Determinants of School Attendance rate for Bolivia: A spatial econometric approach

Mauro Oscar Lua Delboy Céspedes *

* Economist, Universidad Privada Boliviana - Fundación Educativa, Bolivia.
Correspondence: delboymauro@gmail.com

Abstract

The main objective of this paper is to examine the impact of the spatial component on Bolivian school enrollment rate and to identify its main determinants. The contribution of this research is to stablish and identify the key relationships between the school attendance rate and spatial lagged explanatory variables.

Another aim of this research is to make an exploratory analysis of spatial data with reference to the Bolivian municipalities and identify spatial interaction. This objective is reached through the construction of a contiguity matrix using software Geoda. This methodology shows two types of contiguity: the rook contiguity and the queen contiguity. The research evaluates the results of the queen contiguity matrix.

Also, this paper employs Luc Anselin's Spatial methodology using spatial error models as a specification of the spatial regression. Previously running an ordinary least square (OLS) model and make the tests to examine the hypothetical presence of spatial dependence between the data. To achieve this objective, the paper uses Moran's I and Lagrange Multipliers to detect spatial interaction and obtain the correct specification.

This paper uses a municipal-aggregated level data basis built through Bolivian Census in 2012. The percentage rate of school attendance between 6-19 years old is considered as a unit of analysis. Cross-sectional and geo-coded data, as well as shape files and geographic information system (GIS) files of Bolivian municipalities are used for this analysis.The results reveal that there is an existent relationship between school attendance and the spatial component. Spatial diagnostics allow to use a spatial error model to estimate specification's parameters. Hence, the dependent variable can stablish statistically significant relationships with the explanatory variables.

The paper safely concludes that spatial component affects directly the school attendance rate in Bolivia. After the final specification and the spatial analysis, the paper concludes that the relationship between the dependent variable and education and household welfare variables is positive and it is negative with rural percentage of population and poverty variables.

keywords: Spatial Regression, Spatial Dependence, Spatial Error Model, Geo-Coded Data, Contiguity Matrix, School Attendance Rate.

1. Introduction

Spatial Econometrics had an important development over the last years. This branch has highlighted the importance of spatial interaction when cross-sectional data and panel data are modeled (Paelinck and Klaasen, 1979). In this sense, this subfield of Econometrics allows to measure the effect of the spatial component of an economic variable. This effect is analyzed through two fundamental characteristics: (1) Spatial autocorrelation and (2) the spatial heterogeneity (Anselin, 2003).

The importance of specifying, estimate and include the spatial component in modern Econometrics arises as a result of two factors (Anselin, Florax, 1995): (1) increased interest in models that explain similar patterns of behavior between variables that share geographical proximity (contiguity). (2) the need for spatial data management given the existence of georeferenced data associated with availability of geographic information systems (GIS). The development of these factors has represented an innovative methodology that spectrally analyzes the incorporation of spatial effects to the causal relations that are used in regular research.

Spatial regression methods allow to observe spatial dependence between observations when they belong to border regions or neighboring areas (Lesage, 2008). Contiguity property means that observations tend to exhibit values similar to those in nearest locations. Spatial econometrics becomes also important when observed variation in the dependent variable comes as a consequence of unobserved influences. For example, intrinsic or cultural factors could be identified when there are variations in the same variable if it changes from dimension to dimension, Ertur y Koch (2007) use several models that propose physical and human capital externalities as spatially interdependent variables that behave similarly between neighboring regions.

Recent studies have determined that economic phenomena occur not only at a certain time but also in a given space (Haining, 2003). The "new economic geography" (Krugman, 1998) has highlighted the effect that spatial externalities have had on different socioeconomic variables such as international trade and economic growth.

Empirical evidence shows that urban population has a higher probability to attend a school than rural population (Cid, 2011). Given that spatial component has an impact over socio-economic variables, this phenomenon is observed in school attendance rate and municipal poverty (Torres and Franco, 2015).

School attendance has always been an interesting variable for educational economics and also for general economics, given that this is a key variable for development and growth economics (Cetrángolo, Curcio, Caligaro, 2017). Despite the importance of spatial econometrics, the level of this type of research in Bolivia has been low.

This working paper uses a spatial approach to evaluate the determinants of school attendance rate in Bolivia using disaggregated data by municipality for each one of 339 Bolivian municipalities. It seeks to encourage, from a theoretical perspective and an empirical application, the use of spatial lag and/or spatial error models in the field of education and, in this sense, to promote different investigations that may be carried out in the future. In particular, it is important to evaluate the impact of the determinants of school attendance rate, as well as the incidence of the spatial component in this relationship.

This paper's structure is organized as it follows: Section 2 makes a brief review of the basics of spatial econometrics that were used in this investigation. Section 3 shows an exploratory analysis of spatial data and the endogenous and exogenous variables that were used in the investigation. In section 4, we focus on the estimation of the model, as well as the different spatial autocorrelation tests observing the results involved. In section 5 the respective conclusions of this paper are analyzed.

2. *The basics of spatial econometrics*

Spatial econometrics starts from the principle of the first law of geography, which states that all things are related to each other, but the closest things in space have a greater relationship than the distant ones (Tobler, 1979). This postulate constitutes a cornerstone in the presentation of the spatial component as a determining variable in a regression. Luc Anselin (Anselin,1999) defines spatial econometrics as a branch of econometrics that deals with the appropriate treatment of spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) in regression models with cross-sectional and data panel.

Spatial econometrics is concerned with showing the importance that the space component has in the specification of the model, Paelinck and Klaasen (1979) mention five characteristics of spatial econometrics, these are: i) The role of spatial interdependence in this type of models. ii) The asymmetry of spatial relationships. iii) The importance of other explanatory factors located in other spaces. iv) The differentiation between ex-post and ex-ante interactions and v) The explicit modeling of space.

Geographic information is one of the fundamental elements for the use of spatial models. Geo-referenced information data contain important information about spatial interaction and its main characteristic is that they are associated with a location, that is, they can be represented by a map under latitude, longitude and/or distance coordinates. One of the most used forms for the representation of spatial information is to locate the set of polygons in a matrix of specific weights; this matrix is called contiguity and is symbolized with W (weights). Matrix W quantifies connections between regions of a given territory and neighboring regions and helps to identify whether there is spatial autocorrelation between observations. This spatial autocorrelation plays an important role in the spatial analysis of geographic systems (Haining, 2009). Several authors (Pinske and Slade, 1998. Chen, 2008) indicate different methodologies for the construction of a contiguity matrix. The most practical methodology is the construction of a square and binary W matrix that will be later standardized. Formally, spatial relationships are a B subset of a Cartesian product $\mathbb{R}^2 x \mathbb{R}^2 = \{(i,j): i \in \mathbb{R}^2, j \in \mathbb{R}^2\}$ where $i \neq j$ and a set cannot be united to itself: $(i,i) \subseteq \mathfrak{B}$ this relation will be fulfilled for all spatial objects (Tiefelsdorf, 1998). The elements of this matrix will be assigned in the following way:

$$W_{ij} = \begin{cases} 1 & if\ i\ and\ j\ are\ spatially\ linked\ to\ each\ other \\ 0 & otherwise \end{cases} \quad (1)$$

It should be noted that the elements of the main diagonal of this square matrix will be equal to zero (since a region cannot be close to itself). As equation (1) explains more simply, the value of the spatial weights matrix will express a binary relationship with values of one and zero. Each space unit will be symbolized by a row i and its neighboring potentials that will be found in column j being i ≠ j. The value Wij will be one if the region i is considered neighbor to the region j and will take the value of zero when they are not. It should be noted that spatial relationships show a clear difference with temporal relationships, given that temporal relationships show a past-present-future sequence while spatial relationships can be multidirectional and multilateral.

Geoda is a program that provides a graphical interface to methods of exploratory spatial data analysis and employ spatial weights denominated row-standardized form. Row-standardization takes the given weights W_{ij} (e.g, the binary zero-one weights) and divides them by the row sum:

$$W_{ij(s)} = {}^{W_{ij}}\!/_{\sum_j W_{ij}} \quad (2)$$

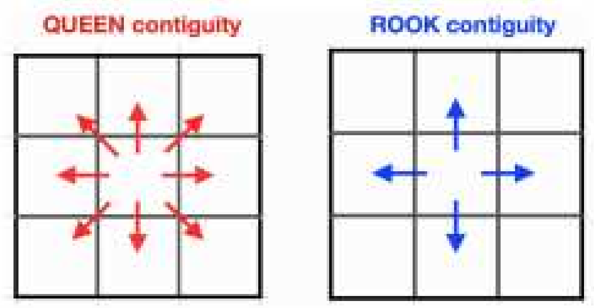
Intuitively, the mentioned standardization is done to obtain percentages of spatial influence (values that go between zero and one) whose sum of each row totals one, equation (3) shows this procedure:

$$S_0 = \sum_i \sum_j W_{ij} \tag{3}$$

As noted above, the spatial weights matrix is important for assigning contiguity between regions. Contiguity is defined as two spatial units that share a common boundary. This paper establishes a difference between two contiguity criteria, the "rock" criteria and the "queen" criteria. These terms refer to an analogy between the movement of chess pieces and the direction of contiguity between neighboring regions.

The rook criterion defines neighbors by the existence of a common edge between two spatial units. The queen criterion is somewhat more encompassing and defines neighbors as spatial units sharing a common edge or a common vertex. Figure 1 expresses this contiguity criterion as it follows:

Figure 1 - Contiguity Criteria



Source: De Bellefon, Loonis and Le Gleut, 2018

Spatial econometrics is a field in which techniques are designed to incorporate and analyze the spatial dependence that exists between observations that are considered contiguous. Spatial dependence is a functional relationship between what happens at a certain point in space and what happens elsewhere (Moreno and Vayá, 2000). This indicates that the dependent variable will have high levels of spatial autocorrelation when its values are affected by both exogenous variables and geographical variables. There are different statistical tests to detect spatial autocorrelation, one of the most common is the Moran I test (Moran,1948), which is a global indicator of spatial dependence.

$$I_i = \frac{\sum_i \sum_j W_{ij} Z_i Z_j}{\sum Z_i^2} \tag{4}$$

In equation (4) Moran's test I is explained where W_{ij} are the terms of a square binary matrix of standardized contiguity, Z_i represents the value of the dependent variable (in our case, the school attendance rate of each region). To determine the presence of spatial autocorrelation, the value of the Moran's I should be close to 1 or, more accurately, the null hypothesis that is the absence of autocorrelation should be rejected in a significant way. However, the Moran's I indicator is not able to detect the observations in which the spatial dependence develops with greater intensity. The identification of spatial clusters is given by two important tools: a) Moran's scatterplot and b) Univariate local Moran's I. Moran's scatterplot is a tool that Geoda software proposes to determine the intensity of spatial autocorrelation in neighboring areas (Anselin, 1993). This scatterplot shows the spatial phenomenon in a more disaggregated way because it analyzes the standardized dependent variable on the abscissa axis, and the spatial lag also standardized on the ordinate axis (y-axis). This graph is divided into four quadrants that show the intensity of the spatial relationship existing between the observations of the dependent variable in order to identify "hot spots". The first and third quadrants show a positive spatial association, while the second and fourth quadrants show negative relationships:

Another tool to obtain additional spatial evidence is Univariate local Moran's I. It is a statistical indicator used to identify the presence of spatial clusters, and this indicator has the null hypothesis of no spatial autocorrelation. These two spatial autocorrelation detection tools, the spatial contiguity matrix W and the data of both endogenous and exogenous variables, are the necessary inputs to represent an autoregressive spatial process and also build a first spatial regression model. The first reference when we want to build a spatial model is the model of Manski (Manski, 1998); this is the basic specification of a spatial model and has the following structure:

$$Y = \rho WY + X\beta + \theta WX + u \tag{5}$$

And the residual equation is:

$$u = \lambda Wu + \varepsilon \tag{6}$$

A parsimonious approach to represent an autoregressive spatial process is the one that shows a relationship between the dependent variables, the independent variables and their respective relationship with the spatial term. If $\lambda = 0$ in equation (6), then we get the first specification that indicates the starting point on the spatial regression methodology, it is known as the Spatial Durbin model (Lesage and Pace, 2009). This expression is given by:

$$Y = \rho WY + X\beta + \theta WX + \varepsilon \tag{7}$$

In the expression (7) the dependent variable is related to the spatial lagged component of the same dependent variable WY with the vector of estimators ρ, the vector of regressors associated to the explicative variables $X\beta$, the spatial lagged component of the independent variables WX with the vector of coefficients θ and the residual term ε.

If $\theta = 0$ then:

$$Y = \rho WY + X\beta + \varepsilon \tag{8}$$

Equation (8) reflects a model that shows an endogenous interaction effect where the dependent variable Y is an Nx1 vector consisting of one observation on the dependent variable for every unit in the sample (i=1, …, N). Also, the dependent variable is in function of the spatial lag vector Wy with the parameter ρ that reflects the strength of spatial dependence; the model is completed by the exogenous/explanatory variables including the before mentioned constant term vector and the respective regression vector of parameters β and the disturbance vector ε that contains independent, normally distributed terms. Note that if the parameter ρ is not statistically significant, the spatial component will not have an impact on the dependent variable, which means that the model becomes in a normally OLS regression, therefore the importance of this spatial parameter. For understanding the impact of the spatial interaction we could premultiplicate equation (8) by the term $(I - \rho W)^{-1}$ and reorganizing terms we obtain expression (9) as it follows:

$$Y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1}\varepsilon \tag{9}$$

Expression (9) shows a traditional regression model with a spatial impact term $(I - \rho W)^{-1}$. This term shares much similarity with the Leontief inverse matrix in the input-output literature. The Leontief matrix shows that each element outside the main diagonal measures the indirect impact that one sector has on the other, while the main diagonal measures the direct impact plus the indirect impact of the sector itself. This argument is necessary to understand the role of the contiguity matrix W in this type of models (Aroca, 2000). If the term $\rho = 0$, equation (5) becomes in a new model called Spatially Lagged X Model like denotes expression (8):

$$Y = X\beta + \theta WX + \varepsilon \tag{8}$$

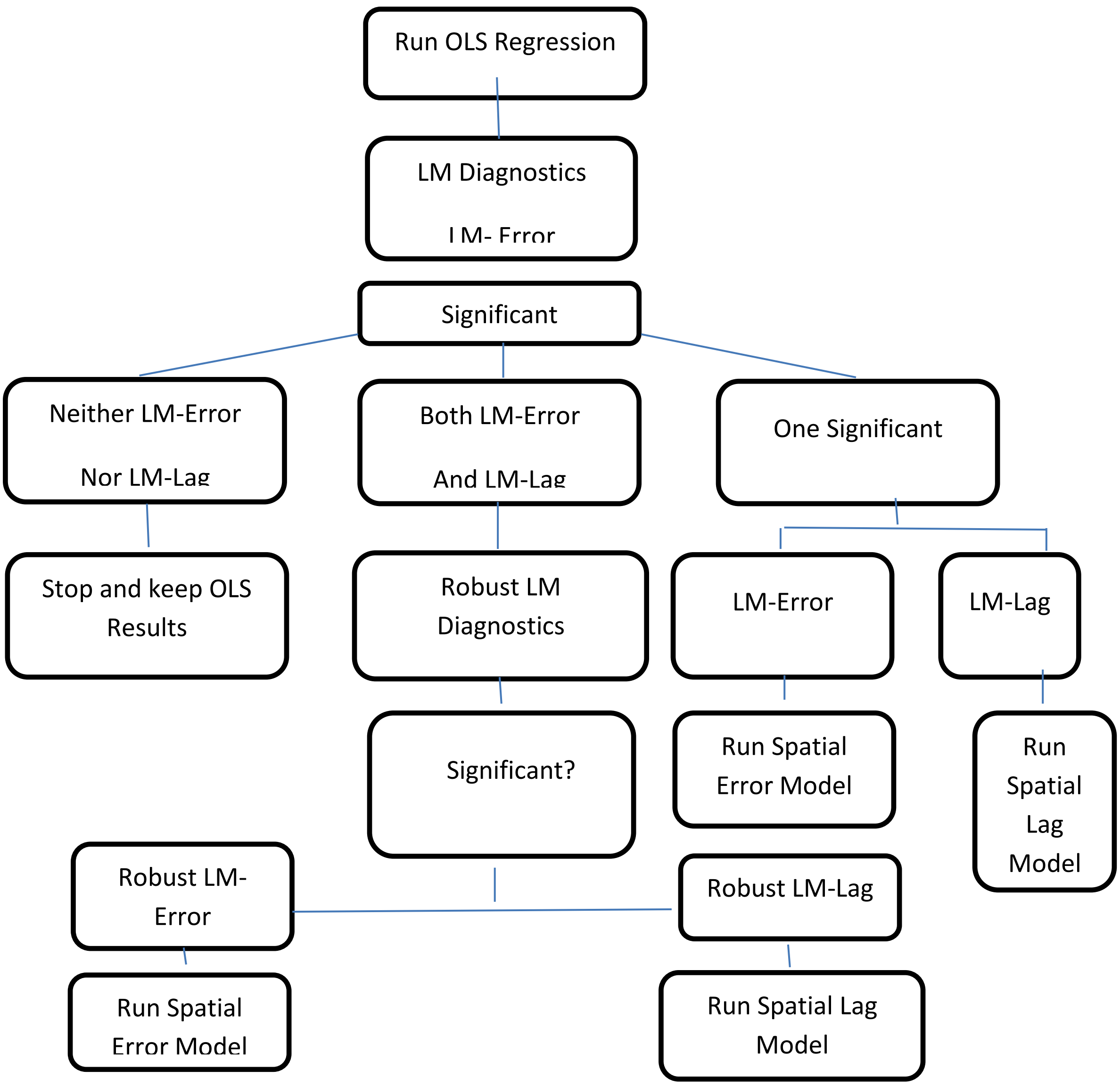
SLX Models are a second way to observe the spatial autocorrelation between neighbors. In this case, the spatial interaction is given when the values of the explicative variables in a region could be related to the value of y in a neighboring region. SLX models produce flexible spatial spillover effects. A Spatial spillover effect is defined as the marginal impact of a change to one explanatory variable in a particular cross-sectional unit on the dependent variable values in another unit, and is derived from the reduced form of a spatial econometric model (Elhorst and Halleck Vega, 2017).

Another case where we can identify spatial correlation would be if $\theta = -\rho\beta$. In this case, the Spatial Durbin Model becomes into a Spatial Error Model that shows spatial interaction on the error term. A demonstration of SEM models could become into regular OLS is showed in Anselin et al.,2003.[1]

Aside of demonstrate spatial interaction with Moran´s I (Local, Global and Scatterplot), this research is based in a methodology of selection criteria to choose which models can be considered for spatial analysis. This methodology shows a path of decision to obtain the most appropriate spatial model (Anselin,1999).

[1] $Let: Y = \rho Wy + X\beta + (-\rho\beta)Wx + \varepsilon$
$(I - \rho W)y - (I - \rho W)X\beta = \varepsilon$
Since: $\lambda = \rho$
$(I - \rho W)y - (I - \rho W)X\beta = (I - \rho W)u$
$Y = X\beta + u$

Figure 2 - Spatial Model Selection Criteria



Source: Adaptation from Anselin, Luc. Exploring Spatial Data with Geoda: A Workbook (2005)

This research will use the model selection criteria exposed in figure 2. It should be noted that if Robust LM-Error and Robust LM-Lag p-values are both significant, the selected model will be the one which has lower p-value on the significance tests.

In a Spatial error model (SEM) the specification is given by equation (5) with no spatial autocorrelation in the endogenous and the exogenous variables ($\rho = \theta = 0$) and equation (6), that also could be written:

$$y = X\beta + (I - \lambda W)^{-1} + \varepsilon \tag{9}$$

The residual term ε is assumed to be independently normally distributed, i.e. $\varepsilon \sim N(0, \sigma^2 I)$. Solving equation (9) for ε we obtain:

$$\varepsilon = (I - \lambda W)(y - X\beta) \tag{10}$$

And the Jacobian:

$$J = \left|\frac{\partial \varepsilon}{\partial y}\right| = |I - \lambda W| \tag{11}$$

Based on equations (9) (10) and (11), it is possible to create a log-likelihood function for the dependent variable. Thus, the Spatial Error Model (SEM) is obtained by adding the term ln $|I - \lambda W|$ to the log-likelihood function of the standard regression model:

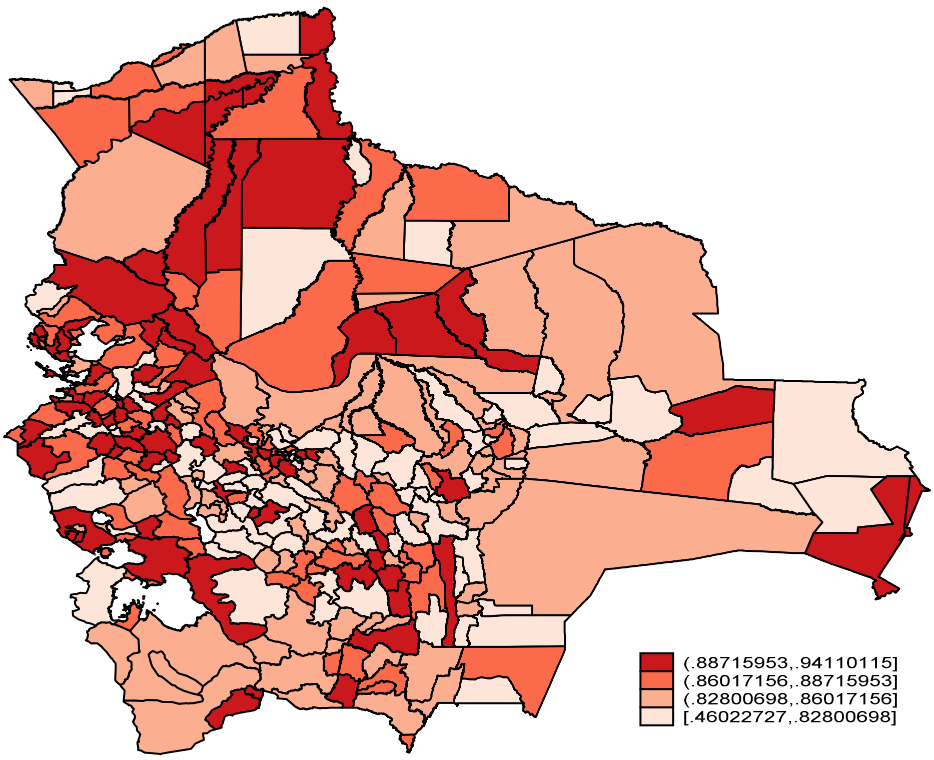
$$\begin{aligned} lnL(\beta, \lambda, \sigma^2 | y, X) \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}ln\sigma^2 + \ln|I - \lambda W| \\ &- \frac{1}{2\sigma^2}(y - X\beta)'(I - \lambda W)'(I - \lambda W)(y - X\beta) \end{aligned} \tag{12}$$

Maximizing the log likelihood function of equation (12) equals to minimizing the sum of transformed squared errors $\varepsilon'\varepsilon$ corrected by log of the Jacobian expression obtained in equation (11). With this procedure, the flowchart of spatial error models resumes in: i) Make the OLS regression and compute the residual vector ii) Maximize the likelihood function to obtain the autoregressive parameter λ. iii) Make a GLS estimation using the estimated parameter λ. iv) Compute the GLS residual vector analyzing the convergence criterion. v) Compute ML estimator for the error variance given the residual vector and the autoregressive vector. To carry out this process, STATA and Geoda software are used on this research.

3. An exploratory analysis of spatial data

In this section we perform an exploratory spatial data analysis (ESDA) that consists in a graphical-statistical representation the endogenous and the exogenous variables of the spatial model. The ESDA represents a scrutiny of classic statistical methods like thematic map, histogram and dispersion diagram both of the exogenous and endogenous variables and is an indispensable instrument when making the first approximations to the study of the structure of socio-spatial information in a given area of study (Buzai and Baxendale, 2009).

Figure 3 - Bolivian School Attendance rate (%)

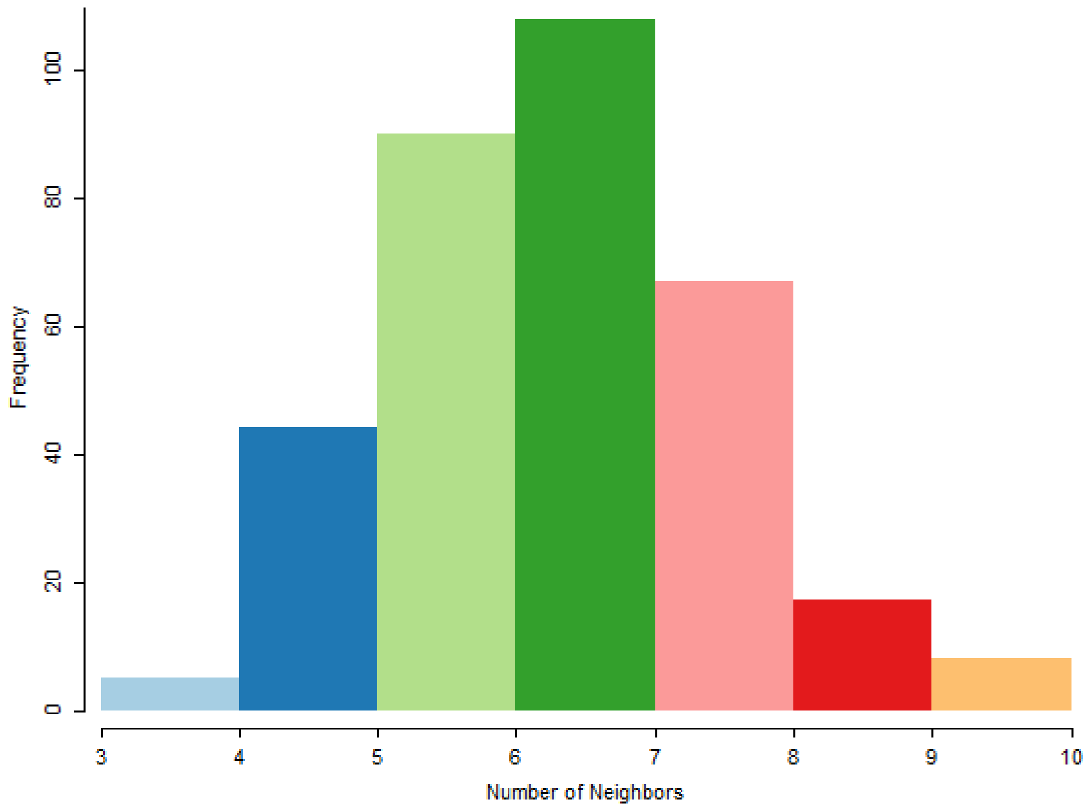


Source: Own elaboration based on INE data, 2012

Thematic maps represent the spatial distribution of the variable. To this purpose, figure 3 shows thematic map of Bolivia with the school attendance rate for each of the 339 Bolivian municipalities. This graphic shows the presence of "clusters" or graphical evidence that make us think that there is spatial correlation on the dependent variable (school attendance rate), but this is only a "clue" of spatial correlation. To be sure, it is important to make another graphical and statistical test.

ESDA suggests to build a histogram of frequencies of the dependent variable. Figure 4 shows the number of neighbors that share each municipality; we have five municipalities sharing borders with at least three neighbor regions, and eight observations sharing borders with 9-10 neighbor regions. It is worth noting that to determine the number of neighbors for each region the queen contiguity matrix is used, as explained in section (2).

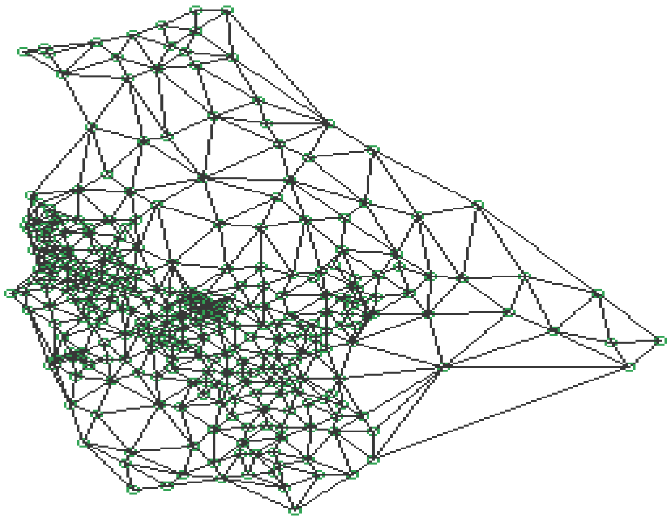
Figure 4 - Histogram number of neighbors (Bolivian Municipalities)



Source: Own Elaboration based on INE municipal data, 2012.

To understand how neighboring connectivity works, we use a connectivity graph that show us precisely how the neighbors are connected. Figure 5 allows to observe the spatial interaction between municipalities.

Figure 5 - Connectivity Map of Bolivian Municipalities



Source: Own Elaboration based on INE municipal data, 2012.

The proposed model on this research uses as explicative variables the level of education, electric power coverage, labor market's rate of participation, child labor rate, percentage of particular households, migration rate and percentage of rural population on every municipality. Briefly, the explanatory power of a variable can be quantified via the correlation coefficient. For this purpose, Table 1 in appendix section shows correlation matrix between the dependent and the exogenous variables.

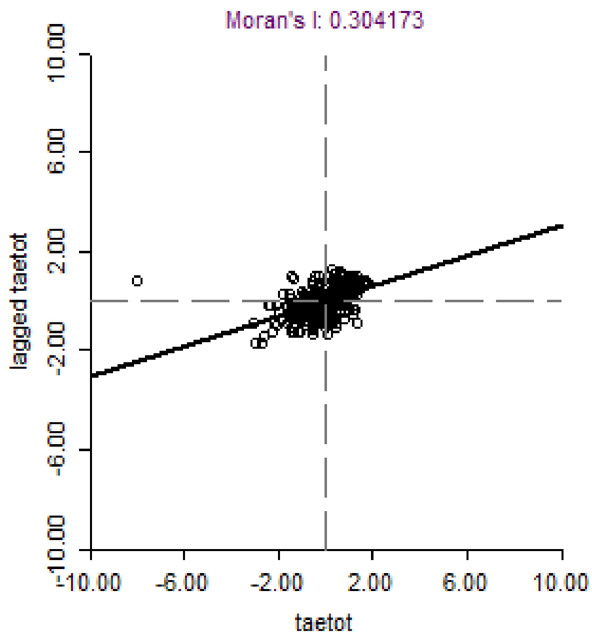
The first explanatory variable is educational level of the municipality; we estimate this variable using the proportion of the population that has higher education. The main hypothesis is that this variable has a positive impact on the dependent variable because a more educated society should have higher percentage of school attendance rate (Cid, 2011).

To measure the urbanization level of the municipality we use two proxy variables: electric power coverage the and percentage of particular households. Both variables are expected to be theoretically positive and statistically significant to the model. Electric power coverage makes it easy to people study during the night or to those who have and allocate more time to that activity. At the same time, it facilitates access to Information and Communication Technologies, ICT (Kanagawa and Nakata, 2008). The other variable used to measure the urbanization level is the percentage of particular households; this variable shows a positive relationship between the urbanization level of a municipality and the percentage of particular households, so the expected sign of this variable should be positive.

Variables like migration rate, child labor rate and percentage of rural population on every municipality should be important to this model, because they represent a measure of the socio-economic dynamic context of the municipality, the impact of these variables on the school attendance should be negative. Higher migration rates show higher risk for scholar absenteeism and being held back (Vargas and Camacho, 2014). Child Labor rate have a negative incidence on school attendance rate, evidence from many countries suggests a close negative association between these two variables (Khanam, 2008. Kumar and Saqib, 2017).

School Attendance rate should show any hint of spatial correlation that will be tested subsequently. For that purpose, figure 6 evaluates Moran's I that could give a clue of spatial correlation (Anselin, 2000). It should be noted that, having a higher (or lower) value than zero in Moran's I does not necessarily implies spatial correlation. For evaluating spatial correlation, this value has to be different than zero and has to be also statistically significant in the spatial model.

Figure 6 - Moran's I School Attendance rate for Bolivia



Source: Own Elaboration based on INE municipal data, 2012.using GeoDa software.

4. Model estimation

In this section, the specification of the model is examined as well as the first results of the OLS and SEM regressions. Following Anselin's methodology, we estimate the model and evaluate the statistical significance of the variables as well as the significance of the spatial component through of Moran's I test. This result will determine if the spatial correlation observed before has an important impact on the model, and also which spatial(s) model(s) is(are) the most appropriate(s) for the research.

Theoretically, as it is explained in previous section, school attendance rate is positively related to educational level of the municipality, electric power coverage, percentage of particular households, and formal market labor participation. The relationship should be negative with the other variables, like migration rate, child labor rate and percentage of rural population. Based on figure 2 that shows spatial model selection criteria, once the LM diagnostics (i.e. both the parameters and the spatial component) are statistically significant then it is possible to pick one spatial model to represent the current relationship between the dependent variable, the independent variable and the spatial component. It should be noted that the diagnostics for spatial dependence showed in table 2 are constructed using row-standardized weights in the spatial matrix, for that reason the evidence that we found is that there exists significant spatial interaction in the school attendance rate of Bolivian municipalities. As the theoretical review underlines and as it is shown in table 2 in the appendix, the coefficient of spatial interaction (Moran's I) is statistically significant which means that the null hypothesis of no spatial correlation is rejected. In other words, the spatial component has a significant impact on the dependent variable.

Also, table 2 shows Lagrange Multiplier tests (Lag and error) and results show that both models can be executed. Nevertheless, Anselin's spatial methodology indicates that if both LM tests are significant, then the robust tests should be observed, and as it is explained before, spatial error model (SEM) and (SARMA) model are the indicated approaches in analyzing the spatial impact on the dependent variable (significant spatial dependence).

Since we found that Lagrange Multiplier test for Spatial Error Model is statistically significant. Anselin's methodology indicates that SEM model should be estimated. Table 3 in the appendix shows the first results of the OLS and SEM regression; we found statistical evidence that lambda spatial component has a significant impact on the school attendance rate. Also, the relationships between the dependent variable and the explanatory variables are in accordance with those indicated by the economic theory. With the exception of fertility rate, all variables are statistically significant to the SEM model and the signs show the expected behavior that is explained on the previous section.

The obtained results show that levels of higher education are important for reducing the school absenteeism into Bolivian municipalities, and a more educated society is associated with a long-run growth economic rate (Hanushek & Woessmann, 2008).

Electric power coverage and the percentage of particular households have a positive and significant relationship with the dependent variable, although, there are different methodologies to measure the level of urbanization (Wen & Ren, 2016. Zhao et al., 2016). These relationships are important to underline that the level of urbanization becomes an important variable and have a significant impact into the school attendance rate.

Labor market participation has a significant impact on the dependent variable. It is important to remark that this is a key variable on the model, because it shows the percentage of (formal) labor market participation; the estimated model shows that there is a positive relationship between this variable and the dependent variable. On the other hand, informality affects the school attendance rate and also economic growth (Urdinola and Semlali, 2010).

Child labor rate is the variable with higher impact in the model, it shows a negative relationship with the dependent variable and its important because it is a widespread phenomenon. Approximately 20% of children between 7-14 years are part of the Bolivian labor force (U.S. Department Labor).

Migration rate have negative impact over school attendance rate; empirical evidence suggests that living in a migrant household lowers the chances of children completing high school (Mackenzie and Rapoport, 2006). The percentage of rural population also has a negative impact on school attendance. This impact shows a measure of social exclusion in rural areas where the school rate attendance is lower.

5. Conclusions

Spatial econometrics has become a prominent topic in the recent scientific literature Emphasis has been made that spatial analysis is a modeling reliable and robust tool which allows to explain the behavior of variables linked to space. Interpretations of estimated parameters and inferences regarding the modeled relationships required steady-state view, where changes in the explanatory variables lead to simultaneous feedback that produce a new steady state equilibrium (Lesage,2008).

As it was clarified in the introduction of this research, the main goal of this paper was to evaluate the impact of the determinants over school attendance rate, as well as the incidence of the spatial component in this relationship. The obtained results proved that spatial component has a significant impact over the school attendance rate in Bolivian municipalities. Also, spatial diagnostics were helpful to determine that Spatial Error Model (SEM) was the most accurate specification to explain the dependent variable's behavior. The explanatory variables used on this research have a significant effect over the dependent variable; the signs and the established relationships were those expected, as the theoretical framework indicted before. It is important to remark the incidence of child labor rate on school rate attendance; this variable is the one that has the higher impact in the model and government should work on effective policies that diminish this variable's percentage in the Bolivian municipalities.

References

Anselin L and Florax R. (1995a) “Introduction. In L. Anselin and R. Florax (eds.) New Directions in Spatial Econometrics”, *Berlin: Springer-Verlag* pp. 3 – 18.

Agudelo G., Franco L.E. and Franco L.C. (2015) “Aplicación de la econometría espacial para el análisis de la miseria en los municipios del departamento de Antioquía” *Semestre Económico*, V. 18 N.37 pp. 103 - 128

Angel-Urdinola D. and A. Semlali (2010) “Labor Markets and School-to-Work Transition in Egypt: Diagnostics, Constraints, and Policy Framework” *Munich Personal RePEc* Archive pp.1 -25

Anselin, L. (1999). “*Spatial econometrics*”, An updated review. Regional Economics Applications Laboratory (REAL), University of Illinois, Urbana-Champaign.

Anselin, L. (2000). Rao’s score test in spatial econometrics. Journal of Statistical Planning and Inference (forthcoming).

Anselin, L. (2003) “Spatial Econometrics” *A companion to Theoretical Econometrics* C.14 pp. 310 - 327

Anselin, L., and H.H. Kelejian (1997). “Testing for spatial error autocorrelation in the presence of endogenous regressors.” *International Regional Science Review* V.20, pp. 82-153.

Aroca P. (2000) “Econometría espacial: Una herramienta para el análisis de la Economía Regional” *IDEAR - Universidad Católica del Norte* pp. 1- 19

Buzai G. and Baxendale C. (2009) “Análisis socio-espacial con sistemas de información geográfica marco conceptual basado en la teoría de la geografía” *Grupo de Estudios sobre Geografía y Análisis Espacial con Sistemas de Información Geográfica (GESIG)* pp. 4-17

Cetrángolo, O., Curcio J. and Calligaro, F. (2017) “Evolución reciente del sector educativo en la región de América Latina y el Caribe” *Publicación de las Naciones Unidas CEPAL – Serie Macroeconomía para el desarrollo* N°191 pp. 9 - 85

Chen Y. (2008) "On the Four Types of Weight Functions for Spatial Contiguity Matrix" Department of Geography, *College of Environmental Sciences,* Peking University, Beijing 100871 pp. 1-8

Cid, J. C. (2011), "Aplicación de un modelo de econometría especial a datos agregados de asistencia escolar en Argentina" Instituto de *Investigaciones Económicas, RD191*. pp. 1-16

Elhorst J.P. and Halleck Vega S.M. (2017) "The SLX model: Extensions and the sensitivity of spatial spillovers to W" *Papeles de Economía Española* 152: pp. 34-50.

Ertur C. and Koch W. (2007), "Growth, technological interdependence and spatial externalities: theory and evidence" *Journal of Applied Econometrics* V.22 pp.1033-1062

Haining, R. P. (2009). "Spatial Autocorrelation and the Quantitative Revolution" *Geographical Analysis* 41, pp. 364 – 374

Haining R. (2003), "*Spatial Data Analysis: Theory and Practice*" University of Cambridge, Cambridge.

Haining, R. P. J. Law, and D. A. Griffith. (2009). "Modelling Small Area Counts in the Presence of Overdispersion and Spatial Autocorrelation." *Computational Statistics and Data Analysis* 53, pp. 2923–37.

Hanushek E. and Woessmann, L. 2008 "The Role of Cognitive Skills in Economic Development" *Journal of Economic Literature*, V. 46:3, pp. 607–668

Paelinck, J. and Klaassen L. (1979), *Spatial Econometrics Volume 1 de Studies in Spatial Analysis,* University of Michigan, Michigan.

Kanagawa M. and Nakata T. (2008) "Assessment of access to electricity and the socio-economic impacts in rural areas of developing countries" *Energy Policy*, V. 36, issue 6, pp. 2016-2029

Kumar, A. and Saqib, N. (2017). "School Absenteeism and Child Labor in Rural Bangladesh. The Journal of Developing Areas" *Tennessee State University*

College of Business. Retrieved June 23, 2019, from Project MUSE database. V. 51(3), pp. 299-316.

LeSage J. P. (2008), "An introduction to Spatial Econometrics", *Revue d'économie indutrielle* pp. 19-42

LeSage J. P. and Pace K. (2004a), "Introduction, « Advances in Econometrics: Volume 18: Spatial and Spatiotemporal Econometrics »", *Oxford: Elsevier Ltd* (eds.1-32)

LeSage, J.P. (2008) "An Introduction to Spatial Econometrics", *Revue d'économie industrielle N°123* pp. 19-44

Loonis V., M.-P- de Bellefon, and R. Le Gleut, (2018). "Codifying the neighbourhood structure. Handbook of Spatial Analysis." *Eurostat. E.* Gros.

Mackenzie D. and Rapoport H. (2006) "Can migration reduce educational attainment? Evidence from Mexico" Journal of Population Economics, vol. 24, issue 4

Moran, P.A.P. (1948). "The interpretation of statistical maps". *Biometrika* V. 35, pp. 60 - 255

Moreno R. and Vayá E. (2000) "*Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial*" Universitat de Barcelona, Barcelona.

Paz, J. A. y Cid, J. C. (2012). "Determinantes de la asistencia escolar de los jóvenes en la Argentina." *Revista Electrónica de Investigación Educativa*, V.14 N°1 pp. 136 -152

Pinske J. and Slade M. (1998) "Contracting in space: An application of spatial statistics to discrete-choice models" *Journal of Econometrics,* V. 85, issue 1, pp. 125-154

Tobler W.R. (1979) "*Cellular Geography. In: Gale S., Olsson G. (eds) Philosophy in Geography. Theory and Decision Library"* (An International Series in the Philosophy and Methodology of the Social and Behavioral Sciences), vol 20. Springer, Dordrecht

Vargas E. and Camacho E. (2014) "Changing Schools? Absenteeism and Educational Lag among Children Recently Migrated from the United States to Mexico"

Centro de Investigaciones sobre América del Norte. Universidad Nacional Autónoma de México

Wen F. & Ren L. (2016) "The Measurement of Urbanization Level Based on Entity Space: A Case Study of JingJinJi Region, China" *EDP Sciences* pp. 1 - 8

ZHAO Min, *CHENG Weiming, LIU Qiangyi, WANG Nan. "Spatiotemporal measurement of urbanization levels based on multiscale units: A case study of the Bohai Rim Region in China[J]". *Journal of Geographical Sciences*, V. 26 (5) pp. 531- 548.

Appendix

Table 1 - Correlation Matrix of the dependent and independent variable

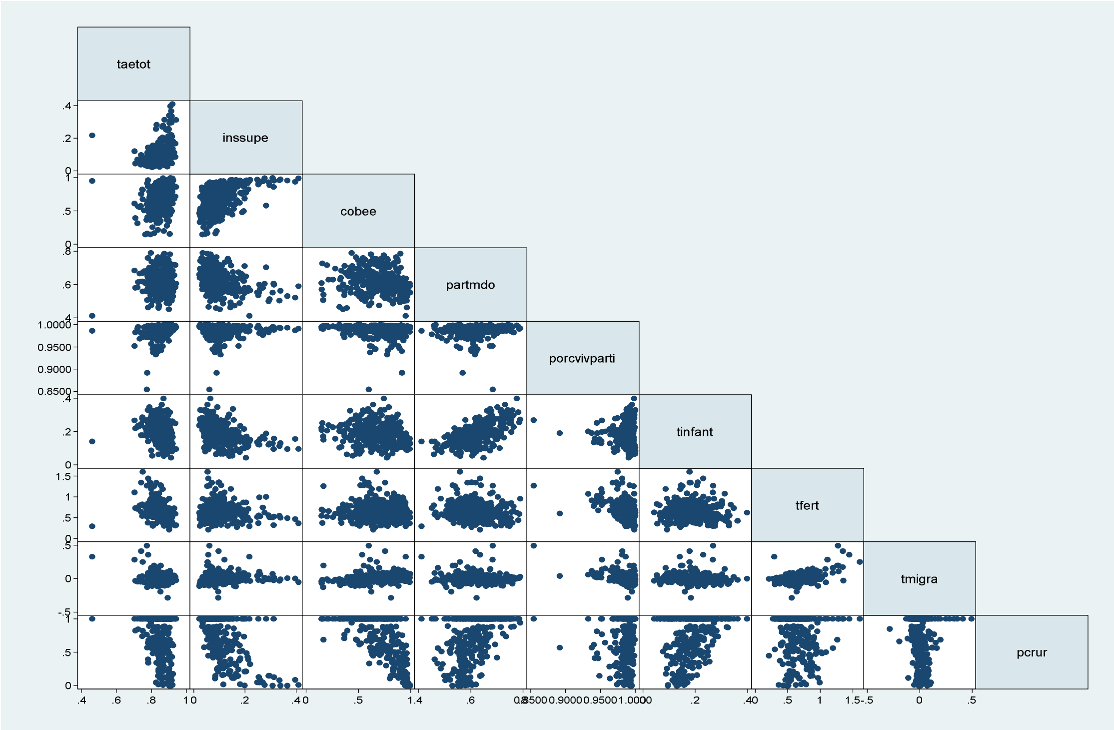


Table 2 - Diagnostics for spatial dependence for weight matrix (row-standardized weights)

Test	Value
Moran’s I	6.1443***
Lagrange Multiplier (lag)	15.5537***
Robust LM (lag)	0.7023
Lagrange Multiplier (error)	31.3940***
Robust LM(error)	16.5426***
Lagrange Multiplier (SARMA)	32.0963***

*** Significance level 1% ** Significance level 5% *Significance level 10%

Table 3 - Econometric Results for dependent variable School Rate Attendance

Variable	MCO Coefficient	SEM Coefficient
R- squared	0.4729	0.5329
Adjusted R- squared	0.4601	
Constant	0.2101	0.3027*
	(0.1620)	(0.1548)
Educational level of the municipality (higher education)	0.0516	0.0787*
	(0.041)	(0.0419)
Electric power coverage	0.0519***	0.0519***
	(0.014)	(0.014)
Labor Market Participation	0.3543***	0.3927***
	(0.0398)	(0.0464)
Percentage of particular households	0.5071***	0.3818**
	(0.1616)	(0.1517)
Migration rate	-0.1467***	-01838***
	(0.0343)	(0.0343)
Fertility Rate	-0.029***	-0.0208
	(0.0117)	(0.0127)
Child labor Rate	-0.4358***	-0.0466***
	(0.0425)	(0.0477)
Percentage of rural Population	-0.0218***	-0.0154*
	(0.0092)	(0.0093)
LAMBDA	-	0.4300***
		(0.0712)

*** Significance level 1% ** Significance level 5% *Significance level 10%