Reception: October 11, 2016 | Acceptance: June 12, 2016

SPATIAL ANALYSIS OF ILLITERACY IN MUNICIPALITIES OF CHIAPAS

—

Gerardo Núñez Medina

Director of research and evaluation of the
Council for Research and Evaluation of the Social Policy of the State of Chiapas
gerardo.nm1@gmail.com

— Abstract—

The purpose of this study is to analyze some determinants of spatial behavior of illiteracy in municipalities of Chiapas to 2010, considering a spatial regression model whose independent variables are: the proportion of speaking population of indigenous language, the percentage of people living in extreme poverty and current family income adjusted by Gross National income. The central hypothesis holds that municipal illiteracy rates are not distributed randomly, but present spatial patterns of concentration and dispersion. The results put on the table the importance of the availability of family income to enable the generation of writers basic literacy skills, as a key to improving school enrollment levels and reduce inequalities between indigenous and mestizo populations.

Keywords

Spatial heterogeneity, autoregressive models, direct and indirect effects, indigenous peoples, extreme poverty.

The objective of this study is to examine the main determinants of the spatial behavior of illiteracy in municipalities in Chiapas in 2010, considering independent variables such as the proportion of indigenous language-speaking population, percentage of the population living in extreme poverty, and municipal family income adjusted according to Gross National Income. The central hypothesis argues that illiteracy rates can be explained as a result of the combination of levels of poverty, indigenism and incomes present in the municipalities, and that these factors are not randomly distributed, but have patterns of concentration and spatial dispersion that strengthen their effects through the interaction between neighboring municipalities.

The importance of this research lies in three fundamental factors: (1) Chiapas is one of the states with the highest incidence of illiteracy in the entire country, according to 2010 census figures; (2) Chiapas is the second entity in the country in terms of the number of indigenous population and the first in indigenous monolingual population; (3) Chiapas is the entity with the highest proportion of the population in poverty and extreme poverty, since measurements began in Mexico in the 1990s, so the analysis of the spatial interaction of these factors results in an essential exercise to know the importance, magnitude and effects of illiteracy at the municipal level in terms of the variables and their spatial interaction.

DATA

The data used for this work were collected from demographic and socioeconomic sources dependent on the Mexican government and from international agencies. The first instance was the general census of population and housing 2010, which was compiled by the National Institute of Statistics, Geography and Informatics (INEGI, 2010), which compiled information on the total municipal population, indigenous population and illiterate population.

The data concerning literacy status were collected based on the guidelines established by INEGI, which defines illiteracy as the population of 15 years or older who declare that they do not know how to read or write a message. Under this criterion, individuals are classified as literate or illiterate. For the purpose of this work, the literate population will be the one who, if 15 years or older, declares that they can read and write a message. That is, they have acquired a basic capacity to access new knowledge, which makes it possible to improve their possibilities of social integration, wealth generation and access to health and education services.

Data on municipal extreme poverty levels were obtained from the estimates published by the National Evaluation Council (CONEVAL, 2014), which were carried out using data obtained from the socioeconomic conditions module of the National Survey of Income and Expenditure Of Households (ENIGH) for all municipalities in Chiapas. While municipal income estimates were obtained from the United Nations Development Program (UNDP, 2014) and are the standardized income component used to calculate the 2010 municipal human development index for the municipalities of Chiapas.

EXPLORATORY ANALYSIS OF SPATIAL DATA

The exploratory analysis of spatial data (AEDE) usually begins with the application of spatial autocorrelation tests to each of the variables involved, to which it is necessary to define an array of spatial contiguities (Chasco, 2003). The spatial contiguity matrix is defined as a binary matrix, whose values depend on whether or not the spatial units are neighbors, so the neighborhood criterion is fundamental. In this case it was considered that two spatial units were neighboring if the distance between their municipal seat was less than or equal to the maximum distance defined between all the municipal seats of the entity, and that two municipalities are neighbors if they are in a smaller or equal radius to 54.6 km away.

Table 1. Spatial autocorrelation, global Moran index.

Variable	Moran I statistic	Expectation	Variance
Analf	0.493	-0.009	0.001
Phlin	0.537	-0.009	0.001
Extremos	0.469	-0.009	0.001
Iingreso	0.260	-0.009	0.001

Source: own elaboration with data of INEGI and CONEVAL

The statistical spatial correlation most commonly used is the Moran index in its global version measures autocorrelation based on the locations and the values of a variable x for all regions simultaneously, i.e., it is a measure of autocorrelation, defined similarly to the Pearson correlation coefficient (Anselin, 1995), with the provision that establishes the hypothesis that the analyzed variable is distributed randomly in space. When the p value of the statistic is significant, one can assume the presence of a pattern of spatial correlations. The global Moran index is estimated from:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

Where W_{ij} represents the elements of the spatial weights matrix, and, $\bar{x}$ the mean of the variable x. It is important to note that the autocorrelation coefficient measures the degree of association of the variable x with respect to their neighbors..

The Moran index in its local version allows the identification of spatial conglomerates in five categories: (1) low-low: spatial units with less than average value, surrounded by units with values below the average of the attribute of interest. These spatial units correspond to clusters called cold zones. (2) low-high: space units with below-average value surrounded by units with values above average. (3) high-low: space units with above-average value surrounded by units with values below average. (4) high-high: space units with above average value, surrounded by units with values above the average. These units correspond to conglomerates called hot zones. And (5) no data: the set of spatial units where the variable of interest is not significantly correlated with neighboring values (Cliff & Ord, 1981).

The results of the estimates concerning the global spatial correlation levels are presented in Table 1 and indicate the presence of high levels of positive spatial autocorrelation for all variables analyzed, indicating the existence of a direct association between municipalities with high levels of illiteracy surrounded by municipalities that in turn present levels of illiteracy above the state average. The same situation occurs in the case of the percentage of indigenous language-speaking populations, extreme poverty and income.

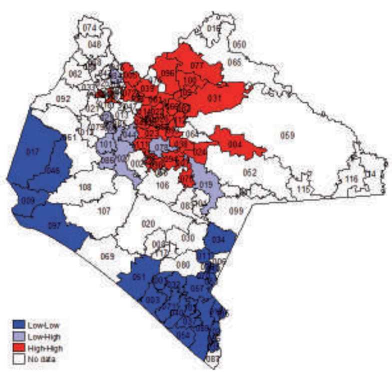
The presence of autocorrelation makes it possible to assume the existence of spatial structures capable of explaining municipal illiteracy levels in terms of possible associations with variables that in turn present high levels of spatial autocorrelation, especially when it occurs in the same regions or conglomerates (Getis & Ord, 1992), as was the case of the indigenous language-speaking population of the state of Chiapas.

Illiterate Population

Literacy can be understood as a process through which individuals acquire the ability to communicate in written form, which constitutes an element that enables the continuous acquisition of abilities and skill of all kinds. Literacy empowers people to develop advantages that will eventually enable them to improve their living conditions. The intrinsic relationship between the ability to read and write properly and the ability to acquire new skills plays an essential role in generating economic growth and reducing inequalities (UNESCO, 2008).

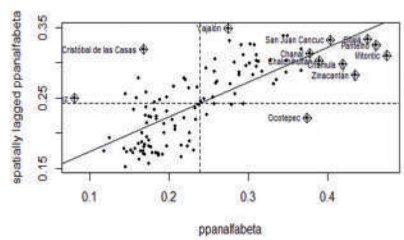
On the other hand, the lack of literacy skills is a key factor in explaining the serious shortcomings associated with extreme poverty, discrimination and social exclusion, in which specific social groups are involved, such as a large proportion of the population that speaks an indigenous language; A population that is presumably the lowest income in the state of Chiapas.

Mapa 1: porcentaje de población analfabeta municipal, Chiapas 2010



Fuente: elaboración propia con datos del INEGI

Gráfica 1: índice de Moran de población analfabeta municipal, Chiapas 2010



Fuente: elaboración propia

Map 1 shows the spatial dispersion of the municipal illiterate population of Chiapas in the year 2010, in the same map the presence of a hot conglomerate (in red) can be observed conformed by the municipalities with the lowest levels of literate population and that in turn are surrounded by municipalities with low levels of literacy. This conglomerate is formed by the municipalities of Simojovel, San Andrés Duraznal, Santiago el Pinar, Bochil and Larrainzar, among others (see map 1).-Municipalities that make up the conglomerate with high concentration of illiterate population in Chiapas.

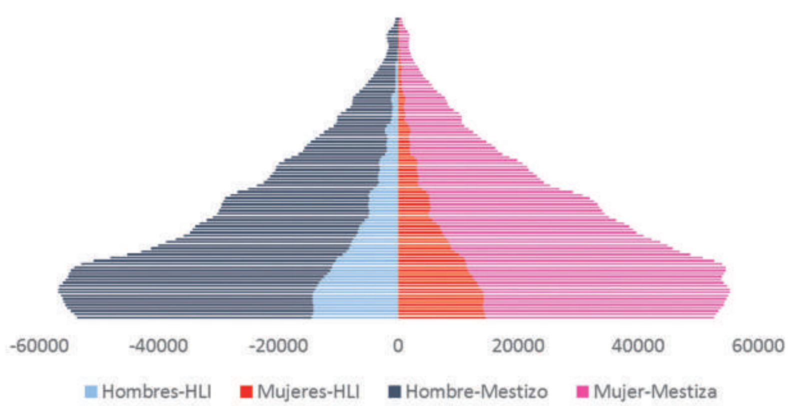
The slope of the regression line observed in figure 1 represents the value of the Moran global spatial autocorrelation index for the proportion of municipal illiterate population, which shows an important level of positive autocorrelation (0.49), which implies that the proportion of the literate population is spatially concentrated. The graph of spatial association in quadrant I show an important grouping of municipalities that correspond to the red zones of the map 1.

The Indigenous Population in Chiapas

The indigenous population can be considered, from a historical and sociocultural perspective, as: "those direct descendants of the peoples who inhabited America since before the arrival of the Spaniards in the fifteenth century, who own a language and culture of their own and share forms of life and particular worldviews, differentiated from Western views "(Bello and Rangel, 2002: 40). In addition to the above, it is important to recognize the presence of large groups of indigenous people who have been culturally assimilated with mestizos with a consequential loss of language, who have also been displaced from their territories and now inhabit the poverty belts of large cities where the vast majority are a poor, marginalized population with no access to formal employment, education and health systems.

Although correct, the above definition is impractical due to the technical difficulties to implement its measurement, so it was decided to use the linguistic criterion established by INEGI, which defines as indigenous population those persons of five years of age or older who answered affirmatively to the question of if they spoke some indigenous language. In this way the indigenous population was identified based on the population census, using the linguistic criterion, which refers to the status of speaker. The indigenous language speaker population (ILSP) is made up of those individuals residing in Chiapas, five or more years of age, who claimed to speak some indigenous language in 2010.

Graph 2. Structure by age and sex of the population of Chiapas and the indigenous language-speaking population, Chiapas 2010

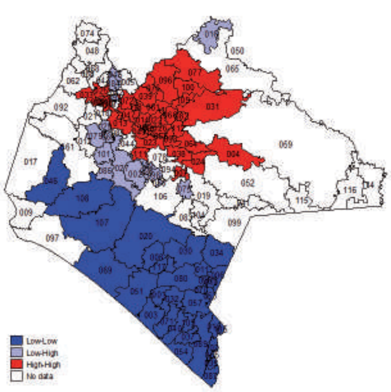


Source: own elaboration with data of census 2010, INEGI

The ilsp of Chiapas represented 23.5% of the total population of the state (see graph 2). Despite the fact that practically one of every four inhabitants can be considered part of the indigenous population, they are still considered a minority and therefore excluded from power and decision-making spaces (INEGI, 2010). Although in many cases, the indigenous population represents more than 60 percent of the population, in a municipality, the president and municipal authorities are usually of mestizo origin. Just over a million inhabitants of Chiapas speak some indigenous language. Of these, 66.2% speak Spanish, however, one third of the indigenous population is monolingual, which directly impacts their ability to access and successfully complete the different levels of the education system, which is reflected in the significant levels of illiteracy, which will reach 21% of the general population by 2010.

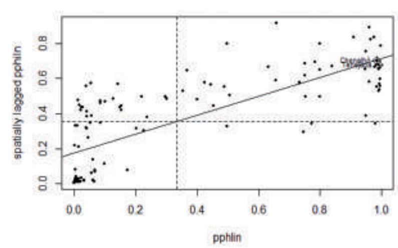
The spatial dispersion of the municipal indigenous population of Chiapas in 2010 can be seen in map 2, the same one where a is a hot conglomerate (red color) can be observed which is formed by the municipalities with the highest levels of indigenous population and that in turn are surrounded by municipalities with high levels of indigenous population. This conglomerate is located in the same area of the state in which the conglomerate of municipalities that presented high levels of concentration of illiterate population is found.

Mapa 2: municipios según condición de indigenismo, Chiapas 2010



Fuente: elaboración propia con datos del INEGI

Gráfica 3: índice de Moran de población indígena municipal, Chiapas 2010



Fuente: elaboración propia

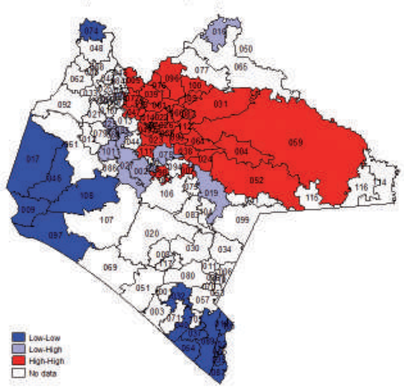
The slope of the regression line observed in Figure 3 represents the value of the global Moran index of spatial autocorrelation for the proportion of municipal indigenous population, which shows a significant level of positive autocorrelation (0.53), which implies that the proportion of Indigenous populations are spatially concentrated. Quadrant 1 of the graph of spatial association shows an important grouping of municipalities that correspond to the red zone of Map 2.

Population in extreme poverty

The definition of population in poverty for Mexico was established by the National Council for the Evaluation of Social Development Policy (CONEVAL, 2014: 26), which states that individuals in multidimensional poverty are those who "... are not guaranteed the exercise of at least one of their rights for social development, and if their income is insufficient to acquire the goods and services they require to meet their needs." For measurement purposes, poverty is quantified in two dimensions: (1) Economic welfare, measured in terms of current income, and (2) Social rights, measured in terms of access to education, health, social security, food, housing and its services.

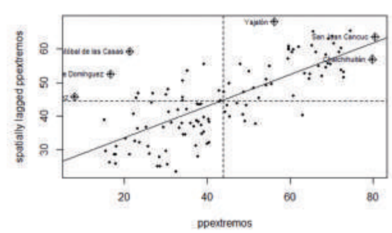
In addition, a person is in extreme poverty when he or she suffers from three or more deficiencies related to their social rights, and their income is below the minimum welfare line. That is, people living in extreme poverty have such low incomes that they cannot acquire the nutrients necessary to maintain a healthy life.

Mapa 3: municipios según distribución de población en pobreza extrema, Chiapas 2010



Fuente: elaboración propia con datos del INEGI

Gráfica 4: índice de Moran de población en pobreza extrema, Chiapas 2010



Fuente: elaboración propia

According to figures from CONEVAL in 2010, at the national level, the population in poverty was 52.1 million people, with 12.8 million of them in extreme poverty. Chiapas occupied the first place in percentage of population in poverty and in extreme poverty; 78.5% of the population of Chiapas was in poverty, of which 1.88 million were in extreme poverty, which represented 38.3% of the total population of the state. At the municipal level, the lowest percentage of extreme poverty was 7.9% in Tuxtla Gutierrez and the highest was for San Juan Cancuc with 80.5%, one of the poorest municipalities in the country.

The dispersion of municipal poverty in Chiapas to 2010 can be seen in Map 3, which shows a clear concentration of levels of extreme poverty in the municipalities of the state where the largest proportion of illiterate population and two cold conglomerates - one around the municipality of Tapachula and another around the municipalities of Arriaga and Tonalá. It is important to consider that the slope of the regression line (graph 4), which represents the value of the overall Moran index for the proportion of the population in extreme municipal poverty, shows a significant positive autocorrelation level of 0.46, which implies that the proportion of the population in extreme poverty is spatially concentrated. Quadrant 1 of Figure 4 shows the importance of the municipalities of San Juan Cancuc, Chalchihuitán and Yajalón that correspond to the red areas of the map 3.

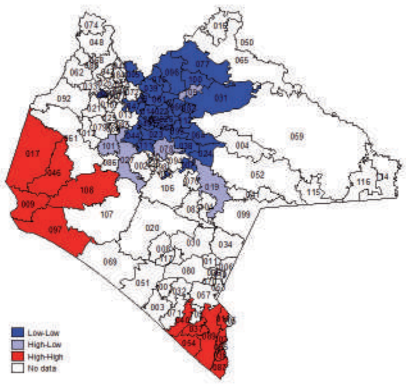
Municipal Income

In order to use a comparable and standardized measure of income at the municipal level, we used the income index, which is one of the three components that make up the Human Development Index (HDI) designed by the research office of the United Nations Program For Development (UNDP, 2014). The income index reflects the ability to access resources that allow individuals to enjoy a decent life. This represents an estimate of the current income available to families at the municipal level, and this is adjusted according to the Gross National Income (GNI). The calculation is based on the estimation of the current income available to families at the municipal level, which is adjusted to the INB from the System of National Accounts of INEGI. This is expressed annually in US dollars in 2010 (UNDP, 2014).

The dispersion of the municipal income in Chiapas in 2010 can be seen in map 4, which shows a clear concentration of high income levels in two areas of the coast, one that surrounds the municipality of Tapachula and the other around the Municipalities of Arriaga and Tonalá. On the other hand, the municipalities in the area of the state where the largest proportion of illiterate and indigenous population in Chiapas is concentrated are, in turn, a conglomerate with a low income concentration.

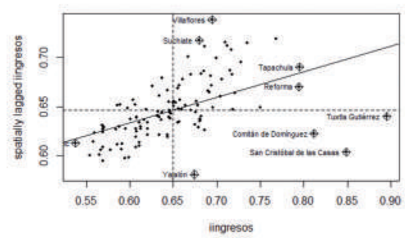
The slope of the regression line observed in Figure 5 represents the value of the Moran global spatial autocorrelation index for the municipal income index, which shows a significant level of positive autocorrelation (0.26), which implies that the proportion of the Illiterate population is spatially concentrated. The graph of spatial association in quadrant IV presents some municipalities whose behavior is interesting, as is the case of Yajalón, an indigenous municipality, located in an area with high levels of extreme poverty and very low income that presents curiously high income levels , in relation to the average of its neighbors.

Mapa 4: municipios según índice de ingreso, Chiapas 2010



Fuente: elaboración propia con datos del INEGI

Gráfica 5: índice de Moran de ingresos municipales, Chiapas 2010



Fuente: elaboración propia

SPATIAL REGRESSION

The interaction between spatial units seeks to be captured from models which consider: (1) the endogenous relationship of the dependent variable (Wy) and spatial units; (2) exogenous relationship between the dependent variable (y) and the independent variables (Wx); and (3) the interaction between the error terms (Wu) (Elhorst, 2014: 8) In addition to the interactions, spatial models seek to capture the spatial heterogeneity via u which captures the effect of omitted variables. The saturated spatial regression model is given by:

$$y_{it} = \rho W y_{it} + X_{it}\beta + W X_{it}\theta + u_{it}$$

$$u_t = \lambda W u_t + \varepsilon_t$$

Wy denotes the effects of endogenous interaction of the spatial units, WX denotes the matrix effects of spatial exogenous interaction, Wu denotes interaction effects of the error terms with the spatial units, ρ represents the autoregressive spatial coefficient, θ and β are vectors of parameters to estimate and W is the spatial weights matrix.

From the saturated model of spatial regression, different models are derived: the autoregressive spatial model (SAR) when $\theta = 0$ and $\lambda = 0$, the spatial error model (ESR) when $\rho = 0$ and $\theta = 0$, the Durbin model (SDM) when $\lambda = 0$; The Durbin Spatial Error Model (SDE); The spatial autoregressive spatial error (SAC) model when $\theta = 0$ and the spatial autoregressive spatial model (SMA).

On the other hand, the model of ordinary regression (OLS) usually ignores the effects of spatial dependence and heterogeneity, that is, it adjusts under the assumption of independence between spatial units. When such assumptions are violated, biased and inconsistent estimates are usually produced, so it is advisable to use a spatial model, especially when there is evidence of the presence of dependence and / or spatial heterogeneity, which occurs frequently in cases where data are collected from space units taken from nearby units, which can show similar patterns.

Table 2. Linear regression model

Variable	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.169	0.088	1.922	0.057
Pphlin	0.061	0.016	3.832	0.000
ppextremos	0.003	0.001	5.264	0.000
iingreso	-0.101	0.110	-0.914	0.363
R-squared	0.780			

Source: own elaboration with data of INEGI and CONEVAL

A fundamental feature of spatial regression models is the simultaneous feedback that emerges from dependency interactions, that is, there are feedback effects between regions resulting from the exchange of stimuli provoked in a unit, by the action of a variable that generates changes in neighboring units, which in turn reverts to the original unit. In addition to the effects generated by the observed variables, spatial heterogeneity can come from latent (unobserved) influences related to cultural, economic, social factors, or a series of factors that can be explained through feedback among neighbors. This type of heterogeneity is captured by the dependent variable (Anselin, 1988; LeSage & Fischer, 2008) and must be treated in the sense that time series do, where dependency is managed through models that adjust the lag of the dependent variable (SMA).

Table 3. Lagrange multiplier test

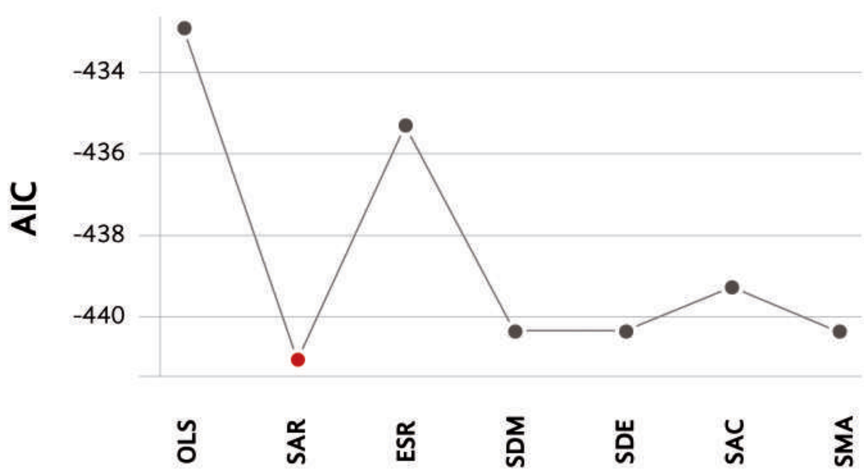
Variable	Statistic	Parameter	p.value
LMerr	2.915	1	0.088
LMlag	11.152	1	0.001
RLMerr	0.004	1	0.950
RLMlag	8.241	1	0.004
SMA	11.126	2	0.004

Source: own elaboration with data of INEGI and CONEVAL

The traditional way of analyzing the validity of simple linear regression models is to review the sign and magnitude of the estimates. That is to say, that the meaning of the estimation of the estimated coefficients is theoretically correct. For example, to increase the average income of the municipalities, it is expected to reduce the levels of municipal illiteracy, or in the opposite direction, by increasing the levels of extreme poverty we would expect an increase in illiteracy levels (see Table 2).

As mentioned above, there is a wide variety of spatial regression models, so one of the crucial problems is choosing the right model (LeSage & Pace, 2009), which is why it is recommended to apply test to specify which one. There are two types of tests: model contrast and data adjustment.

Graph 6. selection of spatial model according to AIC criterion



Most of the model fitting tests are devoted to verifying the existence of spatial correlation. We chose to use the Lagrange multiplier test. It compares the fit of the spatial model with the results from the ordinary linear regression model. The difference is used as a criterion to determine if the relative change of the first derivative of the likelihood function around the maximum affects significantly the autoregressive parameter of the Spatial model. The model with the highest value statistic must be chosen (see Table 3).

For the selection of the model, the Akaike information criterion (AIC) was used in addition to the Lagrange test, which provides a measure of the quality of fit of the model, depending on the data. Given a collection of models, AIC estimates the fit quality of each model and provides a means of selection, based on the value of the maximum likelihood function of the model and the number of estimated parameters. The model with the smallest AIC value should be chosen (see graph 6).

The construction of the spatial model for municipal illiteracy considered three essential elements: spatial heterogeneity among municipalities, spatial autocorrelation of illiteracy, and spatial autocorrelation of the factors that shape illiteracy levels. These elements are present in all six spatial models. As can be seen in Table 4, two spatial autocorrelation models were applied: the spatial lag SAR model, and the autoregressive spatial error model ESR; Two Durbin models: SDM and SDE; The SAC model and the SMA model to verify the existence of autoregressive effects on the errors.

Table 3 presents the results of the Lagrange test, in which it is observed that the model with the highest value is the SAR, however, the difference with the SDM and SMA model may not be large enough, which is why the AIC criterion was also used.

Table 4. AIC criteria for different models

Do not	Model	AIC
1	OLS	-433.00
2	SAR	-441.25
3	ESR	-435.32
4	SDM	-440.40
5	SDE	-440.09
6	SAC	-439.21
7	SMA	-440.07

Source: self-made

The AIC test confirmed that the model with the lowest AIC was SAR, followed by the SDM model and the SMA (see Table 4 and Graph 6). However, the difference with the SDM and SMA models is still small relative to the SAR model, so it was decided to apply the likelihood ratio test (see Table 5), where it is shown that the difference between models is significantly large and therefore the SAR model should be chosen. It is important to note that there is no likelihood test for the SMA model, however, it is clear that the SAR model is a better choice because it is a model with a higher level of parsimony, which presents a higher degree of adjustment of the data.

Table 5. anova test for SAR model vs SDM

Model	n	df	AIC	logLik	Test	L.Ratio	p.value
mod.sar	1	6	-441.253	226.626			
mod.sdm	2	9	-440.404	229.202	1 vs 2	5.151	0.161

Source: own elaboration with data of INEGI and CONEVAL

Another important point to note is that the SAR model adds to the traditional regression a vector of spatial lag that models the effects that municipal illiteracy exerts among its neighboring municipalities, in order to explain the inter-municipal variation. Intuitively, the model establishes how illiteracy levels in each municipality are related to the average illiteracy of its neighbors. The spatial lag vector Wy reflects the average illiteracy levels of the municipalities weighted through the matrix W, while the parameter $\rho = 0.3$ reflects the spatial dependence force, which is also statistically significant (see Table 6).

RESULTS

In order to properly interpret the model it is important to consider that the partial derivative of E (y) with respect to the k^{th} explanatory variable has three fundamental properties: (1) the explanatory variable of spatial unit has an effect on the dependent variable known as effect direct; (2) the change over the dependent variable is not only a function of the kth explanatory variable, but also of the explanatory variables of the neighboring units, and (3) global indirect effects quantify the impact of a variable exogenous change in all neighboring spatial units given the value of the dependent variable (Griffith, 2000).

The adjusted SAR regression model is shown in Table 5. It is important to note that the estimated coefficients have the sign and magnitude expected and that the R-square coefficient is higher than that presented by the ordinary linear regression model (see Tables 2 and 6).

Table 6. Spatial lag model SAR, illiterate population in Chiapas, 2010

Variable	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.176	0.082	2.136	0.033
Pphlin	0.037	0.016	2.273	0.023
pextrema	0.002	0.000	4.490	0.000
lingreso	-0.182	0.105	-1.738	0.082
Rho	0.304	0.089	3.433	0.000
R-squared	0.808			

Source: Based on data from INEGI and CONEVAL

The interpretation of the parameter β in the model space (SAR) expresses the impact of the change in the dependent variable xr on the town i as a combination of direct and indirect influences. This spatial spill originates from effect of the model variables, which basically depends on: (1) the position of the municipality in the territory, (2) the degree of connectivity between the municipalities, defined by the W matrix, (3) force estimated by the parameter ρ spatial dependence, and (4) the magnitude of β estimates coefficients (LeSage & Fischer, 2008; LeSage & Pace, 2009).

While the coefficient β expresses the change of an independent variable, which occurred in the cluster formed by the neighboring municipality i that coincides on the dependent variable of the municipality i, which emerges as a natural consequence of the spatial dependence. Any change in the characteristics of neighboring municipalities, in turn, generate changes that will impact the dynamics of the adjacent municipality i and vice versa. Since the impact of changes in an independent variable differs between regions, it is advisable to define a summary measure for each type of impact, and generally three types are identified: direct effects, indirect effects and total effects.

The direct effect provides a summary measure provoked throughout the state, by changing the variable xr in the municipality i . For example, if in the municipality i extreme poverty levels increase; the average direct effect quantifies their impact on levels of illiteracy in all municipalities in Chiapas. This measure takes into account feedback effects arising from changes in levels of extreme poverty observed in the municipality i, which impacts its neighbors through system modeling spatial dependencies through the matrix W.

Table 7. Effects in order of vicinity, model SAR

Neighbour	Direct	Indirect	Total
Pphlin	0.0370	0.0157	0.0528
Pextrema	0.0022	0.0009	0.0031
Iingreso	-0.1835	-0.0777	-0.2613

Source: Based on data from INEGI and CONEVAL

The SAR model presented in Table 5, indicates that an increase of one percentage point in the proportion of indigenous population of a municipality, will cause an increase of 3.7 percentage points per levels of illiteracy in the state. In the same way it is interpreted as a direct effect of extreme poverty where the increase of one percentage point may be associated with an increase of 0.22 illiteracy points in the state; on the contrary, an increase of an average of one dollar per family income in a municipality would generate a reduction of 18.3 percentage points in illiteracy levels of the state.

The indirect effect is used to measure the impact of the increase of a dependent variable in all neighboring towns, which effects a given municipality i. In the case of SAR fixed effects model, Table 7 shows that the indirect effect of the increase of one percentage point in the proportion of average of speakers of indigenous language, in all of the municipalities in Chiapas, would cause an increase of 1.6 points percentage levels of illiteracy in the municipality i. In the case of extreme poverty, the indirect effects indicate that an increase of one percentage point in the same levels in neighboring municipalities would bring an increase of 0.09 illiteracy percentage points, i.e. ,the indirect effect of extreme poverty among neighboring municipalities only marginally affects the municipality i. When levels of extreme poverty are reduced in neighboring municipalities, illiteracy levels in the municipality i, improve marginally as an indirect effect.

The situation, in the case of municipal income in relation to indirect effect, indicates that when all neighboring municipalities improve their income, literacy levels of the municipality i improve by 7.7 as an indirect effect, whereas if only levels of income improve in the i municipality, literacy level increases by (direct effect) 18.3 points , so that the total effect is less than 26.1 points.

The total effect is the sum of direct and indirect effects, i.e., if all municipalities increase their income by one dollar, the overall effect would reflect, the average impact on levels of illiteracy in a given municipality and the overall effect will include both the impact of indirect and direct effect, which would be less than 26 percentage points.

The main changes in the behavior of municipal illiteracy levels are what can be observed between municipalities. That is, that the greatest inequalities can be found among municipalities that are relatively literate like Tuxtla Gutierrez, and illiterate towns like San Juan Cancuc, Chanal or Mitontic (see Figure 1). The SAR model shows the (direct and indirect) influence of the presence of indigenous peoples, extreme poverty and municipal revenues on the observed municipal levels of the illiterate population. The direct effect of all the variables represents on average 70 percent of the total effects, however, the income effect is essential in terms of direct and indirect impact on municipal levels of observed illiteracy. While poverty has a significant effect on levels of illiteracy, lack of income is the factor that actually modifies the pattern of illiteracy.

CONCLUSIONS

The main objective is to understand the role of spatial heterogeneity in determining levels of illiteracy in Chiapas municipalities. To meet this objective, a strategy based on a methodology of spatial regression, which also quantifies the total effects, direct effects and indirect effects of the determinants of illiteracy in the state. The role of average income in terms of current family income , standardized from Gross National Income, of the indigenous population and the effect of extreme poverty as a proxy for the inability of people to access services to enable them to acquire the minimum skills for literacy is highlighted.

The results showed the existence of a significant spatial pattern to explain the behavior of illiteracy in the territory of Chiapas. The SAR model showed the importance of family income as a key to predict the spatial behavior for determining the levels of illiteracy. The research results put on the table the importance of the availability of family income to enable the generation of the basic skills of reading and writing, as a key to improving levels of school enrollment and reduce inequalities between groups of indigenous people and mestizos. This work seriously questions the role of extreme poverty in a multidimensional way, where social deprivation are marginal to illiteracy, and where changes in levels of extreme poverty are spatially correlated with changes in the levels of municipal illiteracy, which effect is surprisingly low with only 0.2 percent increase in illiteracy per percentage point of extreme Municipal poverty.

The inverse relationship between increasing illiterates and reducing current family income is understandable as an effect on revenues is a factor that even surpasses extreme poverty. Which would mean that even in conditions of extreme poverty, the factor that explains high levels of illiteracy among municipalities is the lack of minimal income which leads them to the status of illiteracy, which precludes the fundamental objective to combat poverty in the State which is to eliminate the intergenerational transmission of poverty through capacity building, where the minimum capacity expected to see increase would be the factor of reading-writing (Lopez and Nunez, 2016).

The relationship between illiteracy and the percentage of speakers of indigenous languages quantified the effect on the levels of illiteracy due to the changes in the ethnic composition of the population, which involves factors related to the physical accessibility and lack of nearby school in indigenous communities, but also factors related to cultural accessibility. The indigenous population does not have the possibilities of real access to the educational system due to social, cultural, and economic differences. In other words, although there is a school nearby, the indigenous population is confronted with barriers linked to language, discrimination, the lack of monetary income or the cost of opportunity- which is a situation that leads to greater levels of illiteracy among the indigenous population. This study makes evident these differences which are the product of non-random spatial distribution patterns.

REFERENCES

Anselin, L. (1988). Spatial econometrics: Methods and Models. *Kluwer Academic Publishers*.

Anselin, L. (1995). Local Indicators of Spatial Associations-LISA. *Geographical Analysis, vol.27:93-115*.

Bello, A. y M. Rangel (2002). La equidad y la exclusión de los pueblos indígenas y afrodescendientes en América Latina y el Caribe, *Revista* CEPAL, (76). Santiago. pp. 39-54.

Chasco, C. (2003). *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales. Consejería de Economía e Innovación Tecnológica,* Comunidad de Madrid. Madrid.

CONEVAL (2014), *Metodología para la medición multidimensional de la pobreza en México* (segunda edición). México, CONEVAL.

Cliff, A. D. y Ord, J. K. (1981). *Spatial processes: Models and Applications.* Pion Limited, London.

Cressie, N. (1993). *Statistics for spatial data John Wiley*. New York.

Elhorst, J. Paul (2014). *Spatial Econometrics, from cross-sectional data to spatial panels*. Springer.

Getis, A. y Ord J.K. (1992). The analysis of spatial association by use of distance statistis. *Geographical Analysis, 24*: 189-199.

Griffith D.A. (2000). *A linear regression solution to the spatial autocorrelation problem*. J Geogr Syst 2(2):141-156.

INEGI (2010). *XIII Censo General de Población y Vivienda 2010. Tabulados Básicos 2010*. Consulta en línea en: < http://www.inegi.org.mx/sistemas/ TabuladosBasicos/Default.aspx?c=16852&s=est>, [consulta septiembre 2016].

LeSage JP, Fischer MM (2008). Spatial growth regressions: model specification, estimation and interpretation. *Spat Econ Anal 3*(3):275-304

LeSage J.P. & Pace R.K. (2009). Introduction to spatial econometrics. *CRC Press Taylor & Francis Group,* Boca Raton.

López, Jorge y Núñez, Gerardo. (2016). *Desigualdad y exclusión social en Chiapas, una mirada a largo plazo.* OXFAM México.

PNUD (2014). *Índice de Desarrollo Humano Municipal en México: nueva metodología. Elaborado por la Oficina de Investigación en Desarrollo Humano del Programa de las Naciones Unidas para el Desarrollo en México*. Programa de las Naciones Unidas para el Desarrollo, México, DF, PNUD.

UNESCO (2008). *El desafío de la alfabetización en el mundo.* Paris. UNESCO.