

Recepción: Octubre 11, 2016 | Aceptación: Junio 12, 2016

ANÁLISIS ESPACIAL DEL ANALFABETISMO EN MUNICIPIOS DE CHIAPAS

—
Gerardo Núñez Medina

Director de investigación y evaluación del
Consejo de Investigación y Evaluación de la Política Social del Estado de Chiapas.
gerardo.nm1@gmail.com



Para citar este artículo:

Núñez, G. (2017) Análisis espacial del analfabetismo en municipios de Chiapas. *Espacio I +D Innovación más Desarrollo* 6 (13) 78-99. doi: 10.31644/IMASD.13.2017.a05

RESUMEN

El trabajo tiene por objeto analizar algunos determinantes del comporta-miento espacial del analfabetismo en municipios de Chiapas al año 2010, considerando un modelo de regresión espacial cuyas variables independien-tes son: la proporción de población hablante de lengua indígena, el porcen-taje de población en pobreza extrema y el ingreso corriente familiar ajustado según Ingreso Nacional Bruto. La hipótesis central sostiene que las tasas de analfabetismo municipal no se distribuyen aleatoriamente, sino que presen-tan patrones de concentración y dispersión espacial. Los resultados ponen sobre la mesa la importancia de la disponibilidad de ingresos familiares para posibilitar la generación de capacidades lecto-escritoras básicas, como un elemento fundamental para mejorar los niveles de inserción escolar y reducir desigualdades entre poblaciones indígenas y mestizas.

Palabras clave

Heterogeneidad espacial, modelos autoregresivos, efectos directos e indirectos, población indígena, pobreza extrema.

SPATIAL ANALYSIS OF ILLITERACY IN CHIAPAS'S MUNICIPALITIES

— 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.

El objetivo de este trabajo es examinar los principales determinantes del comportamiento espacial del analfabetismo en municipios de Chiapas al año 2010, considerando como variables independientes la proporción de población hablante de lengua indígena, el porcentaje de población en pobreza extrema y los ingresos familiares municipales ajustados según Ingreso Nacional Bruto. La hipótesis central sostiene que las tasas de analfabetismo pueden explicarse como resultado de la combinación de los niveles de pobreza, indigenismo e ingresos presentes en los municipios, y que estos factores, no se distribuyen aleatoriamente, sino que presentan patrones de concentración y dispersión espacial que potencializan sus efectos a través de la interacción entre municipios vecinos.

La importancia de esta investigación radica en tres factores fundamentales: (1) Chiapas es una de las entidades federativas con mayor incidencia de analfabetismo en todo el país, según cifras del censo 2010; (2) Chiapas es la segunda entidad del país en cuanto al número de población indígena y la primera en población indígena monolingüe; (3) Chiapas es la entidad con la mayor proporción de población en pobreza y pobreza extrema, desde que iniciaron las mediciones en México en los años 90's, por lo que, el análisis de la interacción espacial de estos factores resulta en un ejercicio, esencial para conocer de la importancia, magnitud y efectos del analfabetismo a nivel municipal en términos de las variables señaladas y de su interacción espacial.

DATOS

Los datos utilizados para este trabajo se recopilaron de fuentes demográficas y socioeconómicas dependientes del gobierno mexicano y de instancias internacionales. La primera instancia fue el censo general de población y vivienda 2010 levantado por el Instituto Nacional de Estadística, Geografía e Informática (INEGI, 2010) de donde se compiló información sobre la población total municipal, población indígena y población analfabeta.

Los datos referentes a la condición de alfabetismo fueron recolectados a partir de los lineamientos establecidos por INEGI que define como analfabeta a la población de 15 años o más que declara no saber leer ni escribir un recado. Bajo este criterio se clasifica a los individuos como alfabetas o analfabetas. Para fines de este trabajo, la población alfabetada será aquella que con 15 años o más declara saber leer y escribir un recado. Es decir, han adquirido una capacidad básica para acceder a nuevos conocimientos, lo que hace posible mejorar sus posibilidades de integración social, de generación de riqueza y de acceso a servicios de salud y educación.

Los datos referentes a los niveles de pobreza extrema municipal fueron obtenidos a partir de las estimaciones publicadas por el Consejo Nacional de Evaluación (CONEVAL, 2014), mismas que fueron realizadas con datos obtenidos del módulo de condiciones socioeconómicas de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) para todos los municipios de Chiapas. Mientras que las estimaciones referentes a los ingresos municipales fueron obtenidas de Programa de Naciones Unidas para el Desarrollo (PNUD, 2014) y son el componente estandarizado de ingresos utilizado para el cálculo del índice de desarrollo humano municipal al año 2010 para los municipios de Chiapas.

ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES

El análisis exploratorio de datos espaciales (AEDE) suelen iniciar con la aplicación de pruebas de autocorrelación espacial a cada una de las variables involucradas, para lo que es necesario definir una matriz de contigüidades espaciales (Chasco, 2003). La matriz de contigüidades espaciales se define como una matriz binaria, cuyos valores dependen de si las unidades espaciales son o no vecinas, por lo que el criterio de vecindad es fundamental. En este caso se consideró que dos unidades espaciales eran vecinas si la distancia entre sus cabeceras municipales era menor o igual a la distancia máxima definida entre todas las cabeceras municipales de la entidad, con lo que dos municipios son vecinos si están en un radio menor o igual a 54.6 km de distancia.

Tabla 1. Autocorrelación espacial, índice de Moran global.

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
ingreso	0.260	-0.009	0.001

Fuente: elaboración propia con datos de INEGI y CONEVAL

El estadístico de correlación espacial más utilizado es el índice de Moran que en su versión global mide la autocorrelación basada en las ubicaciones y los valores de una variable x para todas las regiones simultáneamente, es decir, es una medida de autocorrelación, definida de forma similar al coeficiente de correlación de Pearson (Anselin, 1995), con la salvedad que establece

la hipótesis de que la variable analizada se distribuye aleatoriamente en el espacio. Cuando el valor p del estadístico es significativo, se puede suponer la presencia de un patrón de correlaciones espaciales. El índice de Moran global se estima a partir de:

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

Donde W_{ij} representa los elementos de la matriz de pesos espaciales, \bar{x} la media de la variable x . Es importante señalar que el coeficiente de autocorrelación mide el grado de asociación de la variable x respecto de sus vecinos.

El índice de Moran en su versión local permite identificar conglomerados espaciales en cinco categorías: (1) *low-low*: unidades espaciales con valor inferior al promedio, rodeadas de unidades con valores por debajo la media del atributo de interés. Estas unidades espaciales corresponden a conglomerados llamados zonas frías. (2) *low-high*: unidades espaciales con valor debajo del promedio rodeadas por unidades con valores por encima del promedio. (3) *high-low*: unidades espaciales con valor por arriba del promedio rodeadas por unidades con valores por debajo del promedio. (4) *high-high*: las unidades espaciales con valor por encima del promedio, rodeadas por unidades con valores por arriba de la media. Estas unidades corresponden a conglomerados llamados zonas calientes. Y (5) no data: el conjunto de unidades espaciales donde la variable de interés no se correlaciona significativamente con valores vecinos (Cliff & Ord, 1981).

Los resultados de las estimaciones referentes a los niveles de correlación espacial global se presentan en la tabla 1 e indican la presencia de altos niveles de autocorrelación espacial positiva para todas las variables analizadas, lo que indica la existencia de una asociación directa entre municipios con altos niveles de analfabetismo rodeados por municipios que a su vez presentan niveles de analfabetismo por encima del promedio estatal. La misma situación ocurre para el caso del porcentaje de población hablante de lengua indígena, pobreza extrema e ingresos.

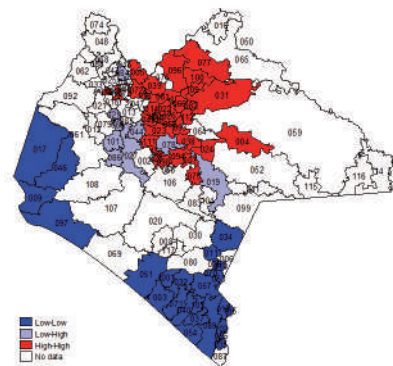
La presencia de autocorrelación permite suponer la existencia de estructuras espaciales capaces de explicar los niveles de analfabetismo municipal en términos de posibles asociaciones con variables que a su vez presentan altos niveles de autocorrelación espacial, sobre todo cuando esta se presenta en las mismas regiones o conglomerados (Getis & Ord, 1992), como fue el caso de la población hablante de lengua indígena del estado de Chiapas.

Población Analfabeta

La alfabetización puede entenderse como un proceso a través del cual los individuos adquieren la capacidad de comunicarse de forma escrita, lo que se constituye como un elemento que posibilita la adquisición continua de habilidades y destrezas de todo tipo. La alfabetización habilita a las personas para desarrollar ventajas que eventualmente le permitirá mejorar sus condiciones de vida. La relación intrínseca entre la capacidad de leer y escribir adecuadamente y la posibilidad de adquirir nuevas destrezas, juega un papel esencial en la generación de crecimiento económico y en la reducción de desigualdades (UNESCO, 2008).

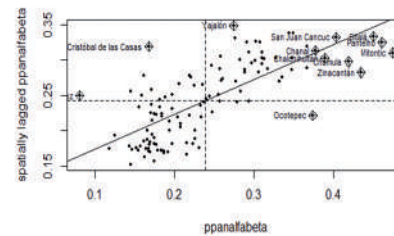
En contra parte, la falta de competencias lecto-escritoras es un factor clave para explicar graves carencias asociadas a la condición de pobreza extrema, discriminación y exclusión social en la que se encuentran inmersos grupos sociales específicos, como es el caso de una amplia proporción de población hablante de lengua indígena; población que además se presume obtiene los ingresos más bajos del estado de 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

En el mapa 1 muestra la dispersión espacial de la población analfabeta municipal de Chiapas al año 2010, en el mismo puede observarse la presencia de un conglomerado caliente (color rojo) conformado por los municipios con los menores niveles de población alfabetizada y que a su vez están rodeados de municipios con bajos niveles de alfabetización, este conglomerado se encuentra conformado por los municipios de Simojovel, San Andrés Duraznal, Santiago el Pinar, Bochil y Larrainzar, entre otros (ver mapa 1). Municipios

que conforman el conglomerado de alta concentración de población analfabeta en Chiapas.

La pendiente de la recta de regresión observada en la gráfica 1, representa el valor del índice de Moran global de autocorrelación espacial para la proporción de población analfabeta municipal, la misma muestra un importante nivel de autocorrelación positiva (0.49) que implica que la proporción de población alfabetizada se encuentra espacialmente concentrada. El gráfico de asociación espacial en su cuadrante I muestra una importante agrupación de municipios que se corresponden a las zonas rojas del mapa 1.

La Población Indígena en Chiapas

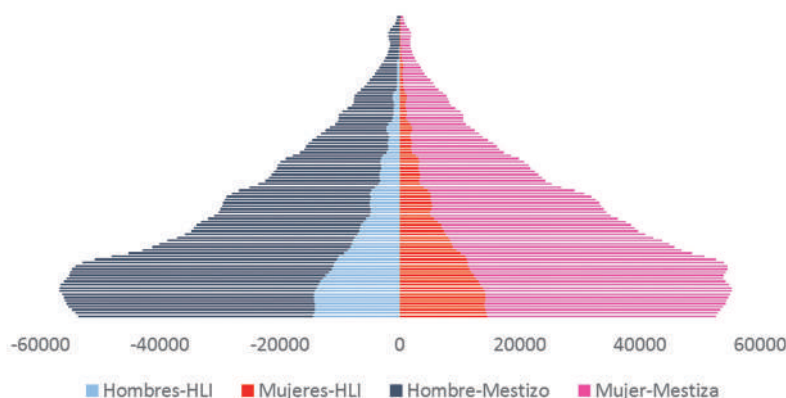
La población indígena puede ser considerada, desde una perspectiva histórica y sociocultural, como: «aquellos grupos descendientes directos de los pueblos que habitaban América desde antes de la llegada de los españoles en el siglo xv, que poseen una lengua y cultura propias y que comparten formas de vida y cosmovisiones particulares, diferenciadas de las occidentales» (Bello y Rangel, 2002: 40). Además de lo anterior es importante reconocer la presencia de grandes grupos de indígenas, que se han asimilados culturalmente a los mestizos con la consecuente pérdida del lenguaje, que han sido también desplazados de sus territorios y que hoy habitan los cinturones de pobreza de las grandes ciudades, donde la gran mayoría de ello es población pobre, marginada y sin posibilidades de acceder a los sistemas formales de empleo, educación y salud.

Aunque acertada, la definición anterior resulta poco práctica, debido a las dificultades técnicas para implementar su medición, de modo que se optó por emplear el criterio lingüístico establecido por el INEGI, mismo que define como población indígena a aquellas personas de cinco o más años que contestaron afirmativamente a la pregunta de si hablaban alguna lengua indígena. De esta forma la población indígena fue identificada con base al censo de población, utilizando el criterio lingüístico, que se refiere a la condición de hablante. La población hablante de lengua indígena (phli) está conformada por aquellos individuos, residentes en Chiapas, de cinco años o más que afirmaron hablar alguna lengua indígena en 2010.

La phli de Chiapas representó 23.5% de la población total del estado (ver gráfica 2). A pesar de que prácticamente uno de cada cuatro habitantes puede ser considerado población indígena, estos siguen siendo considerados una minoría y por ende excluidos de espacios de poder y de la toma de decisiones (INEGI, 2010). Aun cuando en muchos casos, la población indígena representa más del 60 por ciento de la población, de un municipio, el presidente y las autoridades municipales suelen ser de origen mestizo. Poco

más de un millón de habitantes, en Chiapas habla alguna lengua indígena. De ellos 66.2% hablan también español, sin embargo, una tercera parte de la población indígena es monolingüe, lo que impacta directamente sus posibilidades de acceder y cursar con éxito los distintos niveles del sistema educativo, la cual se refleja en los importantes niveles de analfabetismo, mismos que alcanzan el 21% de la población general al año 2010.

Gráfica 2. Estructura por edad y sexo de la población de Chiapas y la población hablante de lengua indígena, Chiapas 2010

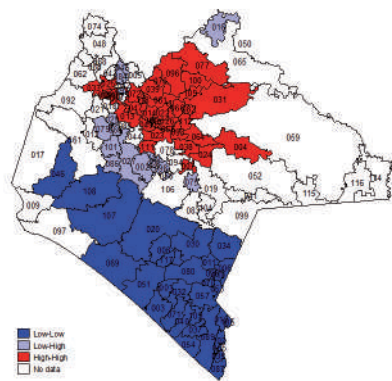


Fuente: elaboración propia con datos de censo 2010, INEGI.

La dispersión espacial de la población indígena municipal de Chiapas al año 2010, puede verse en el mapa 2, en el mismo se observa un conglomerado caliente (color rojo) conformado por los municipios con los mayores niveles de población indígena y que a su vez están rodeados de municipios con altos niveles de población indígena, este conglomerado se encuentra ubicado en la misma zona del estado en la que encontró el conglomerado de municipios que presentaron altos niveles de concentración de población analfabeta.

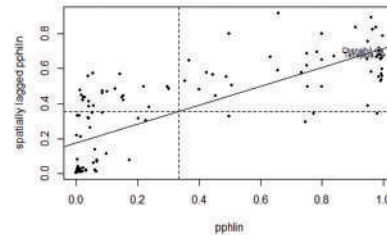
La pendiente de la recta de regresión observada en la gráfica 3, representa el valor del índice de Moran global de autocorrelación espacial para la proporción de población indígena municipal, la misma muestra un importante nivel de autocorrelación positiva (0.53) que implica que la proporción de población indígena se encuentra espacialmente concentrada. El gráfico de asociación espacial en su cuadrante I muestra una importante agrupación de municipios que se corresponden a la zona roja del mapa 2.

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

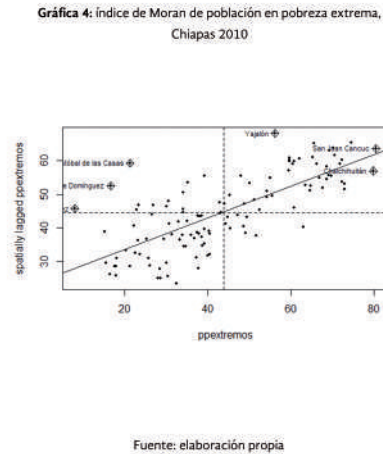
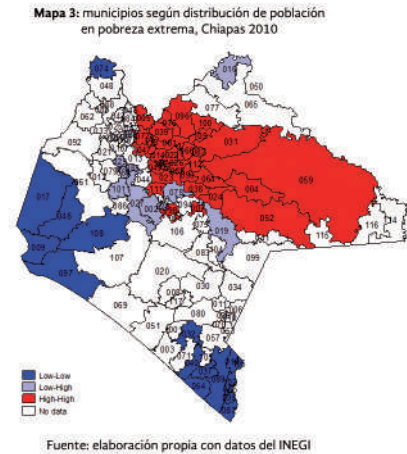
Población en pobreza extrema

La definición de población en pobreza para México fue establecida por el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL, 2014: 26) quien señala que los individuos en condición de pobreza multidimensional son quienes «...no tiene garantizado el ejercicio de al menos uno de sus derechos para el desarrollo social, y si sus ingresos son insuficientes para adquirir los bienes y servicios que requiere para satisfacer sus necesidades». Para efectos de medición, la pobreza se cuantifica en dos dimensiones: (1) Bienestar económico, medido en términos del ingreso corriente, (2) Derechos sociales, medidos en términos de acceso a la educación, salud, seguridad social, alimentación, vivienda y sus servicios.

Adicionalmente, una persona se encuentra en situación de pobreza extrema cuando padece de tres o más carencias relativas a sus derechos sociales, y su ingreso se encuentra por debajo de la línea de bienestar mínimo. Es decir, las personas en situación de pobreza extrema disponen de un ingreso tan bajo que no pueden adquirir los nutrientes necesarios para mantener una vida sana.

Según cifras del CONEVAL en 2010 a nivel nacional la población en pobreza fue de 52.1 millones de personas, 12.8 millones de ellos en pobreza extrema. Chiapas ocupó el primer lugar en porcentaje de población en pobreza y en pobreza extrema; el 78.5% de la población chiapaneca se encontraba en situación de pobreza, de ellas 1.88 millones estaban en condición de pobreza extrema, lo que representó 38.3 por ciento de la población total del estado. A nivel municipal, el porcentaje más bajo de pobreza extrema fue de 7.9%

en Tuxtla Gutiérrez y el más alto para San Juan Cancuc con 80.5%, uno de los municipios más pobres del país.

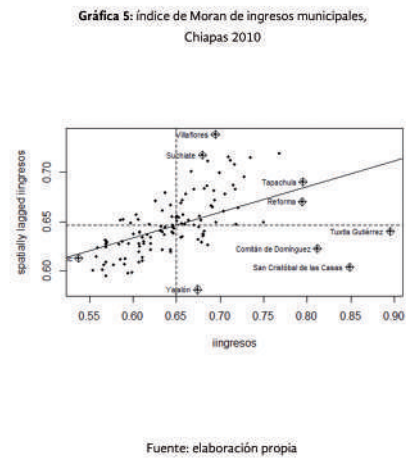
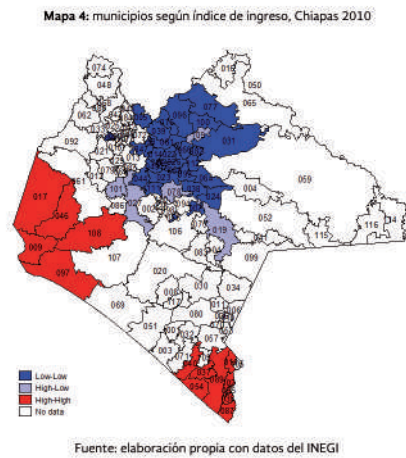


La dispersión de la pobreza municipal en Chiapas al año 2010 puede verse en el mapa 3, donde se muestra una clara concentración de niveles de pobreza extrema en los municipios de la zona del Estado donde se concentra la mayor proporción de población analfabeta y dos conglomerados fríos, uno alrededor del municipio de Tapachula y otro alrededor de los municipios de Arriaga y Tonalá. Es importante considerar que la pendiente de la recta de regresión (gráfica 4), que representa el valor del índice de Moran global para la proporción de población en pobreza extrema municipal muestra un importante nivel de autocorrelación positiva de 0.46, valor que implica que la proporción de población en pobreza extrema se encuentra espacialmente concentrada. El gráfico 4 en su cuadrante I muestra la importancia de municipios de San Juan Cancuc, Chalchihuitán y Yajalón que corresponden a las zonas rojas del mapa 3.

Ingreso Municipal

Con la finalidad de utilizar una medida comparable y estandarizada del ingreso a nivel municipal, se recurrió al índice de ingreso que es uno de los tres componentes que conforman el Índice de Desarrollo Humano (IDH) diseñado por la oficina de investigación del Programa de Naciones Unidas para el Desarrollo (PNUD, 2014). El índice de ingresos refleja la capacidad de acceso a recursos que permiten gozar a los individuos de una vida digna. Este representa una estimación del ingreso corriente del que disponen las

familias a nivel municipal, y éste se ajusta al Ingreso Nacional Bruto (INB). El cálculo se realiza a partir de la estimación del ingreso corriente del que disponen las familias a nivel municipal, mismo que se ajusta al INB proveniente del Sistema de Cuentas Nacionales del INEGI. Éste se expresa anualmente en dólares americanos de 2010 (PNUD, 2014).



La dispersión del ingreso municipal en Chiapas al año 2010 puede verse en el mapa 4, donde se muestra una clara concentración de altos niveles de ingreso en dos zonas de la costa del estado, una que rodea al municipio de Tapachula y la otra alrededor de los municipios de Arriaga y Tonalá. En contraparte, los municipios de la zona del Estado donde se concentra la mayor proporción de población analfabeta e indígena de Chiapas conforman a su vez, un conglomerado de baja concentración de ingresos.

La pendiente de la recta de regresión observada en la gráfica 5, representa el valor del índice de Moran global de autocorrelación espacial para el índice de ingresos municipal, la misma muestra un importante nivel de autocorrelación positiva (0.26) que implica que la proporción de población analfabeta se encuentra espacialmente concentrada. El gráfico de asociación espacial en su cuadrante IV presenta algunos municipios cuyo comportamiento resulta interesante, como es el caso de Yajalón, un municipio indígena, enclavado en una zona con altos niveles de pobreza extrema y de muy bajos ingresos que presenta curiosamente altos niveles de ingreso, en relación al promedio de sus vecinos.

REGRESIÓN ESPACIAL

La interacción entre unidades espaciales busca ser capturada a partir de modelos que consideran: (1) la relación endógena de la variable dependiente (Wy) y unidades espaciales; (2) la relación exógena entre la variable dependiente (y) y las variables independientes (Wx); y (3) la interacción entre los términos de error (Wu) (Elhorst, 2014: 8) Además de las interacciones, los modelos espaciales buscan captar la heterogeneidad espacial a través de u que captura el efecto de variables omitidas. El modelo saturado de regresión espacial está dado por:

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

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

Wy denota los efectos de interacción endógena de las unidades espaciales, WX denota la matriz de efectos de interacción exógena espacial, Wu denota los efectos de interacción de los términos de error con las unidades espaciales, ρ representa el coeficiente espacial autoregresivo, θ y β son los vectores de parámetros a estimar y W es la matriz de pesos espaciales.

A partir del modelo saturado de regresión espacial, se derivan distintos modelos: el modelo espacial autoregresivo (SAR) cuando $\theta=0$ y $\lambda=0$, el modelo de errores espaciales (ESR) cuando $\rho=0$ y $\theta=0$, el modelo Durbin espacial (SDM) cuando $\lambda=0$; el modelo Durbin de errores espaciales (SDE); el modelo espacial autoregresivo y de errores espaciales (SAC) cuando $\theta=0$ y el modelo espacial autorregresivo de medias móviles (SMA).

Por otra parte, el modelo de regresión ordinario (OLS) suele ignorar los efectos de la dependencia y heterogeneidad espacial, es decir, se ajusta bajo el supuesto de independencia entre unidades espaciales, cuando dichos supuestos se violan, suelen producirse estimaciones sesgadas e inconsistentes, por lo que es recomendable utilizar un modelo espacial, en especial cuando se tiene evidencia de la presencia de dependencia y/o heterogeneidad espacial, lo que ocurre frecuentemente en los casos en que se colectan datos de unidades espaciales tomadas de unidades cercanas, por lo que pueden mostrar patrones similares.

Tabla 2. Modelo de regresión lineal

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			

Fuente: elaboración propia con datos de inegi y con eval

Una característica fundamental de los modelos de regresión espacial es la retroalimentación simultánea que surge a partir de interacciones de dependencia, es decir, surgen efectos de retroalimentación entre regiones producto del intercambio de estímulos provocados en una unidad, por la acción de una variable que genera cambios en unidades vecinas, lo que a su vez se revierte a la unidad original. Además de los efectos generados por las variables observadas, la heterogeneidad espacial puede provenir de influencias latentes (no observadas) relacionadas con factores culturales, económicos, sociales, o por una serie de factores que pueden ser explicados a través de la retroalimentación entre vecinos. Este tipo de heterogeneidad es captada por la variable dependiente (Anselin, 1988; LeSage & Fischer, 2008) y debe ser tratada en el sentido que lo hacen las series de tiempo, donde la dependencia es manejada a través de modelos que ajustan el rezago de la variable dependiente (SMA).

Tabla 3. Prueba de multiplicadores de Lagrange

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

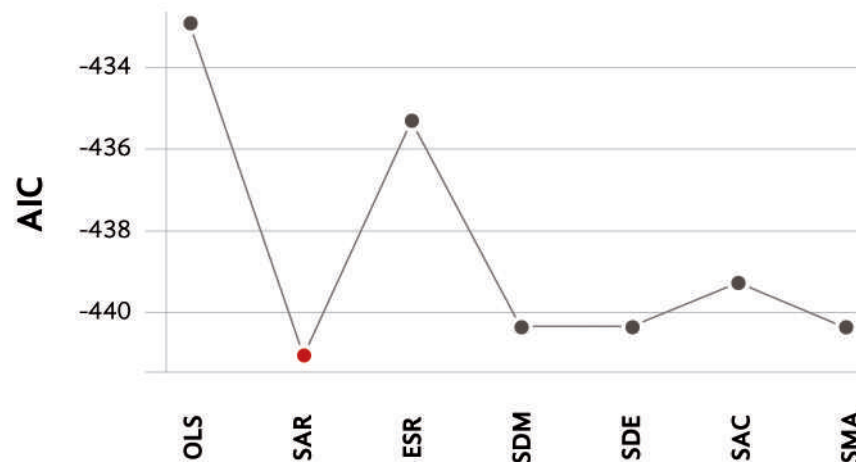
Fuente: elaboración propia con datos de inegi y con eval

La manera tradicional de analizar la validez de los modelos de regresión lineal simple consiste en revisar el signo y la magnitud de las estimaciones. Es decir, que el sentido de la estimación de los coeficientes estimados sea teóricamente correcto, por ejemplo, al incrementar el ingreso medio de los municipios es de esperarse una reducción de los niveles de analfabetismo municipal, o en sentido contrario, al incrementarse los niveles de pobreza extrema esperaríamos un incremento de los niveles de analfabetismo (ver tabla 2).

Como se ha señalado anteriormente, existe una amplia variedad de modelos de regresión espacial, por lo que uno de los problemas cruciales consiste en elegir el modelo adecuado (LeSage & Pace, 2009), para lo que

se recomienda la aplicación de pruebas para la especificación, de las que existen dos tipos: de contraste de modelo y de ajuste de datos.

Gráfica 6. Selección de modelo espacial según criterio AIC



La mayor parte de las pruebas de ajuste de modelo están dedicadas a la verificación de la existencia de correlación espacial, Se optó por emplear la prueba de multiplicadores de Lagrange. La misma compara el ajuste del modelo espacial con los resultados arrojados por el modelo de regresión lineal ordinaria, la diferencia se utiliza como criterio para determinar si el cambio relativo de la primera derivada de la función de verosimilitud alrededor del máximo afecta significativamente el parámetro autoregresivo del modelo espacial. Se debe elegir el modelo que presente el estadístico de mayor valor (ver tabla 3).

Para la selección del modelo se utilizó, adicionalmente a la prueba de Lagrange, el criterio de información de Akaike (AIC) que proporciona una medida de calidad de ajuste del modelo, en función a los datos. Dada una colección de modelos, AIC estima la calidad de ajuste de cada modelo y proporciona un medio de selección, a partir del valor de la función de máxima verosimilitud del modelo y del número de parámetros estimados. Se deberá elegir el modelo que presenté el valor de AIC más pequeño (ver gráfica 6).

La construcción del modelo espacial para el analfabetismo municipal consideró tres elementos esenciales: la heterogeneidad espacial entre municipios, la autocorrelación espacial del analfabetismo y la autocorrelación espacial de los factores que modelan los niveles de analfabetismo. Estos elementos están presentes en los seis modelos espaciales ajustados. Como puede verse en la tabla 4, se aplicaron dos modelos de autocorrelación es-

pacial: el modelo SAR de rezago espacial, y el modelo autoregresivo de error espacial ESR; dos modelos Durbin: SDM y SDE; el modelo SAC y el modelo SMA para comprobar la existencia de efectos autoregresivos sobre los errores.

La tabla 3 presenta los resultados de la prueba de Lagrange, en la misma se observa que el modelo con el valor más alto es el SAR, sin embargo, la diferencia con el modelo SDM y SMA pudiera no ser suficientemente grande, por lo que se utilizó adicionalmente al criterio AIC.

Tabla 4. Criterio AIC para diferentes modelos

No	Modelo	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

Fuente: elaboración propia

La prueba AIC confirmó que el modelo con el menor AIC fue el SAR, seguido por el modelo SDM y el SMA (ver tabla 4 y gráfica 6). Sin embargo, la diferencia con los modelos SDM y SMA continua siendo pequeña en relación con el modelo SAR, por lo que se decidió aplicar la prueba de cociente de verosimilitudes (ver tabla 5), en la misma se muestra que la diferencia entre modelos es significativamente grande y por lo tanto se debe elegir el modelo SAR. Es importante señalar que no existe una prueba de verosimilitudes para el modelo SMA, no obstante, resulta evidente que el modelo SAR es una mejor elección debido a que es un modelo con mayor nivel de parsimonia, que presenta un mayor grado de ajuste de los datos.

Tabla 5. prueba anova para modelo SAR 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

Fuente: elaboración propia con datos de inegi y con eval

Otro punto importante a resaltar es que el modelo SAR añade a la regresión tradicional un vector de rezago espacial que modela los efectos que el anal-

fabetismo municipal ejerce entre sus municipios vecinos, con la finalidad de explicar la variación intermunicipal. Intuitivamente, el modelo establece la forma como los niveles de analfabetismo en cada municipio están relacionados con el analfabetismo promedio sus vecinos. El vector de rezago espacial Wy refleja los niveles de analfabetismo promedio de los municipios ponderados a través de la matriz W , mientras que el parámetro $\rho=0.3$ refleja la fuerza de la dependencia espacial, misma que es además estadísticamente significativa (ver tabla 6).

RESULTADOS

Para interpretar adecuadamente el modelo es importante considerar que la derivada parcial de $E(y)$ respecto de la k -ésima variable explicativa tiene tres propiedades fundamentales: (1) la variable explicativa de la unidad espacial ejerce un efecto sobre la variable dependiente conocido como efecto directo; (2) el cambio sobre la variable dependiente no está en función solo de la k -ésima variable explicativa, sino también de las variables explicativas de las unidades vecinas y (3) los efectos indirectos globales cuantifican el impacto que ejerce el cambio de una variable exógena en todas las unidades espaciales vecinas dado el valor de la variable dependiente (Griffith, 2000).

El modelo ajustado de regresión espacial SAR se muestra en la tabla 5. Es importante señalar que los coeficientes estimados tienen el signo y la magnitud esperadas y que el coeficiente R-cuadrada es superior al presentado por el modelo de regresión lineal ordinario (ver tablas 2 y 6).

Tabla 6. Modelo de rezago espacial SAR, población analfabeta en 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			

Fuente: elaboración propia con datos de inegi y con eval

La interpretación del parámetro β en el modelo espacial (SAR) expresa el impacto del cambio en la variable dependiente x_i sobre el municipio i como

una combinación de influencias directas e indirectas. Esta derrama espacial se origina como efecto de las variables del modelo, lo que depende básicamente de: (1) la posición del municipio en el territorio, (2) el grado de conectividad entre los municipios, definido por la matriz W , (3) la fuerza de la dependencia espacial estimada por el parámetro ρ , y (4) la magnitud de las estimaciones de los coeficientes β (LeSage & Fischer, 2008; LeSage & Pace, 2009).

Si bien, el coeficiente β expresa el cambio de una variable independiente, ocurrido en el conglomerado formado por los vecinos del municipio i , que incide sobre la variable dependiente del municipio i , este surge como una consecuencia natural de la dependencia espacial. Cualquier cambio en las características de los municipios vecinos generará a su vez cambios que impactarán la dinámica del municipio i aledaño y viceversa. Dado que el impacto de los cambios en una variable independiente se diferencia entre regiones, es aconsejable definir una medida resumen para cada tipo de impacto, en general se identifican tres tipos: efectos directos, efectos totales y efectos indirectos.

El efecto directo proporciona una medida resumen del efecto provocado, en todo el Estado, por cambios de la variable x_r en el municipio i . Por ejemplo, si en el municipio i se incrementan los niveles de pobreza extrema, el efecto directo cuantifica su impacto promedio sobre los niveles de analfabetismo en todos los municipios de Chiapas. Esta medida toma en cuenta los efectos de retroalimentación que surgen a partir de cambios en los niveles de pobreza extrema observados en el municipio i , que impacta a sus vecinos a través del sistema de dependencias espaciales modelado a través de la matriz W .

Tabla 7. Efectos según orden de vecindad, modelo SAR

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

Fuente: elaboración propia con datos de inegi y con eval

El modelo SAR presentado en la tabla 5, señala que un incremento de un punto porcentual en la proporción de población indígena de un municipio, provocará un incremento de 3.7 puntos porcentuales promedio en los niveles de analfabetismo del Estado, en el mismo sentido debe interpretarse el efecto directo de la pobreza extrema donde el incremento de un punto porcentual de la misma puede asociarse con el incremento de 0.22 puntos de analfabetismo

en el Estado; por el contrario, el incremento en de un dólar promedio en los ingresos familiares de un municipio generaría una reducción de 18.3 puntos porcentuales en los niveles de analfabetismo del Estado.

El efecto indirecto se utiliza para medir el impacto que el incremento de una variable dependiente, en todos los municipios vecinos, ejerce sobre un municipio *i* particular. En el caso del modelo SAR de efectos fijos, ver tabla 7, se observa que el efecto indirecto del incremento de un punto porcentual de la proporción de población hablante de lengua indígena promedio, de todos los municipios de Chiapas, provocaría un incremento de 1.6 puntos porcentuales en los niveles de analfabetismo del municipio *i*. En el caso de la pobreza extrema, los efectos indirectos señalan que un incremento de un punto porcentual en los niveles de la misma en los municipios vecinos traería un incremento del analfabetismo de 0.09 puntos porcentuales, es decir, que el efecto indirecto de la pobreza extrema entre municipios que son vecinos afecta sólo de forma marginal al municipio *i*. Cuando los niveles de pobreza extrema se reducen en los municipios vecinos, los niveles de analfabetismo en el municipio *i*, mejoran marginalmente como un efecto indirecto.

La situación en el caso del ingreso municipal, en relación al efecto indirecto, indica que cuando todos los municipios vecinos mejoran sus ingresos medios, los niveles de alfabetización del municipio *i* mejoraran en un 7.7 como efecto indirecto, mientras que si sólo mejoran los niveles de ingreso del municipio *i*, su nivel de alfabetización se incrementan 18.3 puntos (efecto directo), por lo que el efecto total es de menos 26.1 puntos.

El efecto total representa la suma de efectos directos e indirectos, es decir que, si todos los municipios incrementan sus ingresos medios en un dólar, el efecto total reflejaría, el impacto promedio sobre los niveles de analfabetismo en un municipio determinado y este efecto total incluirá tanto el impacto del efecto directo como del indirecto, mismo que sería de menos 26 puntos porcentuales.

Los principales cambios en el comportamiento de los niveles de analfabetismo municipal son lo que pueden observarse entre municipios. Es decir, que las mayores desigualdades se pueden encontrar entre municipios relativamente alfabetizados como Tuxtla Gutiérrez y municipios analfabetas como San Juan Cancuc, Chanal o Mitontic (ver gráfica 1). El modelo SAR muestra la influencia (directa e indirecta) de la presencia de población indígena, la pobreza extrema y los ingresos municipales, sobre los niveles observados de población analfabeta municipal. El efecto directo de todas las variables representa en promedio 70 por ciento de los efectos totales, sin embargo, el efecto del ingreso es fundamental en términos del impacto directo e indirecto sobre los niveles observados de analfabetismo municipal. Si bien la pobreza ejerce un efecto importante sobre los niveles de analfabetismo, es la falta de ingresos el factor que realmente modifica el patrón del analfabetismo.

CONCLUSIONES

El objetivo principal es entender el papel de la heterogeneidad espacial en la determinación de los niveles de analfabetismo municipal en Chiapas. Para cumplir con dicho objetivo se planteó una estrategia basada en una metodología de regresión espacial, misma que cuantificó los efectos totales, efectos directos y efectos indirectos de los determinantes del analfabetismo del Estado. Se destacó el papel del ingreso medido en términos del ingreso corriente de las familias, estandarizado a partir del Ingreso Nacional Bruto, la proporción de población indígena y el efecto de la pobreza extrema como un proxi de la incapacidad de las personas para acceder a los servicios que los capaciten para adquirir las habilidades mínimas para alfabetizarse.

Los resultados mostraron la existencia de patrón espacial importante para explicar el comportamiento territorial del analfabetismo en Chiapas. El modelo SAR evidenció la importancia del ingreso de las familias como un determinante fundamental para predecir el comportamiento de los niveles espaciales del analfabetismo. Los resultados de la investigación ponen sobre la mesa la importancia de la disponibilidad de ingresos familiares para posibilitar la generación de capacidades básicas lecto-escritoras, como un elemento fundamental para mejorar los niveles de inserción escolar y reducir desigualdades entre grupos de poblaciones indígenas y mestizos. El trabajo cuestiona seriamente el papel de la pobreza extrema, medida de manera multidimensional, donde las carencias sociales, resultan marginales al analfabetismo, y donde los cambios en los niveles de pobreza extrema si bien, se encuentran espacialmente correlacionados con los cambios en los niveles de analfabetismo municipal, su efecto es sorprendentemente bajo, sólo 0.2 por ciento, de incremento en el analfabetismo por punto porcentual de crecimiento de la pobreza extrema municipal.

La relación inversa entre el incremento de población analfabeta y la reducción del ingreso corriente de las familias es entendible como un efecto en el que los ingresos son un factor que sobre pasa incluso la condición de pobreza extrema. Lo que implicaría que incluso en condiciones de pobreza extrema, el factor para explicar los altos niveles de analfabetismo de los municipios con mayores niveles de incidencia es la carencia de ingresos mínimos quien los lleva a la condición de analfabetismo, con lo que imposibilita el objetivo fundamental de combate a la pobreza en el Estado que es eliminar la transmisión intergeneracional de la pobreza, a través de la formación de capacidades, donde la capacidad mínima que se esperaría se viera incrementada es la lecto-escritora (López y Núñez, 2016).

La relación entre analfabetismo y el porcentaje de población hablante de lengua indígena cuantifica el efecto sobre los niveles de analfabetismo debido a cambios en la composición étnica de la población, donde se en-

cuentran inmersos factores relacionados con la accesibilidad física y la falta de escuelas cercanas a las comunidades indígenas, pero también factores relacionados con la accesibilidad cultural. La población indígena no tiene posibilidades de acceso real al sistema educativo debido a diferencias sociales, culturales o económicas. Es decir, a pesar de la existencia de una escuela cercana, la población indígena enfrenta barreras asociadas al lenguaje, la discriminación, la falta de ingresos monetarios o el costo de oportunidad, situación que lleva a mayores niveles de analfabetismo entre la población indígena. El estudio hace evidente que estas diferencias son producto de patrones de distribución espacial no aleatorios.

BIBLIOGRAFÍA

- 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 statistics. *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.