

Mediciones de Pobreza: Pobreza Energética y Necesidades Básicas Insatisfechas. Un Caso de Estudio desde la Econometría Espacial. Parte 3: Modelo Econométrico Espacial

Poverty Measurements: Energy Poverty and Unsatisfied Basic Needs. A Case Study from Spatial Econometrics. Part 3: Spatial Econometric Model

Presentación: 13 y 14 de septiembre de 2023

Gustavo Schweickardt

CONICET-Universidad Tecnológica Nacional, Facultad Regional Concepción del Uruguay
gaocerin@gmail.com

Resumen

En este trabajo se presentan los pasos para identificar cuál es el Modelo de Regresión Espacial de Corte Transversal, más apto y que incorpore Efectos Espaciales, particularmente el de Dependencia o Autocorrelación Espacial. Se adopta como variable Endógena o Independiente la relacionada con la Pobreza Energética (IDAENE), y como covariable o variable Exógena, dependiente, las Necesidades Básicas Insatisfechas (NBI), para cada Partido de la Provincia de Buenos Aires más la Ciudad Autónoma (CABA). Una vez que la existencia de Autocorrelación Espacial Global en las dos variables pudo ser identificada visualmente en Análisis Exploratorio de Datos Espaciales (AEDE), y confirmada mediante los contrastes del Índice de Moran y el de Geary (Análisis en la frontera entre el Exploratorio y Confirmatorio - AEDE - ACDE -), esta etapa se ubica en el ACDE. Los pasos referidos para la identificación del Modelo requieren introducir nuevos conceptos, para luego pasar a las estimaciones necesarias.

Palabras clave: Autocorrelación Espacial, Análisis Confirmatorio de Datos Espaciales, Modelos de Regresión Espacial

Abstract

This paper presents the steps to identify the most suitable Cross-Sectional Spatial Regression Model that incorporates Spatial Effects, particularly Spatial Dependency or Autocorrelation. The Endogenous or Dependent variable is the index related to Energy Poverty (IDAENE), and the Exogenous or Independent variable is Unsatisfied Basic Needs (UBN) for each county of the Province of Buenos Aires, including the Autonomous City (CABA). Once the presence of Global Spatial Autocorrelation in both variables, it could be identified visually in the Exploratory Analysis of Spatial Data (AEDE) and confirmed by contrasting the Moran Index and the Geary Index (Analysis on the border between the Exploratory and Confirmatory - EEDA -CEDA -), this stage is in the CEDA domain. The steps referred to for the Model's identification require introducing new concepts for proceeding with the necessary estimates.

Keywords: Spatial Autocorrelation, Confirmatory Spatial Data Analysis, Regression Models

Introducción

El Análisis Confirmatorio de Datos Espaciales (ACDE) requiere de una estrategia metodológica, definida por el investigador y soportada en diferentes conceptos, que serán introducidos a continuación (Le Gallo, 2014), (Herrera Gómez, 2015).

La detección de la presencia de Autocorrelación Espacial Global, implica que el espacio es importante. Es decir que existe un Patrón Espacial en los datos. De modo que lo que a continuación debería hacerse, es proponer un Modelo que permita incorporar los Efectos Espaciales, específicamente para la investigación reportada, la Dependencia o Autocorrelación Espacial.

1er Paso de la Estrategia:

En la Estrategia Metodológica aquí propuesta, el primer paso es identificar si la presencia de Autocorrelación detectada, es aparente o real.

Una Autocorrelación Espacial ‘aparente’, implicaría que, en realidad, lo que está generando el Patrón Espacial en la variable Endógena (IDAENE), que es la variable que se busca explicar, son, en general, las covariables o variables Exógenas propuestas en el Modelo (para este estudio, se tiene sólo una: las Necesidades Básicas Insatisfechas, (NBI)); es decir que toda la Autocorrelación Espacial de (IDAENE) será explicada por la variable predictora introducida en el Modelo (NBI). Este Modelo, es entonces, un Modelo Econométrico No Espacial, y responde a un Modelo de Regresión Lineal Simple. Por esta razón el calificativo ‘aparente’: el Patrón Espacial - Autocorrelación Espacial- observado en la variable Endógena, se explica a través de la incorporación de predictoras o Exógenas en el Modelo.

Si la Autocorrelación Espacial fuese ‘real’, no obstante sean introducidas variables predictoras en el Modelo (NBI), sigue existiendo un Patrón Espacial no explicado. Entonces el problema planteado estriba en cómo explicar ese Patrón Espacial (esa Autocorrelación Espacial), que sigue estando presente, pues se debe a la propia estructura de difusión de la variable en estudio (endógena en el Modelo de Regresión Lineal).

El procedimiento para identificar si la Autocorrelación Espacial detectada en la variable de estudio, endógena en el Modelo (IDAENE) es ‘real’ o ‘aparente’, consiste en plantear un Modelo de Regresión Lineal Simple:

$$IDAENE = NBI \beta + u; \text{ con } u \sim N(0, \sigma^2) \quad (1)$$

Donde u , perturbación aleatoria del modelo (error), tiene distribución normal con matriz de varianza σ^2 , bajo los supuestos de Independencia y homocedasticidad de la varianza. El Modelo (1), se estima por Mínimos Cuadrados Ordinarios (MCO).

Siguiendo a Cliff y Ord (1973), entonces, se calculan los ‘residuos’ post estimación, y se examinan mediante el Test de Moran (1950).

La hipótesis H_0 del Test de Moran, afirma la No existencia de Autocorrelación Espacial en los ‘residuos’. La hipótesis alternativa, H_1 , indicaría, sin mayor análisis, que existe estadísticamente Autocorrelación Espacial. De modo que, si no se rechaza H_0 , el Patrón Espacial detectado sobre IDAENE, sería explicado completamente por la introducción de la variable predictora NBI en el Modelo (1).

Entonces la Autocorrelación Espacial resultaría ‘aparente’, y se estaría frente al Modelo más simple, sin Efectos Espaciales, pues no se requiere de su incorporación, y que puede ser estimado por Mínimos Cuadrados Ordinarios (MCO).

En cambio, el rechazo de la hipótesis H_0 -aceptación de la hipótesis alternativa H_1 : ‘Existe Autocorrelación Espacial en los ‘residuos’-, si bien indicaría que la Autocorrelación Espacial resulta ‘real’, no brinda información sobre el posible Modelo, incorporando Efectos Espaciales, a especificar (el contraste de I de Moran ‘se queda a mitad de camino’).

Esta es la limitación del contraste del Índice de Moran, así aplicado (Herrera, 2015): La hipótesis alternativa H_1 , es general y no proporciona una guía clara sobre el tipo de Estructura Espacial que se encuentra en el proceso generador de datos.

Procede enfatizar frente a la presencia de Autocorrelación Espacial ‘real’, que se tiene otra distinción: el Efecto de ‘Dependencia’ o de ‘Autocorrelación’ en la especificación de un Modelo de Regresión Lineal, puede estar presente de dos formas: 1ra) ‘Dependencia Sustantiva’ o 2) ‘Dependencia Residual o del Error’.

La ‘Dependencia Sustantiva’, refiere a la Autocorrelación Espacial en algunas de las variables del modelo (Endógena o Exógenas). Es decir que la misma podría estar presente en la variable Endógena, en las Exógenas, o en ambas.

La ‘Dependencia Residual o del Error’, como su nombre lo sugiere, refiere a la Autocorrelación Espacial en la perturbación aleatoria del Modelo.

A los efectos de la interpretación, resulta de mayor interés cuando existe Autocorrelación Espacial en la variable endógena y/o en las exógenas, es decir la ‘Dependencia Sustantiva’.

2do Paso de la Estrategia:

Una alternativa que supera la limitación del contraste del Índice de Moran, son los Contrastes denominados ‘de los Multiplicadores de Lagrange (LM - *Lagrange Multipliers*)’, resultado de la aplicación del Principio de Máxima Verosimilitud. Estos contrastes tienen la ventaja de que la hipótesis alternativa se encuentra bien definida o restringida (Le Gallo, 2014), (Herrera Gómez, 2015). Desde las hipótesis alternativas asociadas a los Contrastes de los Multiplicadores de Lagrange (LM), puede plantearse diferentes Modelos basados en el Modelo Lineal General, que para el presente trabajo es de la ecuación (1). Los fundamentales se presentan a continuación.

1.- Modelo del Error Espacial (*Spatial Error Model – SEM*)

Una primera hipótesis alternativa, proviene de plantear la presencia de Autocorrelación Espacial en el término de error del Modelo de Regresión Lineal General.

De manera que se tiene un Modelo de Error Espacial Autoregresivo, para el cual, desde (1), se asume que el término de error aleatorio obedece al siguiente proceso:

$$\{IDAENE = NBI \beta + u; u = \rho W u + \epsilon\}; \text{ con } \epsilon \sim N(0, \sigma^2) \quad (2)$$

Donde ρ es el Parámetro (coeficiente) Espacial Autoregresivo, y W es la Matriz de Pesos Espaciales.

Entonces, el contraste de los Multiplicadores de Lagrange (LM) para detectar Autocorrelación Espacial en el Error, establece las siguientes hipótesis (nula, H_0 , y alternativa, H_1): $H_0 \rightarrow \rho = 0$; $H_1 \rightarrow \rho \neq 0$.

2.- Modelo de Rezago Espacial (*Spatial Lag Model – SLM*)

Una segunda hipótesis alternativa, proviene de considerar un Modelo con Estructura Espacial Sustantiva, referido como Modelo de Rezago Espacial (SLM). Este Modelo incorpora en (1) un rezago espacial de la variable dependiente o endógena, ($W Y$), como una nueva explicativa:

$$\{IDAENE = \lambda W IDAENE + NBI \beta + u\}; \text{ con } \epsilon \sim N(0, \sigma^2) \quad (3)$$

Donde λ es el Parámetro (coeficiente) de Rezago Espacial Autoregresivo, y W es la Matriz de Pesos Espaciales, con los demás elementos ya definidos.

El contraste de los Multiplicadores de Lagrange (LM) para la detección de dependencia sustantiva, establece las siguientes hipótesis (nula, H_0 , y alternativa, H_1): $H_0 \rightarrow \lambda = 0$; $H_1 \rightarrow \lambda \neq 0$.

Los Contrastes de los Multiplicadores de Lagrange Robustos (LM*)

Los contrastes LM presentan, como inconveniente, que son sensibles a diferentes tipos de errores de especificación. Por ejemplo, el LMERROR detecta autocorrelación espacial debido a la presencia de un rezago

espacial de la variable endógena ($WIDAENE$), y lo mismo puede decirse del LMLAG, que brinda falsos positivos cuando el término de error contiene un rezago espacial (Wu).

Ante estos inconvenientes, según Herrera (2015), Anselin (1988) y Anselin et al. (1996), se propusieron dos nuevos Multiplicadores de Lagrange diseñados para que su comportamiento sea robusto a dichos errores de especificación:

- El LM*ERROR: analiza la falta de correlación en los residuos, siendo robusto a la omisión del término ($WIDAENE$).
- El LM*LAG: permite detectar la autocorrelación espacial sustantiva en presencia de estructura espacial en el término de error.

Estos contrastes permiten incorporar Elementos o Efectos Espaciales en el Modelo de Regresión Lineal (1) (resuelto por MCO), conforme al rechazo o no de cada una de las hipótesis nulas. De tal modo, puede establecerse una estrategia de especificación, como sigue:

- Si el contraste I de Moran, rechaza $H_0 \rightarrow$ Evidencia a favor de inclusión de Elementos o Efectos Espaciales.
- Si los contrastes LMERROR y LM*ERROR, rechazan $H_0 \rightarrow$ Evidencia a favor de un Modelo de Error Espacial (SEM).
- Si los contrastes LMLAG y LM*LAG, rechazan $H_0 \rightarrow$ Evidencia a favor de un Modelo de Rezago Espacial (SLM).
- Si no se rechaza H_0 bajo ninguno de los contrastes \rightarrow Evidencia a favor del Modelo Lineal General no Espacial (1).
- Si ambos contrastes robustos, LM*ERROR y LM*LAG, rechazan H_0 , entonces se deberán incorporar Elementos o Efectos Espaciales en la parte sistemática ($WIDAENE$) y aleatoria (Wu).

De forma que utilizando estos contrastes conjuntamente, puede especificarse el Modelo más adecuado introduciendo Elementos o Efectos Espaciales.

Desarrollo

Son presentados a continuación, los pasos de la Estrategia Metodológica para este reporte final de la investigación realizada. Se siguen las consideraciones que se establecieron en la Parte 2, respecto de la Matriz de Pesos Espaciales.

Se opta por W Tipo Reina Normalizada por Filas.

Source	SS	df	MS	Number of obs	=	149
Model	121.9802	1	121.9802	F(1, 147)	=	115.09
Residual	155.803276	147	1.05988623	Prob > F	=	0.0000
				R-squared	=	0.4391
				Adj R-squared	=	0.4353
Total	277.783477	148	1.87691538	Root MSE	=	1.0295

IDAENE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
NBI	.252479	.0235348	10.73	0.000	.2059688 .2989892
_cons	3.324969	.1649372	20.16	0.000	2.999015 3.650924

Tabla 1. Resultado de la Estimación del Modelo No Espacial por Mínimos Cuadrados Ordinarios (MCO), reportado por STATA 16

El Modelo (1), resulta, entonces, significativo, y cumple (no se reportan los test correspondientes por cuestiones de espacio), con todos los supuestos (Linealidad, Independencia, Normalidad en la Distribución de residuos, y Homocedasticidad de la varianza).

Test	Statistic	df	p-value
Spatial error:			
Moran's I	7.345	1	0.000
Lagrange multiplier	48.237	1	0.000
Robust Lagrange multiplier	1.358	1	0.244
Spatial lag:			
Lagrange multiplier	55.771	1	0.000
Robust Lagrange multiplier	8.892	1	0.003

Tabla 2. Resultado de los Contrastes de Moran sobre los residuos y de los Multiplicadores de Lagrange en la Estimación del Modelo No Espacial por Mínimos Cuadrados Ordinarios (MCO). Matriz de Pesos Espaciales según el Criterio de Contigüidad, Tipo 'Reina', Normalizada por Filas. Reporte de STATA 16

Los resultados del Tabla 2, indican lo siguiente:

- El Contraste del I Moran sobre los 'residuos' resulta significativo, y hay evidencia para rechazar H0. De modo que la Autocorrelación Espacial persiste, no es explicada por la única predictora (NBI) del Modelo No Espacial, y deberá pasarse a la consideración de un Modelo Espacial (existe evidencia a favor de inclusión de Elementos o Efectos Espaciales);
- Por otro lado, el contraste de Multiplicadores de Lagrange para el Error Espacial es significativo LMERROR, pero el robusto, LM*ERROR, no lo es. De manera que podría aceptarse la H0 ($\rho = 0$), inclinándose a la especificación de un Modelo sin Efectos Espaciales en el término de error;
- El contraste de Multiplicadores para el Rezago Espacial, LMLAG, y el robusto, LM*LAG, son significativos, rechazan H0 ($H_0 \rightarrow \lambda = 0$; $H_1 \rightarrow \lambda \neq 0$). De los resultados en su conjunto, se tendría evidencia a favor de un Modelo del Rezago Espacial (SLM). Se estima, entonces, este Modelo, dado por la ecuación (3), de forma 'robusta', por el Método de Máxima Verosimilitud (*Maximum Likelihood*) o (ML).

Spatial autoregressive model	Number of obs	=	149
Maximum likelihood estimates	Wald chi2(2)	=	224.61
	Prob > chi2	=	0.0000
Log likelihood = -193.04358	Pseudo R2	=	0.4831

IDAENE	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
IDAENE					
NBI	.1503159	.0237251	6.34	0.000	.1038156 .1968162
_cons	1.333855	.2959066	4.51	0.000	.7538883 1.913821
Wc					
IDAENE	.5380194	.0708819	7.59	0.000	.3990935 .6769453
var(e. IDAENE)	.7304387	.0859359			.5800163 .919872

Wald test of spatial terms: chi2(1) = 57.61 Prob > chi2 = 0.0000

Tabla 3. Resultado de la Estimación del Modelo del Rezago Espacial (SLM), por Máxima Verosimilitud (ML). Matriz de Pesos Espaciales según el Criterio de Contigüidad, Tipo 'Reina', Normalizada por Filas. Reporte de STATA 16

Este reporte de STATA 16, confirma la significatividad de las dos variables predictoras incorporadas en la ecuación (3): la Exógena, NBI, y la Endógena rezagada espacialmente, W IDAENE, resultando el Modelo SLM el más apto para explicar la Dependencia o Autocorrelación Espacial detectada.

Conclusiones

Más allá de los resultados reportados, se presentan los siguientes puntos a modo de conclusiones finales:

1ero) Ha sido descripta de la forma más simple posible, la metodología que debe seguirse en la identificación del Modelo Econométrico Espacial que explique la Dependencia o Autocorrelación Espacial en la variable IDAENE, adoptada como Endógena, y asociada a la Pobreza Energética promedio en cada Partido de la Provincia de Buenos Aires y CABA.

2da) La variable Necesidades Básicas Insatisfechas (NBI), resulta ser una predictora estadísticamente significativa del Índice de Dificultad de Acceso a la Energía (IDAENE), así como también, para explicar la Dependencia Espacial, lo es la propia variable IDAENE rezagada espacialmente. El Modelo más adecuado, es referido como SLM.

3era) En cuanto a los objetivos generales, reportados de manera breve, planteados en esta investigación, se resalta, más allá de los resultados obtenidos, el potencial de la Econometría Espacial en diferentes campos del conocimiento.

Referencias

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic. 294 pages.

Anselin, L., Bera, A.K., Florax, R., Yoon, M.J. (1996). "Simple diagnostic tests for spatial dependence". *Regional Science and Urban Economics*, Vol. 26, Issue 1, pp. 77-104.

Cliff, A.D., Ord, J.K. (1973). *Spatial Autocorrelation*. London: Pion. 178 pages.

Herrera Gómez, M. (2015). "Econometría especial usando Stata. Breve guía aplicada para datos de corte transversal". Instituto de Estudios Laborales y del Desarrollo Económico (ielde), Facultad de Ciencias Económicas, Jurídicas y Sociales, Universidad Nacional de Salta (UNSa). Disponible en: <http://hdl.handle.net/11336/7116>

Moran, P. A. (1950). "Notes on continuous stochastic phenomena". *Biometrika*, Vol. 37, No. ½, pp. 17-23.

Le Gallo, J. (2014). "Cross-Section Spatial Regression Models". In: Fischer, M., Nijkamp, P. (eds) *Handbook of Regional Science*. Springer, Berlin, Heidelberg.