

Fenómenos de difusión y determinantes de la pobreza en Colombia: un análisis desde la econometría espacial

Fabio Augusto Rueda Devivero*
Fabio_a_rueda@hotmail.com

RESUMEN

Este documento aborda la problemática de la pobreza, en un contexto de dependencia espacial, y a partir de sus determinantes y la posible existencia de fenómenos de difusión. En primera instancia se verifica la significancia estadística del fenómeno de autocorrelación o dependencia espacial para el periodo de tiempo y la muestra seleccionada, y luego mediante la especificación y estimación de un modelo de determinantes de pobreza, se trata de verificar el tipo de difusión (contagiosa o jerárquica), y las variables que influyen en su nivel. Los resultados no fueron concluyentes en identificar el patrón de difusión, pero apuntaron a que la pobreza guarda relación con el ciclo económico y a que no está influenciada por factores determinísticos asociados a la geografía.

Palabras claves: autocorrelación espacial, difusión espacial, determinantes de la pobreza, datos de panel.

Clasificación JEL: I32, R12, R15, C21.

* Economista, Universidad Tecnológica de Bolívar. Para comentarios o inquietudes favor, dirigirse al teléfono 3008056021. El autor agradece la valiosa y desinteresada colaboración prestada por Javier Pérez Valbuena durante todas las fases de elaboración del presente documento. Igualmente agradece los aportes y comentarios hechos por William Arellano, Daniel Toro, Julio Romero, Jaime Bonet y José Gamarra.

INTRODUCCIÓN

La pobreza es un fenómeno social con innumerables connotaciones, y por ende numerosas son las ramas de las ciencias sociales que la han abordado, entre ellas la economía, catalogándolas todas ellas quizás como el problema más complejo, apremiante y urgente de solucionar. También es abundante lo que de ella se ha escrito en el país, análisis descriptivos se encuentran muchos, ubicación geográfica de las zonas más vulnerables, caracterización de las personas pobres, discusiones metodológicas sobre su medición, y por supuesto, posibles determinantes han sido abordados, todo ello con el propósito de ahondar en su estudio y poder a partir de allí, formular políticas efectivas para combatirla.

Sin embargo, ninguna consideración se había hecho sobre la importancia e implicaciones del espacio en su determinación y dinámica. En el trabajo de PEREZ (2005) se da un primer paso en este sentido, al demostrar que tanto a nivel departamental como municipal, para los años 1985 y 1993, es significativo el fenómeno de dependencia espacial para variables de pobreza (NBI e ICV) en Colombia.

Es decir, que el nivel de pobreza, ya sea que se mida por el NBI o el ICV, de una unidad espacial dada (departamento o municipio), no depende únicamente de las condiciones económicas, sociales, políticas o de cualquier otra índole que imperen en dicha unidad, sino que también se ve afectado por la situación y las condiciones que se afronten en otras unidades.

Ante este nuevo contexto, son muchos los interrogantes que surgen, entre ellos dos que a juicio del autor resulta interesante y pertinente abordar, y son, por una parte el tipo de difusión espacial que en Colombia se presenta y que se desprende de haber sido demostrado que existe dependencia espacial en la distribución de la pobreza, y por otra parte los determinantes que en este contexto influyen sobre el nivel de este fenómeno en nuestro país.

En la literatura sobre el tema, se habla de dos tipos de fenómenos de difusión, contagiosa y jerárquica. La primera se presenta cuando son las características y propiedad intrínsecas del problema las que provocan su expansión, por lo que este debe comenzar a expandirse a las unidades vecinas, mientras que la difusión jerárquica se presenta cuando son los determinantes del problema los

que provocan que este se expanda o surja, lo cual no necesariamente debería ocurrir entre unidades vecinas. Las implicaciones de política ante una situación u otra son completamente diferentes, ya que en presencia de difusión por contagio los programas deben atacar los focos más agudos del problema y a partir de estos se comenzarán a contagiar los vecinos, mientras que ante una situación de difusión jerárquica, se deben intervenir por medio de los determinantes del problema y en cada uno de los lugares donde este se presente. Por lo tanto, de acuerdo al escenario, la estrategia y el costo de atacar la pobreza debería cambiar notablemente.

Igualmente, en el contexto de dependencia espacial de la pobreza en Colombia, sus determinantes cobran relevancia, ya que cualquier ejercicio que haga sin incorporar dicho fenómeno, debe conducir a resultados deficientes, y se requiere por lo tanto del uso de técnicas de análisis espacial para poder abordarlo. En cualquier escenario, tanto de difusión contagiosa como jerárquica, resulta importante conocerlos, así como la magnitud y el sentido en el cual la afectan, si lo que se desea es el diseño y estructuración de programas efectivos para combatirla.

En general, lo que se pretende es, que a partir de haber hallado que el espacio cobra importancia en la distribución de la pobreza, determinar entonces el tipo de difusión espacial que se presenta y los determinantes de esta, los cuales en un contexto u otro resultan relevantes para poder combatirla.

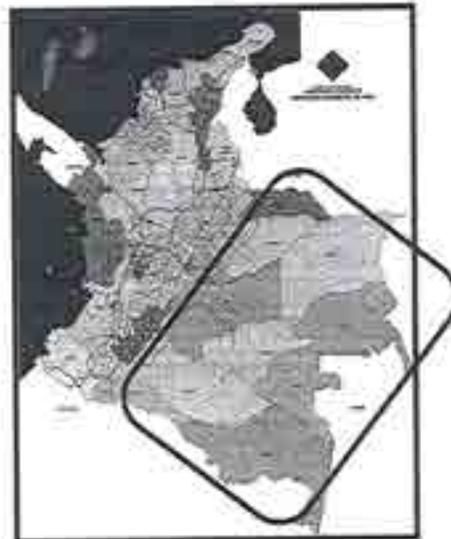
Se hará uso de información departamental para el periodo de tiempo comprendido entre 1996 y 2000, dejando claro de antemano que la elección tanto del número de años como del nivel de agregación estuvo condicionada a la disponibilidad de la información requerida.

Es importante resaltar que serán sustraídos de la muestra un total de nueve (9) departamentos, los cuales corresponden a: Arauca, Vichada, Casanare, Guainia, Guaviare, Vaupes, Putumayo, Amazonas y San Andrés. En el caso de los ocho (8) primeros, se debe básicamente a la inexistencia de la mayoría de las estadísticas requeridas, y en el caso de San Andrés, por el hecho que la técnica cuantitativa que se empleara requiere estricta contigüidad física entre las unidades, y el hecho de ser isla hace que se viole dicho requerimiento. Sin embargo, esto no se debe considerar una dificultad, por dos razones fundamentales, y es que la población de estos departamentos representa

únicamente ente el 5% y 6% de la población nacional, además que corresponde a una zona que no se encuentra conectada por medio de la red vial nacional con el resto del país (Mapa 1), lo cual dificulta la comunicación y el transporte, por lo que inclusive se podría especular, con que de incluirse se crearía un sesgo a rechazar la hipótesis de autocorrelación espacial.

El documento esta organizado de la siguiente manera: después de esta introducción se hará una revisión a nivel teórico de los conceptos de heterogeneidad, autocorrelación y difusión espacial, se citaran aplicaciones empíricas previas tanto a nivel nacional como internacional y se describirá el modelo de determinantes de pobreza a utilizar. En la siguiente sección se explicará la metodología, se mostrarán los resultados de los ejercicios y se dará una interpretación de los mismos, para luego pasar a dar algunas conclusiones y recomendaciones.

Mapa 1
Red vial nacional



Fuente: Tomado de la web del Instituto Nacional de Vías (INVIAS) www.invias.gov.co

1. MARCO TEORICO

Tal como se enmarca en el título y en las líneas previas del presente documento, son dos los escenarios en los cuales será abordado el problema de la pobreza en Colombia. Dichos escenarios son: la difusión espacial y los posibles determinantes de la pobreza. Por ello, a continuación se procederá a desarrollar en el plano teórico cada uno de los dos escenarios, a revisar en la literatura tanto nacional como internacional la forma como estos han sido abordados, y a establecer la relación que entre ellos existe.

1.1. Difusión espacial

Se comenzará por tratar el tema de la difusión espacial, y para ello es necesario hablar de un concepto que resulta muy importante para los propósitos de este trabajo, y es la autocorrelación espacial.

Con respecto a ella, se puede decir que junto con la heterogeneidad espacial son los dos conceptos que dieron vida a la econometría espacial, y permitieron que surgiera como una subdisciplina de la econometría general. En cuanto a la heterogeneidad, en palabras de MORENO y VAYA (2000): "...aparece cuando se utilizan datos de unidades espaciales muy distintas para explicar un mismo fenómeno. En este caso aparecen problemas como la heteroscedasticidad o la inestabilidad estructural, los cuales pueden ser resueltos mediante las técnicas econométricas existentes para series temporales...".

Por su parte, la autocorrelación espacial surge en la medida que el valor de una variable en un lugar del espacio, este relacionado con el valor de dicha variable en otro u otros lugares. En la medida que esto se presente, hace que se viole uno de los principales supuestos con los que se trabaja en ejercicios de corte transversal, y tiene que ver con la independencia de las observaciones muestrales, lo cual conduce a problemas de ineficiencia en cualquier estimación mínimo cuadrática que se lleve a cabo².

² El término ineficiencia en el contexto de estimación econométrica se refiere a que el vector de parámetros que se obtuvo, no es el que tiene la menor varianza dentro del conjunto de vectores que se podría obtener.

³ El estimador de Mínimos Cuadrados Ordinarios (MCO) viene dado por la expresión: $\hat{\beta} = (X'X)^{-1}X'y$; En presencia de autocorrelación espacial ($E(u_i) = 0$; $\text{Var}(u_i) = \sigma^2$) dicho estimador será insesgado ($E(\hat{\beta}) = \beta$) dado que el supuesto hecho sobre su primer momento no se ha violado. Sin embargo, no sería eficiente, tal como se puede observar a continuación:

$$= \sigma^2 \sum_{i=1}^n (1 - \rho_i)^2 (1 - \rho_i^2) \quad (1)$$

$$= \sigma^2 \sum_{i=1}^n (1 - \rho_i^2) X_i' X_i (1 - \rho_i^2) \quad (2)$$

$$= \sigma^2 \sum_{i=1}^n (1 - \rho_i^2) X_i' X_i \rho_i^2 X_i (1 - \rho_i^2) \quad (3)$$

Sin embargo, a diferencia del problema de heterogeneidad, el de autocorrelación no se puede solucionar a partir de las técnicas convencionales, debido a la multidireccionalidad presente en las relaciones de interdependencia de las unidades espaciales. Es decir, a diferencia de los análisis de series de tiempo, donde el problema de autocorrelación implica que una observación en un momento del tiempo, esta influenciada por observaciones de periodos de tiempo previos, en el caso de datos de sección cruzada, las relaciones entre los diferentes agentes o unidades se da en múltiples direcciones. Para ilustrar esta situación, en el mapa 1 es posible observar, de manera hipotética, las múltiples direcciones en que una unidad espacial podría afectar a otras, y a la vez la forma en que dicha unidad podría verse afectada.

Mapa 1
Multidireccionalidad de las relaciones espaciales



El instrumento que se ha desarrollado para captar la interdependencia entre unidades espaciales, es la matriz de contigüidad³. Se trata de una matriz W:

Queda claro entonces la ineficiencia de este estimador, dado que la varianza dada por la expresión (3) es mayor que la que se obtendría sino se violara ningún supuesto, la cual es: $\sigma^2 \cdot \frac{1}{n} \cdot (X'X)^{-1}$

³ Matriz de contigüidad no es el único nombre que recibe, y es muy común referirse a ella como matriz de pesos espaciales.

$$W = \begin{bmatrix} 0 & w_{12} & \dots & w_{1n} \\ w_{21} & 0 & \dots & w_{2n} \\ \dots & \dots & 0 & \dots \\ w_{n1} & w_{n2} & \dots & 0 \end{bmatrix}$$

La cual tiene como características, ser cuadrada y no estocástica, y donde los elementos w_{ij} reflejan la intensidad de la relación existente entre cada par de unidades i y j ⁴.

Para poder verificar la presencia o no de autocorrelación espacial, se debe recurrir a un test, en donde la hipótesis nula plantea la existencia de una distribución aleatoria de la variable bajo estudio (no autocorrelación espacial), y la alterna la existencia de una asociación significativa de valores similares (autocorrelación espacial positiva) o disímiles (autocorrelación espacial negativa). Se debe recurrir para ello, a la construcción de alguno de los estadísticos de dependencia espacial que se han desarrollado, tal como el I de Moran y el C de Geary⁵. Las expresiones de las anteriores medidas son:

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad i \neq j$$

Y

$$C = \frac{N-1}{2S_0} \frac{\sum_i w_i (x_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}$$

Donde, $I \wedge C \rightarrow N(0,1)$

En ambos casos, x_i refleja el valor de la variable cuantitativa x en la unidad i , \bar{x} es su media muestral, w_i los pesos de la matriz W , N es el tamaño muestral y $S_0 = \sum_i \sum_j w_{ij}$.

⁴ Para conocer en detalle la forma como se definen los pesos (w_{ij}) y se construye la matriz W , se recomienda remitirse al Anexo 1 del presente documento.

⁵ Para obtener más información sobre los estadísticos de autocorrelación espacial I-Moran y C-Geary, así como de otros estadísticos que se han desarrollado, se recomienda remitirse al Anexo 2 del presente documento.

Ahora bien, una vez se ha logrado determinar la existencia de autocorrelación o dependencia espacial en la distribución de una variable, es decir, se ha logrado demostrar que el valor de una variable en una unidad dada, esta relacionada con el valor de la misma variable en alguna otra unidad, vale la pena entonces indagar sobre la forma como se relacionan y la manera cómo entre ellas se influyen. Esto último es lo que se denominara difusión espacial, y al respecto se distinguen de dos tipos: contagiosa y jerárquica.

Hay dos elementos que diferencian los tipos de difusión mencionados anteriormente, y son, que la primera implica necesariamente vecindad y es el fenómeno que se este analizando o la variable con que se este trabajando, por sus propiedades intrínsecas las que provocan que se expanda, mientras en la jerárquica, no necesariamente debe haber vecindad y son los determinantes los que provocan que se difunda.

En Colombia, se podría decir que existe un incipiente y creciente interés por el análisis de las relaciones espaciales que diferentes variables presentan a lo largo del territorio, y se encuentran aplicaciones a indicadores de violencia homicida, cultivos ilícitos, descentralización y recientemente a la pobreza. FORMISANO (2002), para el caso de Bogotá, encuentra mediante la aplicación de un análisis exploratorio y confirmatorio de datos espaciales (ESDA y CSDA), que los homicidios se concentran en unas pocas zonas, las cuales se encuentran agrupadas y mantienen todas ellas continuidad geográfica. Las pruebas de autocorrelación espacial validaron lo anterior y logra determinar también que los homicidios presentan patrones de difusión contagiosa.

SÁNCHEZ & DÍAZ (2004), tratando de establecer la relación entre área sembrada con cultivos ilícitos y presencia de actores armados ilegales, encontraron entre otras cosas, dependencia espacial positiva en ambos casos y una relación directa entre ellas. BONET (2004) en su estudio sobre el impacto de la descentralización fiscal en la evolución de las disparidades en el ingreso, encuentra que ninguna de las variables que emplea presentan asociaciones espaciales significativas.

Con el trabajo de PÉREZ (2005), se efectúa una aproximación de las técnicas espaciales al fenómeno de la pobreza, destacándose dentro de sus resultados, asociaciones significativas de varios indicadores de pobreza y bienestar tanto a

nivel departamental como municipal, así como para cada una de las regiones por separado. Las pruebas se efectuaron para los años 1985 y 1993, siendo en este último año la relación más fuerte que en el primero, lo cual apunta a un acentuamiento de la importancia del espacio a través del tiempo.

A nivel internacional, la literatura en este campo naturalmente es más amplia, por lo que se mencionaran solo algunas aplicaciones, tales como las siguientes: ALMEIDA, HADDAD & HEWINGS (2003), mediante la aplicación del ESDA analizan los patrones espaciales que presenta la actividad criminal en el estado de Mina-Gerais (Brasil), encontrando dependencia espacial positiva, clusters tanto de alta como de baja concentración, y la existencia de tendencia espacial en el comportamiento del crimen en 1995.

HADDAD & BUDIC (2004), mediante un ESDA y CSDA encuentran concentraciones significativas dentro del estado de Sao Paulo (Brasil) del desarrollo humano y económico. Dichas concentraciones son de tipo centro-periferia, por lo que en la segunda parte del trabajo se centran en analizar si las inequidades se han ido cerrando o por el contrario ampliando, y si las políticas adoptadas han contribuido a converger o a divergir. Los resultados apuntan a que no hay transferencia del desarrollo del centro hacia la periferia, y las políticas en lugar de contribuir a cerrar la brecha, la han acentuado.

Por mencionar un último ejemplo, REY (2001) incorpora el papel de la dependencia espacial en el análisis de la evolución de la distribución del ingreso regional en Estados Unidos entre 1929 y 1999. Un aporte importante de este trabajo es la incorporación de nuevas medidas de autocorrelación espacial, adicionales y complementarias al tradicional I-Moran.

1.2. Determinantes de la pobreza

Los determinantes de la pobreza, serán abordados por dos importantes razones, las cuales son:

- No se encontraron antecedentes de ejercicios de sección cruzada ni datos de panel, ya que los estudios previos para el caso Colombiano, han sido elaborados en el tiempo y tomando como agregado a todo el país.
- En PÉREZ (2005), se demuestra la presencia de dependencia espacial en la pobreza en Colombia, tanto a nivel departamental como municipal para

1985 y 1993. Esto es un fuerte indicio de la importancia que cobra el espacio en la distribución de este fenómeno a lo largo del territorio nacional, por lo cual la incorporación de esta variable (espacio) es necesaria a la hora de modelarlo.

Los trabajos sobre este tema que se mencionaron antes, y que tomaron como unidad de estudio al agregado nacional, corresponden, el primero a NUÑEZ & SANCHEZ (1999) y el segundo a NUÑEZ & RAMÍREZ (2002). En el primero, los autores hacen uso de variables de tipo demográfico, macroeconómico y de cambio estructural, y llegan a la conclusión que en el largo plazo variables como el desempleo, el tipo de cambio depreciado y el deterioro de los términos de intercambio, están relacionados con los mayores niveles de pobreza, mientras que los niveles de escolaridad y la productividad multifactorial y laboral, guardan relación inversa con ella. En el segundo trabajo, se encuentra que el desempleo y el tamaño del hogar explican, en gran medida, los niveles de pobreza registrados entre 1991 y 2000. Igualmente, la mejora en la distribución del ingreso, la devaluación y el incremento del salario mínimo real, tienen un efecto notable en la disminución de la pobreza.

Trabajos similares a este último, se encontraron para países como México y Malawi, donde haciendo uso de datos microeconómicos obtenidos de encuestas de hogares e ingresos y gastos, estiman probabilidades, para que hogares con ciertas características, estén o caigan en la pobreza. Las características tienen que ver con el tamaño, nivel educativo del jefe de hogar, edad del cabeza de familia, ubicación de la vivienda (en área rural o urbana), entre otras.

El ejercicio que en el presente trabajo se pretende llevar a cabo, esta fuertemente influenciado por las experiencias antes expuestas, sin embargo, con el propósito de adaptarlo a los objetivos y necesidades que en este caso se tienen, será replanteado en importantes aspectos. Adicionalmente, se propondrá un posible sustento teórico para la selección de las variables causales de la pobreza, dado que de los esfuerzos previos, queda la sensación que son esencialmente empíricos.

1.2.1. El modelo

La premisa de la cual se partirá para explicar la pobreza, es que ésta se

encuentra en función del crecimiento económico, así:

$$\text{Pobreza} = f(\text{crecimiento económico})$$

La idea es bastante intuitiva, y diversos son los canales por los cuales el crecimiento puede alterar los niveles de pobreza, tales como el nivel de empleo, el nivel de ingreso, los niveles de productividad, entre otros.

Surge sin embargo, una dificultad derivada de la posible simultaneidad o endogeneidad entre las variables, es decir, la posibilidad que de forma simultánea el crecimiento económico explique la pobreza, y a la vez la pobreza contribuya a explicar el crecimiento⁶, sobre lo cual se encuentra una amplia literatura, en buena medida recogida en un trabajo de LUSTIG, ARIAS & RIGOLINI (2002). Este hecho, impediría el uso del cambio entre (t) y (t-1) del Producto Interno Bruto (PIB) como medida del crecimiento económico, por lo que sería necesario entonces utilizar algún tipo de instrumento⁷.

Para tal fin, se hará uso de un conjunto importante de variables, la mayoría de las cuales fueron utilizadas en su agregado nacional en los trabajos de NUÑEZ (1999 y 2002), que se denominarán variables macroeconómicas, y otras que el autor se atreve a proponer, las cuales serán sustentadas en próximas líneas, y a las que se les denominará variables geográficas y variables institucionales. De esta forma, esta nueva versión del modelo se representaría de la siguiente manera:

$$\text{Pobreza} = f(\text{Var. Macroeconómicas; Var. Geográficas; Var. Institucionales})$$

A continuación se enunciarán, aquellas que fueron seleccionadas dentro de cada categoría, en base al hecho de que son, sobre las que existe tanto en la literatura teórica como empírica más consenso sobre su relación con el crecimiento. Las cuales son:

⁶ En el Anexo 3 del presente documento se encuentra desarrollada en detalle la prueba de endogeneidad de Hausman, con la cual se logra probar la significancia estadística de la simultaneidad entre la pobreza y el crecimiento, para el caso Colombiano y en el período de tiempo relevante para el presente trabajo.

⁷ Instrumentar una variable consiste en: supóngase un modelo lineal II ($y = \beta_0 + \beta_1 x + u$) donde hay simultaneidad entre X e Y. Si existe una variable Z, tal que este fuertemente correlacionada con X y no lo este con el error (u), entonces, Z sirve como instrumento para X y debe romper el problema de endogeneidad o simultaneidad presente en el modelo original.

CATEGORIA	NOMBRE	
Macroeconómica	Inflación	
	Tasa de cambio real	X
	Desempleo	X
	Productividad laboral	X
	Productividad factorial	
	Escolaridad	X
	Tamaño promedio del hogar	
	Distribución del ingreso	X
	Participación del sector servicios en el PIB	X
Geográfica	Productividad agrícola	X
	Condiciones de salud	
	Frecuencia de desastres naturales	
	Acceso a los mercados	
	Urbanización	
	Densidad de la población	
Institucional	Funcionarios públicos sancionados por los entes de control fiscal y disciplinarios	

Las que aparecen marcadas con una X, son las que de forma definitiva fueron seleccionadas como variables explicativas en el modelo, respecto de lo cual cabría decir:

- La inflación, la productividad factorial, el tamaño promedio del hogar y el porcentaje de funcionarios públicos sancionados por los entes de control fiscal y disciplinario, no han sido seleccionados, básicamente por el hecho de que no se encontraron estadísticas sobre ellos en la forma como el presente ejercicio las requiere, que es a nivel departamental.
- Sobre el índice de tasa de cambio real (ITCR) cabe decir, que empíricamente no se ha encontrado un resultado robusto sobre la relación que tiene con la pobreza. A pesar de esto, se esperaría un efecto positivo de las exportaciones netas (XN) sobre el producto, por lo cual, al ser la relación entre ITCR y XN directa, debe entonces ser inversa la relación del ITCR con la pobreza.
- En cuanto al desempleo, el tamaño del hogar y la inequidad en la distribución del ingreso se espera tengan una relación directa con la pobreza, mientras en el caso de la productividad laboral y factorial, la escolaridad y la participación de los servicios en el PIB, se espera sea inversa.
- Dentro de las variables geográficas se ha seleccionado la productividad agrícola. En un trabajo reciente de GALVIS (2001) se demuestra que para los departamentos Colombianos, esta tiene una fuerte correlación, además

de una relación de causalidad con un conjunto importante de factores geográficos. El resto de las variables listadas en esta categoría no son tenidas en cuenta, básicamente por la dificultad o la inexistencia de estadísticas relativas o aproximadas, tanto a nivel nacional, como departamental, siendo estas últimas las requeridas en este caso. Se espera una relación inversa entre esta y la pobreza.

- Con el propósito de medir el papel de las instituciones en el crecimiento, se propone incluir un indicador que involucre el total de funcionarios públicos sancionados por los entes de control fiscal y disciplinario en cada departamento, bajo el argumento que no solo es importante que exista un marco legal y regulatorio definido, sino también que este se cumpla, por lo que cada sanción vendría a representar entonces una perturbación a las reglas. Se espera por lo tanto una relación directa.

Recogiendo la información y los criterios antes expuestos, y expresándolos de manera formal, se tendría una expresión como la siguiente:

$$Y/o \quad Pb_{i,t} = [ITCR_{i,t}, D_{i,t}, Pl_{i,t}, Es_{i,t}, G_{i,t}, Ser_{i,t}, Pag_{i,t}]$$

(+), (-), (-), (+), (-), (-)

$$Pb_{i,t} = \beta_0 + \beta_1 ITCR_{i,t} + \beta_2 D_{i,t} + \beta_3 Pl_{i,t} + \beta_4 Es_{i,t} + \beta_5 G_{i,t} + \beta_6 Ser_{i,t} + \beta_7 Pag_{i,t} + e_{i,t}$$

donde:

i: i-ésimo departamento de Colombia (según la división política definida por la constitución de 1991)

t: corresponde a los años 1996, 1997, 1998, 1999 y 2000

$Pb_{i,t}$: línea de pobreza del departamento *i* en *t*.

$ITCR_{i,t}$: índice de tasa de cambio real del departamento *i* en *t*.

$D_{i,t}$: tasa de desempleo del departamento *i* en *t*.

$Pl_{i,t}$: productividad laboral del departamento *i* en *t*.

$Es_{i,t}$: años promedio de educación de la población mayor de 15 años en el departamento *i* en *t*.

$G_{i,t}$: coeficiente de gini del departamento *i* en *t*.

$Ser_{i,t}$: participación del sector servicios en el PIB del departamento *i* en *t*.

$Pag_{i,t}$: productividad agrícola del departamento *i* en *t*.

$e_{i,t}$: perturbación o término de error aleatorio, que se supondrá ruido blanco.

1.3 Difusión espacial y determinantes de la pobreza: ¿Cómo se integran?

Sobre estos dos conceptos, podría pensarse en primera instancia que no podrían relacionarse o complementarse, sin embargo, no sería muy acertado pensar en esa forma, por lo que a continuación se tratará de argumentar que dicha complementariedad puede existir. Para ello, se comenzará por ahondar en la intuición del fenómeno de difusión espacial, ya que previamente sólo se dio su definición y se dijo a partir de qué se podía deducir.

El fenómeno de difusión se desprende de confirmar la presencia de dependencia espacial, lo cual implica que el nivel de una variable en un lugar dado, no depende exclusivamente de las condiciones imperantes en el, sino que también se ve influido por las situaciones y condiciones vigentes en otros lugares. A partir de lo anterior, el fenómeno debe ser entendido como el mecanismo mediante el cual, las condiciones que rigen en un lugar se transmiten a otro u otros.

En el siguiente esquema es posible observar la clasificación que en la literatura se le ha dado:



Fuente: elaborado por el autor en base a las descripciones hechas en DIAZ & SANCHEZ (2004).

En palabras de SANCHEZ & DIAZ (2004), para el caso de los cultivos ilícitos, se define cada uno de los casos anteriores de la siguiente forma: "La difusión espacial ... se presenta de dos maneras: a saber difusión contagiosa y difusión jerárquica. La difusión contagiosa, como su nombre lo indica, se da cuando un municipio contagia a su municipio vecino. Esta difusión se puede clasificar de dos maneras: relocalización que significa que el cultivo de coca se desplaza hacia los vecinos, pero el municipio local deja de producir ya sea por labores de erradicación u otros factores; difusión de expansión que implica que la presencia de coca se difunde desde el municipio hacia las unidades espaciales contiguas, pero el municipio mantiene producción de coca. La difusión jerárquica, por su parte, se produce sin ningún tipo de contacto espacial, es decir, ocurre por imitación o innovación ..."

FORMIZANO (2002) de forma análoga, complementaria y con carácter más analítico, para el caso de los homicidios en Bogotá, los define tal como sigue: (Difusión jerárquica) "Si el modelo estimado resulta ser un modelo espacial en el error, entonces las variables que conforman el modelo econométrico como tal influyen en la expansión y difusión de los homicidios y estos se dispersan en los diferentes sectores censales dependiendo de las características de cada sector. La forma reducida del modelo econométrico es: ..."

$$\begin{aligned} Y_i &= \beta X_i + \varepsilon_i \\ \varepsilon_i &= \lambda W + \mu_i \\ \mu_i &\rightarrow N(0, \sigma^2 I) \end{aligned}$$

(Difusión contagiosa) "Por el contrario, si el modelo es de rezago espacial, las variables que conforman el modelo econométrico no determinan la expansión o difusión de los homicidios sobre los demás sectores censales, y los homicidios se expanden indiscriminadamente alrededor del foco de violencia; este caso se denomina expansión contagiosa. La forma reducida del modelo econométrico es: ..."

$$\begin{aligned} Y_i &= \rho W Y_i + \beta X_i + \mu_i \\ \mu_i &\rightarrow N(0, \sigma^2 I) \end{aligned}$$

"...Permite explicar el contagio y la difusión de los homicidios; bien sea la tasa de homicidios per se (difusión contagiosa) o las diferentes variables explicativas del modelo econométrico..."

Surge entonces una pregunta crucial, ¿cómo entrelazar los conceptos esbozados por Sánchez, Díaz y Formizano, para el caso que se pretende abordar, que es la pobreza en Colombia? A manera de aproximación, se propondrán los siguientes casos:

- **Difusión contagiosa:** para que se de este tipo de difusión, debe haber necesariamente contacto físico entre las fronteras de las unidades espaciales, y es el fenómeno mismo, por sus características y propiedades lo que provoca su expansión. Varias son las formas y las situaciones, como se podría pensar que esto se puede dar:
 - **Migración de aventura:** en un entorno de pobreza pueden generarse incentivos en los individuos para que emigren a lugares cercanos en busca de nuevas y mejores oportunidades. De esta forma, se estaría contagiando la pobreza a una unidad vecina, ya que estos individuos llegan en busca de recursos y asistencia (expansión) y no a emprender alguna actividad productiva.
 - **Migración oportunista:** una de las implicaciones que tiene la dependencia espacial en el diseño de políticas para la pobreza, es que dentro del costo de ejecutarla, debe tenerse en cuenta la posibilidad que los individuos de unidades vecinas emigren a donde se esta llevando a cabo, con el propósito de captar beneficios de ella. El efecto podría ser de expansión (de menos pobreza) en la medida que la dinámica que genere la puesta en marcha de la política absorba a los inmigrantes, mientras que podría ser de relocalización sino lograra hacerlo.
- **Desincentivos de la pobreza:** en el contexto de un vecindario (de municipios o departamentos) pobre, puede crearse un desincentivo para mejorar la situación ya que lo vecinos se perciben como un ejemplo a seguir o un consuelo a la situación existente, o bien no existe un ejemplo de superación y mejoría que impulse el deseo de mejorar (expansión).
- **Desincentivos oportunistas:** una situación de este tipo se podría dar, en la medida que en un vecindario pobre, una de las unidades emprende un esfuerzo por aliviar la situación de la población, y los vecinos, en lugar de realizar el mismo esfuerzo, optan por esperar que los

beneficios del esfuerzo emprendido se expanda, ya sea vía encadenamientos productivos o por una emigración a donde se este ejecutando la política. En la medida que la política se expanda o logre beneficiar a los inmigrantes, puede haber una expansión de los beneficios, en caso contrario lo que podría darse es una relocalización.

- Difusión jerárquica: en este caso, son los determinantes o causas lo que provoca la difusión. La intuición es sencilla, y es que en todas aquellas unidades con problemas comunes, como alto desempleo, alta desigualdad, baja productividad agrícola, baja escolaridad, baja productividad laboral, alta dependencia de la agricultura, alta corrupción, entre otros, no importa que sean vecinas o no, están expuestas a que surja el problema. Lo que se estaría tratando de encontrar entonces, es un patrón espacial común a lo largo del territorio nacional.

Ahora bien, no importa en cual de los dos escenarios estemos situados, conocer los posibles determinantes de la pobreza es importante, porque permite identificar las variables con las cuales puede ser atacada, así como la magnitud y dirección en que cada una lo hace. Lo que debería cambiar, o por lo menos lo que con este trabajo se propone reflexionar, es que de acuerdo al tipo de difusión que se identifique, las estrategias para combatirla deberían o podrían ser distintas, ya que en un contexto de difusión contagiosa atacando un foco del problema se beneficiarían los vecinos, mientras si es jerárquica, los programas deben dirigirse a cada una de las unidades por separado sin que necesariamente se transmita algún efecto a otras unidades.

2. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES (ESDA) A LAS VARIABLES DEL MODELO

Una vez hecho un recorrido por los conceptos que se pretenden aplicar en este trabajo, corresponde hacer un análisis exploratorio y descriptivo de las variables del modelo a utilizar, lo cual permitirá justificar el uso de las técnicas señaladas previamente, a la muestra y el periodo de tiempo bajo estudio.

El ESDA lo conforman un conjunto de herramientas de visualización, mapeo y

contraste⁸, que pretenden identificar patrones de asociación significativos en la distribución de una variable a lo largo del territorio. Para ello, se requiere disponer de información geoestadística de cada una de las unidades espaciales, tales como su perímetro, área y posición en el globo (distancia a la línea del Ecuador y al meridiano de Greenwich)⁹.

2.1. Variable endógena¹⁰

La variable a explicar en este caso es la pobreza, la cual será medida a través de la "Línea de Pobreza"¹¹ (LP). En esta medida se recoge al porcentaje de individuos cuyo ingreso es inferior al valor monetario de una canasta de bienes básicos de consumo. El criterio para escoger esta medida, y no otra, como puede ser el Índice de Necesidades Básicas Insatisfechas, el Índice de Calidad de Vida o el Índice de Desarrollo Humano, es la disponibilidad, ya que ninguna de ellas pudo obtenerse simultáneamente a nivel departamental y para todos los años que se pretenden abordar.

Durante el periodo en estudio la línea de pobreza presentó para el agregado nacional un fuerte incremento, al pasar de 55.7% a 62.3%, lo que corresponde a un crecimiento de aproximadamente 12%. Este incremento no fue sostenido durante todo el periodo, por el contrario entre 1996 y 1998 hay una reducción de aproximadamente 2 puntos porcentuales, pero a partir de ese año y hasta el 2000, se disparó al pasar de 53.7% a 62.3%.

En cada una de las regiones el comportamiento de la LP fue muy similar, y salvo Bogotá que presenta niveles de pobreza sensiblemente menores al promedio nacional, las demás al año 2000 se sitúan en el orden del 60% (Ver Grafico 1).

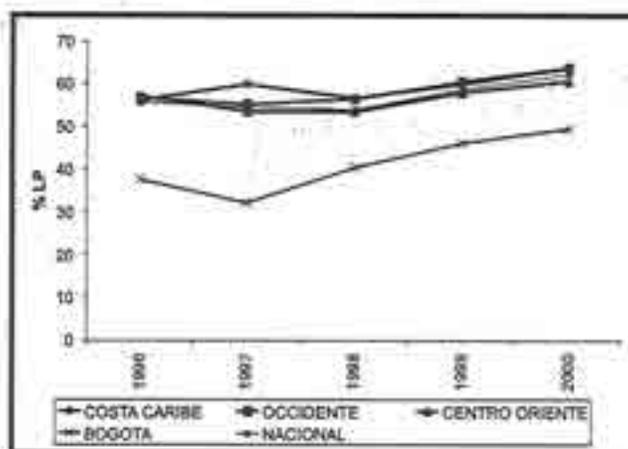
⁸ Para conocer más detalles sobre las herramientas de visualización que se emplean en este trabajo, se recomienda remitirse al Anexo 4 del presente documento.

⁹ El manejo y suministro de esta información en nuestro país está a cargo del Instituto Geográfico Agustín Codazzi (IGAC).

¹⁰ Para conocer la fuente y la forma como se construyeron algunas de las variables, se recomienda remitirse al Anexo 5 del presente documento.

¹¹ Para conocer más detalles sobre esta forma de medir la pobreza, otras medidas alternativas y complementarias, las características, ventajas y desventajas de cada una de ellas, se recomienda remitirse a: EL DESAFÍO DE LA POBREZA, ALVAREZ M, María E, MARTÍNEZ H, Horacio. Fundación Social, Bogotá, 2001.

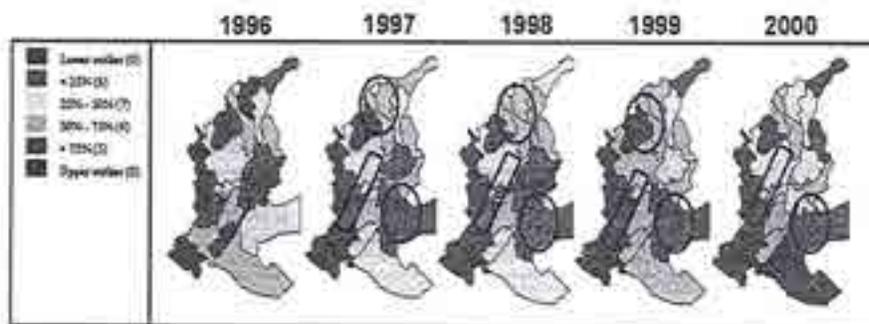
Gráfico 1
Evolución del promedio regional y nacional de la Línea de Pobreza
(1996-2000)



Fuente: SISD DNP, Cálculos del autor.

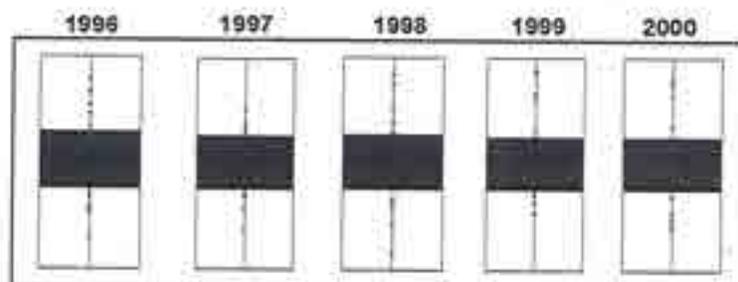
En los cuadros 1 y 2 se presentan los Box Maps y los Box Plots de la línea de pobreza, pudiendo observar que en general la distribución de los datos no parece ser aleatoria, e identificando para cada uno de los años asociaciones de departamentos, las cuales adicionalmente parecen ser persistentes en el tiempo.

Cuadro 1
Box Maps de la Línea de pobreza (1996-2000)



Fuente: Elaborado por el autor en base a estimaciones hechas con el software Geoda 0.9.5-i (Beta)

Cuadro 2
Box Plots de la Línea de Pobreza (1996-2000)



Fuente: Elaborado por el autor en base a estimaciones hechas con el software Geoda 0.9.5-i (Beta)

Dichas asociaciones aparecen señaladas en el cuadro 1, y corresponden una a departamentos del norte, otra a departamentos del centro y sur occidente y otra a departamentos del sur oriente del país. Los departamentos del norte aparecen en la mayoría de los casos en la parte superior de la distribución (franja que está por encima del área gris en los Box Plots), es decir presentan los porcentajes más altos de personas bajo LP, mientras los del centro, sur occidente, y sur oriente del país se ubican en el cuartil más bajo (franja que se encuentra por debajo del área gris en los Box Plots), es decir presentan menores porcentajes de personas bajo LP.

Con el propósito de verificar formalmente si dichas asociaciones de departamentos son significativas, se procedió a efectuar el test de autocorrelación espacial con base a los estadísticos I de Moran y C de Geary. Los resultados de dichos test aparecen en la siguiente tabla (Tabla 1):

Tabla 1
Resultados de los test de autocorrelación espacial

Años	I de Moran	P-value	C de Geary	P-value
1996	-0.084	0.382	0.931	0.321
1997	0.207**	0.031	0.680**	0.018
1998	0.178**	0.049	0.671**	0.014
1999	0.151*	0.073	0.739**	0.040
2000	0.187**	0.043	0.699**	0.022

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor en base Geoda 0.9.5-i (Beta) y STARS

Los resultados del test validan las apreciaciones anteriores, ya que salvo para el año 1996, los esquemas de dependencia espacial resultaron significativos. Para efectuar la prueba se utilizó una matriz de contigüidad binaria, de primer orden, y con criterio "reina" para definir la vecindad.

2.2. Variables exógenas

Las mismas pruebas fueron efectuadas a cada una de las variables exógenas del modelo, obteniendo los siguientes resultados:

Tabla 2
Resultados de los test de autocorrelación espacial

ITCR					Des				
Año	I Moran	P-value	C Geary	P-value	Año	I Moran	P-value	C Geary	P-value
1996	-0.035	0.475	1.111	0.228	1996	-0.115	0.297	1.003	0.492
1997	-0.246*	0.065	1.180	0.141	1997	-0.043	0.498	1.054	0.488
1998	-0.194	0.131	1.063	0.335	1998	0.083	0.172	0.860	0.174
1999	0.043	0.281	0.778*	0.068	1999	0.053	0.235	0.860	0.174
2000	0.055	0.231	0.757*	0.052	2000	0.073	0.182	0.765*	0.057
PL					Es				
Año	I Moran	P-value	C Geary	P-value	Año	I Moran	P-value	C Geary	P-value
1996	0.082	0.175	0.812	0.103	1996	0.004	0.563	0.794*	0.084
1997	0.093	0.154	0.802*	0.091	1997	-0.023	0.438	0.812	0.103
1998	0.049	0.246	0.831	0.128	1998	-0.018	0.425	0.808	0.100
1999	0.032	0.287	0.883	0.179	1999	-0.040	0.489	0.825	0.120
2000	0.039	0.283	0.857	0.168	2000	0.023	0.310	0.827*	0.097
Gru					Ser				
Año	I Moran	P-value	C Geary	P-value	Año	I Moran	P-value	C Geary	P-value
1996	0.047	0.249	0.755*	0.052	1996	-0.012	0.400	0.875	0.201
1997	0.113	0.132	0.713**	0.027	1997	-0.001	0.377	0.849	0.180
1998	-0.015	0.416	1.031	0.417	1998	0.027	0.299	0.805*	0.085
1999	-0.108	0.316	1.069	0.321	1999	0.019	0.319	0.837	0.137
2000	0.054	0.253	0.919	0.282	2000	0.059	0.222	0.819	0.112
Pag									
Año	I Moran	P-value	C Geary	P-value					
1996	0.017	0.328	0.863	0.177					
1997	0.078	0.182	0.779*	0.089					
1998	0.060	0.179	0.757*	0.051					
1999	0.040	0.289	0.811	0.102					
2000	0.007	0.353	0.908	0.269					

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor en base Geoda 0.9.5-i (Beta) y STARS

En general ninguna de las variables explicativas presenta patrones de dependencia espacial, ya que los resultados entre las dos medidas empleadas no son consistentes, y salvo algunos años y en algunas variables, en la generalidad de los casos no resultaron ser significativos. Al igual que con la variable endógena, se utilizó una matriz de contigüidad binaria, de primer orden, y con criterio "reina" para definir la vecindad.

Queda demostrado entonces que existe una estructura de autocorrelación espacial en la LP, por lo cual cualquier intento de modelación que se trate de hacer con esta variable, debe evaluar la necesidad de incorporar explícitamente dicho fenómeno.

3. FENÓMENOS DE DIFUSIÓN Y DETERMINANTES DE LA POBREZA

3.1. Metodología

Tal como se ilustró previamente, una estimación mínimo cuadrática en presencia de dependencia espacial provoca problemas de ineficiencia, y por ende resta potencia a los test de significancia estadística y sobrevalora el valor del R^2 , por lo cual, una vez demostrada la existencia de dependencia espacial, se procederá a evaluar la forma más conveniente de incorporar dicho fenómeno y a especificar la forma funcional más adecuada del modelo a estimar.

En primera instancia se procederá a estimar modelos de corte transversal para cada uno de los años, esto con el fin de construir a partir de los residuos de dichas regresiones los multiplicadores de Lagrange, con los cuales se efectúan los contrastes que permitirán determinar si se requiere especificar un componente espacial dentro del modelo, y la forma como este debe incorporarse.¹²

Previamente y con el fin de prevenir posibles problemas de multicolinealidad, se han construido matrices de correlación contemporánea para las variables exógenas del modelo (Cuadro 3). En estas matrices se puede observar una alta correlación entre la productividad laboral y la escolaridad, razón por la cual esta última se suprime, dado que es bastante intuitiva la idea que se encuentre incorporada o pueda explicar la productividad laboral y no viceversa.

¹² Los test con los multiplicadores de Lagrange están estructurados en base a los principios de máxima verosimilitud, y tienen la ventaja frente a otros como los test ad hoc (L-Moran y K-R), de estar definidos en términos tanto de una hipótesis nula como de una hipótesis alternativa. Para más detalles sobre contrastes de dependencia espacial se recomienda remitirse a: Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial; moreno s rosina, VAYÁ V Esther; edicions universitat de barcelona; pags 75-91.

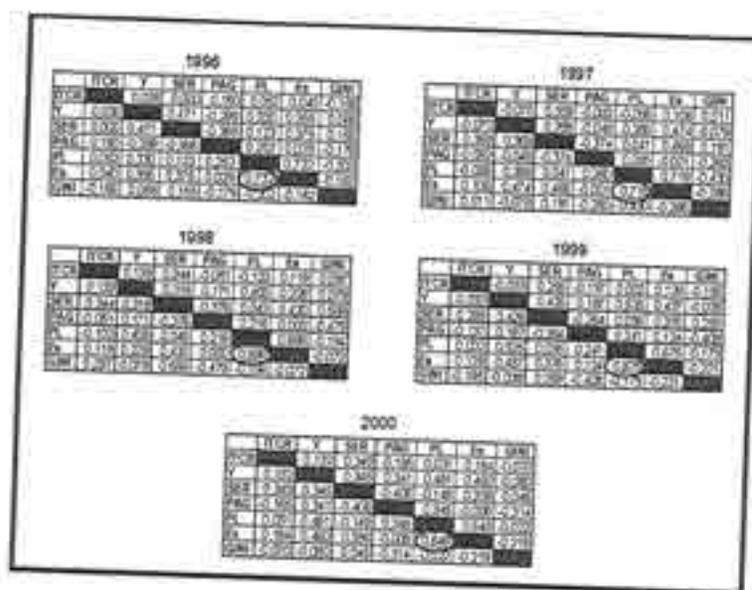
Tomando en consideración lo anterior, el modelo queda especificado de la siguiente forma:

$$Pb_{i,t} = \beta_0 + \beta_1 ITCR_{i,t} + \beta_2 D_{i,t} + \beta_3 Pl_{i,t} + \beta_4 G_{i,t} + \beta_5 Ser_{i,t} + \beta_6 Pag_{i,t} + e_{i,t}$$

$$e_{i,t} \rightarrow N(0,1)$$

conservando cada uno de los términos el significado dado anteriormente.

Cuadro 3
Matrices de correlación contemporánea de las variables exógenas del modelo (1996-2000)



Fuente: Cálculos del autor

Los resultados de las estimaciones hechas al modelo con datos de corte transversal se pueden apreciar en la tabla 3, mientras en la tabla 4 se encuentran los resultados de la evaluación hecha a los residuales de dichas regresiones.

Tabla 3
Estimación del modelo de determinantes de la pobreza (1997-2000)^a

Año	Constant	ITCR	D	PL	G	Ser	Pag
1997	90.150***	-0.368	-0.015	-13.057***	37.338	0.487	-1.980
1998	93.529***	-0.130	-0.457	-9.574***	-8.781	4.590*	-5.113**
1999	66.640*	-0.126	-0.410	-8.952**	48.210	0.362	-3.992
2000	102.02***	-0.190*	-0.280	-5.360**	0.257	-1.313	-8.865**

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor

Tabla 4
Evaluación a los residuales del modelo de determinantes de la pobreza (1997-2000)

Año	Heteroscedasticidad (Breusch-Pagan Test)			Normalidad (Jarque Bera Test)		
	Resultado	Estadístico	P-value	Resultado	Estadístico	P-value
1997	No	4.1812	0.652	No	2.8446	0.241
1998	No	2.1287	0.907	No	1.1145	0.572
1999	No	7.2493	0.265	No	0.7280	0.694
2000	No	3.1449	0.790	No	0.9475	0.6228

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor

A partir de estos residuales homoscedásticos y distribuidos según una normal, se procederá a efectuar los test de dependencia espacial con los multiplicadores de Lagrange. Son tres tipos distintos de especificación que se contrastarán:

Modelo I: Rezago espacial en la variable endógena

En este caso se rezaga espacialmente la variable endógena, y de resultar significativo se estaría encontrando evidencia de difusión por contagio.

$$Pb_{it} = \rho W.Pb_{it} + \beta_0 + \beta_1 ITCR_{it} + \beta_2 D_{it} + \beta_3 PL_{it} + \beta_4 G_{it} + \beta_5 Ser_{it} + \beta_6 Pag_{it} + e_{it}$$

$$e_{it} \rightarrow N(0, \sigma^2 I)$$

^a De este ejercicio y de los siguientes fue excluido el año 1996, dado que en este el fenómeno de autocorrelación espacial no resultó significativo.

De manera rigurosa el test se plantea de la siguiente forma:

H0: $\rho_1 = \rho = 0$ (No dependencia espacial en la variable endógena)

Ha: $\rho_1 \neq 0$ (Dependencia espacial en la variable endógena)

Modelo 2: Rezago espacial en el error

El rezago espacial se incorpora dentro del error, y de resultar la especificación más adecuada, indicaría que la pobreza presenta patrones de difusión jerárquica.

$$Pb_{it} = \beta_0 + \beta_1 ITCR_{it} + \beta_2 D_{it} + \beta_3 Pl_{it} + \beta_4 G_{it} + \beta_5 Ser_{it} + \beta_6 Pag_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \rho_1 W_i \varepsilon_{it} + e_{it}$$

$$e_{it} \rightarrow N(0, \sigma^2 I)$$

En este caso se estaría probando que:

H0: $\rho_1 = 0$ (No dependencia espacial en el error)

Ha: $\rho_1 \neq 0$ (Dependencia espacial en el error)

Modelo 3: SARMA

Simultáneamente se especifica un rezago espacial en la variable endógena y en el término de perturbación, apuntando esto a que de forma simultánea también se presenta difusión por contagio y por jerarquías.

$$y_{it} = \rho_1 W_i y_{it} + \beta_0 + \beta_1 X1_{it} + \beta_2 X2_{it} + \beta_3 X3_{it} + \beta_4 X4_{it} + \beta_5 X5_{it} + \beta_6 X6_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \rho_2 W_i \varepsilon_{it} + e_{it}$$

$$e_{it} \rightarrow N(0, \sigma^2 I)$$

En términos de la hipótesis nula y alternativa se tiene que:

H0: $\rho_1 = \rho_2 = \rho = 0$ (No dependencia espacial simultánea en la variable endógena y el error)

Ha: $\rho_1 \neq \rho_2 \neq \rho \neq 0$ (Dependencia espacial simultánea en la variable endógena y el error)

El estadístico para cada uno de estos casos es:

Modelo 1

$$LM-LE = \begin{bmatrix} e'Wy & e'We \\ S^2 & S^2 \\ RJ_{\rho-1} & -T_1 \end{bmatrix}^2$$

donde $LM-LE \rightarrow \chi_{(1)}^2$

Modelo 2

$$LM-EL = \begin{bmatrix} e'We & -T_1(RJ_{\rho-1})' e'W y \\ S^2 & S^2 \\ [T_1 - T_1'(RJ_{\rho-1})]' & \end{bmatrix}$$

Donde $LM-EL \rightarrow \chi_{(1)}^2$

Modelo 3

$$SARMA = \begin{pmatrix} e'Wy & e'We \\ S^2 & S^2 \\ RJ_{\rho-1} & -T_1 \end{pmatrix}^2 + \begin{pmatrix} e'We \\ S^2 \\ T_1 \end{pmatrix}^2$$

donde $SARMA \rightarrow \chi_{(2)}^2$

Adicionalmente:

E : vector de residuos de la estimación por mínimos cuadrados

W : matriz de contigüidad

S^2 : varianza residual

$$T_1 = tr(W'W + W^2)$$

$$M = I - X(X'X)^{-1}X'$$

$$RJ_{1-\alpha} = \left[T_1 + \frac{(W'X)'M(W'X)}{S^2} \right]$$

X: matriz de diseño del modelo

β Vector de parámetros de la estimación por mínimos cuadrados

Los resultados de estas pruebas se pueden observar en la siguiente tabla:

Tabla 5
Resultados de los test de dependencia espacial con los multiplicadores de Lagrange
(1997-2000)

Año	Modelo 1		Modelo 2		Modelo 3	
	LM-CE	P-value	LM-EL	P-value	SARMA	P-value
1997	1.4388	0.2303	0.2295	0.6326	3.8454	0.1463
1998	1.0130	0.3141	1.4670	0.2211	6.4704**	0.0360
1999	0.6305	0.4271	0.1693	0.6833	2.9362	0.2303
2000	0.6249	0.4282	0.1776	0.6733	2.9572	0.2648

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%

Fuente: cálculos del autor

El criterio para escoger el modelo más adecuado, es aquel que presente el multiplicador de lagrange más alto, entre los que resultaron significativos.¹⁴ Solo resultado significativo el estadístico SARMA para 1998, por lo cual, la estimación obtenida para los años 1997, 1999 y 2000 y que aparece reportada en la tabla 3, resulta ser la más adecuada. Se procederá a estimar nuevamente el modelo para 1998, pero esta vez con una estructura regresiva espacial simultánea para la variable endógena y el término de perturbación, como lo sugiere el resultado del test de dependencia espacial para ese año (Tablas 6 y 7).

¹⁴ Econometría espacial: características de la violencia homicida en Bogotá, formisano, Michel. Bogotá, Colombia, documentos cede 02-10, septiembre de 2002, pag. 28.

Tabla 6
Estimación del modelo de determinantes de pobreza con estructura SARMA (1998)

Año	WPL	Constad	ITCR	D	PL	G	Ser	Pag	We
1998	0.454***	75.285***	-0.346***	-0.175	-12.359**	18.511	5.373***	-3.807***	-0.449***

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor

Tabla 7
Evaluación a los residuales del modelo de determinantes de la pobreza con estructura SARMA (1998)

Año	Heteroscedaticidad (Breusch-Pagan Test)			Normalidad (Jarque Bera Test)		
	Resultado	Estadístico	P-value	Resultado	Estadístico	P-value
1998	No	2.276	0.942	No	2.234	0.327

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor

Adicionalmente, con el propósito de poder captar de forma simultánea en la estimación la posible heterogeneidad que puede haber entre los departamentos a través del periodo de tiempo considerado, se estimo el modelo de determinantes de pobreza mediante un panel de datos.

Para efectuar esta estimación no se especificó ningún tipo de estructura espacial, dado que las pruebas de los multiplicadores de lagrange no permitieron identificar un patrón determinado, y para la mayoría de los años se acepto la hipótesis de no autocorrelación espacial en la o las variables correspondientes.

Se efectuó la prueba de especificación de Hausman para determinar cual debía ser la estimación (con efectos fijos o aleatorios) mas adecuada, arrojando como resultado que debe ser con efectos fijos (Tabla 8), lo cual implica que los elementos no observables se encuentran correlacionados con las variables explicativas.

Tabla 8
Estimación del modelo de determinantes de pobreza mediante un panel de datos con
efectos fijos (1997-2000)

Variable	Coficiente	P-value
Constant**	66.557	0.012
ITCR	0.1094	0.169
D*	0.4023	0.089
Ser	-21.0895	0.517
PL**	-21.8099	0.016
Pag	1.3612	0.774
G**	30.4719	0.032
Hausman test**	15.39	0.0174
R^2 Within	0.4456	
R^2 Overall	0.6015	

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor

3.2. Interpretación e implicaciones del ejercicio econométrico

La primera implicación que se puede extraer del ejercicio econométrico, es que no se logró determinar el patrón de difusión espacial que presenta la pobreza. Podría pensarse que al no resultar significativos los multiplicadores de Lagrange, la conclusión inmediata sería que no hay tal patrón de difusión. Podría también cuestionarse la aparente contradicción que surge al demostrarse el fenómeno de dependencia espacial, pero no haberse podido determinar la forma como se expande o difunde la pobreza a lo largo del territorio.

Respecto a lo anterior, varias explicaciones pueden ser exploradas:

- **Variables explicativas:** podría argumentarse que no se logró identificar dentro de las variables explicativas, aquella que lograra transmitir el fenómeno de dependencia a la variable endógena. Sin embargo, de ser así el error hubiese estado correlacionado espacialmente, lo cual, tal como puede apreciarse en la tabla 9 no ocurre, además se estaría de esta forma ignorando la posibilidad de difusión por contagio, es decir, que sea la variable endógena por sus características intrínsecas, la que provoca la expansión.

Tabla 9
Resultados de los test de autocorrelación espacial a los residuales de la estimación
hecha con datos de corte transversal (1997-2000)

Año	I-Moran	P-value
1997	-0.0709	0.4180
1998	-0.1777	0.1830
1999	-0.0605	0.4620
2000	-0.0626	0.4410

*** Significativo al 1%; **Significativo al 5%; *Significativo al 10%
Fuente: cálculos del autor

- **Débil dependencia espacial:** pese a ser significativo el fenómeno de autocorrelación espacial en la LP, dicha correlación tal como puede observarse en la tabla 1 es baja (inferior a 0,2 en la mayoría de los casos), lo cual estaría impidiendo obtener un resultado concluyente a la hora de identificar el modelo espacial mas adecuado.
- **Número de unidades espaciales:** los test de autocorrelación espacial (I-Moran, LM-EL, LM-LE, SARMA) están contruidos bajo propiedades asintóticas, por lo cual su aplicación en muestras finitas debe hacerse con cautela. Al respecto, diferentes ejercicios de simulación han sido adelantados para analizar la potencia y tamaño de estos estadísticos ante diferentes especificaciones de matrices de contigüidad, distribución del error y tamaño muestral¹⁵. Los resultados apuntan a que pese a presentar los estadísticos buenas propiedades, son sensibles principalmente a una mala especificación de la matriz de pesos y a la no normalidad del término de perturbación, sin dejar de serlo también ante muestras muy pequeñas.

Por lo anterior, y pretendiendo guardar estricto rigor y moderación, se debería afirmar, no que no haya un patrón de difusión espacial en la pobreza, es decir ignorar los resultados del test I de Moran y C de Geary, sino que la evidencia empírica no permite ser concluyentes en la identificación del mecanismo de

¹⁵ Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial; Morozzo S Rosma, Vaya V Esther; ediciones universitat de Barcelona; pags 91-97.

expansión o contracción en el espacio de la variable o fenómeno bajo estudio, debido posiblemente a deficiencias estadísticas dadas principalmente por la información muestral. Este se constituye en un limitante importante para análisis o extensiones posteriores, dado que la manera de subsanar esta dificultad sería trabajar con un mayor nivel de desagregación de los datos (municipios), lo cual, dada la naturaleza de la mayor parte de las variables utilizadas es bastante complicado de lograr.

En cuanto a los resultados obtenidos de la estimación de los modelos de sección cruzada y de panel de datos, cabe resaltar:

- Las relaciones de causalidad entre las variables explicativas que resultaron significativas y la pobreza, son las esperadas. El desempleo y el coeficiente de gini afectan positivamente, mientras la productividad laboral lo hace en sentido contrario.
- La productividad laboral es de las variables significativas, la que aparece como más robusta.
- El coeficiente de gini es la variable que por la magnitud del coeficiente presenta mayor impacto sobre la pobreza, seguida de la productividad laboral y el desempleo.
- Otras variables como el índice de tasa de cambio real, la productividad agrícola y la participación de los servicios en el PIB no resultaron significativas en el ejercicio de panel, sin embargo para algunos años de las estimaciones con datos de corte transversal si lo fueron, reportando en esos casos los signos esperados.

Quizás el resultado más importante y que engloba los antes descritos, es que la pobreza guarda una relación de causalidad con variables que tienen relación con el ciclo económico, por lo que se podría afirmar entonces que esta (la pobreza) también guarda relación con este ciclo. Adicionalmente, el resultado sobre la productividad agrícola (variable influida por factores como la calidad del suelo, altitud, grado de precipitación, humedad, entre otros) implicaría que la pobreza no se ve afectada por elementos determinísticos como la geografía.

4. CONCLUSIONES Y RECOMENDACIONES

Con el presente trabajo se perseguía determinar la existencia de fenómenos de difusión espacial, así como identificar variables que permitieran explicar el nivel de pobreza, medida por la línea de pobreza departamental, para el periodo comprendido entre 1996 y 2000. Como conclusiones inmediatas cabe destacar que no se logró identificar un patrón de difusión (entre contagiosa y/o jerárquica) pese a que las medidas de autocorrelación espacial para la LP resultaron significativas.

En cuanto a los determinantes, aparecen como significativas dentro de todas las variables contempladas, la tasa de desempleo, el coeficiente de gini y la productividad laboral, presentando las dos primeras una relación directa, mientras la última una relación inversa con la LP.

Queda demostrado y medido de esta manera, que la pobreza se encuentra relacionada con el ciclo económico. Adicionalmente, al no resultar significativa la productividad agrícola, se encuentra evidencia de que factores determinísticos, como la geografía, no influyen en ella.

Uno de los propósitos que se tenía con el ejercicio, era propiciar una reflexión sobre la orientación que debían tener los programas para combatir la pobreza a partir de determinar la presencia de difusión por contagio o jerarquías, en términos de las diferentes implicaciones que en la operatividad y costo tendría cada una. Sin embargo, la evidencia empírica no apoyo este objetivo y no hay por lo tanto argumentos para poder abordar la discusión.

Las principales implicaciones que se derivan de los determinantes encontrados, son que cualquier política adoptada para reactivar o impulsar el aparato productivo, ya sea a nivel nacional o departamental debe tener impacto sobre la pobreza. Por otro lado, tiene importantes implicaciones de política, el hecho de resultar la productividad laboral la variable con el efecto más robusto, ya que como también fue demostrado en este trabajo, esta tiene una alta correlación con la escolaridad, por lo tanto, una manera de combatirla estructuralmente es haciendo un gran esfuerzo por llevar cada vez a mayor número de personas, especialmente a la más pobres, al sistema educativo.

Las políticas redistributivas también resultan relevantes para combatir la

pobreza, por lo tanto debe existir el mayor compromiso tanto del gobierno nacional como de los departamentales por hacer efectivas dichas políticas.

En síntesis, a la luz de la evidencia empírica la mejor forma de combatir la pobreza a lo largo del territorio nacional es estimulando la actividad productiva, aumentando, bajo principios de calidad y pertinencia, la cobertura educativa, y haciendo un esfuerzo por mejorar la distribución del ingreso.

5. BIBLIOGRAFIA

- ALMEIDA, Eduardo; HADDAD, Eduardo; HEWING, Jeffrey. The spatial pattern of crime in Minas Gerais: an exploratory análisis. Núcleo de economía regional y urbana de la Universidad de Sao Paulo, TD Nereus 22, 2003.
- BONET, Jaime. Descentralización fiscal y disparidades en el ingreso regional: la experiencia Colombiana. Documentos de trabajo sobre economía regional, Banco de la Republica, 2004.
- DLAZ, Ana Maria; SANCHEZ, Fabio. Geografía de los cultivos ilícitos y conflicto armado en Colombia, Bogota, Colombia, Documento CEDE 2004-18, Marzo de 2004.
- FLOREZ, Carmen Elisa; DUREAU, Francoise; GUZMÁN, Juan Carlos; CORRALES, Martha Lucia; PARDO, Campo Elias. Movilidad espacial y pobreza en el contexto de un mercado de trabajo regional, Bogota, Colombia, Documento CEDE 99-02, Febrero de 1999.
- FORMISANO, Michel. Econometría espacial: características de la violencia homicida en Bogota, Bogota, Colombia, Documento CEDE 02-10, Septiembre de 2002.
- GALVIS, Luis Armando. ¿Qué determina la productividad agrícola departamental en Colombia? Documentos de trabajo sobre economía regional, Banco de la Republica, 2001.
- GALVIS, Luis Armando. Determinantes de la migración interdepartamental en Colombia, 1988-1993. Documentos de trabajo sobre economía regional, Banco de la Republica, Junio de 2002.
- GARZARODRÍGUEZ, Jorge. The determinants of poverty in México.
- GUJARATI, Damodar. Econometría. Tercera edición, McGraw Hill, 1997.
- HADDAD, Monica; BUDIC, Zorica. Reducing Intra-urban Inequalities: Spatial analysis of Public Intervention. Working paper, 2004.
- HERNANDEZ, Roberto; FERNANDEZ, Carlos; BAPTISTA, Pilar. Metodología de la investigación. McGraw Hill, 1997.
- JALAN, Jyotsna; RAVALLION, Martin. ¿Spatial poverty traps?. The World Bank, Development research group, Diciembre de 1997.
- LUSTIG, Nora; ARIAS, Omar; RIGOLINI, Jamele. Reducción de la pobreza y crecimiento económico: la doble causalidad. Documento presentado en el seminario "La teoría del desarrollo en los albores del siglo XXI", Santiago de Chile, 2001.
- MAYORGA, Mauricio; MUÑOZ, Evelyn. La técnica de datos de panel: una guía para su uso e interpretación. Documento de trabajo del Banco Central de Costa Rica, Costa Rica.

Septiembre, 2000.

MORENO SERRANO, Rosina; VAYA VALCARCE, Esther. Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial, Edicions Universitat de Barcelona, 2000.

NÚÑEZ, Jairo; RAMÍREZ, Juan Carlos. Determinantes de la pobreza en Colombia. Años recientes, Bogota, Colombia, Documento CEDE 02-19, Noviembre de 2002.

NÚÑEZ MENDEZ, Jairo; SÁNCHEZ TORRES, Fabio José. Estimaciones trimestrales de la línea de pobreza y sus relaciones con el desempeño macroeconómico Colombiano (1977-1997). Bogota, Colombia, Archivo de Macroeconomía, Departamento Nacional de Planeación, Documento 110, mayo de 1999.

PÉREZ VALBUENA, Gerson Javier. Dimensión espacial de la pobreza en Colombia (1985,1993). Documentos de economía regional, Centro de estudios económicos regionales del Banco de la Republica, 2005.

REY, Serge. Spatial Dependence in the Evolution of Regional Income distributions. Working paper, REAL, 2001.

ROCHA, Ricardo; MARTINEZ, Hermes. Pobreza, crimen y crecimiento regional en Colombia, Bogota, Colombia, Archivos de Macroeconomía, Departamento Nacional de Planeación, Documento 230, agosto de 2003.

The determinants of poverty in Malawi, 1998. The national economic council, Lilongwe, Malawi; The national statistical office, Zomba, Malawi; The international food policy research institute, Washington, DC, USA, junio de 2001.

SAENZ CASTRO, Jorge Enrique; HERRERA SAAVEDRA, Juan Pablo; GUZMAN SILVA, Oscar Enrique. Contracción leves y profundas: efectos asimétricos sobre la pobreza, el caso colombiano 1984-2000. Bogota, Colombia, Archivos de Macroeconomía, Departamento Nacional de Planeación, documento 231, agosto de 2003.

WOOLDRIDGE, Jeremy M. Introducción a la econometría. Thomson Learning, México, 2001.

6. ANEXOS

6.1. Anexo 1: Matriz de contigüidad o pesos espaciales¹⁸

La matriz de contigüidad o pesos espaciales, es el instrumento desarrollado para captar las relaciones entre unidades espaciales, y consiste en definir una matriz W , tal que:

$$W = \begin{bmatrix} 0 & W_{12} & \dots & W_{1n} \\ W_{21} & 0 & \dots & W_{2n} \\ \dots & \dots & 0 & \dots \\ W_{n1} & W_{n2} & \dots & 0 \end{bmatrix}$$

donde W es una matriz cuadrada y no estocástica, y los elementos W_{ij} reflejan la intensidad de la relaciones entre las unidades i y j . El único consenso con respecto a las características que debe tener cada uno los elementos de W que están por fuera de la diagonal principal, es que sean no negativos y finitos (Serrano y Valla, 2000).

La forma mas común de construirla es aplicando el concepto de contigüidad física de primer orden, en base a lo cual, tomara valor de 1, si las unidades i y j son físicamente adyacentes o 0 en caso contrario (por definición $W_{ii} = 0$). La adyacencia física también puede ser definida de varias maneras, entre las que se encuentran las siguientes:

Criterio de vecindad	Número de vecinos	Definición
Lineal	2	Serán vecinos de i las unidades que se encuentren a su lado izquierdo y a su lado derecho
Torre	4	Serán vecinos de i las unidades que compartan con ella alguno de sus lados
Arb	4	Serán vecinos de i las unidades que compartan con ella alguno de sus vértices
Reina	8	Serán vecinos de i las unidades que compartan con ella alguno de sus lados o vértices

Las principales limitaciones de aplicar el concepto de contigüidad física tienen que ver, por un lado con uno de los principios básicos de la econometría

¹⁸ Ibid. Pags 23-26.

espacial, que es la asimetría, lo cual se estaría violando dado que la matriz claramente sería simétrica. Por otro lado, no tiene en cuenta la posibilidad que elementos diferentes a la adyacencia física ejerzan influencia en la interrelación entre unidades.

Otro criterio que se ha sugerido para la construcción de W , es el de proximidad física propuesto por Cliff y Ord (1973, 1981), según el cual:

$$w_{ij} = d_{ij}^{-\alpha} \beta_{ij}$$

Donde d_{ij} es la distancia que separa a las unidades i y j y β_{ij} la longitud relativa de la frontera común entre i y j con relación al perímetro de i , mientras α y β son parámetros a estimar. De forma similar Dacey (1968) propuso la siguiente definición:

$$w_{ij} = \gamma_{ij} \beta_{ij} \alpha_i$$

Allí β_{ij} mantiene la definición dada anteriormente, γ_{ij} es un factor de contigüidad binario y α_i es la proporción del área que ocupa i dentro de todo el sistema. Otra definición fue dada por Anselin (1980), según la cual:

$$w_{ij} = \frac{1}{d_{ij}^2}$$

O lo que sería lo mismo, que la matriz W sea la inversa de las distancias al cuadrado, con lo cual la intensidad de la interdependencia entre dos unidades disminuye a medida que la distancia aumenta.

Las anteriores definiciones, pese a que conducen a la construcción de matrices asimétricas, siguen muy apegadas a relacionar las unidades espaciales por aspectos puramente físicos o geográficos. Tratando de subsanar este hecho, Bodson y Peeters (1975) propusieron la utilización de un criterio de accesibilidad general que combina en una función logística la influencia de varios canales de comunicación en la relación entre unidades, como podrían ser las carreteras, los ferrocarriles, puertos u otros medios de transporte.

$$w_{ij} = \sum_{n=1}^N K_n \left\{ \frac{a}{1 + b e^{(c_j d_{ij})}} \right\}$$

En la anterior expresión, K_{ij} representa la importancia relativa del medio de comunicación n , d_{ij} la distancia entre dos unidades i y j y N el número de medios de comunicación, mientras que a , b y C_j son parámetros a estimar. En la misma dirección, Case (1993) propuso tener en cuenta características socioeconómicas ($w_{ij} \frac{1}{|x_i - x_j|}$), y Lopez-Bazo (1999) el grado de intercambio comercial.

Sin importar cual de los criterios antes expuestos sea el que se siga para definir la matriz de pesos espaciales, deberá siempre garantizarse su exogeneidad a la hora de efectuar la estimación y contrastación econométrica. Adicionalmente, la práctica más usual, es estandarizar por filas una vez se tenga conformada a W , de modo que cada fila sume 1 y con ello se pueda tener una aproximación al peso que cada unidad tiene dentro del sistema.

6.2. Anexo 2: Medidas de autocorrelación espacial¹⁷

Las medidas de autocorrelación espacial más conocidas, son el I-Moran, C-Geary y G(d) de Getis y Ord, aunque son las dos primeras las que han tenido mayor aplicación computacional, y se encuentran programadas en los paquetes estadísticos más populares en el campo de la econometría espacial. Dichos estadísticos fueron diseñados con el propósito de poder efectuar contrastes de autocorrelación espacial, en los cuales se estaría probando en la nula una distribución aleatoria de la variable bajo estudio, y en la alterna una asociación significativa de valores similares o disímiles. Las expresiones para dichos estadísticos son:

$$I = \frac{N}{S_0} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad ; \quad i \neq j$$

$$C = \frac{N-1}{2S_0} \frac{\sum_{i=1}^N w_{ij} (x_i - x_j)^2}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

Y

$$G(d) = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij}(d) x_i x_j}{\sum_{i=1}^N \sum_{j=1}^N x_i x_j} \quad i \neq j$$

¹⁷ Ibid. Pag. 33-44

Donde, $I \wedge C \wedge G(d) \rightarrow N(0,1)$

En dichas expresiones, x_i refleja el valor de la variable cuantitativa x en la unidad i , \bar{x} es su media muestral, w_{ij} los pesos de la matriz W , N es el tamaño muestral y.

$$S_0 = \sum_i \sum_j w_{ij}$$

En el caso del I-Moran, un resultado positivo (negativos) implica la asociación de valores similares (disímiles), con la C-Geary, es todo lo contrario, ya que un valor positivo (negativo) implica un esquema de dependencia espacial negativa (positiva), y con la G(d)-Getis y Ord, la relación también es directa, solo que en este caso un resultado positivo (negativo) implica una asociación significativa de unidades con valores altos (bajos) de la variable en estudio.

Por otra parte, en cuanto al I-Moran y C-Geary, puede usarse en su cálculo cualquiera de los criterios de contigüidad descritos en el anexo 1 para la construcción de la matriz W , lo cual no aplica para la G(d), dado que en este caso se requiere estrictamente de una matriz simétrica.

Las medidas anteriores, tienen la característica que arrojan un resultado global, es decir donde se considera conjuntamente a todas las unidades espaciales, por lo que usualmente también son llamados contrastes o medidas de autocorrelación espacial global. Existen también un conjunto de medidas que consideran a cada unidad por separado, y tienen por propósito verificar la posible existencia de clusters o asociaciones significativas de unidades en áreas específicas o delimitadas dentro del sistema que se está trabajando. De este modo, es posible verificar si las relaciones espaciales son más fuertes en determinadas zonas o se reparten homogéneamente a lo largo del territorio, o si en un contexto donde globalmente se rechazó la hipótesis de dependencia espacial, sea posible que en alguna porción del espacio que se está analizando dicha hipótesis sea válida.

Dentro de este tipo de medidas se encuentran las siguientes:

$$I_i = \frac{z_i}{\sum_{j \in J} z_j} \sum_{j \in J} w_{ij} z_j$$

$$New-G_i = \frac{\sum_{j \in J} w_{ij} x_j - W_i \bar{x}(i)}{\left\{ \frac{[(N-1)S_{ii} - W_i^2]}{(N-2)} \right\}^{1/2}}$$

Y

$$New-G_i^* = \frac{\sum_{j \in J} w_{ij} x_j - W_i^* \bar{x}}{\left\{ \frac{[(N)S_{ii}^* - W_i^{*2}]}{(N-1)} \right\}^{1/2}}$$

Donde z_i es el valor de la variable normalizada correspondiente a la unidad i , J , el conjunto de regiones vecinas a i , w_{ij} los valores de la matriz W , x_i el valor de la variable correspondiente a la unidad i , y:

$$W_i^* = W_i + w_{ii} \cdot z(i)^2 = \frac{1}{N-1} \sum_{j \in J} \left(\frac{I_j}{x_j - \bar{x}(i)} \right)^2 \cdot \bar{x}(i) = \frac{1}{N-1} \sum_{j \in J} x_j$$

$$S_{ii} = \sum_{j \in J} w_{ij}^2 \text{ y } S_{ii}^* = \sum_{j \in J} w_{ij}^{*2}$$

Adicionalmente, las tres medidas se distribuyen normales con media cero y varianza uno ($N(0,1)$), y en cuanto a la matriz W , puede esta definirse bajo cualquiera de los criterios que han sido expuestos previamente en otros apartados del presente documento.

En el caso particular del I_i , este tiene la propiedad de ser un estadístico LISA (por sus siglas en inglés: "Local indicador of spatial autocorrelation"), lo cual implica que la suma de las medidas locales debe aproximarse al resultado global, tal como se expresa en la siguiente condición:

$$I = \sum_i I_i \gamma = \sum_i I_i \left[S_0 \left(\frac{\sum_i (x_i - \bar{x})^2}{N} \right) \right]^{-1}$$

6.3. Anexo 3: Prueba de endogeneidad de Hausman

Se parte del siguiente sistema de ecuaciones²⁸

$$Pb_{i,t} = \beta_0 + \beta_1 C_{i,t} + u_{i,t} \quad (1)$$

$$C_{i,t} = \alpha_0 + \alpha_1 Pb_{i,t} + \alpha_2 Es_{i,t} + w_{i,t} \quad (2)$$

Donde:

i: *i*-ésimo departamento de Colombia (según la división política definida por la constitución de 1991)

T: corresponde a los años 1996, 1997, 1998, 1999 y 2000

$Pb_{i,t}$: línea de pobreza del departamento *i* en *t*.

$C_{i,t}$: variación del PIB entre *t* y *t*-1 del departamento *i* en *t*.

$Es_{i,t}$: población del departamento *i* en *t*.

$u_{i,t}, w_{i,t}$: términos aleatorios, que se supondrán ruido blanco.

La prueba consiste en verificar si el crecimiento económico ($C_{i,t}$) se encuentra correlacionado con el término de error de la ecuación (1). De estarlo, se probaría que hay endogeneidad o simultaneidad entre $Pb_{i,t}$ y $C_{i,t}$.

El procedimiento para llevar a cabo la prueba es el siguiente:

1. Encontrar a partir del sistema de ecuaciones planteado, una expresión para la pobreza ($Pb_{i,t}$) que no contenga al crecimiento ($C_{i,t}$), y otra para $C_{i,t}$ que no contenga a $Pb_{i,t}$. Esto se logra reemplazando recursivamente las expresiones

²⁸ A partir de las condiciones de orden y rango se logra determinar que la ecuación (1) se encuentra exactamente identificada, mientras la (2) se encuentra subidentificada.

(1) y (2).

2. Estimar la expresión que se obtuvo para el crecimiento.
3. Tomar los residuos de la regresión efectuada en el paso anterior, para incluirlos como variable explicativa ($Z_{i,t}$) en la ecuación (1).
4. Efectuar una prueba de significancia estadística sobre el parámetro de $Z_{i,t}$. De ser significativo se aceptaría la hipótesis de endogeneidad entre pobreza y crecimiento económico.

Las expresiones que se obtuvieron para $Pb_{i,t}$ y $C_{i,t}$ fueron las siguientes:

$$Pb_{i,t} = \Pi_0 + \Pi_1 ES_{i,t} + v_{i,t} \quad (3)$$

$$C_{i,t} = \Pi_2 + \Pi_3 ES_{i,t} + z_{i,t} \quad (4)$$

Donde:

$$\Pi_0 = \frac{\beta_0 + \beta_1 \alpha_0}{1 - \beta_1 \alpha_1}, \quad \Pi_1 = \frac{\beta_1 \alpha_2}{1 - \beta_1 \alpha_1}, \quad \Pi_2 = \frac{\alpha_0 + \beta_0}{1 - \beta_1}, \quad \Pi_3 = \frac{\alpha_2}{1 - \beta_1}$$

$$v_{i,t} = \frac{\beta_1 w_{i,t} + u_{i,t}}{1 - \beta_1 \alpha_1} \quad \text{y} \quad z_{i,t} = \frac{u_{i,t} + w_{i,t}}{1 - \beta_1}$$

Una vez se ha estimado (4), se toman sus residuos y se incorporan en (1), de tal forma que esta última queda de la siguiente forma:

$$Pb_{i,t} = \beta_0 + \beta_1 C_{i,t} + \beta_2 z_{i,t} + \beta_3 v_{i,t}$$

Formalmente el test de endogeneidad o simultaneidad de Hausman plantea que:

H0: Ausencia de endogeneidad ($\beta = 0$)

Ha: Presencia de endogeneidad ($\beta \neq 0$)

Los resultados de las estimaciones para (5) se pueden observar en la siguiente tabla¹⁹:

¹⁹ Se estimaron modelos de corte transversal para cada uno de los años en estudio.

Año	$ES_{i,t}$	t-statistic	p-value
1996*	-5.667643	-1.945018	0.0653
1997***	-6.839898	-2.997959	0.0069
1998	-2.686641	-1.497099	0.1492
1999*	-46.04547	-1.789757	0.0879
2000**	5.714577	2.238376	0.0362

*** Significativo al 1%; ** Significativo al 5%; * Significativo al 10%
Fuente: cálculos del autor

El mismo ejercicio se realizó utilizando como variable endógena el NBI, obteniendo los resultados que se exponen a continuación:

Año	$ES_{i,t}$	t-statistic	p-value
1996**	-9.149387	-2.575153	0.0176
1997**	-7.028238	-2.609449	0.0164
1998*	-4.584438	-2.022505	0.0560
1999**	-84.57947	-2.362962	0.0279
2000**	7.887551	2.238525	0.0361

*** Significativo al 1%; ** Significativo al 5%; * Significativo al 10%
Fuente: cálculos del autor

Resulta claro entonces que: EN GENERAL SE RECHAZA LA HIPOTESIS NULA y por ende SE DEMUESTRA FORMALMENTE LA ENDOGENEIDAD ENTRE $Pb_{i,t}$ Y $C_{i,t}$.

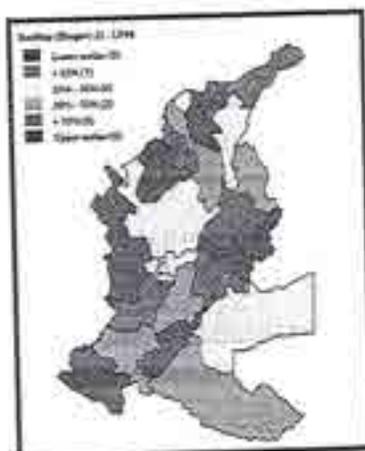
6.4. Anexo 4: Herramientas de visualización y mapeo del Análisis Exploratorio de Datos Espaciales (ESDA)²⁸

Las herramientas de visualización y mapeo que en el presente trabajo se han empleado, son aquellas que tienen aplicación en los software más populares en econometría espacial, tales como Arcview, Geoda y Stars. Dichas herramientas

²⁸ En base a: GEODA 0.9.5-4 (BETA) USER GUIDE. ANSELIN, Luc. 2004.

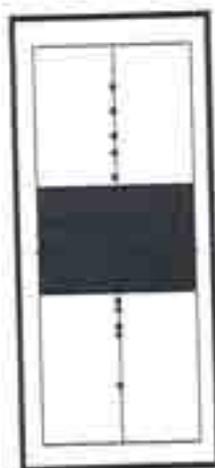
son:

BOXMAPS



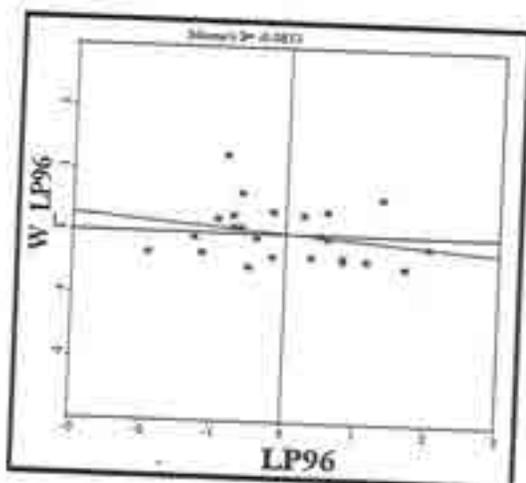
Permite visualizar en un mapa la distribución de los datos, dividiéndola en cuatro partes iguales y asignando a cada uno los cuartiles un color diferente. Identifica la presencia de valores atípicos (outliers) tanto en la parte alta como en la parte baja de la distribución.

BOXPLOT



Permite visualizar la distribución de los datos. El área color gris corresponde a la distribución entre el 26% y el 75%, la franja que se ubica por encima corresponde al cuartil más alto (76% - 100%), mientras la franja que se ubica por debajo corresponde al cuartil más bajo de la distribución (0% - 25%).

SCATTERPLOT



Permite visualizar en un plano cartesiano la ubicación de la variable bajo estudio con respecto a su rezago espacial. En el eje X se distribuye el valor de la variable para cada una de las unidades, y en el eje Y su rezago espacial. La pendiente de la línea que se ajusta a la distribución de los datos, corresponde al valor del estadístico I de Moran.

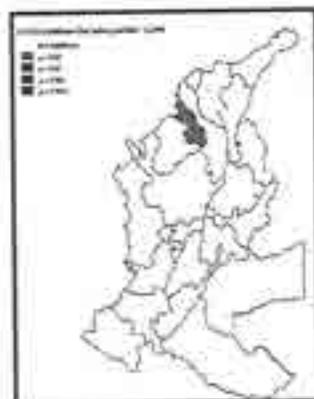
LISA CLUSTER MAP



Permite visualizar en un mapa los estadísticos LISA I de Moran que resulten significativos, así como la característica del cluster que se forme alrededor de dicha unidad, entre alto-alto, alto-bajo, bajo-alto y bajo-bajo, marcando con un

color diferente cada una de las anteriores opciones.

LISA SIGNIFICANCE MAP



Permite visualizar en un mapa el nivel de significancia de los LISA I de Moran que aparecen en el LISA Cluster Map.

ISA BOX PLOT



Posee la misma interpretación y utilidad del Box Plot expuesto anteriormente, salvo que en este caso representa la distribución de los LISA I de Moran. Permite identificar aquellas unidades espaciales que más aportan al I de Moran global. Los LISA I de Moran que resulten significativos aparecerán en la distribución como outliers.

6.5. Anexo 5: Fuentes y operacionalización de las variables del modelo

En la siguiente tabla aparecen resumidas las fuentes de donde se obtuvieron las variables que se emplean en el trabajo.

Variable	Fuente	Descripción
Línea de pobreza departamental (LP)	Departamento Nacional de Planeación (DNP)	Tomado del Sistema de información sociodemográfica del DNP
Índice de Tasa de Cambio Real departamental (ITCR)	PEREZ (2003)	Tomado de: "La tasa de cambio real regional y departamental en Colombia, 1990 - 2002" de PEREZ Javier, Documento de trabajo sobre economía regional del Banco de la República, Cartagena, Septiembre de 2003.
Tasa de desempleo departamental (D)	DNP	Tomado del Sistema de información sociodemográfica del DNP
Productividad laboral departamental (PL)	Cálculos del autor	Cálculos del autor en base a las cuentas departamentales del Departamento Nacional de Estadística (DANE) y proyecciones de la población de esa misma entidad.
Educación departamental (Es)	DNP	Tomado del Sistema de información sociodemográfica del DNP
Coefficiente de Gini departamental (G)	DNP	Tomado del Sistema de información sociodemográfica del DNP
Participación del sector servicios de cada departamento en el PIB departamental (Ser)	Cálculos del autor	Cálculos del autor en base a las cuentas departamentales del DANE
Productividad agrícola departamental (Pag)	Cálculos del autor	Cálculos del autor en base a las cuentas departamentales del DANE y a las proyecciones de población rural de esa misma entidad

Los casos en que se debió recurrir a las cuentas departamentales, estas se tomaron a precios constantes de 1994. En el caso del ITCR, la estimación no se encontraba disponible para los departamentos de Caquetá y Chocó, por lo que para obtener una aproximación a los valores de esta variable en estos departamentos, se efectuó una interpolación espacial. Dicha interpolación se realiza de la siguiente forma:

$$ITCR_i = \sum_j \left\{ (ITCR)_j \left[\frac{dist_i}{\sum_j dist_i} \right] \right\}$$

Donde:

$$dist_i = \sqrt{(CoordX_{Dpto_j} - CoordX_{Dpto_i})^2 + (CoordY_{Dpto_j} - CoordY_{Dpto_i})^2}$$

J son las unidades espaciales vecinas.

El procedimiento es bastante sencillo, ya que básicamente el resultado será el promedio ponderado de los ITCR de las unidades vecinas, donde la ponderación vendrá dada por la distancia que separa a las unidades vecinas de la unidad donde se desconoce el valor de la variable, de modo que a mayor distancia menor ponderación y viceversa.