Investigación Económica, vol. LXV, 258, octubre-diciembre, 2006, pp. 129-160

Econometría espacial y ciencia regional

JORGE A. PÉREZ PINEDA*

INTRODUCCIÓN

El análisis económico convencional ha dado, usualmente, mayor importancia al papel del *tiempo* como dimensión clave de estudio, sin valorar adecuadamente el factor *espacial*. Después de un período de estancamiento, particularmente de la ciencia regional, recientemente, en la década de los noventa, autores como Krugman (1991 y 1998), Fujita *et al.* (2000), entre otros, renovaron el interés por estos temas en obras como su "Economía espacial".[1]

Manuscrito recibido en enero de 2005; aceptado en marzo de 2006.

* Doctor en Economía por la Universidad Complutense de Madrid, Departamento de Economía Aplicada I <japerpe@yahoo.com>. El autor agradece al CONACYT que financió la estancia doctoral bajo la cual se realizó este trabajo, así como a dos dictaminadores anónimos por sus valiosos comentarios y aportaciones a la primera versión de este artículo.

[1] Un recorrido muy claro y profundo del papel del espacio en la economía y de las distintas escuelas vinculadas a su estudio puede consultarse en Asuad (2001).

Así, el resurgimiento de la ciencia regional a través de la reconsideración y revalorización del espacio en el análisis económico, ha traído la aparición de un nuevo campo teórico que intenta abarcar lo que pioneros de la teoría de la localización, la geografía económica y la economía regional propusieron en su momento, agrupados bajo el epígrafe de la "Nueva Geografía Económica" (Fujita *et al.*, 2000).[2] Dicho enfoque destaca el papel de los rendimientos crecientes a escala, las fuerzas centrípetas (interpretadas como economías externas marshallianas) y centrífugas, en un contexto de causación circular acumulativa.

Paralelamente al interés renovado por estos temas, se dio también un amplio desarrollo de técnicas y métodos de medición, respaldados por la creación de nuevos *software* que permitirían hacer ciertos cálculos y mapear información de tipo geográfica y regional.

En este contexto surgen metodologías como la econometría espacial (EE) y programas vinculados a ésta como el SpaceStat (Anselin, 1991), así como otros programas que también permiten su uso, como la librería de EE que incorpora MATLAB (LeSage, 1999) y, de reciente aparición (y aún en fase experimental), programas que vinculan el análisis regional con series de tiempo, como el STARS[3] (Space-Time Analisis of Regional Systems). Este tipo de herramientas facilitaría el análisis con datos de corte transversal y temporal referenciados geográfica y espacialmente.

En forma complementaria, el desarrollo del Sistema de Información Geográfica (GIS, por sus siglas en inglés) permitiría que otros programas, más o menos vinculados a la ciencia geográfica, regional y espacial, pudieran ofrecer también opciones de interés para trabajar con bases de datos cada vez más sofisticadas a nivel espacial, tales como ArcView, el Mappoint de Microsoft o las versiones de Excell anteriores a 1998, que contaban con

[2] Cabe decir que tal resurgimiento no ha estado exento de críticas, en el sentido de que puede ser cuestionable que la llamada "Nueva Geografía Económica" tenga algo de nuevo y de geográfico. Al respecto, una crítica reciente se encuentra en Martin (1999).

[3] Para mayor detalle e información sobre este programa se recomienda consultar: S.J. Rey y M.V. Janikas, "STARS: Space-Time Analysis of Regional Systems", *Economics Working Paper Archive at WUSTL Series "Urban/Regional"* n° 0406001, 2004.

herramientas para generar mapas; así como la hoja de cálculo de Lotus, que también contó con estas facilidades, entre otros.

En ese sentido, consideramos importante exponer en este artículo los aspectos más relevantes de la EE como una técnica alternativa en investigaciones de temas regionales, porque una gran parte de los estudios de este tipo, para el caso mexicano en los últimos años, no la han empleado.[4]

Muestra de ello son las técnicas identificadas en algunos de los trabajos más recientes y representativos para la economía mexicana en materia de crecimiento y desequilibrios regionales, entre los que destacan: el análisis clásico de la convergencia regional con o sin variantes geográficas en Arroyo (2001), Esquivel (1999 y 2000), Messmacher (2000), Cermeño (2001); el análisis sectorial-estatal para evaluar el impacto regional de la apertura económica en estudios comparativos como el realizado por Chamboux-Leroux (2001) en que se confrontan análisis de equilibrio general (Krugman y Livas, 1996) frente a modelos gravitacionales como Polese y Pérez (1996) o estudios específicos como en Hanson (1998); el papel de la especialización flexible y las mejores prácticas (*best practices*) en la búsqueda de un desarrollo regional y local más equilibrado en Dussel y Ruiz (1999) y Ruiz (2000).[5]

El desarrollo del presente artículo incluye los siguientes aspectos: una breve descripción de los antecedentes de la EE (parte II), resaltando las principales características de esta metodología a través de los "efectos espaciales" que surgen en los modelos con datos geográficamente referenciados (parte III), donde se definen dichos efectos y se mencionan las principales causas de éstos, sus formas de detección y especificación.

[4] En el nivel internacional, pasa algo similar en cuanto a la subutilización de estas técnicas, en Vaya y Moreno (2000) se hace un recuento de este tipo de trabajos al menos hasta el 2000. *Vid.* cuadro 1.1, pp.16-17, que da cuenta del escaso uso de esta econometría. Adicionalmente, estas autoras presentan uno de los pocos tratamientos breves y concisos del uso de estas técnicas en lengua española.

[5] A pesar de ello es de interés destacar, en la línea de los estudios de crecimiento y desequilibrios con matices espaciales en otros países, los trabajos de S.J. Rey (2001), S.J. Rey y M. Janikas (2005), J. Villaverde y A. Maza (2003), Bueno y Alañón (2000), M.T. Ramírez y A.M. Loboguerrero (2002), entre otros. Para el caso mexicano los trabajos vinculados a la economía espacial se han relacionado más a temas relacionados con la migración como en G. Lara Ibarra e I. Soloaga (2005).

Seguidamente se plantean las formas de especificación de modelos espaciales más comunes a partir de la exposición del modelo de regresión clásico: modelos de rezago espacial (Lag), de error espacial (Error) y modelos simultáneos (parte IV). Por último se incluyen algunas consideraciones finales (parte V) y un anexo con los principales estadísticos de correlación y contrastes de dependencia (pruebas de detección de efectos espaciales).

La econometría espacial: antecedentes

El desarrollo de la ciencia regional, bautizada originalmente como economía espacial por Isard (1956) planteó la necesidad de diferenciar no sólo la economía estándar (o no espacial) de la economía espacial (o regional) sino también de diferenciar las herramientas analíticas que se utilizaban y aquéllas que con el tiempo se iban perfeccionando.[6]

Este proceso de diferenciación ha sido permanente, así por ejemplo, en sus orígenes el análisis regional moderno reconocía la existencia de cuatro raíces teóricas: 1) la teoría de la localización; 2) la teoría del multiplicador internacional e interregional; 3) análisis de insumo-producto interindustrial; 4) la programación matemática (Meyer, 1966). En esa misma línea Isard (1960) señalaba ya la importancia de contar con marcos analíticos regionales integrados, que permitieran una síntesis de las diferentes herramientas de análisis regional y técnicas vinculadas a subsistemas regionales (distintos procesos y escalas regionales) bajo un mismo marco teórico.

En esa evolución se desarrolla una serie de trabajos que comienzan a estimar y evaluar problemas que surgían en modelos econométricos vinculados con análisis regionales y multiregionales, con lo cual se generaron, por un lado, avances en el desarrollo de herramientas de análisis regional de forma independiente; por el otro, avances en el análisis de modelos integrados regionales.

En la primera categoría, se encontraría la econometría espacial, término que acuñaría Paelinck a principios de los setenta para referirse a modelos

[6] Como referencia a esos primeros desarrollos de herramientas vinculadas a aspectos espaciales se suelen reconocer los trabajos de Moran (1948) y Geary (1954) sobre autocorrelación espacial.

econométricos multiregionales, mientras que en la segunda (modelos integrados regionales) se encontrarían modelos como los de econometría e insumo producto, que combinan más de una metodología para un mismo tipo de análisis (conocidos como EC+IO models, por sus siglas en inglés).[7]

La EE es de gran utilidad cuando se considera el uso de variables vinculadas al espacio (datos referenciados geográfica o espacialmente, por ejemplo, datos a nivel estatal), ya que este tipo de datos suelen presentar (por su naturaleza) relaciones multidireccionales, traducidas como dependencia en el espacio o autocorrelación espacial (similar a los problemas de dependencia temporal o autocorrelación serial, presente en series de datos temporales) que pueden invalidar el uso de la econometría clásica.

La EE se presenta así, como una alternativa útil a la econometría clásica para tratar con datos referenciados espacialmente, y por otro, contrastar ciertos fenómenos económicos a través de la modelación de relaciones entre observaciones.

Posteriormente, Paelinck y Klaassen (1979) definirían la EE en función de cinco características: 1) el papel de la interdependencia espacial; 2) la asimetría en las relaciones espaciales; 3) la importancia de factores explicativos localizados en otros espacios; 4) diferenciación entre interacción *ex-ante* y *ex-post*; 5) modelación explícita del espacio.

En vías de una construcción más formal de una definición y un desarrollo metodológico más completo, la obra de Anselin (1988) sienta las bases de un cuerpo analítico más sólido; conceptualmente plantea que: "La colección de técnicas que tratan con las peculiaridades causadas por el espacio en el análisis estadístico de los modelos de la ciencia regional se consideran el dominio de la econometría espacial".

---

7 Sobre esta última e importante línea de desarrollo metodológico regional y multiregional, véase S.J. Rey, "Integrated Regional Econometric + Input-output Modeling: Issues and Opportunities", *Papers in Regional Science*, 79, 2000, pp. 271-292. Aquí se reconoce, después de un interesante tratamiento de los tipos y variantes de modelos integrados en la ciencia regional, la escasa relación entre la EE y los modelos EC+IO, a pesar de existir temas tangenciales como la dependencia espacial, por lo que se proponen interesantes ideas para investigaciones futuras en aras de integrar más estas metodologías.

La EE identifica principalmente dos tipos de fenómenos en los datos utilizados, conceptualizados como efectos espaciales: la dependencia espacial entre las observaciones y la heterogeneidad espacial que puede surgir en los datos analizados.[8]

Adicionalmente la EE permite:

- Contrastar la presencia de dependencia espacial (¿existe?, ¿de qué tipo es?).
- Especificar una estructura de dependencia espacial (¿quiénes interactúan?).
- Estimar modelos con dependencia (¿qué estructura la recoge mejor? por ejemplo, Lag o Error).
- El uso de matrices de pesos o rezagos para incorporar la dependencia espacial, y métodos de estimación basados en máxima verosimilitud y variables.
- Instrumentales (como los más utilizados) cuando mínimos cuadrados .
- Ordinarios (MCO) presentan desventajas.

LOS EFECTOS ESPACIALES

En el primero de los efectos espaciales, la dependencia espacial, profundizamos un poco más, dado que la modelación y su detección son parte central en el cuerpo analítico de la EE. En cuanto al segundo efecto, la heterogeneidad espacial, puede manejarse aplicando técnicas basadas en la econometría clásica, por lo que se mencionan sólo los aspectos más relevantes en torno a éste.

La dependencia espacial

Cuando se trabaja con series de tiempo en econometría, uno de los supuestos del modelo lineal clásico, plantea que no debe haber autocorrelación o correlación serial en las perturbaciones del modelo. Esto implica que el

[8] En la revisión de la econometría clásica realizada por exponentes de la EE como Anselin (1988) o Vaya y Moreno (2000) se identifica la ausencia del tratamiento de estos efectos espaciales. Las últimas autoras por ejemplo citan el caso de Novales (1997, pp. 161-162) quien al hablar de autocorrelación menciona sólo de forma anecdótica que: "datos basados en criterios geográficos pueden mostrar correlación en los términos de error de dichas observaciones, lo cual se denomina *autocorrelación espacial*"; sin embargo, no plantea ni propone ningún método para su tratamiento a pesar de reconocer la existencia de procedimientos concretos. Algo similar pasa en Gujarati (1995, p. 293) que reconoce que

término de perturbación de una observación no está asociado al término de perturbación asociado a otra observación:

$$E(ui,uj) = 0 \quad i \neq j \qquad [1]$$

Siguiendo a Gujarati (1995) la autocorrelación se puede definir como "la correlación existente entre los miembros de una serie de observaciones ordenadas en el tiempo o espacio". Este autor como muchos otros, reconoce que los conceptos de correlación y autocorrelación serial se suelen tomar como sinónimos, si bien existen diferencias.

Diferenciando correlación y autocorrelación serial, Gujarati (1995, pp. 288-289) señala que, la autocorrelación serial plantea la correlación de rezagos de una serie dada, consigo misma, rezagada en un número de unidades de tiempo (u1,u2,…,u10 y u2,u3,…, u11), mientras que la correlación serial plantea la correlación de rezagos entre dos series diferentes (u1,u2,…,u10 y v1,v2,…,v10).

Cuando se traslada este concepto a datos transversales y con algún criterio geográfico, sus términos de error podrán también estar relacionados entre sí, encontrándonos entonces con el concepto de correlación espacial (Novales, 1997, p. 162) o autocorrelación espacial (Gujarati, 1995, p. 292), para hablar de aquella correlación en el espacio en vez de la temporal. La presencia de autocorrelación plantea así la relación, o correlación, o la falta de independencia entre los errores de distintos periodos u observaciones espaciales, por lo que también habrá referencias de la autocorrelación como dependencia, si bien existen diferencias conceptuales referentes a los momentos de la distribución de probabilidad conjunta de las variables, Vaya y Moreno (2000, p. 21).[9]

---

la correlación no sólo se presenta en series temporales, sino también en datos de corte transversal que tratan con observaciones de regiones, a lo que varios autores llaman autocorrelación espacial, pero como el caso anterior no se propone ni se especifica cómo tratar el problema.

[9] Estas autoras señalan: "En un sentido estricto, los conceptos de dependencia y autocorrelación espacial no son sinónimos, siendo la autocorrelación espacial una expresión más débil de la dependencia espacial, relativa únicamente a los primeros momentos de la distribución conjunta de una variable". (Vaya y Moreno 2000, p. 21, nota al pie 2). A pesar de ello en la literatura de la EE se les trata indistintamente

Una vez hechas estas matizaciones conceptuales de la violación a uno de los supuestos de los modelos de MCO, podemos abundar en otros aspectos de interés de la autocorrelación o dependencia espacial en la EE.

Definición, tipos y causas

Definida como, "la existencia de una relación funcional entre un punto dado en el espacio y lo que ocurre en cualquier otro", es una situación que suele reflejar la ausencia de independencia en observaciones de conjuntos de datos de tipo transversales (Anselin, 1988, p. 11; LeSage, 1999, p. 3). En términos formales se define como:

$$\begin{gathered}\text{dado } yi=f(yj),\ i=1,\ \ldots,n,\ j\neq i\\ \mathrm{Cov}[yi,\ yj]=E[yi\ yj]-E[yi]-E[yj]\neq 0,\ \text{para } i\neq j\end{gathered} \qquad [2]$$

El que una observación asociada a una localización (i) se relacione con otra observación en una localización j≠i, tal relación se expresa por el momento condicional de la covarianza entre ambas localizaciones.[10]

Se identifican dos tipos de autocorrelación espacial, *positiva* si la existencia de un fenómeno determinado en una región dada propicia su expansión a otras regiones circundantes y dicha expansión genera la concentración del mismo. De manera opuesta, la autocorrelación espacial *negativa* se refiere a la existencia de fenómenos en una región que impiden u obstaculizan la aparición de estos en otras regiones vecinas.

Las causas de la autocorrelación espacial se identifican en dos hechos: la existencia de errores de medida para observaciones en unidades espaciales contiguas y la existencia de varios fenómenos de interacción espacial (Anselin, 1988, pp. 11-13):[11]

---

[10] *Vid.* Anselin (1999, pp. 4-7) y Anselin (1988, pp. 11-13) para aspectos más formales.

[11] Los motivos de la autocorrelación en series temporales o correlación serial son similares a los que propician la autocorrelación en series transversales bajo autocorrelación espacial como: variables excluidas y sesgo de especificación (o errores de medida); por lo que será el ámbito que tratamos el que defina el tipo de datos y, por tanto, el tipo de fenómenos que esperamos surjan en nuestro análisis.

- Errores de medida (también conocida como autocorrelación residual): derivados de la posibilidad de escasa correspondencia entre el ámbito espacial del fenómeno de estudio y las unidades espaciales de observación, los errores de medición serán muy probables.
- Fenómenos de interacción espacial (conocida como autocorrelación sustantiva): es decir, lo que ocurre en un territorio afecta a otros territorios, lo que ocurre en un punto en el espacio está determinado por lo que pasa en otro(s) punto(s), debido a interdependencias en tiempo y espacio de unidades espaciales como: efectos de difusión (*spillovers effects*), de transferencia, procesos de dispersión, interacciones, externalidades, jerarquías, etcétera.

La existencia de estos fenómenos causantes de la dependencia espacial llevan a considerar instrumentos que permitan incorporarles dentro de modelos econométricos y estadísticos para detectar la presencia y tipo de dependencia espacial.

Instrumentos para incorporar la dependencia espacial

En torno a la dependencia espacial, se observa que su existencia tiene orígenes similares a la autocorrelación serial; sin embargo, una diferencia importante y clave en la ciencia regional y la EE es que, mientras la autocorrelación serial vincula variables en una sola dirección, entre un periodo y el siguiente o el anterior (años, días, trimestres, meses, etc.), la autocorrelación espacial vincula variables no homogéneas agrupadas de formas muy diversas en el espacio.

Es decir, plantea relaciones multidireccionales, donde cada observación posee distintas características (distintos tamaños, distintas ubicaciones, distintas distancias entre ellas, etc.).

La EE reconoce usualmente dos instrumentos a través de los cuales se puede expresar la dependencia espacial y con ello resolver el problema de la multidireccionalidad en la modelación de este efecto espacial: las matrices de pesos y los retardos espaciales.

Matrices de pesos espaciales

Un instrumento que fusiona el hecho de la interdependencia y las relaciones multidireccionales son las matrices de pesos espaciales, de retardos o contactos, definida con la letra W (por la palabra inglesa *weight*, peso) y representada de la siguiente forma:

$$W = \begin{bmatrix} 0 & w_{12} & . & w_{1N} \\ w_{21} & 0 & . & w_{2N} \\ . & . & . & . \\ w_{N1} & w_{N2} & . & 0 \end{bmatrix} \quad [3]$$

Esta matriz deberá ser simétrica, transpuesta y, por tanto, cuadrada, sus elementos w_{ij} representan la interdependencia existente entre las regiones i y j, y serán no estocásticos y exógenos al modelo (Vaya y Moreno, 2000, p. 23; Anselin, 1999, p. 6).

Aunque se reconoce que no hay una forma estándar de construir esta matriz, la propuesta por Moran y Geary, basada en la noción de contigüidad binaria entre unidades espaciales (de primer orden), sigue siendo vigente y de gran utilidad.

Así, los elementos w_{ij} toman los valores binarios de 0 si las regiones i y j no son vecinas (o adyacentes) y 1 si lo son (es decir el valor de sus pesos serán no negativos y finitos) en función de la definición que se adopte en su construcción. Por ejemplo, si se considerán distancias en lugar de vecindades para contigüidad entre ciudades (o estados, etc.), lejos podría valorarse con 0 y cerca con 1.[12]

Dado que no existe una definición generalizada de lo que debe ser la matriz de peso espacial (salvo por el reconocimiento que sus valores sean no negativos y finitos), en la literatura existente se puede encontrar otro tipo de

[12] En Anselin (1988, cap. 3), LeSage (1999) y Vaya y Moreno (2000) se puede profundizar sobre los aspectos formales y conceptuales de las matrices de pesos espaciales y sus variedades, basados en los trabajos de Cliff y Ord, Moran y Geary, entre otros.

matrices de pesos que puede emplearse de la misma forma que la descrita, en función de los objetivos de investigación o criterios de análisis, entre las más usuales se encuentran las siguientes (Vaya y Moreno, 2000, pp. 24-25):

1. Matriz de distancias: basada en la distancia entre regiones se define como $w_{ij}=d_{ij}^{-a}\beta_{ij}^{b}$, donde d_{ij} es la distancia entre las unidades i y j, y β_{ij} la longitud relativa de la frontera común entre i y j, relacionada al perímetro i, con parámetros a estimar a y b. Otra medida similar es $w_{ij} = \gamma_{ij}\beta_{ij}\alpha_i$ donde a diferencia del anterior γ_{ij} es un factor de contigüidad binario y α_i el área de la región i con relación a la total.
2. Matriz inversa de distancias: en ésta la intensidad de la interdependencia entre dos regiones disminuirá con la distancia que separa los centros respectivos. ($w_{ij}=1/d)^{n}$.
3. Matrices alternativas: evaden el tratamiento de la contigüidad o proximidad física, sobresalen matrices basadas en la distancia económica entre regiones, o que capturan el grado de intercambio comercial entre regiones analizadas.[13]

Adicionalmente, todas estas matrices pueden estandarizarse, dividiendo cada elemento w_{ij} por la suma total de la fila a la que pertenece, donde la suma de cada fila será igual a la unidad ($w_{ij}/\Sigma w_{ij}$). Con ello se busca ponderar por igual la influencia total de cada región con sus vecinas; sin embargo, la estandarización no siempre es adecuada para matrices basadas en el concepto de distancia, por los problemas de asimetría que pueden surgir y, por tanto, de interpretación o significado (Vaya y Moreno, 2000).

Retardos espaciales

A diferencia del análisis de series temporales en que las observaciones de variables independientes y dependientes presentes y pasadas encuentran correspondencia, tal situación no se cumple de la misma manera entre variables espaciales.[14]

---

[13] Para una descripción más formal de la manera de construir estas matrices se pueden consultar Anselin (1988, pp. 19-22), Vaya y Moreno (2000, pp. 24-25), LeSage (1999, p. 13), basados en los trabajos de Cliff y Ord, Dacey, Bodson y otros.

[14] En el análisis temporal esto indicaría la relación entre variable dependiente e independiente que muchas veces no es instantánea (sino rezagada), lo cual plantea una relación de forma unidireccional,

Por ello, la noción de cambio (entre periodos *shift*) es asimilada en la EE a través de la noción de cambio espacial, pero en análisis de tipo regional, basado en la utilización de mapas como fuente de información; este concepto pierde fuerza por la multiplicidad de relaciones existentes (multidireccionalidad, falta de grados de libertad, etc.) que no permiten el cálculo de parámetros eficientes.

Como solución a ello, la EE plantea el operador de retardos espaciales, el cual consiste en un promedio ponderado de variables aleatorias en localizaciones vecinas –con ponderaciones fijas y exógenas– (Anselin, 1999, p. 5; Vaya y Moreno, 2000a, pp. 26-27).

El operador de rezagos espaciales se forma definiendo para cada localización yi a su vecino en la correspondiente columna como un elemento distinto de cero w_{ij} en una matriz de pesos espaciales W positiva y no estocástica, siendo cada elemento de una variable retardada espacial (igual al promedio ponderado de los valores de la variable en el subgrupo de observaciones vecinas llamado Si, dado w_{ij}=0 para j≠Si) de la siguiente manera:

$$[Wy]i = \Sigma j\ w_{ij} * y_j \quad \text{con } j=1,\ldots,N \qquad [4]$$

Definido así el producto de la matriz de pesos por el vector de observaciones de una variable aleatoria (Wy), en que w_{ij} son pesos espaciales y *y* un vector de observaciones de la variable aleatoria Nx1.[15]

A pesar de su utilidad, se debe ser cuidadoso con el tipo de matriz de pesos espaciales que se trabaje, ya que se plantea que una matriz estandarizada, por ejemplo, no siempre arrojará información económicamente interpretable o significativa (Anselin, 1988, pp. 23-24; Vaya y Moreno, 2000, pp. 27-28).

---

usualmente vinculados a los modelos autoregresivos, modelos de retardos distribuidos o de medias móviles, en que el operador de retardos $Lx_t=x_{t-1}$ es de gran utilidad, mientras que en EE las relaciones son multidireccionales, razón por la que se requiere una alternativa en el ámbito espacial.

[15] *Vid.* Anselin (1988, pp. 22-24), Anselin (1999, pp. 5-6) y Vaya y Moreno (2000, pp. 26-27) para aspectos formales de definición.

La heterogeneidad espacial

La estructura de los datos que se suelen utilizar en análisis que incorporan variables regionales, urbanas o consideraciones espaciales concretas, suelen presentar inestabilidad sobre el espacio, ya que mediante datos espaciales diferentes (heterogéneos) se trata de explicar el mismo fenómeno.

Es decir, las observaciones de variables utilizadas vinculadas con regiones, o la agregación seleccionada, no poseen las mismas características, a decir: las regiones no poseen el mismo tamaño, ni la misma densidad de actividad (de población, de empresas, etc.), ni el mismo nivel de ingreso, ni la misma distribución de recursos.

En torno a las causas de la heterogeneidad espacial, se identifican dos problemas en que ésta se manifiesta; la inestabilidad estructural y la heteroscedasticidad:

- La inestabilidad estructural: está vinculada a la estructura espacial de las observaciones, es decir cuando no existe homogeneidad en éstas, se tiene inestabilidad espacial, que repercute en las formas funcionales y en la variación de los parámetros de regresión.
- La heteroscedasticidad: vinculada a los procesos espaciales, surge ante la omisión de variables o formas de especificación erróneas, lo cual provoca errores de medición.

Respecto a la detección de ambos problemas, se considera que utilizando los métodos existentes en la EC se pueden tratar la heteroscedasticidad e inestabilidad estructural; la primera con pruebas como White o Breusch-Pagan y la segunda con la prueba de cambio estructural de Chow.[16]

Para especificar y estimar modelos en presencia de heterogeneidad espacial se proponen típicamente cuatro opciones:[17] 1) variación de coeficientes aleatorios; 2) *Switching regressions* en contexto espacial (también conocido como regímenes espaciales); 3) expansión espacial; 4) regresiones ponderadas geográficamente.

---

[16] *Vid.* Vaya y Moreno (2000, pp. 133-135) para aspectos formales de estas pruebas en su entorno clásico y espacial.

[17] Para aspectos formales y mayor profundización de estos métodos por separado e incorporando la dependencia espacial, pueden consultarse: Vaya y Moreno (2000, pp. 137-141) y Anselin (1988).

Dado que para contrastar y detectar la heterogeneidad espacial se pueden utilizar los mismos métodos convencionales que se usan para detectar y corregir la heteroscedasticidad en la econometría clásica, es que no se le da un tratamiento profundo en la EE. La excepción es el caso en que se presentan simultáneamente ambos efectos espaciales (heterogeneidad y dependencia espacial); por estas mismas razones, a la dependencia espacial y sus formas de tratarla se le dedica más atención dentro de este campo.

TIPOLOGÍAS DE MODELOS ESPACIALES: ESPECIFICACIÓN, ESTIMACIÓN Y DETECCIÓN

Después del análisis de dependencia espacial, el siguiente paso es analizarla en modelos de regresión espaciales basados en las formas generales en que se puede representar e incorporar la dependencia espacial.

Para ello partimos de la presentación "formal" del modelo lineal clásico en el cuadro 1, a fin de diferenciarlo de las estructuras espaciales, así como de las consecuencias e implicaciones de introducir la dependencia espacial en cada estructura espacial.

CUADRO **1**
Características del modelo de regresión lineal clásico

Modelo lineal clásico	*Supuestos*
1. $y = XB+e$	El modelo de regresión es lineal, es decir, la variable dependiente es una función lineal de un conjunto específico de variables independientes, más una perturbación.
2. $E[e/X] = 0$	El valor esperado de la perturbación aleatoria debe ser cero para cualquier observación.
3. $Var[e_j/X]=\sigma^2$, para $i=1,\ldots,n$ $Cov[e_i,e_j/X]=0$, para $i\neq j$ $\therefore E[ee'/X]=\sigma^2 I$	Las perturbaciones tienen una varianza uniforme y no están correlacionadas (implica homoscedasticidad y no autocorrelación).
4. X fija en muestras repetidas	Las observaciones en variables independientes se pueden considerar fijas en muestras repetidas.
5. X es una matriz con rango K	No existe relación lineal exacta entre las variables.

Los estimadores del modelo serán MELI (BLUE), teorema Gauss-Markov.
Fuente: P. Kennedy (2003, p. 50).

Derivado de lo anterior, podemos establecer de manera muy sucinta que la diferencia entre la econometría clásica (o estándar) y la EE se basa en el tratamiento e incorporación de los dos efectos espaciales comentados en el tercer apartado, en términos de LeSage (1999): "la econometría tradicional ha ignorado estos dos elementos que violan los supuestos de Gauss-Markov usados en la modelación de regresiones."

Esto implica, por un lado, que la dependencia espacial viola el supuesto de que en un muestreo repetitivo las variables explicativas son fijas e independientes y, por tanto, ni éstas ni sus residuos seguirán una distribución normal e independiente. Por el otro, la heterogeneidad espacial viola el supuesto de que existe una sola relación lineal a lo largo de observaciones de una muestra de datos, lo cual evidencia problemas de heteroscedasticidad.

Siguiendo a Novales (1997, p. 161), podemos recordar que una de las propiedades del estimador de MCO de los coeficientes de un modelo lineal plantea que el término de error tiene una matriz de covarianzas escalar (todos sus elementos son cero, excepto los de la diagonal principal que son igual a σ_u^2). Existen dos situaciones en que esta matriz posee una estructura distinta: una de ellas es cuando la varianza del término de error es distinta de unas observaciones a otras (heteroscedasticidad, vinculada al segundo efecto espacial) $Var(u_t) = \sigma^2_t$, con $\sigma^2_t \neq \sigma^2_s$ para $t \neq s$, en datos de sección cruzada ocurre si $Var(u_i) = \sigma_i^2$, con $\sigma_i^2 \neq \sigma_j^2$, para $i \neq j$.

Una segunda situación ocurre cuando, con datos de series temporales, los términos de error correspondientes a distintos periodos no son independientes entre sí; es decir, $Cov(u_t, u_{t-k}) = E(u_t, u_{t-k}) \neq 0$ para algún $K>0$. Esto hace que los elementos fuera de la diagonal de la matriz de covarianzas no sean todos nulos, por lo que ésta no será diagonal. Situación denominada autocorrelación, que refleja el hecho de que el término de error está correlacionado consigo mismo a través del tiempo.

Cuando dicha situación se presenta en muestras de sección cruzada se le denomina autocorrelación espacial antes definida y explicada, como el primer efecto espacial. La econometría convencional presenta pruebas para tratar con problemas de heteroscedasticidad, y autocorrelación serial, pero no así para tratar la autocorrelación espacial, de ahí la importancia de la EE.

La dependencia espacial (según el tipo) en modelos de regresión aparece como consecuencia de autocorrelación en variables dependientes o independientes y por la aparición de un esquema de dependencia espacial en las perturbaciones.

Estas variantes vinculadas al tipo de dependencia espacial se pueden traducir a su vez en distintas formas de incorporación a los modelos de regresión espacial a través de matrices de peso para su especificación como antes se señaló, tales como:

- Variables dependientes espacialmente rezagadas: Wy.
- Variables explicativas espacialmente rezagadas: Wx.
- Términos de error espacialmente rezagados: Wu.

Dado que nuestra intención es centrarnos en una visión global de estas técnicas, por motivos de espacio nos enfocaremos a la modelación general y descriptiva estándar de estas variantes.[18]

Las matizaciones señaladas, están detrás de la determinación o elección del tipo de modelo espacial a especificar más adecuado para tratar con los efectos espaciales detectados, cuyas variantes más usuales son:

- Modelos Error (de error espacial) si la dependencia espacial es residual.
- Modelos Lag (de rezago espacial) si la dependencia espacial es sustantiva.

Estos dos tipos de dependencia (residual y sustantiva) determinarán el tipo de especificación del modelo a seguir (en tres especificaciones generales). Una característica común de las formas generales de procesos espaciales es la consideración de formas autorregresivas, que de esa manera podrán captar el efecto multidireccional de las variables espaciales.

[18] Para ver en detalle y a profundidad aspectos específicos e implicaciones de estas técnicas sobre la econometría estándar o variantes en las técnicas de EE se recomienda consultar Anselin, Florax y Rey (2004) y Anselin (1988). Asimismo, este tipo de especificaciones espaciales posee un referente directo con los modelos de retardos distribuidos (finitos no restringidos, polinomiales y geométrico) estudiados en series de tiempo, *Vid.* Greene (2000, cap. 17) abunda sobre el funcionamiento de estas metodologías.

Así, la primera estructura (especificación) general (simultánea) que se propone en forma matricial, derivada de la estructura del modelo lineal (y=XB+e), se denomina modelo espacial autorregresivo de primer orden que plantea la siguiente forma:

$$y = \rho Wy + u \qquad [5]$$

Donde *y* y u son vectores de variables y el término de error, respectivamente, y W la matriz de pesos espaciales.[19] El término ρWy es la estructura autorregresiva en la variable dependiente que intenta explicar el modelo (el rezago espacial en la variable dependiente). Dentro de esta especificación se puede derivar una amplia gama de modelos, que pueden incluir la combinación de estructuras autorregresivas y procesos de medias móviles, adicionalmente la disponibilidad de información también influye sobre la variedad de estructuras modelables.

En ese sentido, las especificaciones más usuales de modelos espaciales que se proponen se caracterizaran por su estructura lineal autorregresiva; así se identifican dos tipos como punto de partida de las derivaciones que puedan existir: modelos de regresión espaciales con datos de corte transversal (tipo1) y modelos de regresión lineal para datos que combinan tiempo–espacio (tipo 2).

Retomando la definición de dependencia espacial como "la situación en que la variable dependiente o el término de error en cada localidad está relacionada con observaciones de la variable dependiente o valores del término de error en otras localidades", el caso general se puede mostrar formalmente como:

$$E[y_i, y_j] \neq 0 \quad \text{o} \quad E[e_i, e_j] \neq 0, \text{ para las localidades i y j} \qquad [6]$$

[19] En Anselin (1988, cap. 4), se plantea con mayor detenimiento la diferencia y la importancia de los procesos simultáneos y condicionales espaciales sobre una correcta especificación y validación. Asimismo se menciona también que, dentro de la ciencia regional lo más usual es la modelación basada en la forma simultánea (expresada como probabilidad conjunta).

La estructura que requiere la dependencia espacial para estimar en un caso de este tipo viene dada por la matriz de pesos W, que reduce el número de parámetros desconocidos a uno (el coeficiente de asociación espacial del proceso espacial autorregresivo o en el proceso espacial de medias móviles).

Como señala Anselin (1992) las consecuencias de ignorar autocorrelación espacial en un modelo de regresión cuando está presente, dependen de la forma de la hipótesis alterna. Habrá dos casos básicamente, como se mencionó, que se derivan de la primera especificación y que recogen algún tipo de dependencia. En el primero, la autocorrelación espacial ignorada se vincula a la variable dependiente (y), identificando tal caso con un modelo de rezago espacial (modelo Lag) o de dependencia espacial sustantiva.

Así, la segunda especificación (basada en el modelo base del tipo 1), se reflejaría en un modelo mixto regresivo-autorregresivo espacial de primer orden (conocido como modelo Lag o de retardo espacial) que tendrá la siguiente forma (rezago más otras variables explicativas):

$$y = \rho Wy + X\beta + u \;\rightarrow\; y = (1-\rho W)^{-1} X\beta + u \qquad [7]$$
$$u \sim N(0, \sigma^2 I)$$

Donde y es un vector Nx1, Wy el retardo espacial de la variable dependiente y X una matriz de K variables exógenas, u un término de perturbación ruido blanco, y ρ el parámetro espacial autorregresivo, con $H_0 = \rho = 0$. Si tal forma de autocorrelación espacial se ignorara, los estimadores de MCO serían insesgados y la inferencia basada en el modelo de regresión clásica (un modelo sin Wy) sería incorrecta.

La tercera especificación, relacionada con la autocorrelación espacial vinculada al término de error, o autocorrelación residual (error espacial), plantea el modelo base tipo 2[20] cuya estructura formal plantea un proceso

[20] Este caso plantea la incorporación del tiempo en un modelo espacial, con ello se consideran patrones de dependencia y heterogeneidad de corte transversal, con lo que a partir de este modelo también se pueden realizar varias especificaciones que consideran una amplia variedad en que se manifiesta la dependencia tiempo-espacio y formas de heterogeneidad.

espacial para los términos de error, también conocido como modelo con perturbaciones espaciales autorregresivas de primer orden:

$$\begin{gathered} y_{it} = X_{it}\beta_{it} + \varepsilon_{it} \\ \varepsilon_{it} = \lambda W\varepsilon + u \;\rightarrow\; y_{it} = X_{it}\,\beta_{it} + (1-\lambda W)^{-1}\,u_{it} \\ u \sim N\,(0,\sigma^2 I) \end{gathered} \tag{8}$$

Donde X_{it} es un vector renglón de observaciones para la unidad espacial i en el periodo t, β_{it} es un vector de parámetros específicos tiempo-espacio y ε_{it} es un término de error con las siguientes condiciones: $E[\varepsilon_{it}] = 0$ y $E[\varepsilon_{it}, \varepsilon_{js}] \neq 0$.

La segunda condición posibilita la existencia de varias opciones de modelar las dependencias tiempo-espacio y patrones de heterogeneidad espacial si la varianza se considera constante, o la existencia de algún efecto espacial.[21]

Con $W\varepsilon$ como el rezago especial en el término de error, λ como el coeficiente autorregresivo y u con un término de error bien comportado (homoscedástico e incorrelacionado). En este caso la hipótesis nula seria $H_0{=}\lambda{=}0$, y las consecuencias de ignorar dependencia espacial en el error son las mismas que para la heteroscedasticidad, los estimadores de MCO permanecen insesgados, pero no eficientes, al ignorar la correlación entre términos de error.

El análisis y detección del tipo de dependencia espacial del modelo que se desarrolle o con el que se trabaje, es clave para la correcta especificación y aplicación de un método de estimación que permita inferir sobre la teoría que desde la ciencia regional se proponga.

De acuerdo con lo anterior, el origen de la autocorrelación espacial determina el tipo de especificación del modelo de regresión a seguir que incorpore este efecto espacial en su estructura; así, los siguientes modelos mostrados

---

[21] Sobre la variedad de modelos basados en esta forma general véase Anselin (1988, pp. 36-39). En este caso la variación se puede dar en el espacio sobre el índice i, el tiempo en el índice t, o sobre tiempo-espacio (it). Asimismo, $E[\varepsilon\varepsilon'] = \sigma^2[(I-\lambda W)^{-1}(I-\lambda W)^{-1'}]$, lo cual implica que cada localidad esta relacionada con cualquier otra localidad en el sistema, vinculándonos con la dependencia espacial (en el error o en la variable dependiente).

en los cuadros 2 y 3 son los más utilizados en la EE, al incorporar variantes en las restricciones (Vaya y Moreno, 2000b; Anselin, 1999).[22]

CUADRO **2**
Variantes de modelos espaciales 1

Tipo	*Estructura*
Caso 1: modelo mixto regresivo-autorregresivo espacial con perturbaciones espaciales autoregresivas, cuya estructura presenta la siguiente forma: $$y = \rho W_1 y + X\beta_1 + \varepsilon$$ $$\varepsilon = \lambda W_2 \varepsilon + u$$ $$u \sim N(0, \Omega)$$ Matriz de varianza-covarianza $\Omega_{ii} = h_i(z\alpha)$, $h_i > 0$	Donde β es un vector Kx1 de parámetros asociado con la matriz Nxk de variables exógenas X (no a variables dependientes rezagadas); ρ es el coeficiente de la variable dependiente espacialmente rezagada; λ el coeficiente en la estructura espacial autorregresiva para la perturbación ε; u es el término de ruido blanco distribuido normalmente con una Ω en la diagonal general de la matriz de convarianzas. Los elementos de la diagonal permiten heteroscedasticidad como una función P+1 de variables exógenas z, que incluye un término constante. Los P parámetros α se asocian a los términos no constantes, teniendo así que para $\alpha=0$ entonces $h=\sigma^2$ (presencia de homoscedasticidad). Finalmente, las matrices W_1 y W_2 NxN son matrices de pesos espaciales estandarizadas o no estandarizadas, asociadas con el proceso espacial autorregresivo en la variable dependiente y el término de perturbación respectivamente. Esto permite a los dos procesos ser conducidos por estructuras espaciales distintas, planteando que en general, el modelo tendrá 3+K+P parámetros desconocidos, vectorialmente: $$\theta = [\rho, \beta`, \lambda, \sigma^2, \alpha]`$$

[22] En estos modelos es necesaria la utilización de algún tipo de matriz espacial W para capturar el efecto de la autocorrelación espacial. Cabe destacar también que los últimos dos casos mencionados son los más generales que incorporan el efecto espacial en la variable endógena y en el término de perturbación, pero no son las únicas formas de especificación. Como se mencionó arriba, según el

CUADRO **2, continuación...**

Tipo	*Estructura*
Caso 2: una segunda especificación espacial más general que incorpora elementos de éstos ([7] y [8]), denominado modelo mixto regresivo espacial con perturbaciones autorregresivas y heteroscedásticas, con la siguiente estructura: $y = \rho W_1 y + X\beta_1 + W_2 R\beta_2 + \varepsilon$ $\varepsilon = \lambda W_3 \varepsilon + u$ $u \sim N(0, \Omega);\ \Omega_{ii} = h_i(Z\alpha),\ h_i > 0$	Donde y es el vector Nx1 de la variable endógena, X una matriz NxK_1 de variables exógenas y R una matriz NxK_2 de variables exógenas retardadas espacialmente que no necesariamente coinciden con las de la matriz X. El término ε incorpora una estructura de dependencia espacial autorregresiva bajo un esquema de Markov de primer orden, u representa un vector distribuido normalmente con matriz de varianza y covarianza diagonal heteroscedastica (cuyos elementos de la diagonal principal estarán en función de P+1 variables exógenas Z). Finalmente, ρ es el coeficiente de la variable retardada espacialmente, λ es el coeficiente en la estructura autorregresiva espacial de ε, las β_1 y β_2 son vectores k_1x1 y k_2x1 asociados a las variables exógenas y exógenas retardadas respectivamente, y α es un vector Px1 asociado a los términos no constantes de Z.[23]

Fuente: elaboración propia basada en Vaya y Moreno (2000).

Cabe destacar que el común denominador de las especificaciones antes descritas es que la dependencia espacial se ha incorporado al término de error en una estructura autorregresiva de primer orden que es lo comúnmente usado, aunque se puede modelar de manera distinta (con órdenes superiores). Para ello se sugieren algunas especificaciones como las siguientes (casos 4 y 5):

valor que tome algún parámetro en [7] o la forma de la matriz de varianzas-covarianzas de [8] se tendrá un modelo que capture las dependencias y patrones de heterogeneidad de forma distinta. *Vid.* Anselin (1988, pp. 34-39) y Vaya y Moreno (2000, pp. 70-75).

[23] En este caso, a pesar de los subíndices en las matrices de contigüidad (W) no hay consenso sobre su especificación y se sugiere la matriz de contigüidad binaria, para evitar relaciones espurias.

CUADRO **3**
Variantes de modelos espaciales 2

Tipo	*Estructura*
Caso 4: $y = X\beta_1 + \varepsilon$ $\varepsilon = \theta W_2 u + u$ $u \sim N(0, \sigma^2 I)$	Tal especificación plantea la utilización de un proceso de medias móviles espacial SMA(1) donde θ es el coeficiente de medias móviles espacial y u el término de perturbación no correlacionado.
Caso 5: $y = \rho W_1 y + X\beta_1 + \varepsilon$ $\varepsilon = \theta W_2 u + u$ $u \sim N(0, \sigma^2 I)$	Basado en el anterior es posible reespecificar otro modelo que incorpora la estructura de dependencia espacial media móvil, llamado modelo mixto regresivo espacial autorregresivo con perturbaciones espaciales cuya estructura SARMA (1,1).

Elaboración propia basado en Vaya y Moreno (2000).

Finalmente, se debe señalar que las especificaciones [7] y [8] están relacionadas directamente con los conceptos de autocorrelación espacial sustantiva y residual respectivamente, las cuales se conceptualizan de la siguiente forma:

- Autocorrelación espacial sustantiva (autocorrelación causada por fenómenos de interacción espacial): en el primer caso, al omitir erróneamente un retardo espacial de la variable endógena (o exógena) la dependencia espacial se trasladaría directamente al término de perturbación, el cual estaría correlacionado espacialmente. Su solución implica la inclusión de una variable sistemática correlacionada espacialmente en el modelo de un retardo espacial (modelo Lag) como las expresiones [6] y [7].
- Autocorrelación espacial residual (autocorrelación causada por errores de medida): cuando la dependencia espacial residual no está causada por la omisión errónea de un retardo de alguna de las variables sistémicas, se considera la existencia de autocorrelación espacial residual como *nuisance* o perturbación. La forma de solucionar tal problema se basa en la especificación de un modelo espacial en términos del error, que considere explícitamente esquemas de dependencia espacial basados en el error, tal y como se incorpora en la expresión [8] (Vaya y Moreno, 2000, p. 69).[24]

[24] También se puede modelar la dependencia en la variable exógena, pero es poco usual, en tal caso habría una especificación como los llamados modelos espaciales Durbin, o de factor común, con expresiones como: $y = \lambda Wy + X\beta - \lambda WX\beta + u$ (las restricciones o factor común $\lambda\beta = -\lambda\beta$). Si la restricción no se cumpliera se complicaría la estructura, quedando fuera de la familia de modelos autoregresivos *Vid.* Anselin (2002).

Métodos de estimación

Una vez detectado el tipo de dependencia espacial (sustantiva o residual) podemos analizar los métodos más adecuados para estimar bajo dependencia espacial, sea cual sea ésta, dado que MCO deja de ser del todo útil.[25]

En concreto y de forma general se consideran dos métodos de estimación, el primero y más utilizado es el método de máxima verosimilitud (MV) y el segundo, el método de variables instrumentales (VI). Como procedimiento alternativo, para muestras grandes y para tratar la dependencia espacial simultánea, también se pueden usar métodos condicionales como el *coding method* basado en MCO sobre un subconjunto de observaciones no contiguas.[26]

Respecto a máxima verosimilitud, siguiendo a Vaya y Moreno (2000), obsérvese el cuadro 4.

Por otra parte, el método de VI consiste muy sucintamente en eliminar la correlación entre la variable endógena rezagada (espacialmente) y el término de error (por lo que se usa principalmente en modelos tipo Lag), para lo cual se busca una (o varias) variable(s) o instrumento(s) *Q,* aproximada(s) a la variable dependiente original (Z = [yL, X]) y correlacionadas con ésta pero no con las perturbaciones, resolviendo para el parámetro θ, que es el estimador de VI θ_{VI}. Matricialmente:

$$y = Z\theta + \varepsilon \quad \text{con} \quad \theta_{VI} = [Q'Z]^{-1}Q'y \qquad [9]$$

La técnica de VI se presenta como una alternativa a la de MV; sin embargo, su utilización genera estimadores consistentes e insesgados pero no eficientes, cuya pérdida de eficiencia se vincula al tamaño de la muestra. Ambos casos

---

[25] Usualmente para modelos Lag, los métodos de estimación comúnmente utilizados son el de máxima verosimilitud y variables instrumentales, y para modelos tipo Error se suelen utilizar los métodos de máxima verosimilitud y el de momentos generalizados dadas sus características, véase Anselin (1992, V).

[26] Para más detalle véase Anselin (1995, pp. 49-50). Otros métodos menos usados son: estimación bayesiana (LeSage, 1999) y métodos de estimación robustos basados en la técnica de *boostrap o jacknife* (Anselin, 1988).

plantean limitaciones que son importantes considerar en el desarrollo de una investigación, sin embargo en términos generales estas metodologías son las más utilizadas.[27]

CUADRO **4**
Estimación por máxima verosimilitud

Estimación bajo máxima verosimilitud para un retardo espacial en la variable endógena	*Estimación bajo máxima verosimilitud en presencia de correlación espacial en el término de error*
Como en el contexto temporal, los estimadores MV se obtendrán a partir del logaritmo de la función de verosimilitud asociada al modelo especificado,[28] igualando a cero el vector de derivadas parciales de este (Ln L(θ), con θ= [ρ, β\`, $\lambda$, $\sigma^2$, $\alpha$ ]\`). Los pasos que se sugieren son: 1. Estimación por MCO del modelo $y=X\beta o+u$ y obtener el estimador b_o de MCO para βo. 2. Estimación por MCO del modelo $Wy=X\beta_L+u$ obtener el estimador b_L MCO para β_L.	Para el caso en que existe la presencia de una estructura espacial autorregresiva en el término de perturbación la obtención del estimador MV de λ no es tan directo, y se recomienda un proceso iterativo: 1. Estimación por MCO del modelo $y=X\beta+u$ y obtener el estimador b de MCO para β. 2. Obtención del conjunto inicial de residuos MCO, e_{MCO}. 3. Con lo anterior, obtener el valor λ que maximiza la función de verosimilitud para este caso (distinta al anterior).

[27] Para profundizar sobre limitaciones de estas metodologías y formas de superarlas, se recomienda consultar Vaya y Moreno (2000) y Anselin (1988), principalmente, dado que la intención aquí sólo es ilustrar los principales aspectos de estas metodologías.

[28] Por ejemplo para un modelo mixto regresivo espacial autorregresivo de primer orden como el presentado en la expresión [7] tendría una función de verosimilitud como la siguiente:

$$\mathrm{LnL}_c = \mathrm{cte} + \ln|I - \rho W| - \frac{N}{2}\ln\frac{(e_o - \rho e_l)'(e_o - \rho e_l)}{N}$$

Donde el término de error es homoscedástico y hay un retardo espacial en la variable endógena, y de la cual buscamos un estimador del parámetro ρ bajo MV del cual dependen todos los términos de esta función de verosimilitud. Para abundar sobre aspectos formales de este caso y del modelo con perturbaciones espaciales véanse Anselin (1988, cap. 6) y Vaya y Moreno (2000a, cap. 4).

CUADRO **4**, **continuación...**

Estimación bajo máxima verosimilitud para un retardo espacial en la variable endógena	*Estimación bajo máxima verosimilitud en presencia de correlación espacial en el término de error*
3. Obtener los residuos e_o y e_L. 4. Con los residuos anteriores, se obtiene el estimador de MV ρ maximizando la función de verosimilitud. 5. Con la estimación ρ, mediante MV obtener β con $b_{MV} = b_o - \rho_{MV} b_L$ y $\sigma^2_{MV}=(1/N)(e_o-\rho_{MV}b_L)`(e_o-\rho_{MV}b_L)$.	4. Obtenida λ, se obtiene b_{MCGE}. 5. Obtención de un nuevo conjunto de residuos con la estimación MCGE de β del punto 4. 6. Con lo obtenido en 5, regresar al punto 3 y repetir el proceso hasta alcanzar criterios de convergencia. 7. Con las estimaciones definitivas de β y λ se calcula $\sigma^2_{MV} = (1/N)e'B'Be$ donde $B=(I - \lambda_{MV}W)$.

Fuente: Vaya y Moreno (2000a, pp. 102-103). Se abunda en este método dado que es el más utilizado y también es accesible para trabajar con él mediante programas econométricos como SpaceStat.

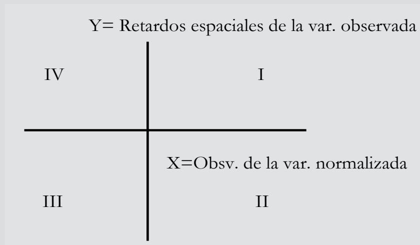
Detección de la dependencia espacial

En concreto se identifican tres estadísticos de contrastación de dependencia espacial dentro del análisis exploratorio, que son complementarios: la I de Moran, la C de Geary y el Contraste G(d) de Getis y Ord, en los que se puede usar cualquier definición de W, usualmente estandarizada (salvo en el caso de la G(d) que requiere una matriz simétrica) aunque la definición de W influirá en los resultados (los respectivos estadísticos se definen en el cuadro A1 del anexo final).

Un instrumento adicional que podemos utilizar para la contrastación del grado de dependencia espacial de las variables, es el diagrama de dispersión (*scatter plot*) *de Moran*, ilustrado en el diagrama 1.

La forma en que se encuentren los puntos sobre el diagrama de Moran indicará el tipo de autocorrelación. Si el diagrama no muestra uniformidad en las observaciones representadas, indicará ausencia de autocorrelación, si los datos agrupados describen una diagonal, dependiendo de la pendiente (igual al valor del contraste de la I de Moran) habrá autocorrelación positiva (pendiente positiva, cuadrantes I y III) o negativa (pendiente negativa, cuadrantes II y IV).

DIAGRAMA **1**
Scatterplot *de Moran*



Fuente: basado en Vaya y Moreno (2000).

Siguiendo las estructuras en que se puede especificar la autocorrelación en los modelos de regresión (Lag o Error), existe una serie de estadísticos propuestos para la dependencia espacial sustantiva y residual dentro del análisis de regresión. Originalmente en Anselin (1999) se sugieren sólo tres estadísticos:un estadístico global de autocorrelación Moran I, una para estructuras de rezago LM-LAG y una para estructuras de error LM-ERR.

En concreto, para evaluar la omisión errónea de un retardo espacial de la variable endógena (dependencia espacial sustantiva), ésta se puede contrastar con pruebas basadas en el principio del multiplicador de Langrange (LM por sus siglas en inglés) o Rao Score (RS). (Véase cuadro A2 del anexo final para los contrastes de dependencia espacial sustantiva.)

En cuanto a la autocorrelación espacial residual, se contrasta con dos tipos de pruebas. Las primeras y más comunes; la I de Moran y Kelejian-Robinson (K-R), y las segundas, basadas en los principios de máxima verosimilitud y el multiplicador de Lagrange, las pruebas LM-ERR y LM-EL (véase cuadro A3).

CONSIDERACIONES FINALES

Dada la similitud de la dependencia espacial en modelos espaciales con los modelos de series temporales, se puede pensar que las propiedades de los estimadores con variables dependientes rezagadas o correlación serial residual se pueden trasladar al caso espacial, pero no es así debido a la naturaleza multidireccional de los modelos espaciales usados en la ciencia regional.

Así, la utilización de la econometría clásica, y métodos convencionales de ésta como mínimos cuadrados ordinarios, y algunas de las pruebas usadas para detectar y tratar la dependencia temporal, pueden quedar invalidadas en algunos de los casos aquí especificados de existir dependencia espacial sustantiva o residual.

En ese sentido, la EE se presenta como una alternativa interesante dentro del contexto actual, en el que muchas investigaciones de corte regional, geográfico o espacial enfatizan la importancia de lo local y regional sobre lo global o agregado. Se deja constancia así de los aspectos más sobresalientes de esta técnica para futuras investigaciones en el caso mexicano, como complemento de otras en proceso y como aportación a la difusión de la EE en lengua española, esperando con ello una mayor utilización y profundización de aspectos tanto teóricos como empíricos en la materia.

Referencias

Anselin, L.,"A Model of Growth through Creative Destruction", *Econometrica,* 60(2), March, 1992.

———,"Spatial Externalities, Spatial Multipliers and Spatial Econometrics", *Regional Economics Applications Laboratory*, USA, Department of Agricultural and Consumer Economics, University of Illinois, Urban-Champaign Urbana, 2002, pp. 1-13

———, "Spatial Effects in Econometric Practice in Environmental and Resource Economics", Paper to be presented at the *Allied Social Science Associations*, 2001 Annual Convention, New Orleans, LA, January 5-7, 2001, pp. 1-14.

———, "Spatial Econometrics", *Bruton Center*, School of Social Sciences, Richardson, TX 75083-0688, University of Texas at Dallas, 1999, pp. 1-30.

———, *Spatial Econometrics: Methods and Models*, The Netherlands, Kluwer Academic Publishers, 1988.

———, *SpaceStat Version 1.80 User´s Guide*, Urban Champaign, Urban IL 61801, University of Illinois, 1995, pp. 1-62.

Anselin, L., Raymond J.G.M. Florax, S.J. Rey (eds.), *Advances in Spatial Econometrics Methodology, Tools and Applications*, Series: Advances in Spatial Science XXII, 2004, pp. 513.

Arroyo, F., "Dinámica del PIB de las entidades federativas de México, 1980-1990", *Comercio Exterior*, 51(7), México, Bancomext, julio de 2001, pp. 583-599.

Asuad, N., *Economía regional y urbana*, México, Benemérita Universidad Autónoma de Puebla, Colegio de Puebla, AEFE-UNAM, 2001.

Bueno, J. y A. Alañón, *Regional Growth and Regional Imbalances: Spain and USA*, Documento de trabajo 2000-18, España, Facultad de Ciencias Económicas y Empresariales, Universidad Complutense de Madrid, 2000, pp. 1-19

Cermeño, R., "Decrecimiento y convergencia de los estados mexicanos: un análisis de panel", *El Trimestre Económico*, México, FCE, LXVIII(4), 272, octubre-diciembre 2001, pp. 603-629.

Chamboux-Leroux, J., "Efectos de la apertura comercial en las regiones y la localización industrial en México", *Comercio Exterior*, 51(7), México, Bancomext, julio de 2001. pp. 600-609.

Christaller, W., *Central Places in Southern Germany*, Jena Fischer, 1933. Traducción inglesa de C.W. Baskin, Londres, mayo de 1966.

Dussel, E. y C. Ruiz, *Dinámica regional y competitividad industrial*, México, Editorial Jus/UNAM- Fundación Friedrich Ebert, Facultad de Economía, 1999.

Esquivel, G., "Geografía y desarrollo económico en México", Documento de trabajo de la red de centros del BID, R. 389, Banco Interamericano de Desarrollo, abril del 2000, pp. 1-45.

———, "Convergencia regional en México, 1940-1995", *El Trimestre Económico*, LXVI(4), 264, México, FCE, octubre-diciembre de 1999, pp. 725-761.

Fujita, M., P. Krugman y A. Venables, *Economía espacial*, 1era. edición, España, Editorial Ariel, 2000.

Geary, R.C., "The Contiguity Ratio and Statistical Mapping", *The Incorporated Statistician,* 5, 1954, pp. 115-145.

Greene, W.H., *Análisis econométrico*, tercera edición, España, Prentice Hall, 1999.

Gujarati, D., *Econometría*, 2da. edición, México, McGraw-Hill, 1995.

Hanson, G., "North American Economic Integration and Industry Location", *Oxford Review of Economic Policy*, 14(2), 1998, pp. 30-44.

Isard, W., *Location and Space Economy*, MIT Press, Cambridge, 1956.

Isard, W., D.F. Bramhall, G.A. Carrothers, J.H. Cumberland, L.N. Moses, D.O. Price, E.W. Schooler, *Methods of Regional Analysis*, Cambridge, MA, MIT Press, 1960.

Kennedy, P., *A Guide to Econometrics*, Fifth edition, United Kingdom, The MIT Press, 2003.

Krugman, P., "What's New about the New Economic Geography?", *Oxford Review of Economic Policy*, 14(2), 1998, pp. 7-17.

———, *Geography and Trade*, USA, Published jointly by Leuven University Press and The MIT Press, 1991.

Krugman, P. y R. Livas, "Trade Policy and the Third World Metrópolis", *Journal of Development Economics*, 49, 1996, pp. 137-150.

Lara Ibarra, G. e I. Soloaga, *Determinants Of Migration in Mexico: A Spatial Econometrics Approach*, Puebla, Universidad de las Américas, 2005.

LeSage, J., *Spatial Econometrics*, USA, University of Toledo, 1999.

Lösch, A., *The Economics of Location*, Jena Fischer, 1940. Traducción inglesa de New Haven (CT), Yale University Press, 1954.

Martin, R., "The New 'Geographical Turn' in Economics: Some Critical Reflections", *Cambridge Journal of Economics*, 23, 1999, pp. 65-91.

Messmacher, M., "Desigualdad regional en México. El efecto del TLCAN y otras reformas estructurales", México, *Banco de México,* Documento de Investigación n° 2000-4, D.G.I.E., 2000, pp. 1-25.

Meyer, J.R., "Regional Economics: A Survey", en *Surveys of Economic Theory*, vol. II, American Economic Association and Royal Economic Society, 1966.

Moran, P.A., "The Interpretation of Statistical Maps", *Journal of the Royal Statistical Society*, Series B, 10, 1948, pp. 243-251.

Novales, A., *Econometría*, 2da. edición, Madrid, España, McGraw-Hill, 1997.

Paelinck, J. y L. Klaassen, *Spatial Econometrics*, Saxon House, Farnborough, 1979.

Polese, M. y S. Pérez Mendoza, "Integración económica norteamericana y cambio regional en México", en *Comercio Exterior*, 45(2), febrero de 1995.

Ramirez, M.T. y A.M.Loboguerrero, "Spatial Dependence and Economic Growth: Evidence from a Panel of Countries", *Borradores de Economía*, Banco de la República de Colombia, 206, 2002, pp. 1-37.

Rey, S.J., "Integrated Regional Econometric+Input-output Modeling: Issues and Opportunities", *Papers in Regional Science*, 79(3), 2000, pp. 271-292.

———, "Spatial Dependence in the Evolution of Regional Income Distributions", *Department of Geography and Regional Economics Applications Laboratory*, 2001.

Rey, S.J. y M.V. Janikas, "STARS: Space-Time Analysis of Regional Systems", *Economics Working Paper Archive at WUSTL series "Urban / Regional"* n° 0406001, 2004.

———, "Regional Convergence, Inequality, and Space", *Journal of Economic Geography*, 5, 2005, pp. 155-176.

Ruiz, C., "Mejores prácticas para el desarrollo industrial local", *El Mercado de Valores,* México, Nacional Financiera, octubre de 2000, pp. 26-34.

Vaya, E.y R. Moreno, *Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial*, Edicions Universitat de Barcelona, UB44 Manuals, 2000.

Villaverde J. y A. Maza, "Desigualdades regionales y dependencia espacial en la Unión Europea", *CLM-Economía* n° 2, primer semestre de 2003, pp. 109-128.

ANEXO

CUADRO **A1**
Estadísticos de correlación global

Estadístico	*Características*
I de Moran[a] $$I=\frac{N}{So}\frac{\sum_{ij}^{N} w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{N}(x_i-\bar{x})^2},\ i\neq j$$ Si la muestra fuese demasiado grande la I de Moran se estandariza adoptando la siguiente forma: $$Z(I)=\frac{I-E(I)}{[V(I)]^{1/2}} \sim N(0,1)$$	Donde xi es la variable cuantitativa en la región i, $\bar{x}$ la media muestral, w_{ij} los pesos de la matriz W, N el tamaño de la muestra y So= $\sum_i\sum_j w_{ij}$. Donde E(I) y V(I) son la esperanza y la varianza de I. El estadístico posee una distribución asintótica normal. Con Z(I) no significativa no se rechazará (acepta), Ho=no autocorrelación. Un valor significativo positivo (negativo), implicará autocorrelación positiva (negativa). Bajo autocorrelación positiva (negativa) se indica presencia de concentración de valores similares (disímiles) de las observaciones (x) entre regiones vecinas.
La C de Geary $$C=\frac{N-1\sum_{ij}^{N} w_{ij}(x_i-x_j)}{2So\sum_{i=1}^{N}(x_i-\bar{x})^2},\ i\neq j$$	Donde xi es la variable cuantitativa en la región i, $\bar{x}$ la media muestral, w_{ij} los pesos de la matriz W, N el tamaño de la muestra y So= $\sum_i\sum_j w_{ij}$. Este estadístico se contrasta de la misma forma que el anterior, con Ho= No autocorrelación vs. H1=dependencia espacial, sin embargo se interpreta con otro significado. Así una C de Geary estandarizada negativa (positiva) indicará dependencia positiva (negativa).
G(d) de Getis y Ord [b] $$G(d)=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(d)x_i x_j}{\sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j},\ i\neq j$$	Definido como medida de concentración espacial de una variable x. Se consideran 2 regiones i y j como vecinas si están a una distancia (d) determinada, así w_{ij}(d) tomará valor de 1, de lo contrario 0. La G(d) también se puede estandarizar, convirtiéndola en Z(G), que al igual que las anteriores se distribuye asintóticamente normal N(0,1). Se contrastará Ho para Z(G)=ausencia de autocorrelación, si es positiva y significativa indicará tendencia a la concentración de valores (similares) elevados de x en el espacio analizado, si es negativa y significativa, todo lo contrario.

Notas: a/ En la I de Moran las variables se introducen en desviaciones, por lo que se puede interpretar como medida de correlación de cada xi con las regiones a las que se vincula.
b/ Se considera una medida del grado de asociación o concentración de la variable en el espacio, dado que está basada en la suma de productos de la variable no normalizada. Por el tipo de datos que deben utilizarse para este estadístico (positivos y naturales) no puede aplicarse a los residuos de una regresión.
Fuente: Vaya y Moreno (2000, pp. 33-38).

CUADRO **A2**
Contrastes de dependencia espacial sustantiva en modelos de regresión

Contrastes de dependencia espacial sustantiva	*Características*
Tests LM-LAG $\text{LM-LAG}=\dfrac{\left[e'Wy/S^2\right]^2}{RJ_{\rho-\beta}}$	Donde *e* es el vector de residuos MCO del modelo lineal comentado, Wy el término espacial rezagado, S^2 la estimación de la varianza residual del mismo modelo y $RJ_{\rho\text{-}\beta} = [T1+(WX\beta)'M(WX\beta)/S^2]$. Esta prueba se distribuye como una χ^2 con un grado de libertad. Se ha impuesto $\lambda=\alpha_\rho=0$ con $Ho= \rho=0$
Test LM-LE $\text{LM-LE} = \left[\dfrac{e'Wy/S^2-e'We/S^2}{RJ_{\rho-\beta}-T1}\right]^2$	Con el mismo significado en los elementos definidos en el estadístico anterior y con T1= Traza $(W'W+W^2)$. Se distribuye como una χ^2 con un grado de libertad. A diferencia de la prueba anterior, esta es robusta ante posibles especificaciones erróneas locales, como la existencia de un término de perturbación correlacionado espacialmente.

Fuente: Vaya y Moreno (2000, pp. 88-89).

CUADRO **A3**
Contrastes de dependencia espacial residual en modelos de regresión

Contrastes de dependencia espacial residual	*Características*
La I de Moran $I = \dfrac{N}{S}\dfrac{e'We}{e'e}$	Donde *e* es el vector de residuos de MCO de un modelo lineal, N el tamaño muestral y $S=\sum_i\sum_j w_{ij}$ (la suma de los elementos w_{ij} de la matriz de pesos). En un contexto asintótico y con el supuesto de residuos incorrelacionados, la I de Moran estandarizada se distribuye como una normal estándar.

CUADRO **A3**, continuación...

Contrastes de dependencia espacial residual	*Características*
Test K-R (Kelejian- Robinson) $K\text{–}R = \frac{\gamma' Z' Z\gamma}{\alpha'\alpha/h_R}$	Expresión derivada de una regresión auxiliar (cuya variable dependiente es $C=e_i\, e_j$) en que se usan los productos cruzados de los residuos de las observaciones potencialmente correlacionadas espacialmente (según la matriz de contactos) y los productos cruzados de las variables explicativas de tales observaciones. Las variables explicativas de la regresión auxiliar Z_h, están formadas por productos cruzados X_i y X_j. γ es el vector de coeficientes obtenido de MCO en una regresión de C en Z, y α el vector asociado a los residuos.
Test LM-ERR $\text{LM-ERR} = \left[\frac{e'We/S^2}{T1}\right]^2$	Donde *e* es el vector de residuos de MCO, T1= Traza ($W'W+W^2$), S^2 la estimación de la varianza residual de dicho modelo. A diferencia de los dos estimadores anteriores éste se distribuye como una χ^2 con un grado de libertad.
Test LM-EL LM-EL = $\frac{\left[e'We/S^2 - T1(RJ_{\rho-\beta})^{-1} e'Wy/S^2\right]^2}{\left[T1 - T1^2(RJ_{\rho-\beta})^{-1}\right]}$	Donde M es una matriz idempotente y $RJ_{\rho-\beta}$= $[T1+(WX\beta)'M(WX\beta)/S^2]$. e es el vector de residuos MCO del modelo lineal comentado, Wy el término espacial rezagado, S^2 la estimación de la varianza residual y T1= Traza ($W'W+W^2$)

Fuente: Vaya y Moreno (2000, pp. 79-83).