The central mission of the Regional Economics and Spatial Modeling (REASM) laboratory at the University of Arizona is education and research. The purpose of the Discussion Papers is to circulate intermediate and final results of this research among readers within and outside the University of Arizona. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of REASM or of the University of Arizona.

All requests and comments should be directed to Dr. Sandy Dall’erba, Regional Economics And Spatial Modeling (REASM) laboratory, University of Arizona, Harvill Building, Room 435B, P.O. Box 210076, Box #2, Tucson, AZ 85721-0076, Phone (520) 626-8054, Fax (520) 621-2889, dallerba@email.arizona.edu

The Regional Economic Impact Simulator: a WebGIS Tool for Regional Policy Analysis

Dall’erba Sandy, Janezic Melissa, Morris K. Richard and Neasham Fred

REASM 12-01 August 2012
The Regional Economic Impact Simulator: a WebGIS Tool for Regional Policy Analysis

by Sandy Dall’erba¹, Melissa Janezic², K. Richard Morris² and Fred Neasham²

¹ Director, Regional Economics and Spatial Modeling laboratory, University of Arizona
² Student in MSGIST (Master of Science in Geographic Information Systems Technology, School of Geography and Development, University of Arizona)

Abstract: An important aspect of academic research is to be able to transmit complicated material to an audience of both specialists and non-specialists. This is the goal of the Regional Economic Impact Simulator. Based on a webGIS platform, it allows anyone to build a regional policy scenario of his choice and to visualize on a map, in a matter of seconds, how regional economic growth is modified as a result of it. Because of interactions between regions, it is not only the locality where the scenario is implemented that will experience a change in growth, but the entire system.
An example of regional policy could be: what if the federal government were to double funding to support businesses in Pima county (Arizona)? How would this change affect the economic growth of Pima and of other counties? The answer can be found here, on the WebGIS site of the Regional Economic Impact Simulator.
Section 1: Statement of objectives

An important aspect of academic research is to be able to transmit complicated material to an audience of both specialists and non-specialists. This is the goal of the Regional Economic Impact Simulator. Based on a webGIS platform, it allows anyone to build a regional policy scenario of his choice and to visualize on a map, in a matter of seconds, how regional economic growth is modified as a result of it. Because of interactions between regions, it is not only the locality where the scenario is implemented that will experience a change in growth, but the entire system.

Section 2: Data, model description and estimation

All the variables we use in this webGIS are measured for 3,076 counties of the conterminous US States (Alaska and remote islands such as Hawaii and Puerto Rico are thus excluded). The time period covered is described in the model below. All the data come from two well-known statistical offices: the Bureau of Economic Analysis and the Census Bureau.

The regional policy scenarios implemented in the frame of our webGIS rely on the estimation of an economic growth model called the conditional beta-convergence model. This model has been widely discussed and used in the literature (Solow, 1956; Durlauf and Johnson, 1995; Barro and Sala-i-Martin, 1995; Mankiw et al., 1992; Dall’erba and Le Gallo, 2008). In essence, it measures the degree to which the growth rate of per capita GDP (Gross Domestic Product) of a county is related to its initial level of per capita GDP and to a set of additional conditioning explanatory variables including federal government spending (as in Dall’erba and Le Gallo, 2008, in the case of Europe). The impact of the latter variables on growth has never been formally tested in the literature, hence this webGIS offers a significant contribution to the understanding of the role of public spending on economic growth in the US.

In addition, the model accounts for the presence of spatial dependence, also called spatial autocorrelation, across counties. This phenomenon captures interregional interactions such as, among others, backward and forward trade linkages, technology spillovers and migration (Dall’erba and Le Gallo, 2008) and has been highlighted numerous times in the economic growth literature (Dall’erba, 2005a, 2005b; Le Gallo and Dall’erba, 2006; Celebioglu and Dall’erba, 2010). More precisely, the type of spatial autocorrelation modeled here is called a spatial lag model where one of the explanatory variables \( W_{gT} \) represents the spatially weighted average of the growth rates of the neighbouring counties.

As usual in the spatial econometric literature, the definition of neighboring counties is based on pure geographical distance, as exogeneity of the latter is unambiguous. More precisely, we use the great circle distance between regional centroids, defined as

\[
\begin{align*}
    w_{ij}^*(k) = 0 & \quad \text{if} \quad i = j, \forall k \\
    w_{ij}^*(k) = d_{ij} & \quad \text{if} \quad d_{ij} \leq 100 \text{ miles} \quad \text{and} \quad w_{ij} = w_{ij}^* / \sum_j w_{ij}^* \\
    w_{ij}^*(k) = 0 & \quad \text{if} \quad d_{ij} > 100 \text{ miles}
\end{align*}
\]

where \( w_{ij}^* \) is an element of the unstandardized weights matrix; \( w_{ij} \) is an element of the standardized weights matrix \( W \); \( d_{ij} \) is the great circle distance between centroids of region \( i \) and \( j \). Several distance cut-offs have been tested. The choice of 100 miles is based on the maximization of the Moran’s I statistics of global spatial autocorrelation applied to the distribution of the initial level of income. Its value is statistically significant. Here, the cutoff parameter, 100 miles, indicates that the neighbors of each county are those located within 100 miles.

Based on all the elements above, the economic growth model we estimate can be written as follows:
\[
g_i = \rho W g_i + \alpha + \beta_0 y_0 + X \beta_2 + \text{fed. funding} \beta_3 + \varepsilon \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)
\]  

(2)

where \( g_i \) is the \((N \times 1)\) vector of average growth rates of per capita income between 2000 and 2008 (in log); \( W g_i \) is the \((N \times 1)\) vector of the average value of \( g_i \) among the neighbors of county \( i \) based on the neighborhood definition specified in (1); \( \alpha \) is a constant term; \( y_0 \) is the \((N \times 1)\) vector of per capita income in 2000 (in log); \( X \) is a \((N \times 16)\) matrix of explanatory variables measured in 2000 and described in table (1) below; \( \text{fed. funding} \) is a \((N \times 3)\) matrix of various types of per capita federal funding (measured as an average over 2000-2008) also described in table (1). There is conditional beta-convergence if the estimate of \( \beta_1 \) is negative and statistically significant. The coefficient estimates \( \beta_2 \) and \( \beta_3 \) capture the impact of the other explanatory variables on growth. Due to the presence of spatial dependence across observations, the above model is estimated by Maximum Likelihood (ML). The results appear in table 1 below.

Finally, note that we decided to allow our webGIS to display and run simulation experiments (see part 3 below) on the statistically significant variables only.
Table 1- ML estimation results of the economic growth model (2)

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Coefficient estimate and p-value</th>
<th>Explanatory variables</th>
<th>Coefficient estimate and p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_g T )</td>
<td>0.649 (0.000)</td>
<td>Services (share of workers in)</td>
<td>-0.005 (0.059)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.125 (0.000)</td>
<td>Agriculture (share of workers in)</td>
<td>-0.019 (0.000)</td>
</tr>
<tr>
<td>Per capita income in 2000 (in log)</td>
<td>-0.014 (0.000)</td>
<td>Density (population / square mile)</td>
<td>2.327×10^{-7} (0.031)</td>
</tr>
<tr>
<td>Population growth (average over 2000-2008)</td>
<td>-0.007 (0.000)</td>
<td>Minority (share in population)</td>
<td>0.006 (0.004)</td>
</tr>
<tr>
<td>Employment rate</td>
<td>0.008 (0.083)</td>
<td>Female head of household (share)</td>
<td>-0.019 (0.059)</td>
</tr>
<tr>
<td>Manufacturing (share of workers in)</td>
<td>-0.026 (0.000)</td>
<td>Education (share of people with a bachelor degree or higher)</td>
<td>0.009 (0.018)</td>
</tr>
<tr>
<td>Construction (share of workers in)</td>
<td>-0.575×10^{-3} (0.933)</td>
<td>Share of people with a high school degree</td>
<td>0.002 (0.503)</td>
</tr>
<tr>
<td>Retail trade (share of workers in)</td>
<td>-0.024 (0.000)</td>
<td>Company size (share below 5-9 employees)</td>
<td>0.013 (0.070)</td>
</tr>
<tr>
<td>Wholesale trade (share of workers in)</td>
<td>-0.024 (0.020)</td>
<td>Federal spending for farmers (per capita)</td>
<td>-5.432×10^{-6} (0.000)</td>
</tr>
<tr>
<td>Transportation (share of workers in)</td>
<td>-0.008 (0.300)</td>
<td>Federal spending for the retirees (per capita)</td>
<td>7.066×10^{-6} (0.000)</td>
</tr>
<tr>
<td>Finance (share of workers in)</td>
<td>-0.002 (0.762)</td>
<td>Federal spending for businesses (per capita)</td>
<td>2.054×10^{-5} (0.000)</td>
</tr>
</tbody>
</table>

Notes: Calculations performed in Geoda 0.9.5. There are \( N = 3,076 \) observations. p-values are in brackets. Log likelihood: 9857.91, AIC: -19671.8, SC: -19539.1. The result of the BP (Breusch-Pagan) test indicates the significant presence of heteroskedasticity (p-value = 0.000). The LR (Likelihood Ratio) test on spatial lag dependence indicates that the spatial lag model outperforms the OLS (Ordinary Least Squares) model (p-value = 0.000).

Section 3: Scenario simulation process

As any spatial lag model, model (2) above can be rewritten under the following form:

\[
g_T = (I - \rho W)^{-1} X^* \beta + (I - \rho W)^{-1} \varepsilon
\]

(3)

where \( X^* \) is a \( (N \times 22) \) matrix of all explanatory variables (including the constant). Since \( \rho \) and the elements of the standardized weights matrix \( W \) are less than one, the matrix \( (I - \rho W)^{-1} \) can be expanded as follows: \( (I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \ldots + \rho^n W^n \). As a result, two types of global spillover effects are relevant in the spatial lag model (Anselin, 2003):

1. A multiplier effect for the explanatory variables: the growth rate of county \( i \) is not only affected by a marginal change of the explanatory variables of county \( i \) but also is affected by...
marginal changes of the explanatory variables in the other counties, more importantly so for closer counties.

(2) A diffusion effect for the error process: a shock in the regression residuals at any location will propagate to all the other regions of the sample. This diffusion effect also declines with distance.

In our webGIS, we use the property of the multiplier effect by allowing users to modify the value of any statistically significant variables of model (2) and visualize how it changes growth not only in the county where the change took place but also in all the other counties. Formally, the relationship that is mapped is as follows:

\[ g_T = (I - \rho W)^{-1} X^\top \beta + (I - \rho W)^{-1} \epsilon \]

(3)

\[ \partial g_T = (I - \rho W)^{-1} \partial X_i \hat{\beta}_i \]

(4)

where \((I - \rho W)^{-1}\) is a matrix of which elements do not change with the shock, \(\partial X_i\) is the value of the shock implemented on variable \(i\), and \(\hat{\beta}_i\) is the coefficient estimate associated to variable \(i\) in table 1. Based on this relationship, the result of any simulation shows the change in each county’s growth rate, not the new growth rate after the shock has been implemented. While the latter element can easily be calculated, it is left, in conjunction with other extensions, for future developments of our webGIS.

References: