: a evolução das diferenças regionais no Brasil. Pesquisa e Planejamento Econômico 20(3) 415-478.

Rosenstein-Rodan, P., 1971. Problems of industrialization of Eastern and Southeastern Europe, Economic Journal 53 204-207.

Savedoff, W., 1990. Os diferenciais regionais de salário no Brasil: segmentação versus dinamismo da demanda. Pesquisa e Planejamento Econômico 20(3) 521-556.

Servo, L., 1999. Diferenças de salários no Brasil: uma análise para as regiões metropolitanas. Proceedings of XXVII Encontro Nacional de Economia. Salvador, BA: ANPEC.

Silveira-Neto, R., Campelo, A., 2003. O perfil das disparidades regionais de renda no Brasil: evidências a partir de regressões quantílicas para os anos de 1992 e 2001. Proceedings of XXXI Encontro Nacional de Economia. Porto Seguro, BA: ANPEC.

Souza, C., 2007. A Nova Geografia Econômica: três ensaios para o Brasil. Doctorate Thesis. Belo Horizonte: Universidade Federal de Minas Gerais.