TEXTO PARA DISCUSSÃO Nº 1111

ESTIMATION OF MULTIEQUATION CROSS-SECTION MODELS IN THE PRESENCE OF SPATIAL AUTOCORRELATION

Alexandre Carvalho
Daniel da Mata
Kenneth M. Chomitz

Brasília, agosto de 2005

TEXTO PARA DISCUSSÃO Nº 1111

ESTIMATION OF MULTIEQUATION CROSS-SECTION MODELS IN THE PRESENCE OF SPATIAL AUTOCORRELATION*

Alexandre Carvalho**
Daniel da Mata**
Kenneth M. Chomitz***

Brasília, agosto de 2005

O presente texto, em língua inglesa, não foi objeto de revisão editorial.

* Agradecimentos/Acknowledgments
This paper is part of a joint research program between the World Bank and the Instituto de Pesquisa Econômica Aplicada (Ipea), Brasília. We are grateful to Jonny Andersson, Piet Buys, George da Guia, Waldir Netto, Edinaldo Tebaldi and Pedro Albuquerque, for assistance, advice, data construction, and literature reviews. We have benefited from discussions with José Raimundo Carvalho, Somik Lall and Mansueto Facundo. All errors are the authors'. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of Ipea or the views of the World Bank, its Executive Board of Directors, or the countries they represent.
** Directorate of Regional and Urban Studies, Institute of Applied Economic Research (Ipea)/Diretoria de Estudos Regionais e Urbanos (Dirur), Instituto de Pesquisa Econômica Aplicada.
*** Development Research Group, World Bank.

ipea Instituto de Pesquisa Econômica Aplicada

Fundação pública vinculada ao Ministério do Planejamento, Orçamento e Gestão, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

URL: http://www.ipea.gov.br

Ouvidoria: http://www.ipea.gov.br/ouvidoria

ISSN 1415-4765

JEL E24, O18, R11, R12, R58

TEXTO PARA DISCUSSÃO

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

SUMÁRIO

SINOPSE

Neste artigo, nós descrevemos técnicas econométricas para tratar autocorrelação espacial em modelos multiequacionais, com dados em *cross-section*. Os procedimentos abordados aqui se baseiam no método de momentos generalizados espacial (GMM espacial) proposto em Conley (1999). Uma extensão para estimação com variáveis instrumentais com informação plena é apresentada. Nós empregamos simulações de Monte Carlo para verificar as propriedades assintóticas dos estimadores descritos. As simulações sugerem que, mesmo na presença de heterogeneidade espacial, o GMM espacial apresenta erros padrões apropriados. Além disso, estatísticas *t* usuais parecem seguir a distribuição normal padronizada. Finalmente, nós apresentamos uma aplicação, em que são estimadas equações de salário para estudar crescimento e desenvolvimento regional nos municípios brasileiros, entre 1991 e 2000.

ABSTRACT

We describe econometric techniques to treat spatial autocorrelation in multiequation cross-section models. The cross-section approaches discussed here are heavily based on the spatial GMM procedure, proposed by Conley (1999). An extension for full-information instrumental variable models is presented. Monte Carlo simulations are employed in order to verify some asymptotic properties of the Spatial GMM approach. The simulations suggest that, even in the presence of spatial nonstationarity, the spatial GMM still delivers valid standard errors. Besides, usual *t*-statistics appear to have a standard normal distribution. An application for estimating labor and wage equations to study regional growth and development of the Brazilian municipalities, between 1991 and 2000, is presented.

1 Introduction

Studies in regional development are increasingly oriented towards understanding social, environmental, and economic outcomes at a fine spatial scale. This trend is facilitated by an explosion in the availability of spatially referenced data, and a revolution in the ease of handling this data through cheaper and more user-friendly geographic information systems. As a result, analyses formerly undertaken at the state level can now be pursued at the level of the county or even census tract. The finer spatial data resolution allows for more cross-variation in explanatory variables and permits greater attention to spatial processes.

Finer spatial units of analysis and more explicit treatment of spatial processes bring with them a need to allow for the spatial dependence in econometric analysis. Several practical difficulties have impeded the widespread application of standard spatial econometric techiques (see, e.g. Anselin 1988) to fine-scale regional development studies. First, at least until Kelejian and Prucha (2004) described spatial two and three stage generalized least squares procedures, it has been difficult to apply spatial methods to models involving instrumental variables or multiple equations. Second, standard spatial econometric techniques are sensitive to misspecification of the spatial weights matrix (Bell and Bockstael 2000). This is a serious problem, since the weights matrix is in general unobservable and unknown. Third, it may be possible that the spatial dependence is heterogeneous accross the region of study. In fact, Silva e Resende (2005) study municipality growth in two different states in Brazil, and find a strong spatial autocorrelation in one of them, and a nonsignificant autocorrelation in the other. Fourth and perhaps most important, application of standard spatial techniques (including those of Kelejian and Prucha) becomes computationally challenging when applied to county-level data sets with thousands of observations, because of the need for inversion of $n \times n$ matrices, where n is the number of observations.

Against this backdrop, the spatial GMM technique of Conley (1999) is an attractive alternative to more standard techniques. The idea behind Conley's GMM is to proposed a consistent estimator for the GMM moment conditions covariance matrix. His spatial dependence consistent estimator follows the same idea as the Newey and West (1987) heteroskedasticity and autocorrelation consistent covariance matrix estimator, and therefore is also positive semi-definite. Because the matrix estimator is calculated by a simple summation of cross-products for neighboring vectors, its computation avoids high-dimensional matrices manipulations, in contrast to estimation techniques based on spatial contiguity matrices. Furthermore, since the spatial GMM framework does not assume any parametric form, it is robust to model misspecification, such as the misspecification of weight matrices.

The goal of this paper is to enhance the usefulness of Conley's estimator in two ways. First, we use Monte Carlo methods to test the validity of the asymptotic approximations to the standard test statistics, and to compare the performance of Conley's GMM estimator to other estimators. In the simulations, we introduce an explicit spatial heterogeneity in the data generating process. The results suggest the validity of the spatial GMM estimator even in the presence of the spatial dependence heterogeneity. Second, we show how Conley's technique can be employed for estimation of multiple-equation, full information models. We illustrate the extended procedure through the estimation of a simultaneous equation describing spatial labour supply and demand in Brazil.

Section 2 describes the simultaneous equation model estimation problem in the presence of spatial dependence. Section 3 discusses the spatial GMM estimation procedure, proposed by Conley (1999). We introduce Conley's general framework and discuss its application to several estimation situations. In Section 4, we present a simulation experiment to investigate the validity of the asymptotic approximations for common test statistics. Section 5 presents an empirical application, where we estimate a system of labor supply and demand equations for Brazilian municipalities over the period 1991-2000[1]. Section 6 concludes the paper.

2 Problem description

Consider a system of K linear regression models, which are part of a possibly more general system of equations. We can then represent the estimated models in the form

$$y_{k,i} = \mathbf{x}'_{k,i}\beta_k + \varepsilon_{k,i}, \tag{1}$$

where $k \in \{1, 2, \ldots, K\}$ indexes the equations and $i \in \{1, 2, \ldots, N\}$ indexes the observational cross-section units (municipalities, for example). The vector $\mathbf{x}_{k,i}$ contains the endogenous and exogenous right-hand-side variables. The coefficient vector β_k is equation specific and does not necessarily have the same dimension in all K regressions. The unobservable components $\varepsilon_{k,i}$ have mean zero and are not necessarily are homoskedastic and uncorrelated. In fact, given the spatial structure of the data, it is expected that $\varepsilon_{k,i}$ is correlated with $\varepsilon_{k,j}$, when municipalities i and j are proximate. It is expected that the residuals

[1]Municipalities are the third level in Brazilian government units. They are equivalent to counties in the United States.

across equations $\varepsilon_{k_1,i}$ and $\varepsilon_{k_2,i}$ are also correlated. The presence of correlation between $\varepsilon_{k_1,i}$ and $\varepsilon_{k_2,i}$, for $k_1 \neq k_2$ suggests the usage of limited or full information estimation techniques.

Section 3 handles the estimation problem in the presence of endogeneity for some of the right-hand-side variables and when the equation residuals are spatially autocorrelated, by employing a generalized method of moments (GMM) estimation procedure.

3 Spatial generalized method of moments

The spatial GMM estimator has nice frequentist properties. Conley (1999) provided conditions that guarantee consistency and asymptotic normality of the proposed estimator. One of the advantages of the estimator is the robustness of the parameter estimates' covariance matrix to misspecification of the distance between observations. Therefore, it avoids the drawback faced by most spatial econometrics techniques, which are based on somehow arbitrary contiguity matrices.

We initially review the general spatial GMM framework, with focus on the estimation of the covariance matrix for the moment conditions. We then discuss how this general procedure can be employed in two particular cross-section situations. The first one corresponds to the estimation of single equation spatial models, in the presence of endogeneity, by using intrumental variables. The second situation corresponds to the simultaneous estimation of a system of spatial models, allowing for endogeneity in each one of them.

3.1 The spatial GMM estimator

We now present the estimation framework, which can be employed to estimate the parameters in the multiequation model in (1). Because of the endogeneity in some of the right-hand-side variables, we will have to use instruments to identify the parameters. In the real examples presented in Section 5, we describe the list of instruments employed in each equation. Based on the chosen instruments, we can then write the general form for the moment conditions below.

$$\mathrm{E}\{g(W_i;\delta)\} = \mathbf{0}. \tag{2}$$

These moment conditions may correspond to each equation in (1) separately or to the whole set of equations simultaneously. Estimation of each equation separately will be treated in Subsection 3.2, while estimation of all equations simultaneously will treated in Subsection

3.3. In (2), when estimating each equation separately, for example, $g(W_i;\delta) = \left[y_{k,i} - \mathbf{x}'_{k,i}\beta_k\right]\mathbf{z}_{i,k}$ is a m_k-dimensional column vector, where k is the equation of interest and m_k is the corresponding number of instruments in the column vector $\mathbf{z}_{i,k}$. The vector $W_{k,i} = [y_{k,i}\ \mathbf{x}'_{k,i}\ \mathbf{z}'_{i,k}]'$ contains all observable variables, and the unknown parameter vector is then $\delta = \beta_k$.

Based on (2), the GMM estimator $\hat{\delta}$ for δ is the argument that minimizes

$$Q(d) = \left[\frac{1}{N}\sum_{i=1}^{N} g(W_i;d)\right]' \Psi \left[\frac{1}{N}\sum_{i=1}^{N} g(W_i;d)\right], \tag{3}$$

for all $d \in \Theta$, where $\Theta \in \Re^m$ is the set of allowed values for the coefficient vector. The component Ψ is a $m \times m$-dimensional positive definite weighting matrix. If we make $\Psi = \widehat{\Omega}^{-1}$, where $\widehat{\Omega}$ is a consistent estimator for the covariance matrix Ω of the moment conditions $g(W_i;\delta)$, the GMM estimator $\hat{\delta}$ is efficient[2].

The main challenge in using the GMM approach in a spatial context is finding a consistent estimator $\widehat{\Omega}$ of Ω. Conley (1999) suggests a procedure on the same line as the Barlett window estimator used by Newey and West (1987). In a spatial context, the general formula[3] for $\widehat{\Omega}$ would be

$$\widehat{\Omega} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} K(i,j)\mathbf{Y}_i(d^*)\mathbf{Y}_j(d^*)', \tag{4}$$

where d^* is an initial estimate for δ, obtained by minimizing $Q(d)$ in (3), with Ψ replaced by an arbitrary initial positive definite matrix. We can make Ψ equal to the identity matrix $\mathbf{I}_m$, for example. We define the vector $\mathbf{Y}_i(d^*)$ as $\mathbf{Y}_i(d^*) = g(W_i;d^*)$, $i = 1,\ldots,N$. To guarantee a consistent and positive semi-definite $\widehat{\Omega}$, the weight $K(i,j)$ can be defined as

$$K(i,j) = \begin{cases} (1 - D_H(i,j)/L_H)(1 - D_V(i,j)/L_V) & \text{for } D_H(i,j) < L_H \text{ and } D_V(i,j) < L_V, \\ 0 & \text{otherwise.} \end{cases}$$

Functions $D_H(i,j)$ and $D_V(i,j)$ correspond to the horizontal and vertical distances, respectively, between cross-section units (municipalities, for example) i and j, and L_H and L_V correspond to the maximum horizontal and vertical distances, for which cross-section units

[2]For more details on GMM estimation, see Matyas (1999).

[3]To simplify the model description, the exposition in this paper is a little different from the more formal treatment given in Conley (1999).

i and j present some correlation in the moment conditions $g(W_i; \delta)$. Note that the covariance matrix Ω estimate is calculated as a summation of sample moments, where the cross-products $\mathbf{Y}_i(d^*)\mathbf{Y}_j(d^*)'$, $i \neq j$, are included only when the municipality j is located inside the rectangle, centered at municipiality i, and with dimensions $2 \times L_H$ and $2 \times L_V$. The linearly decaying weights $K(i,j)$ imply the well desired positive semi-definite property for $\widehat{\Omega}$.

The form for the weight function $K(i,j)$ raises the issue of choosing the cutoff values. In Section 5, we perform a sensitivity analysis, by choosing different cutoffs and comparing the results. In general, the parameter estimates do not seem to change much when the cutoffs vary.

After calculating $\widehat{\Omega}$, one can make $\Psi = \widehat{\Omega}^{-1}$ in (3), and obtain an efficient estimator $\hat{\delta}_{\text{GMM}}$ for δ. Under certain regularity conditions, the estimator $\hat{\delta}_{\text{GMM}}$ is consistent and has normal asymptotic distribution. The asymptotic covariance matrix can be estimated by

$$\widehat{\mathbf{C}} = N\left[\left[\sum_{i=1}^{N} \frac{\partial}{\partial \delta'} g(W_i; \delta)\Big|_{\delta=\hat{\delta}_{\text{GMM}}}\right]' \widehat{\Omega}^{-1} \left[\sum_{i=1}^{N} \frac{\partial}{\partial \delta'} g(W_i; \delta)\Big|_{\delta=\hat{\delta}_{\text{GMM}}}\right]\right]^{-1}. \tag{5}$$

For an over-identified problem, where the number m of instrumental variables is greater than the number k of regressors, one can test for the null hypothesis of valid over-restricting instruments. Rejecting this hypothesis indicates the model is misspecified. Hansen (1982) proposes testing for an over-restricted model using the test statistic

$$J_N(\hat{\delta}_{\text{GMM}}) = N\left[\frac{1}{N}\sum_{i=1}^{N} g(W_i; \hat{\delta}_{\text{GMM}})\right]' \widehat{\Omega}^{-1} \left[\frac{1}{N}\sum_{i=1}^{N} g(W_i; \hat{\delta}_{\text{GMM}})\right]. \tag{6}$$

Under the null hypothesis, $J_N(\hat{\beta}_{\text{GMM}})$ has an asymptotic χ^2_{m-k} distribution, with $m - k$ degrees of freedom. For situations where the null hypothesis is rejected, Matyas (1999) describes procedures to identify subsets of problematic instrumental variables.

3.2 Limited information instrumental variable estimation

Even when the equation residuals in the system (1) are correlated, one can estimate each equation separately; this corresponds to a limited information instrumental variable approach. The moment condition in this case is

$$\mathrm{E}\{\left[y_{k,i} - \mathbf{x}'_{k,i}\beta_k\right]\mathbf{z}_{i,k}\} = \mathbf{0}. \tag{7}$$

Because $\mathbf{z}_{i,k}$ is a m_k-dimensional column vector, k, $k = 1, 2, \ldots, K$, we have m_k moment conditions in (7). Note that we have a separate set of moment conditions for each system equation k. One can show that, specifically for the instrumetal variable estimation problem, the closed-form solution for the minimization problem in (3) is given by

$$\hat{\beta}_k = [X_k' Z_k \Psi_k Z_k' X_k]^{-1} [X_k' Z_k \Psi_k Z_k' \mathbf{y}_k], \quad \text{for} \;\; k = 1, 2, \ldots, K \tag{8}$$

where $X_k = [\mathbf{x}_{k,1} \;\; \mathbf{x}_{k,2} \;\; \ldots \;\; \mathbf{x}_{k,N}]'$ is obtained by stacking the covariate vectors $\mathbf{x}_{k,i}'$, $i = 1, \ldots, N$. Analagously, $Z_k = [\mathbf{z}_{k,1} \;\; \mathbf{z}_{k,2} \;\; \ldots \;\; \mathbf{z}_{k,N}]'$ and $\mathbf{y}_k = [y_{k,1} \;\; y_{k,2} \;\; \ldots \;\; y_{k,N}]'$. The estimation procedure consists of first obtaining b_k^* by replacing Ψ_k by the identity matrix in (8), and then employing expression (4) to estimate the covariance matrix Ω_k of the moment conditions in (7). The second stage consists of finding $\hat{\beta}_k$ by replacing Ψ_k by $\widehat{\Omega}_k^{-1}$ in (8).

Using equation (5), we can estimate the covariance matrix for $\hat{\beta}_k$. In this case, expression (5) simplifies to

$$\widehat{\mathbf{C}}_k = N \big[X_k' Z_k \widehat{\Omega}_k^{-1} Z_k' X_k \big]^{-1}. \tag{9}$$

In summary, applying (8) and (9) for all K equations in the system (1), one can obtain parameters estimates and implement statistical inference for each equation separately.

3.3 Full information instrumental variable estimation

We now treat the simultaneous estimation of all K equations in (1), recognizing the possible correlation between the equation residuals $\varepsilon_{k_1,i}$ and $\varepsilon_{k_2,i}$, for $k_1 \neq k_2$, $k_1, k_2 \in \{1, 2, \ldots, K\}$. In this case, we replace the moment conditions in (7) by the condition below

$$\mathrm{E}\{g(W_i; \gamma)\} = \mathrm{E}\left\{ \begin{bmatrix} \big[y_{1,i} - \mathbf{x}_{1,i}'\beta_1\big]\mathbf{z}_{1,i} \\ \big[y_{2,i} - \mathbf{x}_{2,i}'\beta_2\big]\mathbf{z}_{2,i} \\ \cdots \\ \big[y_{K,i} - \mathbf{x}_{K,i}'\beta_K\big]\mathbf{z}_{K,i} \end{bmatrix} \right\} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \cdots \\ \mathbf{0} \end{bmatrix}.$$

Observe that the moment condition above corresponds to stacking all K conditions represented in (7). All observed variables are included in $W_i = [W_1' \;\; W_2' \;\; \ldots \;\; W_K']'$, and all coefficients are included in $\gamma = [\gamma_1' \;\; \gamma_2' \;\; \ldots \;\; \gamma_K']'$. The GMM estimation problem becomes to find $\hat{\gamma}$ in

$$\hat{\gamma} = \arg\min_{\gamma \in \Theta} \left[\frac{1}{N} \sum_{i=1}^{N} g(W_i; \gamma) \right]' \Psi \left[\frac{1}{N} \sum_{i=1}^{N} g(W_i; \gamma) \right]. \tag{10}$$

The closed-form solution to the minimization problem im (10) can be expressed as

$$\hat{\gamma} = \left[\mathrm{M}'\Psi\mathrm{M}\right]^{-1}\mathrm{M}'\Psi\mathrm{V}, \tag{11}$$

where

$$\mathrm{M} = \begin{bmatrix} \frac{1}{N}\left[Z_1'X_1\right] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{1}{N}\left[Z_2'X_2\right] & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{1}{N}\left[Z_K'X_K\right] \end{bmatrix} \quad \text{and} \quad \mathrm{V} = \begin{bmatrix} \frac{1}{N}\left[Z_1'Y_1\right] \\ \frac{1}{N}\left[Z_2'Y_2\right] \\ \dots \\ \frac{1}{N}\left[Z_K'Y_K\right] \end{bmatrix}.$$

Initially, we make Ψ equal to the identity matrix, for example. After estimating the covariance matrix of the moment conditions Ω, using (4), we can obtain the final estimate for γ by replacing Ψ in (11) by $\widehat{\Omega}^{-1}$. Finally, the covariance matrix for the coefficient estimates $\hat{\gamma}$ is given by

$$\widehat{\mathbf{C}} = \frac{1}{N}\left[\mathrm{M}'\widehat{\Omega}^{-1}\mathrm{M}\right]^{-1}. \tag{12}$$

In Section 5, we present an application of both limited information and full information spatial GMM approaches to study municipality growth in Brazil.

4 Monte Carlo Simulations

In this section we describe the Monte Carlo experiment to study the performance of the spatial GMM estimator. The data describe the municipalities of Brazil in 1991[4]. The latitude and longitude of each municipio's centroid are used as its coordinates for the purposes of estimation. The data are apt for this experiment for three reasons. First, with 4,267 observations, the dataset has a dimension large enough to be computationally challenging for non-GMM spatial techniques that require inversion of spatial weight matrices. Second, the municipalities are very unevenly spaced (see the Figure 1). Even though this feature of the data is contemplated in Conley's analytical results, where he assumes general sampling schemes, allowing for clusters of points in space, we tried to address how the diposition of the Brazilian municipalities centroids affect the estimates distribution. Finally, the simulated data generating process, as discussed below, assumes a heterogenous structure for the regression errors, violating the stationarity hypothesis in Conley (1999). Some authors advocate that the stationarity assumption is too restrictive for most of applications. Silva e

Resende (2005) study the municipalities growth in two states in Brazil, and found significant spatial autocorrelation in only one of them, suggesting a strong spatial nonstationarity in the data generating process. As will be shown in the results below, even with the nonstationarity hypothesis the spatial GMM approach still works, delivering better results than the usual OLS estimator or the nonspatial GMM estimator.

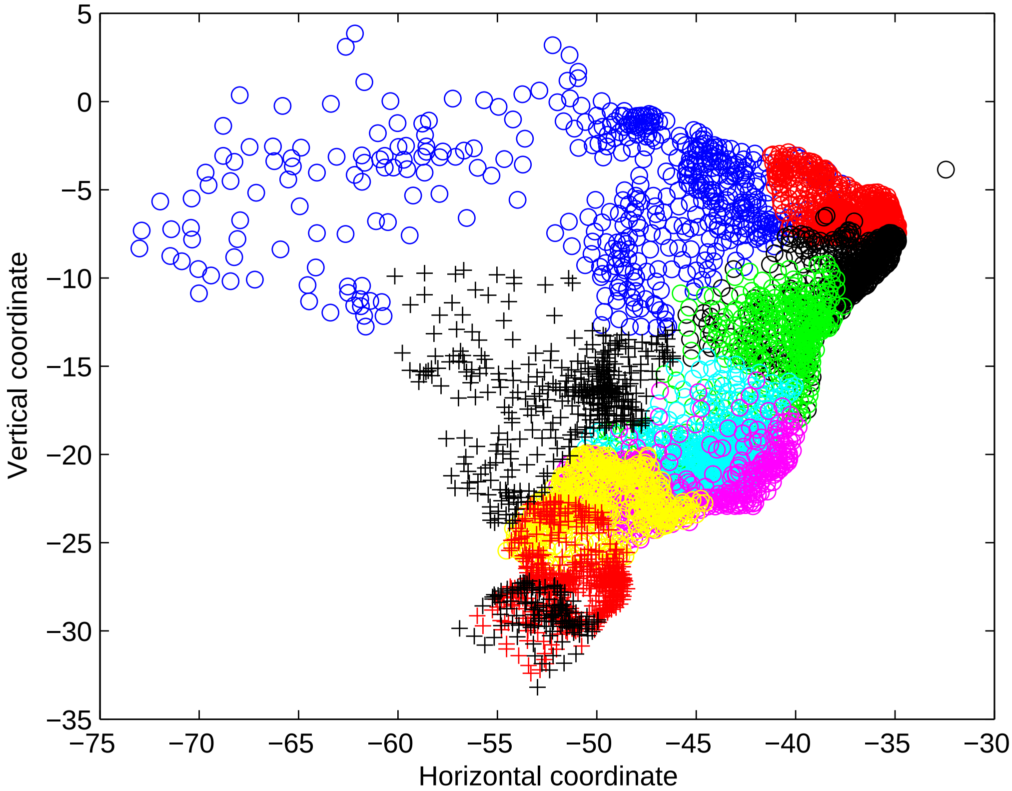

Figure 1: Centroids for the 4,267 municipalities in 1991.

4.1 Data generating process

The data generating process used in the simulations was the following

$$Y = X\beta + \varepsilon, \tag{13}$$

where Y is a $N \times 1$ response vector, X is a $N \times k$ covariate matrix, and ε is a N random

[4] To handle the different numbers of municipality between the years 1991 and 2000, we had to use a geographic defition called minimum comparable area (MCA). In this definition, we map the 2000 municipalities into the 1991 municipalities. Nonetheless, hereafter we will call MCA's and municipalities indistinguishably.

error vector. We have chosen a fixed X matrix, equal to the intrumental variable matrix in the labor supply equation (see Section 5). Each row in equation (13) corresponds to a municipality in the 1991 map. We then used the 1991 map centroid coordinates. To introduce some spatial correlation into the data generating process, we assumed the spatial autorregressive model for the errors

$$\varepsilon = \rho W_N \varepsilon + u, \tag{14}$$

where u is a vector of independent normal random variables with zero mean and variance structure described below. The constant ρ is the spatial autocorrelation coefficient, and W_N is a $N \times N$ neighborhood relationship matrix, with zeros in the main diagonal. The off-diagonal elements are calculated as $W_{i,j} = 1/(1 + d_{i,j})^{\alpha_i}$, with $d_{i,j}$ the distance between the centroid of municipality i and the centroid of municipality j, and α_i a decaying rate, which depends on the municipality i.

To introduce nonstationarity into the data generating process, we used different variances σ_i^2 for the components of u (forcing heteroscedasticity) and different values for the decaying rate α_i. We then divided the municipalities in nine non-overlaping groups, and specified different values for the error variance and for the rate α_i accross these groups. Figure 1 presents the disposition of the municipalities in each of the nine groups. Because the element $W_{i,j} = 1/(1 + d_{i,j})^{\alpha_i}$ depends specifically on municipality i and not on j, where i preceeds j in the database order, we are implicitly imposing a non-isotropic model, since spatial correlation will depend on the direction for the vector joining municipalities i and j. By construction, the covariance matrix of u is a diagonal matrix, with different elements σ_i^2 in each diagonal subgroup.

In the simulations, we tried different combinations of values for α_i and σ_i^2. Tables 2, 3, 9 and 10 present the values for σ^2 and α_i in each subgroup, together with the simulation results. For a matter of comparison with usual contiguity models, we added a column to the tables, containing the corresponding spatial autocorrelation estimated coefficient $\hat{\lambda}$ in a first-order spatial autoregressive model (vide LeSage and Pace, 2004)[5]. Finally, to have a graphical idea about the spatial autocorrelation created in our block diagonal approach, we present in Figures 2, 3, 6 and 8, the scatter plots of the vector of simulated residuals r, versus

[5]To estimate the first-order spatial model, we used the programs provided by the authors, available on the website www.spatial-econometrics.com. The estimate $\hat{\lambda}$ was obtaining for only one replication of the simulated error vector, so as to have an idea about the spatial correlation in terms of first-order neighborhood effects.

the neighbohrs residual average vector $W_{[0 \text{ or } 1]} \times r$, where $W_{[0 \text{ or } 1]}$ is a neighbohrood matrix, where the element (i,j) is 1 when municipalities i and j are neighbohrs and 0 otherwise. For comparison, we obtained a statistically significant estimated autocorrelation coefficient equal to 0.467, for the errors obtained from the regression using real data.

To simulate the error structure in equation (14), we can simulate a the sequence of iid normal random variables u first and then pre-multiply it by $(I - \rho W_N)^{-1}$ to obtain ε. For a reasonable small N, inverting the matrix $(I - \rho W_N)$ is not a difficult task. However, for high N, this numerical inversion can become very cumbersome and time consuming.

Because in the simulation experiment the generation of the error vector ε has to be done repeatedly, we had to introduce some simplification into the error process. In order to do that, we assumed a block-diagonal structure for the neighborhood relation matrix. Instead of using W_N as described above, we used a modification W_N^*, which consists of blocks of size K_b along the main diagonal of W_N. Also, for elements of W_N smaller than a threshold l_b, in absolute value, we made the corresponding element in W_N^* equal to zero. For example, for $N = 6$, consider the matrix W_N below

$$\mathrm{W_N} = \begin{bmatrix} 0.0 & 0.8 & 0.3 & -0.1 & 0.05 & -0.02 \\ 0.8 & 0.0 & 0.6 & 0.08 & 0.1 & -0.04 \\ 0.3 & 0.6 & 0.0 & 0.04 & 0.01 & 0.05 \\ -0.1 & 0.08 & 0.04 & 0.0 & 0.8 & -0.1 \\ 0.05 & 0.1 & 0.01 & 0.8 & 0.0 & 0.5 \\ -0.02 & -0.04 & 0.05 & -0.1 & 0.5 & 0.0 \end{bmatrix}.$$

For $K_b = 3$ and $l_b = 0.2$, the modified matrix W_N^* becomes

$$\mathrm{W_N^*} = \begin{bmatrix} 0.0 & 0.8 & 0.3 & - & - & - \\ 0.8 & 0.0 & 0.6 & - & - & - \\ 0.3 & 0.6 & 0.0 & - & - & - \\ - & - & - & 0.0 & 0.8 & - \\ - & - & - & 0.8 & 0.0 & 0.5 \\ - & - & - & - & 0.5 & 0.0 \end{bmatrix}.$$

Because W_N^* is block diagonal with blocks of size K_b, so is $(I - \rho W_N^*)$. Therefore, inverting $(I - \rho W_N^*)$ consists of inverting each block separately, which is much more attractive computationally. Obviously, using smaller blocks implies less intense computational work, but a lower degree of spatial dependence in the simulated data. In the simulations presented in this paper, we used blocks of size $K_b = 400$, for $N = 4,267$ (the last block ended up having 267 elements). In the final step for constructing the neighborhood relation matrix, we calculated the sum of the elements in each row of W_N^* and then found the maximum sum s_{max}. Finally, the neighborhood matrix $\widetilde{W}_N$ actually used was obtained by dividing each row in W_N^* by s_{max}.

As we discussed above, using a block diagonal instead of a full W_N^* matrix has the advantage of reducing the computational work, which is important in this experiment, provided we have to simulate the data generating process repeatedly. On the other hand, regardless the true resulting process, the block diagonal structure still guarantees the existence of spatial autocorrelation in the data, which is the main feature we want to simulate, in order to study the properties of different estimation schemes. Finally, as we can note by looking at the estimated first-order spatial autocoefficients $\hat{\lambda}_i$ in Tables 2, 3, 9 and 10, and the scatter plots in Figures 2, 3, 6 and 8, our block diagonal procedure was able to reproduce a reasonable spatial autocorrelation structure.

To guarantee model stability in the spatial process in (14), the autocorrelation coefficient ρ has to be between $1/\lambda_{min}$ and $1/\lambda_{max}$, where λ_{min} and λ_{max} are the smallest and the highest eigenvalues of the matrix $\widetilde{W}_N$. The eigenvalue λ_{min} is negative whereas the eigenvalue λ_{max} is positive[6]. By using different values for ρ we can regulate the degree of autocorrelation. In the simulation results presented below, we used $\rho = 0.95 \times 1/\lambda_{max}$.

Some care must be taken when interpreting the spatial dependence structure implied by the construction above. By dividing each row of W_N^* by the maximum row sum s_{max} we are automatically causing most of the elements of $\widetilde{W}_N$ to be very small. Only the elements of $\widetilde{W}_N$ corresponding to very close pairs of municipalities i and j (so that $d_{i,j}$ are significantly small) will have a reasonable size. An immediate consequence is that big municipalities will be less influenced by their neighbors than small municipalities. This approach is different for example from the usual contiguity matrix approach, where the element (i, j) in W_N is 1 if municipalities i and j are neighbors and 0 otherwise. In this case, some authors advocate standardizing W_N, by dividing each of its rows by the sum of row elements. By construction this strucuture implies that each municipality is equally influenced by the conjunction of its neighbors, regardless the municipality size[7]. Therefore, even when $\rho = 0.95 \times 1/\lambda_{max}$, the spatial corrrelation dependence can still be very mild for most of the municipalities, specially when the dispersion in the municipalities sizes is high.

[6]Because of the block structure, finding the eigenvalues of $\widetilde{W}_N$ consists simply of finding the eigenvalues of each block separately.

[7]See LeSage and Pace (2004), Anselin (2002 and 2004), for more details.

4.2 Parameter estimation

By construction, the data generating process used in the simulations does not imply any endogeneity in the right-hand-side variables. Therefore, we can use ordinary least squares (OLS) as one of the estimation methods. Additionally, for each generated data set in the simulations, we estimated the parameters using nonspatial GMM (without any covariance matrix spatial correction) and using spatial GMM (described in Section 3.1). When using the GMM, the instruments are the right-hand-side variables. Both OLS and simple GMM do not account for the spatial autocorrelation in the residuals.

Our goal in the experiment is to address the true levels in the Wald tests for parameter significance[8], as well as the validity of the asymptotic approximation for the distributions of the t-statistics. In the data generating process, we set the parameters equal to the values in Table 1. These values were obtained by a regression of wages on eight variables. For the estimated regression, we obtained the estimate $\hat{\sigma}^2 = 0.04237$ for the residuals variance. Three extra variables with coefficient equal to 0 were included in the equation estimated at each iteration in the experiment, as a test of the validity of the estimation method. We also set the intercept to be zero.

In each simulation experiment, we generated 400 data sets, obtaining 400 vectors of estimated parameter and standard errors for each estimation method. Because the instrumental variables are the same as the right-hand-side variables, the point parameter estimates for OLS, simple GMM and spatial GMM coincide. On the other hand, the standard errors can be quite different, implying different hypothesis test results. Table 1 presents the true parameter values in the simulations and the corresponding mean estimated values in the experiment shown in Table 2.

Table 2 presents the estimated Wald test levels, for nominal levels of 10%, 5% and 1%. For the results presented, we set the autocorrelation coefficient to be $\rho = 0.95 \times 1/\lambda_{max}$, the threshold value $l_b = 0.01$, the block size $K_b = 400$, and both cutoff values for the spatial GMM equal to 3.0. The different values for the decay rate α_k and the variance σ_k^2 are also presented in the table. The variances σ_k^2 are presented as multiples of $\hat{\sigma}^2 = 0.04237$, so as, for group 2 in Table 2, the true considered variance was 2.0×0.04237.

Observe that, in general, the estimated levels for Wald test are closer to the nominal levels in the spatial GMM approach. Both the OLS and the simple GMM underestimate

[8] Because we are testing the significance of a single parameter, the Wald test correspond simply to the usual t-test, assuming the normal asymptotic approximation holds.

Table 1: Parameter values in the simulation experiment (dependent variable: variation of logarithm of employement between 1991 and 2000)

Explanatory variable	True Parameter	Mean parameter estimates (OLS, simple GMM and spatial GMM)
Intercept	0.0000	-0.0068
Proportion 5-15 over 15-55 yrs old	0.2696	0.2686
Proportion native in 1991	0.1657	0.1658
Teacher qualification in 1991	-0.0070	-0.0074
Logarithm of homicides per capita in 1991	-19.077	-19.649
Proportion farmers in 1991	-0.5549	-0.5518
Illiteracy rate in 1991	-0.0022	-0.0022
Logarithm of wage in 1991	0.1199	0.1205
Bank Dummy	-0.0216	-0.0218
Fertility rate in 1991	0.0000	0.0004
Logarithm of population in 1991	0.0000	1.8e-005
Delta logarithm of market potential	0.0000	0.0034

the parameter estimate standard deviation, delivering estimated true levels much higher than the nominal ones. For the coefficient of log population in 1991, for example, for a nominal level of 10%, both the OLS and the simple GMM resulted in rejecting the null hypothesis $H_0 : \beta_{\text{log pop 1991}} = 0$ for more than 20% of the 400 simulated datasets, while the spatial GMM presented a estimated true level of 12.25%. Similarly, for a level of 1%, inference with spatial GMM rejects the null hypothesis of zero intercept 0.75% of the 400 replicates, whereas inference with OLS and with simple GMM rejects the null hypothesis 6.0% and 6.75% respectively of the 400 simulated datasets. In general, the standard errors obtained via spatial GMM seem to be valid in average, and estimated true levels in the tests are close to the nominal levels.

We replicated the Monte Carlo experiment for different configurations of decay rates, variances, threshold values and block sizes, and the results seemed to be quite robust. Table 3 presents the simulation results for an experiment with smaller variances, and the spatial GMM still presents estimated true levels close to the nominal ones. Figures 4 and 5 present histograms of the p-values for the 400 simulated datasets, for the Wald tests of parameter significance. Figure 4 corresponds to the experiment in Table 2, while Figure 5 corresponds to the configuration in Table 3. The histograms support the discussion above about the validity of the normal approximation for the distribution of the t-statistics in the spatial GMM framework. Note that the distribution of the p-values in the spatial GMM framework looks reasonably close to a uniform distribution in the interval $(0, 1)$. For both the

Table 2: Estimated levels for Wald tests for Experiment 1 ($\rho = 0.95$ and $l_b = 0.01$)

Heterogeneity group	σ_k^2	α_k	$\hat{\lambda}_k$
Group 1	1.0	7.0	0.1999
Group 2	2.0	9.0	0.4239
Group 3	1.5	12.0	0.3339
Group 4	0.5	9.0	0.2989
Group 5	1.5	8.0	0.3559
Group 6	2.0	10.0	0.5699
Group 7	1.5	7.0	0.3929
Group 8	2.5	11.0	0.6189
Group 9	2.0	9.0	0.2929

Explanatory variable	Nominal test level	Estimated true test level (%)		
		Simple OLS	Simple GMM	Spatial GMM
Intercept	1%	0.0600	0.0675	0.0075
	5%	0.1475	0.1575	0.0625
	10%	0.2350	0.2425	0.1225
Fertility in 1991	1%	0.0475	0.0475	0.0250
	5%	0.1125	0.1325	0.0850
	10%	0.1725	0.1900	0.1275
Population in 1991	1%	0.0600	0.0600	0.0100
	5%	0.1600	0.1625	0.0675
	10%	0.2250	0.2250	0.1225
Delta market potential	1%	0.0500	0.0650	0.0200
	5%	0.1075	0.1325	0.0550
	10%	0.1625	0.1975	0.1025

OLS and the simple GMM estimators, the distributions of the p-values appear more concentrated in the region close to zero, which is a consequence of the underestimated standard errors. The Appendix presents the simulation results for other data generating processes.

The simulation exercise supports the use of the spatial GMM when we are suspicious about the presence of spatial dependence. Besides, even in the presence of heterogeneity in the spatial dependence, using the spatial GMM seems to guarantee that the distribution of the p-values, and hence the true test levels, will be closer to what is expected according to the asymptotic approximations. Observe that the spatial GMM estimator did not make any assumption about the parametric form of the spatial dependence in the data generating process.

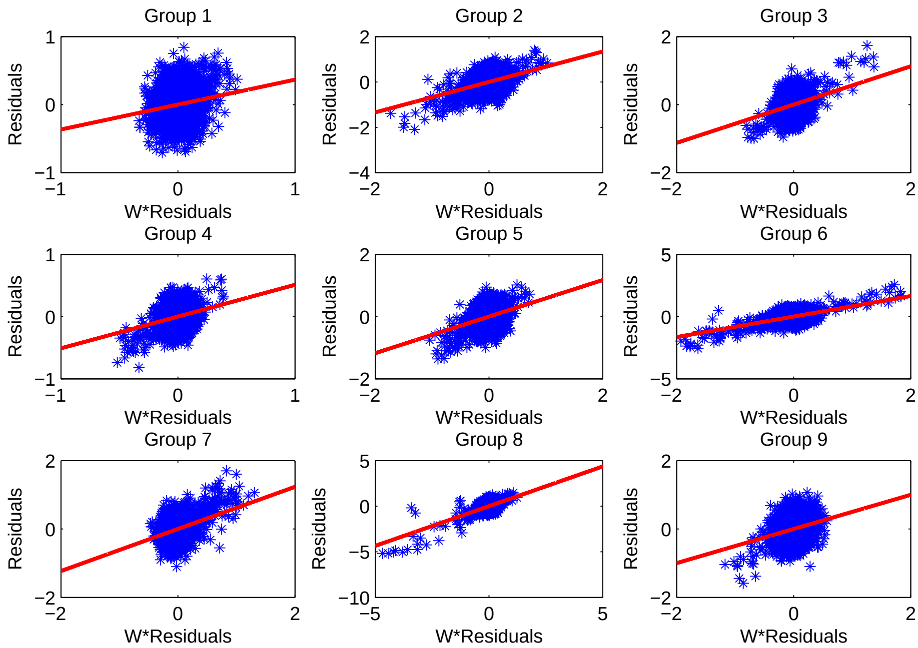

Figure 2: Scatter plot of simulated residual vector r versus neighbohrs residual average vector for every heterogeneous group (Experiment 1).

5 Application to municipality growth in Brazil

The proposed methodologies can be employed in a wide set of situations, ranging from studies on municipal economic growth, intracity analysis of housing markets to state spending spillovers (see, for example, Baicker, 2005). In this section, we present an application of the limited information spatial GMM and the full information spatial GMM to study the determinants of changes in labor supply and demand in Brazilian municipalities. In the original database, some of the municipalities presented missing values for some of the variables. We replaced the missing observations in each variable by the overall variable average. A more detailed and elaborated study can be found in Chomitz, da Mata, Carvalho and Magalhães (2005). For the purposes of the current paper we use a simpler specification, and so the results here do not necessarily correspond to the final findings of that paper.

The Brazilian economy is marked by the presence of spatial disparities and inequalities. The income per capita ratio between the richest and the poorest federation units was 8.9 in 1960 and 6.2 in 1996 (Azzoni *et. al.*, 2000). Several studies pose the question of regional development and inequality and proposed some insights (see, for instance, Azzoni *et. al.*, 2000, Ferreira and Diniz, 1994, Ferreira and Ellery, 1996, and Silveira Neto, 2001).

Table 3: Estimated levels for Wald tests for Experiment 2 ($\rho = 0.95$ and $l_b = 0.01$)

Heterogeneity group	σ_k^2	α_k	$\hat{\lambda}_k$
Group 1	1.0	7.0	0.1959
Group 2	1.2	9.0	0.4249
Group 3	0.8	12.0	0.3379
Group 4	1.5	9.0	0.2989
Group 5	1.5	8.0	0.3579
Group 6	0.8	10.0	0.5669
Group 7	1.5	7.0	0.3839
Group 8	1.1	11.0	0.6279
Group 9	0.9	9.0	0.2949

Explanatory variable	Nominal test level	Estimated true test level (%)		
		Simple OLS	Simple GMM	Spatial GMM
Intercept	1%	0.0500	0.0525	0.0075
	5%	0.1300	0.1300	0.0725
	10%	0.2075	0.2125	0.1225
Fertility in 1991	1%	0.0400	0.0400	0.0225
	5%	0.1100	0.1050	0.0700
	10%	0.1725	0.1725	0.1200
Population in 1991	1%	0.0475	0.0500	0.0075
	5%	0.1500	0.1500	0.0750
	10%	0.2225	0.2275	0.1350
Delta market potential	1%	0.0475	0.0575	0.0175
	5%	0.1175	0.1400	0.0625
	10%	0.1700	0.1875	0.1075

We employ the spatial GMM procedures to study the municipalities performance between 1991 and 2000. With this purpose, a labor demand-supply two equation model was estimated. In our study, the growth of employment is a proxy for the growth of the municipality size and the growth of wage is for the growth in productivity, the two ways one can look at local level growth[9].

We now provide a brief description of the estimated model. Assume that each municipality i has a production function $f(K, L; A)$ for a composite output, where K is the aggregate capital that represents both physical and a vector measuring the quantity and quality of worker education, L is the number of workers; A is a vector of productivity shifters, including transport connectivity to markets, local governance quality, and agroclimate[10].

[9]Hereinafter demand equation and wage equation will be synonyms, as well as supply and labor/employment equation.

[10]To simplify the notation, we initially drop the subscript i in the economy quantities for municipality i.

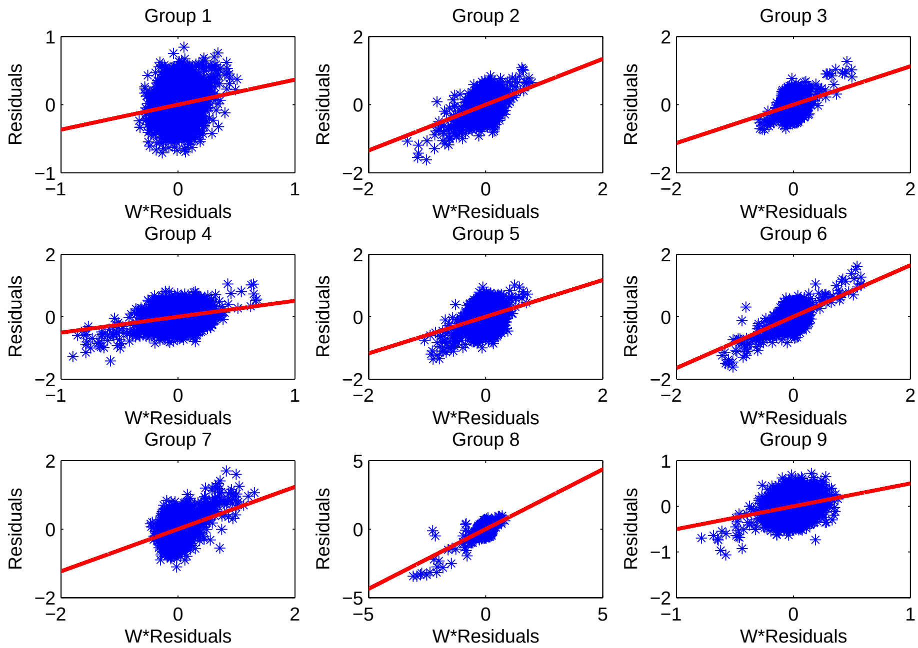

Figure 3: Scatter plot of simulated residual vector r versus neighbohrs residual average vector for every heterogenous group (Experiment 2).

The municipality faces a price which is a function $P(MP, GT)$ of local market potential or demand and of government transfers to individuals. The measure MP is operationalized as an inverse-distance-weighted function of the total incomes of neighboring municipalities. A labor demand equation expresses the wage rate as the value of the marginal product

$$w = P(MP, GT)\left[\frac{\partial}{\partial L} f(K, L; A)\right]. \tag{15}$$

We model labor supply as an upward sloping supply curve anchored at a prior period's workforce

$$L_t = L_t(w, \text{Educ}, COL, WF_{t-1}, \text{Amenities}), \tag{16}$$

where WF_{t-1} is the size of the cohort at period $t-1$, which corresponds to period t work force. Current employment is larger or smaller than that cohort (reflecting in or out migration, and labor force participation), depending on the wage, that could be adjusted for education and for the cost of living. Alternatively, let $w*$ be the adjusted wage, so that

$$L_t = L_t(w^*, WF_{t-1}, \text{Amenities}). \tag{17}$$

Note that this labor supply model presumes relatively slow adjustment to changing demand

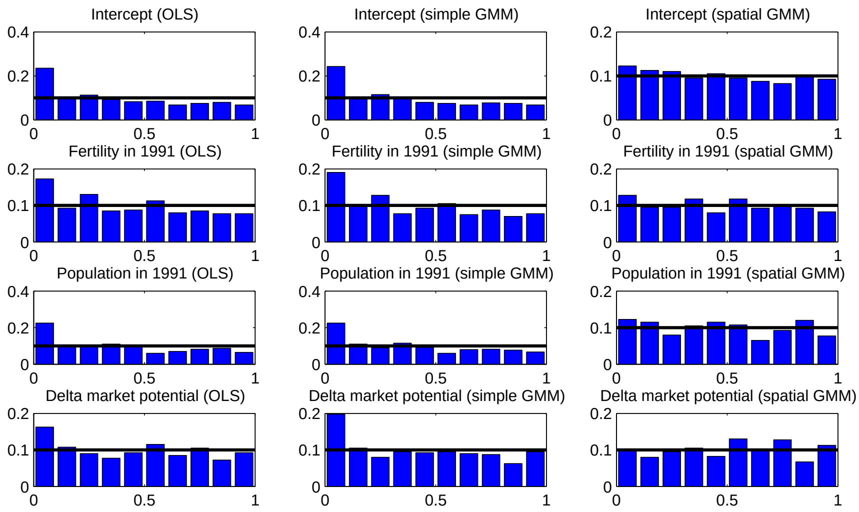

Figure 4: Histogram for p-values in the Wald tests for the significance of the tested parameters (Experiment 1).

conditions. The specification for the wage equation is therefore

$$\begin{aligned} E\{\Delta \log w\} =& \beta_0 + \beta_1 \Delta \log L + \beta_2 \Delta \log MP + \beta_3 \Delta \log GT \\ & + \beta_4 \Delta \log K_{\text{physical}} + \beta_5 \Delta \log K_{\text{human}} + \beta_6 \text{RAINFALL}. \end{aligned} \tag{18}$$

We replace $\Delta \log K_{\text{physical}}$ and $\Delta \log K_{\text{human}}$ by functions of initial human capital, infrastructure and governance variables such as years of schooling of the employed individuals, percentage of primary school teachers with tertiary degree (hereinafter denoted teacher qualification), transportation cost to São Paulo, transportation cost to the nearest federation unit capital and a dummy for governance, which has value 1 if the local government uses computers for financial accounting, and zero otherwise.

The supply equation has the following specification

$$E\{\Delta \log L\} = \log(w_{\text{t-1}}) + EWF_{\text{t-1}} + \log(MP_{t-1}) + \text{Amenities}_{t-1} + \text{Agroclimate}, \tag{19}$$

where EWF is the relative size of the entering work force, and MP captures the spillover effect of nearby areas.

Endogeneity is a big concern: most of the right-hand-side variables are to some extent endogenous, so that an intrumentation scheme has to used. Time-lagged or space-lagged variables provide potential instruments. For instance, in the wage equation, we instrumented $\Delta \log L$ with initial period demographic variables instrument, $\Delta \log MP$ with spatially lagged Δ log(Mean education) and transportation cost variables in 1995 with their

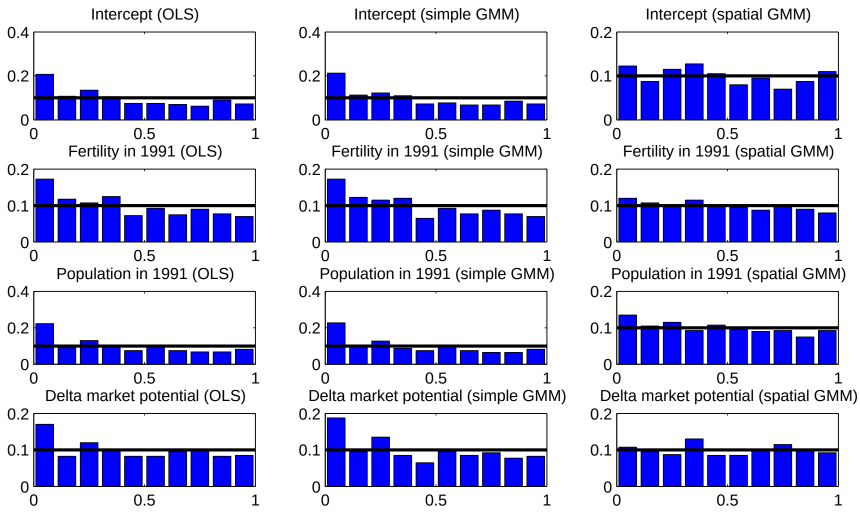

Figure 5: Histogram for p-values in the Wald tests for the significance of the tested parameters (Experiment 2).

values in 1968. We treated the initial human capital and governance variables as exogenous. For more details on the right-hand-side variables and their instruments see Tables 4 and 5. We estimated the model in differences, so as to avoid the presence of fixed effects related to growth in wages and employment in the municipalities. Note that both equations are identified.

Table 4: Labour equation - instruments and right-hand-side variables

Variable	Instrument(s)
Employment rate in 1991	Illiteracy rate in 1991 Fertility rate in 1991
Logarithm of wage in 1991	Exogenous
Proportion 5-15 over15-55 yrs old	Exogenous
Proportion native in 1991	Exogenous
Delta logarithm market potential	Exogenous
Teacher qualification in 1991	Exogenous
Logarithm homicides per capita in 1991	Exogenous
Proportion farmers in 1991	Exogenous
Bank Dummy	Exogenous
Logarithm of population in 1991	Exogenous

To estimate the equations in first differences, we employed data for 1991 and 2000. The main source are the 1991 and 2000 IBGE Census, 1991 IBGE Municipality data, DATASUS and NEMESIS's transportation cost data. The municipalities were aggregated in terms of Minimum Comparable Areas (MCAs) in order to deal with the municipalities

Table 5: Wage equation - instruments and right-hand-side variables

Variable	Instrument(s)
Teacher qualification in 1991	Exogenous
Years of schooling in 1991	Exogenous
Total precipitation	Exogenous
Government with Accountability	Exogenous
Delta logarithm of employment	Employment rate in 1991 Soil quality First and third principal components of monthly rainfall
Logarithm of transport cost to São Paulo in 1995	Logarithm of transport cost to São Paulo in 1968
Logarithm of transport cost to nearest state capital in 1995	Logarithm of transport cost to nearest state capital in 1968
Delta logarithm of government transference	Dependency ratio in 1991 Illiteracy rate in 1991 Repetition rate in 1991 Proportion of elderly persons in 1991 Logarithm of population in 1991 Proportion 5-15 over 15-55 yrs old in 1991
Delta logarithm of Market Potential	Delta logarithm of education Market Potential Urbanization rate in 1991 Population density in 1991

splits mostly due to Brazilian 1988 Constitution. Therefore, the estimation were performed for 4,267 MCAs instead of 5,507 municipalities in 2000. Estimations were performed using our codes written in Ox[11].

The results for the wage equation are in Table 6, 11 and 12; for the labor equation, the results are in Tables 7, 13 and 14. For both demand and supply sides we ran the Spatial GMM regressions using several cutoffs in order to test the sensitivity of the results. We used cutoff values of 0.3, 1.0, 2.0, 3.0 and 5.0. For each demand and supply we present the results for the 2SLS and then for the several spatial GMM regressions. The 2SLS demand result displays that most of the variables are significant at 1%.

The wage equation Spatial GMM results do not support the ones for 2SLS. For the cutoff equal to 0.3 degrees (about 30 kilometers)[12], the signs of the estimates did not change. Nonetheless, some variables became not significant. According to the GMM results, only years of schooling, precipitation, delta employment, delta government transference and

[11]A free academic version of Ox can be downloaded from www.dornik.com. Interested users can request the spatial GMM codes employed here to the corresponding author. Codes in Stata for the limited information spatial GMM are available at http://gsbwww.uchicago.edu/fac/timothy.conley/research/gmmcode/statacode.html.

[12]Each unit in the cutoff measure is equivalent to a hundred kilometers.

Table 6: Wage equation - spatial GMM results

Explanatory Variable	Two-Stage Least Squares		Spatial GMM Cutoff = 0.3	
	Estimate	Standard Error	Estimate	Standard Error
Intercept	-0.1917	0.0680	-0.3824	0.3596
Teacher qualification in 1991	0.0055	0.0023	0.0049	0.0074
Years of schooling in 1991	0.0620	0.0043	0.0611	0.0144
Government with Accountability	0.0234	0.0129	0.0137	0.0436
Total precipitation	0.0000	0.0000	0.0001	0.0000
Delta employment	-0.4701	0.0517	-0.5225	0.2512
Transport cost to São Paulo	-0.0353	0.0058	-0.0304	0.0255
Transport cost to nearest state capital	-0.0517	0.0060	-0.0320	0.0242
Delta government transference	0.3129	0.0305	0.2485	0.1326
Delta Market Potential	0.5191	0.0920	0.7850	0.3872
crit. fn. test of overid. restrictions	——	——	15.381	——

market potential matters. It is useful to note that the only variable with drastic changes is the government transference, when we applied different cutoffs. With cutoff values of 1.0, 2.0, 3.0 and 5.0, this variable is no longer significant, and for a cutoff=5.0, its estimated coefficient changes sign.

One can observe a quite similar pattern for the labor equation results. Concerning the 2SLS supply equation estimation, we have that again all variables are significant at 1% except for the Bank dummy (significant at 10%) that reflects the presence or not of a bank in the municipality. For the Spatial GMM results, one can see that the Bank dummy at remains not significant for different cutoff values.

In general, all other variables are significant when we use spatial GMM estimation, with different cutoffs. Besides, both the signs and the dimension of the coefficient estimates seem to be robust to the use of spatial GMM with several cutoff values. The exception is the variable homicides per capita. Even though this variables is quite significant for different cutoffs, its coefficient presents unstable estimates when the cutoff increases. In fact, for cutoff = 0.3, the estimated parameter is -31.5, and this value drops to -79.2 for cutoff = 5.0.

The conclusions are that human capital, neighborhood level of development and natural advantage matters for growth of wage, i.e, for the improvement of productivity. Related to the size increase, population age composition and amenities matters. The composition of economic activity is also important, the more the specialization in agricultural sector,

Table 7: Labor equation - spatial GMM results

Explanatory Variable	Two-Stage Least Squares		Spatial GMM Cutoff = 0.3	
	Estimate	Standard Error	Estimate	Standard Error
Intercept	-0.4986	0.1400	-0.4694	0.1832
Wage in 1991	0.1986	0.0136	0.1982	0.0195
Proportion 5-15 over 15-55 yrs old	0.4359	0.0415	0.4369	0.0558
Proportion native in 1991	0.3198	0.0385	0.3332	0.0517
Delta market potential	0.2289	0.0269	0.2266	0.0397
Teacher qualification in 1991	-0.0122	0.0025	-0.0125	0.0035
Homicides per capita in 1991	-28.8330	4.5671	-31.4990	10.0180
Proportion farmers in 1991	-1.3984	0.1639	-1.4525	0.2230
Bank Dummy	-0.0162	0.0099	-0.0143	0.0131
Population in 1991	-0.0310	0.0054	-0.0319	0.0074
Employment rate in 1991	0.8289	0.1508	0.8668	0.2047
crit. fn. test of overid. restrictions	——	——	6.0126	——

the less is the growth. In this analysis, growth of neighboring municipalities is beneficial: growth of market potential increases the labor supply (depressing wage), but also boosts labor demand, increasing the wage. The net effect is both higher employment and higher wages since the direct effect of market potential on wage is bigger than the indirect one (market potential - employment, employment - wage).

Some quite similar patterns appear when it comes to the application of the full information spatial GMM method. The general results for both wage and labor equations, employing full information spatial GMM, remain very similar to the results using limited information procedures, as shown in Tables 8, 15 and 16. Note that, in the wage equation, government accoutability becomes significant with increasing cutoff values, differently to what was observed in the limited information case. As for the supply side, note that the estimates for the variable homicides per capita still present a very unstable behavior, in spite of its statistical significance: when cutoff = 5.0, the estimate drops to -140.9. For this equation, limited and full-information procedures appear to perform somewhat similarly except for two points: in the full information case, with cutoff equal to 5.0, population and teacher qualification lose statistical significance.

Finally, as mentioned in Section 3, we can use the J-statistics to test for over-identified restrictions. For the wage equation, the tests do not reject the null hypothesis of orthogonality between the instruments and the equation residuals, for all used cutoff values. On the other hand, for the labor equation, the J-statistics resulted significant for cutoff = 0.3,

Table 8: System of equations - full information spatial GMM results (cutoff value = 0.3)

Explanatory Variable	Estimate	Standard Error	P-value
Labor equation			
Intercept	-0.5682	0.1780	0.0014
Wage in 1991	0.2019	0.0194	0.0000
Proportion 5-15 over 15-55 yrs old	0.4302	0.0536	0.0000
Proportion native in 1991	0.3351	0.0495	0.0000
Delta market potential	0.2273	0.0392	0.0000
Teacher qualification in 1991	-0.0120	0.0035	0.0006
Homicides per capita in 1991	-32.4240	9.9646	0.0011
Proportion farmers in 1991	-1.3719	0.2177	0.0000
Bank Dummy	-0.0160	0.0130	0.2183
Population in 1991	-0.0283	0.0070	0.0001
Employment rate in 1991	0.8233	0.1979	0.0000
Wage equation			
Intercept	-0.4377	0.3661	0.2318
Teacher qualification in 1991	0.0033	0.0075	0.6630
Years of schooling in 1991	0.0607	0.0148	0.0000
Government with Accountability	0.0288	0.0448	0.5210
Total precipitation	0.0000	0.0000	0.1666
Delta employment	-0.6198	0.2611	0.0176
Transport cost to São Paulo	-0.0318	0.0260	0.2219
Transport cost to nearest state capital	-0.0321	0.0247	0.1930
Delta government transference	0.2230	0.1363	0.1018
Delta Market Potential	1.0047	0.4020	0.0124

cutoff = 1.0 and cutoff = 2.0. For cutoffs equal to 3.0 and 5.0, the orthogonality hypothesis is not reject, even for significance level of 10%.

6 Final Comments

In this paper, we addressed the estimation of multiequation models, where the observations are spatially correlated. The framework presented is quite general, and by using generalized method of moments estimation, we were able to handle the problem of endogeneity in some of the right-hand-side variables. We based our analysis heavily on Conley's (1999) work.

The idea behind Conley's GMM is to build a consistent estimator for the GMM moment conditions covariance matrix. His spatial dependence consistent estimator is base on

Newey and West (1987) heteroskedasticity and autocorrelation consistent covariance matrix estimator. Because the spatial GMM framework does not assume any parametric form, it is robust to model misspecification, such as the misspecification of weight matrices. The finite sample properties of the spatial GMM estimator are investigated via Monte Carlo simulations, where the spatial dependence is heterogeneous by construction. We find evidence that, even in the presence of spatial heterogeneity, standard asymptotic approximations for the test statistics still seem to hold.

We employed the proposed framework to the spatial estimation of labor supply and demand equations, in order to study the determinants of Brazilian municipality economic growth between 1991 and 2000. The growth of employment is a proxy for the growth of municipalility population and the growth of wage is a proxy for the growth in productivity. For public policies prescription, the results found here are that investment in education matters. We found also spillover effects of one municipality's growth onto its neighbors. On the other hand, policies such as to improve local level governance and transportation network did not appear as significant. We estimated the model for all Brazilian municipalities, what may not be a reasonable strategy, given the enormous heterogeneity across city hierarchy in the country. For a more thorough investigation, see Chomitz, Da Mata, Carvalho and Magalhes (2005).

Several questions remain open regarding the spatial GMM approach discussed in this paper. Among these questions, we can mention: (1) analytical results to study the asymptotic distributions for the parameter estimates, as well as functions of the parameter estimates, in the presence of nonstationarity, since Conley (1999) showed these results for stationary random fields; (2) extending the analytical results for the spatial panel data GMM, both for the stationary as well as the nonstationary case; (3) applications of the spatial GMM technique to nonlinear regression models, in the presence or not of endogeneity; (4) extension of the spatial GMM framework to panel data, with an analitical investigation of the asymptotic properties of the estimators and a extensive simulation experiment to investigate the finite sample properties of the adapted estimators. These and other extensions are under current investigation by the authors.

APPENDIX

Table 9: Estimated levels for Wald tests for Experiment 3 ($\rho = 0.95$ and $l_b = 0.005$)

Heterogeneity group	σ_k^2	α_k	$\hat{\lambda}_k$
Group 1	1.0	4.0	0.0689
Group 2	1.2	6.0	0.2899
Group 3	0.8	10.0	0.4389
Group 4	1.5	6.0	0.1599
Group 5	1.5	4.0	0.3119
Group 6	0.8	8.0	0.5289
Group 7	1.5	5.0	0.4389
Group 8	1.1	7.0	0.5399
Group 9	0.9	4.0	0.1809

Explanatory variable	Nominal test level	Estimated true test level (%)		
		Simple OLS	Simple GMM	Spatial GMM
Intercept	1%	0.0350	0.0375	0.0050
	5%	0.0950	0.0950	0.0575
	10%	0.1600	0.1575	0.1125
Fertility in 1991	1%	0.0200	0.0200	0.0150
	5%	0.0675	0.0700	0.0650
	10%	0.1275	0.1325	0.1050
Population in 1991	1%	0.0300	0.0350	0.0075
	5%	0.1100	0.1175	0.0625
	10%	0.1725	0.1825	0.1125
Delta market potential	1%	0.0300	0.0375	0.0250
	5%	0.1025	0.1100	0.0700
	10%	0.1650	0.1650	0.1225

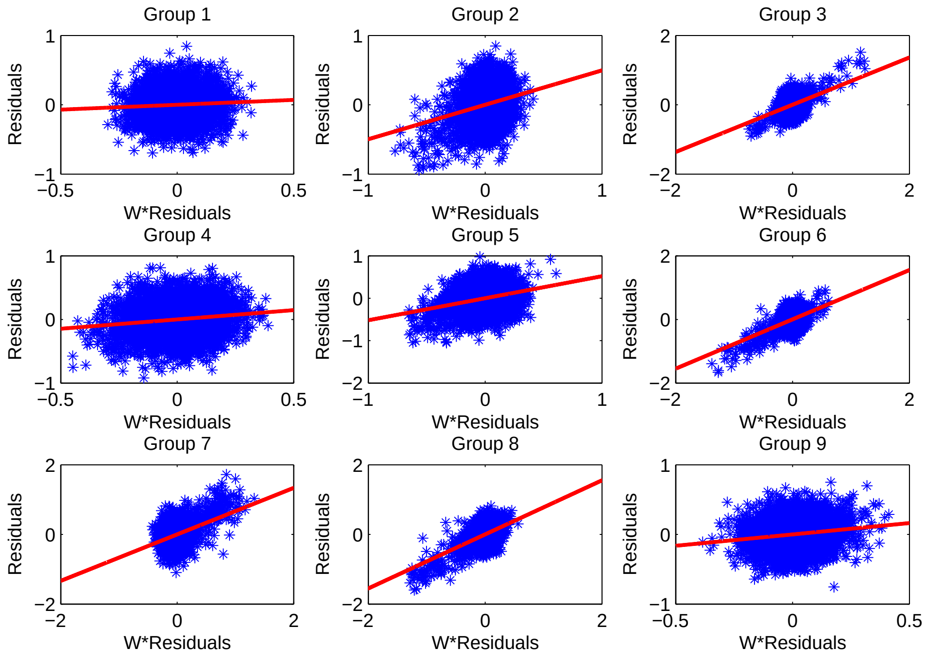

Figure 6: Scatter plot of simulated residual vector r versus neighbohrs residual average vector for every heterogenous group (Experiment 3).

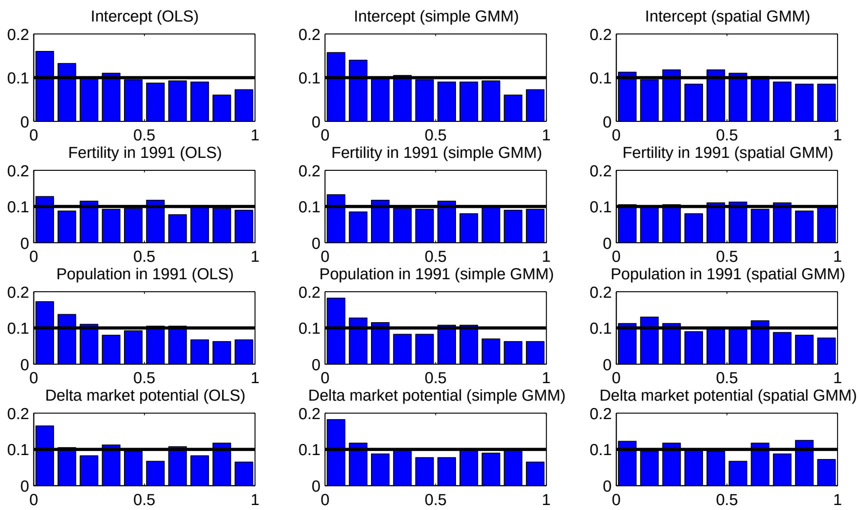

Figure 7: Histogram for p-values in the Wald tests for the significance of the tested parameters (Experiment 3).

Table 10: Estimated levels for Wald tests for Experiment 4 ($\rho = 0.95$ and $l_b = 0.005$)

Heterogeneity group	σ_k^2	α_k	$\hat{\lambda}_k$
Group 1	1.0	4.0	0.0729
Group 2	2.0	6.0	0.2889
Group 3	1.5	10.0	0.4389
Group 4	0.5	6.0	0.1609
Group 5	1.5	4.0	0.3059
Group 6	2.0	8.0	0.5329
Group 7	1.5	5.0	0.4379
Group 8	2.5	7.0	0.5399
Group 9	2.0	4.0	0.1839

Explanatory variable	Nominal test level	Estimated true test level (%)		
		Simple OLS	Simple GMM	Spatial GMM
Intercept	1%	0.0325	0.0375	0.0100
	5%	0.1100	0.1200	0.0600
	10%	0.1775	0.1900	0.1075
Fertility in 1991	1%	0.0175	0.0225	0.0200
	5%	0.0725	0.0775	0.0650
	10%	0.1200	0.1375	0.1050
Population in 1991	1%	0.0425	0.0375	0.0100
	5%	0.1025	0.1050	0.0550
	10%	0.1875	0.1850	0.0975
Delta market potential	1%	0.0400	0.0500	0.0250
	5%	0.0975	0.1300	0.0700
	10%	0.1625	0.1875	0.1250

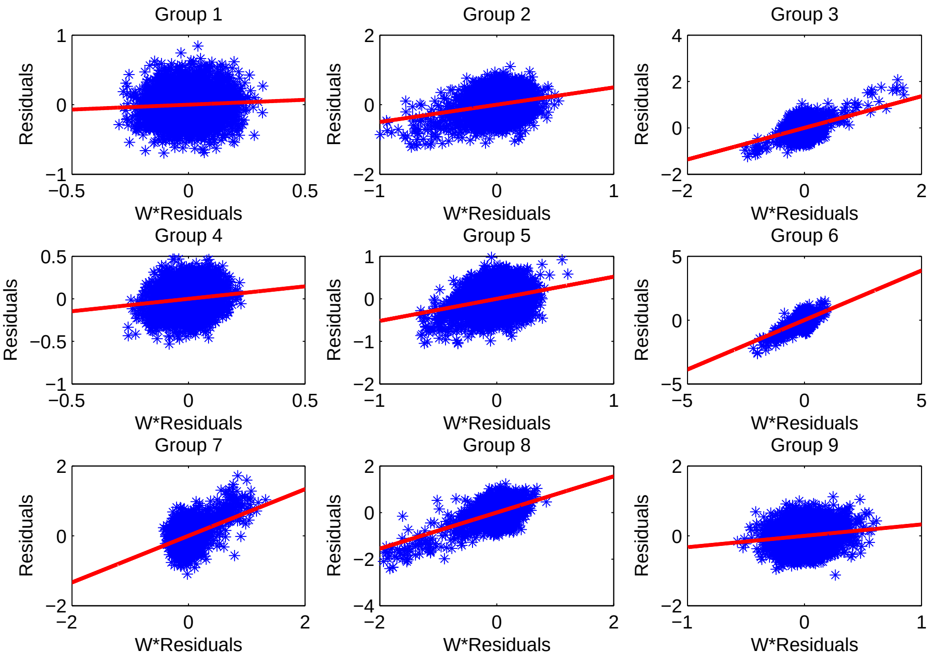

Figure 8: Scatter plot of simulated residual vector r versus neighbohrs residual average vector for every heterogenous group (Experiment 4).

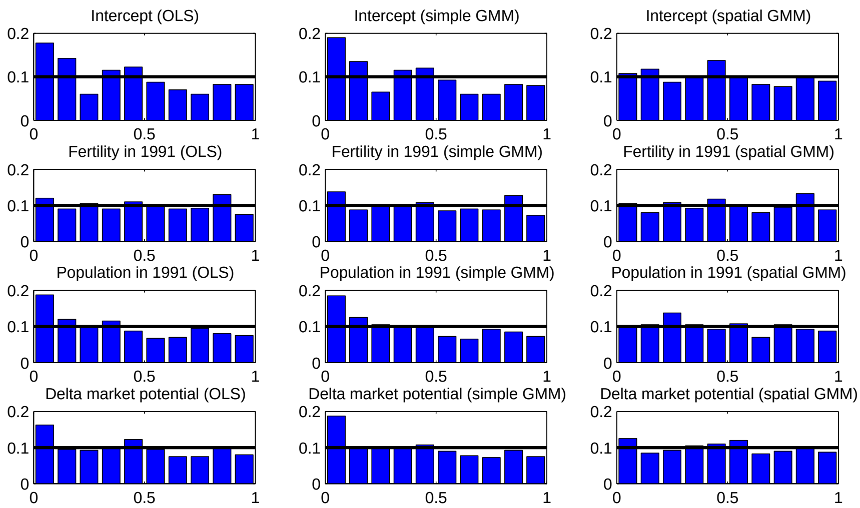

Figure 9: Histogram for p-values in the Wald tests for the significance of the tested parameters (Experiment 4).

Table 11: Wage equation - spatial GMM results

Explanatory Variable	Spatial GMM Cutoff = 1.0		Spatial GMM Cutoff = 2.0	
	Estimate	Standard Error	Estimate	Standard Error
Intercept	-0.4137	0.5594	-0.4799	0.6733
Teacher qualification in 1991	0.0036	0.0077	0.0017	0.0084
Years of schooling in 1991	0.0597	0.0161	0.0600	0.0176
Government with Accountability	0.0194	0.0452	0.0279	0.0481
Total precipitation	0.0001	0.0000	0.0001	0.0000
Delta employment	-0.6188	0.3167	-0.7185	0.3606
Transport cost to São Paulo	-0.0282	0.0382	-0.0248	0.0464
Transport cost to nearest state capital	-0.0238	0.0346	-0.0145	0.0385
Delta government transference	0.1880	0.1682	0.1069	0.1964
Delta Market Potential	0.8510	0.5558	0.9892	0.6548
crit. fn. test of overid. restrictions	12.032	——	9.4067	——

Table 12: Wage equation - spatial GMM results

Explanatory Variables	Spatial GMM Cutoff = 3.0		Spatial GMM Cutoff = 5.0	
	Estimate	Standard Error	Estimate	Standard Error
Intercept	-0.5767	0.7046	-1.1599	0.7240
Teacher qualification in 1991	-0.0007	0.0085	-0.0065	0.0082
Years of schooling in 1991	0.0640	0.0178	0.0747	0.0165
Government with Accountability	0.0381	0.0515	0.0842	0.0557
Total precipitation	0.0001	0.0000	0.0001	0.0000
Delta employment	-0.8589	0.3927	-1.3215	0.4586
Transport cost to São Paulo	-0.0184	0.0478	0.0136	0.0475
Transport cost to nearest state capital	-0.0056	0.0401	0.0352	0.0411
Delta government transference	0.0252	0.2241	-0.2070	0.2733
Delta Market Potential	1.1405	0.7025	1.8211	0.7628
crit. fn. test of overid. restrictions	8.5672	——	8.2167	——

Table 13: Labor equation - spatial GMM results

Explanatory Variable	Spatial GMM Cutoff = 1.0		Spatial GMM Cutoff = 2.0	
	Estimate	Standard Error	Estimate	Standard Error
Intercept	-0.4264	0.2233	-0.3972	0.2395
Wage in 1991	0.2064	0.0262	0.2102	0.0329
Proportion 5-15 over 15-55 yrs old	0.4555	0.0687	0.4622	0.0792
Proportion native in 1991	0.3599	0.0668	0.3764	0.0808
Delta market potential	0.2315	0.0492	0.2352	0.0582
Teacher qualification in 1991	-0.0143	0.0047	-0.0149	0.0059
Homicides per capita in 1991	-39.4940	18.0560	-47.9240	27.3860
Proportion farmers in 1991	-1.6002	0.2960	-1.6993	0.3405
Bank Dummy	-0.0151	0.0137	-0.0157	0.0151
Population in 1991	-0.0346	0.0091	-0.0350	0.0108
Employment rate in 1991	0.9966	0.2727	1.0838	0.3210
crit. fn. test of overid. restrictions	4.0326	——	2.9306	——

Table 14: Labor equation - spatial GMM results

Explanatory Variable	Spatial GMM Cutoff = 3.0		Spatial GMM Cutoff = 5.0	
	Estimate	Standard Error	Estimate	Standard Error
Intercept	-0.3891	0.2502	-0.3621	0.2798
Wage in 1991	0.2130	0.0377	0.2071	0.0453
Proportion 5-15 over 15-55 yrs old	0.4588	0.0883	0.4300	0.1034
Proportion native in 1991	0.3950	0.0899	0.4257	0.1014
Delta market potential	0.2329	0.0649	0.2196	0.0758
Teacher qualification in 1991	-0.0148	0.0068	-0.0124	0.0083
Homicides per capita in 1991	-57.2070	34.9390	-79.1860	45.4530
Proportion farmers in 1991	-1.7670	0.3682	-1.8346	0.4142
Bank Dummy	-0.0162	0.0172	-0.0155	0.0188
Population in 1991	-0.0347	0.0126	-0.0308	0.0154
Employment rate in 1991	1.1419	0.3523	1.1721	0.4058
crit. fn. test of overid. restrictions	2.6071	——	2.8928	——

Table 15: System of equations - full information spatial GMM results (cutoff value = 3.0)

Explanatory Variable	Estimate	Standard Error	P-value
Labor equation			
Intercept	-0.4502	0.2269	0.0473
Wage in 1991	0.2141	0.0355	0.0000
Proportion 5-15 over 15-55 yrs old	0.3937	0.0784	0.0000
Proportion native in 1991	0.4614	0.0833	0.0000
Delta market potential	0.1957	0.0583	0.0008
Teacher qualification in 1991	-0.0123	0.0064	0.0566
Homicides per capita in 1991	-100.5600	27.9930	0.0003
Proportion farmers in 1991	-1.8587	0.3401	0.0000
Bank Dummy	-0.0212	0.0165	0.1982
Population in 1991	-0.0256	0.0103	0.0127
Employment rate in 1991	1.2160	0.3329	0.0003
Wage equation			
Intercept	-0.6796	0.7158	0.3424
Teacher qualification in 1991	-0.0058	0.0082	0.4813
Years of schooling in 1991	0.0661	0.0182	0.0003
Government with Accountability	0.0712	0.0510	0.1631
Total precipitation	0.0001	0.0000	0.0260
Delta employment	-1.1274	0.3972	0.0045
Transport cost to São Paulo	-0.0197	0.0487	0.6852
Transport cost to nearest State Capital	-0.0021	0.0408	0.9591
Delta government transference	-0.0945	0.2258	0.6756
Delta Market Potential	1.5181	0.7310	0.0378

Table 16: System of equations - full information spatial GMM results (cutoff value = 5.0)

Explanatory Variable	Estimate	Standard Error	P-value
Labor equation			
Intercept	-0.3779	0.2284	0.0981
Wage in 1991	0.1872	0.0404	0.0000
Proportion 5-15 over 15-55 yrs old	0.3236	0.0840	0.0001
Proportion native in 1991	0.4961	0.0917	0.0000
Delta market potential	0.1888	0.0636	0.0030
Teacher qualification in 1991	-0.0066	0.0074	0.3708
Homicides per capita in 1991	-140.8800	31.0600	0.0000
Proportion farmers in 1991	-1.8823	0.3609	0.0000
Bank Dummy	-0.0114	0.0174	0.5128
Population in 1991	-0.0137	0.0107	0.2010
Employment rate in 1991	1.1840	0.3721	0.0015
Wage equation			
Intercept	-1.4167	0.7155	0.0477
Teacher qualification in 1991	-0.0110	0.0079	0.1658
Years of schooling in 1991	0.0810	0.0161	0.0000
Government with Accountability	0.1250	0.0507	0.0136
Total precipitation	0.0001	0.0000	0.0114
Delta employment	-1.7002	0.4333	0.0001
Transport cost to São Paulo	0.0244	0.0476	0.6078
Transport cost to nearest State Capital	0.0500	0.0410	0.2226
Delta government transference	-0.3932	0.2626	0.1342
Delta Market Potential	2.3127	0.7637	0.0025

REFERENCES

ANDREWS, D. W. K. Heteroskedasticity and Autocorrelation consistent covariance matrix estimation, Journal of Econometrics 59, 817--58. 1991.

ANDREWS, D. W. K. and MONAHAN, J. C. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator, Econometrica 60, 953-66. 1992.

ANSELIN, L. Under the hood. Issues in the specification and interpretation of spatial regression models. Department of Agricultural and Consumer Economics. University of Illinois, Urbana-Champaign. mimeo. 2002.

________. Advances in Spatial Econometrics: Methodology, Tools and Applications. Advances in Spatial Science. Springer. 2004.

AZZONI, C., MENEZES, T., MENEZES, N., NETO, R. Geografia e convergência da renda entre os estados brasileiros. In: HENRIQUES, R. (Org.). Desigualdade e pobreza no Brasil. Rio de Janeiro: IPEA, p.299-343. 2000.

BAICKER, K. The spillover effects of state spending. Journal of Public Economics 89, p. 529-544. 2005.

BELL, K. P. and BOCKSTAEL, N. E. Applying the generalized-moments estimation approach to spatial problems involving microlevel data, Review of Economics and Statistics 87, 72--82. 2000.

CHOMITZ, K., da MATA, D., CARVALHO, A., MAGALHÃES, J.C. Spatial patterns of development in non-metropolitan Brazil. Brasília: IPEA, mimeo. 2005

CLIFF, A. and ORD, J. Testing for Spatial Autocorrelation among Regression Residuals, Geographical Analysis 4, 267-284. 1972.

________. Spatial Autocorrelation, Pion, London. 1973.

________. Spatial Processes, Models and Applications, Pion, London. 1981.

CONLEY, T. G. GMM estimation with cross sectional dependence. Journal of Econometrics 92, 1-45. 1999.

DRUSKA, V. and HORRACE, W. C. Generalized moments estimation for spatial panel data. NBER. Technical working paper 291. https://www.nber.org/papers/TO291. 2003.

FERREIRA, P. C. G., ELLERY JR., R. G. Convergência entre a renda per capita dos estados brasileiros. Revista de Econometria, v.16, n.1, p.83-103, apr. 1996.

FERREIRA, A. H. B., DINIZ, C. C. Convergência entre as rendas per capita estaduais no Brasil. Belo Horizonte: UFMG/CEDEPLAR. 27p. (Texto para Discussão, n.79). 1994.

FREES, E. W. Assessing cross-sectional correlation in panel data. Journal of Econometrics 69, 393--414. 1995.

HANSEN, L. P. Large sample properties of generalized method of moments estimators, Econometrica 50, 1029--54. 1982a.

HANSEN, L. P. and SINGLETON, K. J. Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269--86. 1982b.

KELEJIAN, H. H. and PRUCHA, I. R. Estimation of simultaneous systems of spatially interrelated cross sectional equations. Journal of Econometrics 118, 27-50. 2004.

LESAGE, J. P. and PACE, K. Spatial and Spatiotemporal Econometrics. Advances in Econometrics. JAI Press. 2004.

MATYAS, L. Generalized Method of Moments Estimation. Themes in Modern Econometrics. Cambridge University Press. 1999.

NEWEY, W. K. and WEST, K. D. A simple, positive semi-definite, heteroskedastic and autorrelation consistent covariance matrix. Econometrica 55, 703-708. 1987.

SILVA, A. M. A., RESENDE, G. M. Crescimento economico comparado dos municipios alagoanos e mineiros: uma analise espacial. Brasília/DF: IPEA (Instituto de

Pesquisa Econômica Aplicada), p.1-25. (mimeo). 2005.

SILVEIRA NETO, R. M. Crescimento e spillovers: a localização importa? Evidências para os estados brasileiros. Revista Econômica do Nordeste, Fortaleza, v.32, N.ESP., p.524-545. 2001.

WHITE, H. Asymptotic Theory for Econometricians. Academic Press. 2000.

.

EDITORIAL

Coordenação
Silvânia de Araujo Carvalho

Supervisão
Iranilde Rego

Revisão
Renata Frassetto de Almeida Rose
Naiane de Brito Francischetto (estagiária)
Raquel do Espírito Santo (estagiária)

Editoração
Aeromilson Mesquita
Elidiane Bezerra Borges

Brasília
SBS – Quadra 1 – Bloco J – Ed. BNDES, 9º andar
70076-900 – Brasília – DF
Fone: (61) 3315-5090
Fax: (61) 3315-5314
Correio eletrônico: editbsb@ipea.gov.br

Rio de Janeiro
Av. Presidente Antônio Carlos, 51, 14º andar
20020-010 – Rio de Janeiro – RJ
Fone: (21) 3804-8118
Fax: (21) 2220-5533
Correio eletrônico: editrj@ipea.gov.br

COMITÊ EDITORIAL

Secretário-Executivo
Marco Aurélio Dias Pires

SBS – Quadra 1 – Bloco J – Ed. BNDES,
9º andar, sala 904
70076-900 – Brasília – DF
Fone: (61) 3315-5374
Fax: (61) 3315-5314
Correio eletrônico: madp@ipea.gov.br

Tiragem: 130 exemplares