



Munich Personal RePEc Archive

Fundamentals of Applied Spatial Econometrics

Herrera Gómez, Marcos

CONICET-IELDE, National University of Salta

2017

Online at <https://mpra.ub.uni-muenchen.de/80871/>
MPRA Paper No. 80871, posted 19 Aug 2017 11:59 UTC

Fundamentos de Econometría Espacial Aplicada*

Marcos Herrera

CONICET-IELDE, Universidad Nacional de Salta

Resumen

La creciente disponibilidad de información georreferenciada necesita de herramientas econométricas particulares como las desarrolladas por la econometría espacial. Esta rama econométrica se dedica al análisis de la heterogeneidad y dependencia espacial en modelos de regresión. En este trabajo se revisan los desarrollos más consolidados del área relativos a la especificación e interpretación de la dependencia espacial en corte transversal y datos de panel. El trabajo se completa con dos ejemplos empíricos clásicos.

Código JEL: C12, C21, C23.

Palabras Clave: Modelos espacio-temporales, Dependencia espacial, Matriz espacial.

Abstract

The growing availability of Geo-referenced information needs particular econometric tools such as those developed by Spatial Econometrics. This econometric branch is dedicated to the analysis of heterogeneity and spatial dependence in regression models. In this paper, I review the most consolidated developments in the area related to the specification and interpretation of spatial dependence in cross-section and panel data. The work is completed with two classic empirical examples.

JEL Code: C12, C21, C23.

Keywords: Spatio-temporal models, Spatial dependence, Weighting matrix.

*Este trabajo es una versión preliminar de un capítulo en elaboración basado con material del curso de Econometría Espacial dictado en UdeSA (2016). Agradezco profundamente a Jesús Mur por ser mi mentor en estos temas, además de un gran amigo. El trabajo ha sido parcialmente financiado por el proyecto del Ministerio de Economía y Competitividad de España (ECO2015-65758-P). Errores u omisiones son de exclusiva responsabilidad del autor. E-mail: mherreragomez@gmail.com.

1. Introducción

La econometría espacial es una rama econométrica dedicada al análisis de *efectos espaciales* en modelos de regresión en datos de corte transversal y de panel. Anselin (2010) ubica el nacimiento del área en el año 1979 producto de una serie de publicaciones que sentaron las primeras bases metodológicas (Bartels y Ketellapper, 1979; Bennett, 1979; Hordijk, 1979; Paelinck y Klaassen, 1979). Prácticamente una década más tarde, la publicación del libro *Spatial econometrics: Methods and models* (Anselin, 1988) consolida las principales líneas metodológicas que marcaron la primera agenda de investigación dentro del área. Desde los inicios, el principal método de estimación de los modelos espaciales fue por máxima verosimilitud (*ML*, *Maximum Likelihood*). Sin embargo, los aportes realizados por Kelejian y Prucha (1998, 1999) y Lee (2003) impulsaron la estimación por el variables instrumentales (*IV*, *Instrumental Variables*) y método de momentos generalizados (*GMM*, *Generalized Method of Moments*). Adicionalmente, el libro de LeSage y Pace (2009) resume los aportes por estimación bayesiana.

Tradicionalmente, los modelos que incorporaban explícitamente el espacio (topología geográfica) estaban confinados a temas de ciencia regional, economía geográfica o urbana. Tal es así que la primera vez que se empleó el término *econometría espacial* fue para brindar fundamentación metodológica a modelos regionales y urbanos (Jean Paelinck, *Annual Meeting of the Dutch Statistical Association*, Tilburg, 1974). Sin embargo, el actual entendimiento del espacio no está restringido a cuestiones geográficas ampliando el uso de las técnicas al espacio económico y social (distancia socio-económica). La diversidad de aplicaciones que utilizan herramientas de econometría espacial incluye temas de redes sociales, criminología, finanzas públicas, ciencia política, psicología, salud, educación, historia, transporte, marketing, entre muchos otros.

En este capítulo se revisan y organizan los desarrollos más consolidados de econometría espacial enfatizando un rango de tópicos relativos a la especificación e interpretación de los efectos espaciales. Estos efectos pueden dividirse en: (1) *dependencia o interacción espacial*, generalmente entendida por su versión más débil de *autocorrelación espacial* y (2) *heterogeneidad espacial* que es un tipo de inestabilidad estructural proveniente del término de error (heteroscedasticidad) o de los coeficientes de las variables del modelo. Entre estos efectos, la heterogeneidad espacial puede ser resuelta mediante herramientas econométricas estándares. En cambio, la dependencia espacial es el efecto que mayor dificultad presenta y por eso nos dedicaremos casi exclusivamente a su tratamiento. Debido a la amplia literatura sobre el tema, el interés se centrará en modelos con variable dependiente continua que buscan incorporar dependencia espacial.

Dentro del área económica, la econometría espacial adquirió mayor relevancia a partir de la década del 90' por diferentes desarrollos teóricos que involucraron cuestiones geográficas y de interacción social (por ejemplo, Krugman, 1991; Manski, 1993; Brock y Durlauf, 1995, entre otros avances). Puede afirmarse que esta demora obedece a cuestiones históricas, la teoría

económica minimizó la influencia del espacio o lo supuso neutral al comportamiento de los mercados y a los procesos productivos y distributivos de bienes. Un simple repaso de los principales libros de texto utilizados en las carreras de ciencias económicas permite comprobar que las principales corrientes teóricas no mencionan casi ningún aspecto de la localización de los individuos o de los mercados. Dicho olvido se traslada directamente a la enseñanza de la econometría. El análisis econométrico de datos con información georreferenciada no está presente en los libros de econometría de base y el problema de dependencia es relegado casi con exclusividad a series temporales. Este último punto impulsa el desarrollo del presente capítulo con la intención de cubrir una temática prácticamente ausente en cursos de econometría intermedia y avanzados.

El resto del capítulo está organizado de la siguiente manera. En la sección 2 se presentan los principales elementos de econometría espacial. La sección 3 desarrolla los principales modelos espaciales en corte transversal y en datos de panel. En la sección 4 se discute la interpretación de los modelos que incluyen efectos espaciales endógenos. La sección 5 ilustra los desarrollos teóricos mediante dos ejemplos empíricos. Finalmente, los comentarios finales son incluidos en la sección 6.

2. Conceptos y elementos espaciales

Esta sección contiene conceptos y temas relacionados a la dependencia espacial en su versión más débil de autocorrelación espacial. Antes de explorar la autocorrelación espacial, se presenta una discusión sobre los datos espaciales y sus particularidades. Este conocimiento nos permitirá hablar de autocorrelación espacial de una forma más precisa y entender la necesidad de la matriz de contactos espaciales, elemento clave en el área.

2.1. Sobre la naturaleza de los datos

El primer aspecto a destacar, y no siempre mencionado, es sobre la naturaleza de los datos espaciales. Existen diferentes tipos de datos espaciales y esto conlleva desarrollos estadísticos particulares para los mismos.

Cressie (1993) realiza una taxonomía de datos según los supuestos del proceso estocástico espacial (*spatial random field*). Sin una excesiva formalización, este proceso estocástico puede definirse como un conjunto de variables aleatorias \mathbf{y} , indexadas por localización i , $\{y_i, i \in D\}$, donde D es un índice proveniente de un subconjunto del espacio euclidiano ($D \subset \mathbb{R}^2$). Dependiendo del supuesto realizado sobre D , los datos pueden clasificarse como geoestadísticos (*geostatistical data*), areales o regionales (*lattice data*) y de patrón de puntos (*point pattern data*).

Los datos geoestadísticos son datos provenientes de una superficie continua ocupando un

subconjunto fijo D . Por ejemplo, D puede ser el área de Argentina y los datos representan la superficie de precipitaciones. Los datos regionales provienen de un índice D fijo, siendo una colección finita, regular o irregular, de puntos o áreas que particionan al espacio euclidiano. Ejemplo de este tipo es el área de Argentina particionada con datos del producto bruto geográfico por provincias. En el caso del patrón de puntos son datos en donde D es un proceso puntual en \mathbb{R}^2 . La diferencia de este último tipo de datos respecto a los anteriores proviene del hecho de que los puntos en D son el resultado de un proceso aleatorio, siendo el principal interés la localización de un evento en particular. Este último tipo de dato espacial es rara vez analizado en nuestra área. Para mayor detalle sobre estos tipos de datos una excelente referencia es Cressie (1993).

Las diferencias existentes entre los datos geoestadísticos y regionales pueden relacionarse con la noción de modelos de datos (*data models*) que son abstracciones de la realidad en forma manejable y almacenada por medio de los sistemas de información geográfica (*GIS, Geographic Information Systems*). El modelo raster almacena campos (mallas o cuadrículas regulares) pertenecientes a distribuciones espaciales continuas y representadas como superficies. En el modelo vectorial se almacenan objetos discretos representados por líneas, puntos y polígonos. En economía, estos objetos pueden corresponder a jurisdicciones, viviendas, radios censales, etc. Consecuentemente, los datos regionales y el modelo vectorial se encuentran asociados de una forma natural y son típicamente analizados en econometría espacial.

Sin importar el tipo de dato espacial a analizar, existen problemas que son comunes a todos. Uno de ellos es el problema de unidad de área modificable (*Modifiable Areal Unit Problem, MAUP*), versión geográfica de la falacia ecológica, en donde las conclusiones basadas en una agregación particular de las áreas o regiones pueden cambiar si se agregan los mismos datos en un conjunto diferentes de áreas o regiones (Waller y Gotway, 2004). El efecto de agregaciones geográficas diferentes fue investigado por Openshaw y Taylor (1979), con datos electorales de 99 *counties* en Iowa, incorporando el término *MAUP* para describir el problema. Estos autores construyeron todos los posibles agrupamientos en grandes distritos alcanzando diferentes e inexplicables correlaciones con rangos de valores desde -0.97 hasta +0.99. El problema tiene dos componentes distinguibles: uno es el efecto escala o agregación y el otro es el efecto agrupamiento o zonificación (Openshaw y Taylor, 1979; Openshaw, 1984). El primer efecto radica en la obtención de distintas inferencias estadísticas cuando el mismo conjunto de datos es agregado en diferentes resoluciones espaciales, especialmente cuando se agregan pequeñas áreas en grandes regiones (por ejemplo, datos de familias agregados para radios censales o para comunas de la Ciudad Autónoma de Buenos Aires, *CABA*). El segundo efecto se refiere a la variación en el resultado debido al agrupamiento alternativo de las regiones en la misma o similar escala espacial (por ejemplo, datos de familias agregados por comunas o por distritos escolares de *CABA*). Ambos efectos pueden generar errores que afectarán la validez de los resultados. El problema, igual que en el caso de la falacia ecológica, no tiene solución y esto

implica la necesaria cautela que debe hacerse al extraer conclusiones inferenciales, más aún si son base para la toma de decisiones políticas.

Otro problema relacionado a la naturaleza de los datos espaciales, o debido al proceso de medición, ocurre al integrar diferentes tipos de datos en una única base de información. Por ejemplo, es común que la toda información no esté disponible al nivel de escala deseada: imaginemos que deseamos investigar el impacto de la delincuencia o contaminación ambiental en un análisis hedónico de precio de vivienda. Los datos de ubicación de las viviendas son una nube de puntos (modelo vectorial) pero los datos de contaminación es una superficie (campo continuo, modelo raster) o en el caso del índice de delincuencia puede estar construido a nivel barrial (polígonos, modelo vectorial). La necesaria transformación de los datos espaciales a la escala de interés genera el problema de cambio de soporte (*COSP, Change Of Support Problem*). El término “soporte” hace referencia al tamaño y volumen de cada base pero también incluye la forma, el tamaño y la orientación espacial de los objetos o campos representados. El paso de un soporte involucra generar nuevas variables relacionadas a las originales pero con diferencias espaciales y estadísticas. Existen alternativas para atenuar el problema aunque no lo resuelven completamente. Para conocer el detalle de estas alternativas puede consultarse el artículo de Gotway y Young (2002) y referencias allí indicadas.

Una vez revisada la naturaleza de los datos espaciales discutiremos la definición de dependencia en estos datos y el enfoque más utilizado para introducir este elemento en la especificación econométrica.

2.2. Definiendo autocorrelación espacial

Es típico en econometría establecer independencia de las observaciones permitiendo así un tratamiento estadístico manejable. Una implicación de este supuesto es que los valores observados y_i , para un individuo i , son estadísticamente independientes de otros valores y_j , para el individuo j , $E(y_i y_j) = E(y_i) E(y_j) = 0$ donde $E(\cdot)$ es el operador de esperanza. Sin embargo, cuando los individuos representan a periodos temporales o a regiones en el espacio este supuesto es demasiado restrictivo y deberá relajarse permitiendo alguna forma de dependencia.

La definición más habitual de dependencia espacial es por medio de la autocorrelación espacial que puede definirse, vagamente, como la similitud de valores en localizaciones similares. La autocorrelación puede ser positiva, en el sentido que un valor alto (bajo) de una variable aleatoria en una posición espacial es acompañado por un vecindario de valores altos (bajos). También es posible una autocorrelación espacial negativa con un valor elevado (bajo) en una posición geográfica y valores bajos (elevados) en el vecindario. Ya sea si la autocorrelación espacial es positiva o negativa, su existencia implica que una muestra de datos contiene menor información que una muestra no correlacionada. Esta pérdida de información debe ser explícitamente considerada en la estimación y este es el problema esencial en temas de econometría

aplicada con datos espaciales.

De manera más formal, la autocorrelación espacial es una versión re-escalada de la siguiente condición de momentos (covarianzas):

$$\text{Cov}(y_i, y_j) = E(y_i y_j) - E(y_i) E(y_j) \neq 0, \quad \forall i \neq j, \quad (1)$$

donde y_i, y_j son observaciones de una variable aleatoria en la localización i y j en el espacio, es decir, todo par (i, j) posee información geográfica específica medida por latitud y longitud.

Dada una muestra de n observaciones en un corte transversal, existen $(n^2 - n)$ covarianzas a ser estimadas. La estimación de todas las covarianzas es imposible de realizar y no puede resolverse incrementando el tamaño muestral, siempre habrá más covarianzas que datos disponibles. Este es un problema de parámetros incidentales y para resolverlo se necesita imponer restricciones sobre la forma en que las observaciones interactúan entre si.

Existen al menos dos enfoques utilizados para estimar la covarianza espacial. En geostatística, todos los pares de localizaciones son clasificados según la distancia que los separa entre si y la covarianza es expresada como una función continua de dicha distancia, comúnmente conocida como variograma o semi-variograma. Esta alternativa supone que el proceso estocástico es continuo en el espacio (datos geoestadísticos). La segunda alternativa supone que el proceso espacial es discreto (datos regionales) construyendo una medida de similitud para la variable aleatoria y otra medida de similitud para la localización. La similitud de la localización se materializa en la construcción del vecindario para cada observación bajo una matriz de pesos espaciales. Esta alternativa es la elección mayoritaria en econometría espacial y será desarrollada a continuación.

2.3. Importancia de la matriz de pesos espaciales

El enfoque preponderante en econometría espacial utiliza la matriz de pesos espaciales para brindar una medida de similitud de las localizaciones. Esta matriz comúnmente se denominada \mathbf{W} (es similar a la matriz \mathbf{G} en redes sociales) y describe la conectividad entre n unidades que se encuentran localizadas en un espacio bidimensional. La construcción de \mathbf{W} está basada en al menos dos supuestos clave sobre la estructura espacial: (1) un criterio de conectividad que define qué unidades pueden ser consideradas vecinas entre si, y (2) un supuesto de pesos espaciales que operacionaliza cómo los vecinos se afectan entre si.

Para entender cómo la matriz \mathbf{W} permite simplificar el problema de dependencia espacial formalizaremos la discusión. Supóngase un proceso autorregresivo espacial *SAR* (*Spatial Autoregressive*) tal que la variable de estudio y se distribuye espacialmente en tres regiones de la

siguiente forma:

$$\begin{aligned}
y_i &= \alpha_{ij}y_j + \alpha_{ik}y_k + u_i, \\
y_j &= \alpha_{ji}y_i + \alpha_{jk}y_k + u_j, \\
y_k &= \alpha_{ki}y_i + \alpha_{kj}y_j + u_k, \\
u_i; u_j; u_k &\sim i.i.d. (0; \sigma^2).
\end{aligned} \tag{2}$$

es decir, en la primera ecuación, el valor y en la región i depende del valor de y en la región j y k , más un término aleatorio que se distribuye de manera idéntica e independiente entre las localizaciones. Lo mismo puede decirse de las regiones j y k . En términos matriciales, la expresión (2) se reduce a:

$$\begin{aligned}
\begin{bmatrix} y_i \\ y_j \\ y_k \end{bmatrix} &= \begin{bmatrix} 0 & \alpha_{ij} & \alpha_{ik} \\ \alpha_{ji} & 0 & \alpha_{jk} \\ \alpha_{ki} & \alpha_{kj} & 0 \end{bmatrix} \begin{bmatrix} y_i \\ y_j \\ y_k \end{bmatrix} + \begin{bmatrix} u_i \\ u_j \\ u_k \end{bmatrix}, \\
\mathbf{y} &= \mathbf{A}\mathbf{y} + \mathbf{u},
\end{aligned} \tag{3}$$

donde

$$\mathbf{A} = \begin{bmatrix} 0 & \alpha_{ij} & \alpha_{ik} \\ \alpha_{ji} & 0 & \alpha_{jk} \\ \alpha_{ki} & \alpha_{kj} & 0 \end{bmatrix}.$$

El problema con el sistema (3) es que, bajo un corte transversal, hay más parámetros que observaciones (3 observaciones y 6 parámetros). Puede imponerse simetría, reduciendo el número de parámetros pero, de todas formas, no podrá estimarse. Por lo tanto, es necesario buscar una solución alternativa para poder estimar la dependencia subyacente en los datos.

La solución radica en imponer un conjunto de restricciones sobre las relaciones de dependencia. De esta forma, la estructura de la matriz \mathbf{A} es reparametrizada de la siguiente manera:

$$\mathbf{A} = \rho\mathbf{W},$$

donde ρ es un parámetro a estimar que captura el efecto promedio de interacción entre todas las regiones y \mathbf{W} es la matriz de pesos espaciales. Los elementos de \mathbf{W} son:

$$\mathbf{W} = \begin{bmatrix} 0 & w_{ij} & w_{ik} \\ w_{ji} & 0 & w_{jk} \\ w_{ki} & w_{kj} & 0 \end{bmatrix}, \tag{4}$$

tal que se ha intercambiado a los parámetros α 's (parámetros originales del modelo) por w_{ij} (coeficientes exógenos al modelo) y el modelo queda expresado como $\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{u}$, donde $\mathbf{W}\mathbf{y}$

es interpretado como el rezago espacial de \mathbf{y} por analogía a series temporales.¹

Generalizando, la matriz \mathbf{W} será de orden $n \times n$, siendo n el tamaño muestral. Cada elemento de \mathbf{W} es denominado peso espacial, w_{ij} . Los pesos espaciales capturan la vecindad, siendo diferentes de cero cuando las regiones i y j son consideradas vecinas. Por convención, ninguna región puede ser vecina de si misma, dando como resultado que la diagonal principal de \mathbf{W} contenga todos sus elementos iguales a cero, $w_{ii} = 0$.

La matriz de pesos espaciales ocupa una posición clave y su elección condicionará al resto del análisis. El problema central radica en cómo se construyen los pesos de la matriz \mathbf{W} . En la práctica econométrica, la matriz de contactos es construida mediante diferentes criterios. Estos criterios van desde el uso de la posición geográfica hasta el uso de flujos que capturan interacciones sociales y utilizan otras fuentes de información socio-económica.

La forma más tradicional para construir la matriz es bajo criterios geográficos, siguiendo la primera ley de la geografía (Tobler, 1970): “Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes.” Entre los criterios geográficos, puede definirse vecindad por contigüidad, por alguna función de distancia, por k vecinos más cercanos o alguna combinación de estas opciones.

El criterio contigüidad considera como vecinos a los polígonos que comparten límites. Hay varias opciones dentro de este criterio: tipo torre (vecinos al norte-sur y este-oeste) o tipo reina (todos los polígonos que comparten límite), nombres alusivos a los movimientos del ajedrez. Cuando la información es de tipo puntual (como el caso de viviendas), la aplicación de este criterio puede realizarse mediante la creación regiones artificiales (polígonos Voronoi o Thiessen).

Una opción más natural para la información de tipo puntual es utilizar algún criterio de distancia. Es posible construir diversas medidas, por ejemplo: vecinos son aquellas unidades que se encuentran a una distancia inferior a un determinado umbral, definido por un radio d^* . También es posible usar este criterio cuando se cuenta con datos de áreas, en este caso la distancia entre polígonos se calculará utilizando los centroides de los polígonos. Dependiendo de la aplicación empírica, es posible cambiar el centroide por otro punto más relevante del polígono. Por ejemplo, si se trabaja con información de países puede usarse la ubicación de cada ciudad capital.

El problema con los criterios de contigüidad y de distancia por umbral es que pueden dar lugar a regiones aisladas que no tienen vecinos. Esto puede suceder debido a que la densidad de los puntos no es regular en el mapa o, bajo contigüidad, cuando se trabaja con mapas que contienen islas que naturalmente no comparten límites con otras regiones.

En el caso de contigüidad, el problema de regiones sin vecinos puede salvarse asignando

¹En series temporales, un proceso autorregresivo de orden 1, AR(1), puede representarse como $\mathbf{y} = \alpha \mathbf{L} \mathbf{y} + \mathbf{u}$, siendo \mathbf{L} el operador de rezagos temporales. La diferencia entre \mathbf{L} y \mathbf{W} radica en la multidireccionalidad espacial y esto genera que las técnicas estadísticas temporales no tengan las propiedades deseadas en el espacio.

manualmente vecinos a estas unidades. Bajo el criterio de distancia por umbral, puede determinarse un umbral cuya amplitud asegure que cada observación contenga al menos un vecino. Este criterio es conocido como max-min, el umbral será la máxima distancia del vecino más cercano de todas las unidades. Un umbral elevado puede ocasionar que existan regiones con una excesiva cantidad de vecinos dando lugar a un gran número de conexiones irrelevantes.

Otro criterio de uso habitual es el de k vecinos más cercanos. Utilizando latitud y longitud, se va eligiendo como vecino al punto más cercano hasta obtener el número de vecinos establecidos, k . Bajo este criterio todas las regiones poseerán la misma cantidad de vecinos evitando el problema de unidades aisladas o unidades con excesiva cantidad de vecinos.

Una vez definido el criterio de vecindad, resta operacionalizar cómo los individuos se afectan entre sí. Una forma simple es utilizando una elección binaria, con $w_{ij} = 1$ cuando i y j cumplen con el criterio de vecindad elegido y $w_{i,j} = 0$ cuando no. Si el criterio utilizado es distancia, los pesos binarios pueden reemplazarse por alguna función de distancia entre unidades espaciales. Por ejemplo, utilizando la función la inversa $w_{ij} = f(d_{ij}) = 1/d_{ij}$, tal que a mayor distancia menor relación entre i y j , cumpliendo de esta forma con la ley de Tobler. La literatura aplicada también ha combinado los criterios de distancia o contigüidad con otra información geográfica como el perímetro del polígono, la extensión de la frontera en común, y otras características geográficas.

Otras especificaciones de pesos espaciales combinan información socio-económica y con información geográfica. Continuando con el ejemplo de distancia podría definirse cada peso como $w_{ij} = [|z_i - z_j| + 1]^{-1} \times 1/d_{ij}$, donde z es una variable que captura una distancia socio-económica (PBI, flujo comercial, etc.). En otras aplicaciones, directamente se aplica un concepto de distancia socio-económica: $w_{ij} = [|z_i - z_j| + 1]^{-1}$ con z siendo el ingreso per cápita o el porcentaje de población con alguna característica (tasa de pobreza, porcentaje de algún grupo étnico). El problema con la distancia socio-económica es que la exogeneidad del peso es dudosa en numerosas ocasiones.

Definidos los pesos espaciales, lo habitual es trabajar con alguna transformación de la matriz debido a que mejora las propiedades estadísticas de los estimadores. La transformación más utilizada es la normalización por fila, en donde los nuevos pesos se obtienen como $w_{ij}^* = w_{ij}/\sum_j w_{ij}$, tal que la suma de cada fila de W será sea igual a la unidad: $\sum_j w_{ij}^* = 1$. Esta transformación tiene implicancias a la hora de interpretar el coeficiente espacial ρ : medirá el impacto promedio de los vecinos sobre una región. Otra opción es normalizar por columna, $w_{ij}^* = w_{ij}/\sum_i w_{ij}$, en cuyo caso el coeficiente espacial capturará el impacto promedio de la región sobre los vecinos.

Respecto al orden de vecindad, puede ser de interés conocer el impacto de los vecinos de los vecinos. Por ejemplo, bajo el criterio de distancia por umbral, el vecindario de segundo orden estará compuesto por las regiones que se encuentren a una distancia $d^* < d_{ij} \leq 2d^*$. Para estos casos, los ordenes superiores a 1 se identifican con un supra-índice: $\mathbf{W}^{(j)}$, $\forall j \geq 2$, siendo j el

orden de vecindad.

Por lo general la elección del criterio de vecindad es un a priori del investigador. Aunque existen un conjunto de lineamientos útiles para la elección de la matriz establecidos por Griffith (1996): (i) Es mejor utilizar una especificación razonable de la matriz de pesos geográficos que considerar todas las conexiones nulas; (ii) modelos con bajos ordenes deben ser preferidos por sobre los modelos de altos ordenes; (iii) en general, es mejor emplear una matriz de pesos sub-identificada que una sobre-identificada.

Sin importar cuál sea el caso empírico, el desarrollo econométrico supone que \mathbf{W} contiene elementos estrictamente exógenos a la especificación econométrica y la elección de los vecinos no debería realizarse mediante variables consideradas en el modelo analizado, ya sea relacionadas a la variable dependiente o a las variables explicativas.

3. Especificaciones espaciales alternativas

La creación de la matriz de pesos espaciales brinda la oportunidad de contrastar e incluir nuevas variables al modelo econométrico. A nivel empírico, la inclusión de estos efectos espaciales ha sido dirigida por dos tipos de enfoques. Por un lado, los modelos espaciales pueden especificarse siguiendo una teoría sobre la interacción entre los agentes, enfoque conocido como *theory-driven*. Hay numerosas investigaciones que utilizan teorías bien establecidas e incorporan externalidades en forma de *spillovers* espaciales. En contraposición a este tipo de enfoque, gran parte de la literatura aplicada ha sido típicamente dirigida por problemas que no tienen un sustento teórico formal. Este enfoque, denominado *data-driven*, encuentra justificación por problemas ya mencionados: los datos no se encuentran disponibles en la misma escala y localización al proceso espacial de interés, existen errores de medición u omisión de variables que generan dependencia espacial. El enfoque *data-driven* busca obtener estimadores consistentes, insesgados y eficientes controlando por la estructura espacial presente en los datos. En esta sección exploraremos las alternativas de especificación que han sido utilizadas bajo ambos enfoques.

3.1. Modelos espaciales de corte transversal

Un punto de partida muy utilizado en investigaciones empíricas es un modelo no-espacial, usualmente estimado por mínimos cuadrados ordinarios (*MCO*). Este modelo lineal no-espacial es un modelo lineal general con la siguiente especificación:

$$\mathbf{y} = \alpha\boldsymbol{\iota} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad (5)$$

siendo la variable dependiente \mathbf{y} un vector de dimensión $(n \times 1)$, $\boldsymbol{\iota}$ es un vector $(n \times 1)$ de unos asociados al término constante α , \mathbf{X} es una matriz de variables explicativas de orden $(n \times K)$,

β es un vector de parámetros desconocidos de orden $(K \times 1)$ y $\mathbf{u} = (u_1, u_2, \dots, u_n)'$ es un vector de términos aleatorios de dimensión $(n \times 1)$, donde u_i es independiente e idénticamente distribuido para todo i con media cero y varianza constante σ^2 . Se supone que el modelo cumple con los supuestos habituales en la parte sistemática.

La detección de efectos espaciales en la ecuación (5) permite introducir efectos espaciales de diversas formas, tal como puede observarse en la Figura 1. Un recorrido habitual es introducir efectos espaciales secuencialmente usando como punto de partida el modelo lineal, estrategia “de lo particular a lo general” (*Stge, Specific-to-general*). Para aplicar esta estrategia puede utilizarse una batería de tests estadísticos de multiplicadores de Lagrange, *LM*, simples y robustos, que permiten detectar las posibles estructuras espaciales omitidas (consúltese Anselin et al., 1996, para mayor detalle). En muchas aplicaciones, previo al desarrollo de la estrategia *Stge*, se realiza un análisis exploratorio de datos espaciales (mapas, detección de agrupamientos, test espaciales univariados, detección de *outliers* espaciales, etc.) que permite establecer conjeturas sobre la importancia del espacio, generalmente bajo el enfoque *data-driven*. Anselin (2005) muestra cómo realizar dicho análisis con el programa GeoDa, Bivand (2010) ofrece similares alternativas bajo R y Herrera (2015) presenta las opciones disponibles en Stata.

Otra estrategia de especificación, una vez detectada la importancia del espacio, es comenzar por un modelo espacial lo más complejo posible, estrategia “de lo general a la particular” (*Gets, General-to-specific*) (Hendry, 1979). En la literatura espacial el modelo más complejo es conocido como modelo general de Cliff-Ord (modelo 1, Figura 1):

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \alpha \mathbf{1} + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \mathbf{u}, \\ \mathbf{u} &= \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \end{aligned} \tag{6}$$

donde $\mathbf{W}\mathbf{y}$ representa el efecto de interacción espacial endógeno o efecto espacial sustantivo, $\mathbf{W}\mathbf{X}$ son los efectos de interacción exógenos y $\mathbf{W}\mathbf{u}$ es el efecto de interacción espacial en el término de error o efecto espacial residual. El parámetro ρ es denominado coeficiente de rezago espacial, λ es el coeficiente de autocorrelación espacial y $\boldsymbol{\theta}$ son los coeficientes espaciales de dependencia local.

La ecuación (6) posee una estructura similar al conocido modelo de Manski (1993) de efectos sociales. En el modelo de Manski las interacciones sociales son especificadas como el resultado esperado de un equilibrio social. La similitud entre ambos modelos ha sido destacada por Lee (2004, 2007) entre otros. La especificación del modelo es:

$$\begin{aligned} \mathbf{y} &= \rho_m E[\mathbf{y}|\mathbf{a}] + \alpha \mathbf{1} + \mathbf{X}\beta + E[\mathbf{X}|\mathbf{a}]\boldsymbol{\theta} + \mathbf{u}, \\ \mathbf{u} &= \lambda_m E[\mathbf{u}|\mathbf{a}] + \boldsymbol{\varepsilon}, \end{aligned} \tag{7}$$

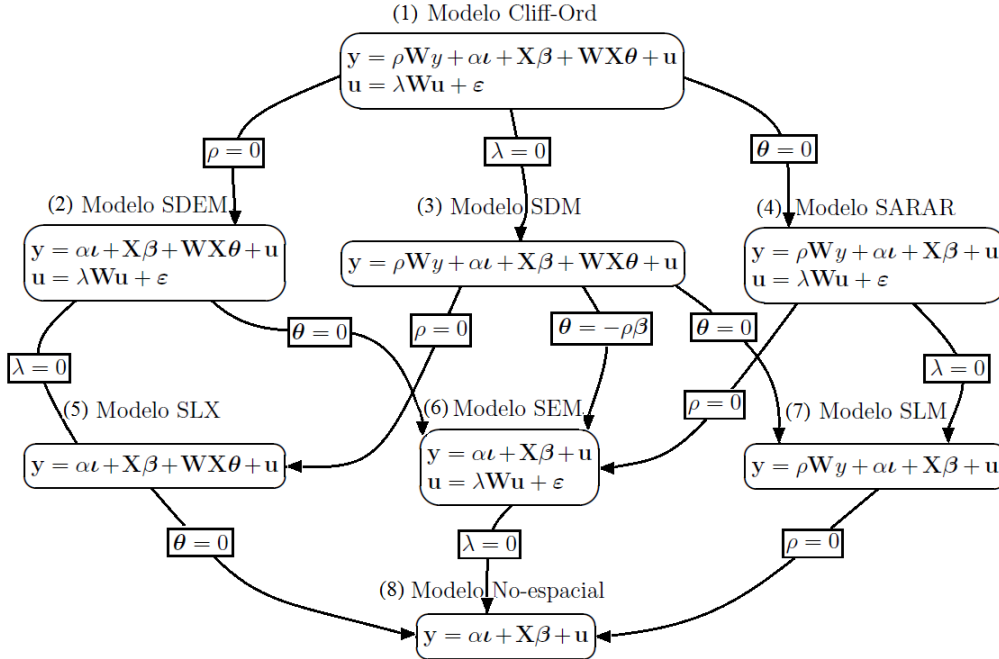
donde \mathbf{a} indexa unidades no-solapadas de vecinos. En la literatura de redes sociales, ρ_m cap-

tura el efecto endógeno del resultado esperado de los vecinos sobre la variable dependiente, θ captura el efecto contextual o exógeno de las características grupales y λ_m captura el efecto correlacionado de características no observables. Luego de algunas manipulaciones, la forma reducida del modelo puede escribirse como:

$$\mathbf{y} = \alpha\mathbf{1} + \mathbf{X}\boldsymbol{\beta} + E[\mathbf{X}|\mathbf{a}] \frac{(\rho_m\boldsymbol{\beta} + \boldsymbol{\theta})}{1 - \rho_m} + \frac{\rho_m}{1 - \rho_m} E[\mathbf{u}|\mathbf{a}] + \mathbf{u}, \quad (8)$$

y esta última ecuación permite visualizar el problema de identificación planteado por Manski, “problema de reflexión”: no es posible identificar el efecto endógeno del efecto exógeno a menos que se impongan restricciones adicionales. Los modelos espaciales se diferencian de esta literatura en el sentido que no se busca estimar el efecto “vecindario” ya que cada unidad observada se encuentra conectada por una estructura que se supone conocida por medio de la especificación de \mathbf{W} .² Salvando esta diferencia, el modelo general de Cliff-Ord presenta similares problemas de identificación imposibilitando su uso en la estrategia *Gets*. Es decir, el modelo espacial más complejo solo puede incorporar dos efectos espaciales de los tres posibles (ρ , λ o θ) necesitando imponer restricciones en el efecto restante.

Figura 1: Modelos espaciales alternativos en corte transversal



²Aquí también puede destacarse la centralidad de la matriz de contactos como aquella que refleja la exactitud de las verdaderas relaciones entre las unidades espaciales y todos los resultados inferenciales serán condicionales a la calidad de esta matriz.

Las especificaciones alternativas según la restricción sobre el modelo de Cliff-Ord se presentan en la Figura 1, modelos (2), (3) y (4). Si se impone la restricción $\boldsymbol{\theta} = \mathbf{0}$, el modelo resultante es denominado *SARAR* (*Spatial AutoRegressive with AutoRegressive error*), con dependencia espacial y sustantiva, impulsado principalmente por Kelejian y Prucha (1998) en sus desarrollos teóricos. Este modelo es frecuentemente utilizado en teoría econométrica ya que contiene los dos términos espaciales más complejos de estimar. Sin embargo, a nivel empírico ha sido escasamente utilizado ya que proporciona cierta protección contra la falta de especificación en una parte del modelo pero contamina las estimaciones de otras partes (Pace y Zhu, 2012).

LeSage y Pace (2009) proponen al modelo espacial de Durbin (*SDM*, *Spatial Durbin Model*) como la mejor alternativa de especificación general (modelo 3, Figura 1):

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \alpha \boldsymbol{\iota} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \mathbf{u}. \quad (9)$$

Los motivos que llevan a esta elección son comprensibles: (1) la omisión de un término relevante en el error solo genera pérdida de eficiencia; pero (2) los costos de ignorar efectos espaciales en la variable dependiente y/o la omisión de variables explicativas espaciales relevantes generan estimadores sesgados e inconsistente de los coeficientes restantes. Elhorst (2014b) reconoce la ventaja del modelo *SDM*, sin embargo, no descarta al modelo *SDEM* (*Spatial Durbin Error Model*) como una especificación de interés y competitiva en términos predictivos (modelo 2, Figura 1). Los análisis comparativos entre los modelos *SDM* y *SDEM* muestran que son prácticamente indistinguibles cuando no existen motivos teóricos para la incorporación de un rezago espacial endógeno.

Es interesante mencionar que el modelo *SDM* anida la mayor cantidad de modelos: si $\boldsymbol{\theta} = \mathbf{0}$ el modelo se reduce a un *SLM* (*Spatial Lag Model*), si $\rho = 0$ el modelo alcanzado es el *SLX* (*Spatial Lag in X's*), y mediante una combinación no-lineal ($\boldsymbol{\theta} = -\rho\boldsymbol{\beta}$) el *SDM* puede reducirse a un modelo *SEM* (*Spatial Error Model*). Este último resultado se alcanza reemplazando $\boldsymbol{\theta} = -\rho\boldsymbol{\beta}$ en la ecuación (9):

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \alpha \boldsymbol{\iota} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}(-\rho\boldsymbol{\beta}) + \mathbf{u}.$$

Reagrupando términos en la última ecuación y reemplazando la notación ρ por λ se obtiene el modelo *SEM*:

$$\mathbf{y} = \alpha^* \boldsymbol{\iota} + \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda \mathbf{W})^{-1} \boldsymbol{\varepsilon}.$$

Debido a la dificultad de estimar modelo más complejos, en la primera etapa de desarrollo de econometría espacial los modelos insignia eran el *SLM* y el *SEM*. El modelo de rezago espacial tuvo una presencia preponderante en esta etapa inicial por dos razones. Por una parte, bajo un esquema económico de teoría de juegos entre agentes, con funciones de pago cuadráticas en sus

variables de elección, el *SLM* puede ser visto como la condición de primer orden o función de reacción que surge del problema de optimización simultáneo de n individuos (Brueckner, 2003). Otra razón estaba justificada por cuestiones de estimación, la inclusión de un término espacial endógeno evita la inconsistencia del estimador y, además, el modelo de rezago espacial obtiene una estimación del *spillover* espacial que es muy atractiva a nivel empírico. El otro modelo con buena presencia en los inicios fue el modelo *SEM* que permite una ganancia de eficiencia respecto al modelo no-espacial. Una crítica empírica al modelo *SEM* sucede cuando la autocorrelación espacial es débil: los resultados tienden a ser similares a los obtenidos por una estimación no-espacial robusta.

El modelo espacial más simple en términos de estimación es el *SLX*:

$$\mathbf{y} = \alpha\mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \mathbf{u}, \quad (10)$$

donde $\boldsymbol{\theta}$ captura el efecto espacial de las variables explicativas de los vecinos a cada observación. Estos efectos son conocidos como contextuales en la literatura de modelos sociales y tienen un impacto local.³

Dentro de la literatura, el *SLX* no ha sido muy utilizado en aplicaciones empíricas y menos aún en trabajos teóricos. La falta de interés en econometría teórica proviene del hecho que no ofrece dificultad en su estimación, el modelo puede estimarse por *MCO*. Pero en cuestiones aplicadas Halleck Vega y Elhorst (2015) aconsejan no descartar los resultados que ofrece este modelo.

3.2. Modelos espaciales en datos de panel

La taxonomía de los modelos presentados en la Figura 1 puede extenderse fácilmente a modelos espacio-temporales estáticos. Para pasar de información de corte transversal a información de datos de panel solo es necesario incorporar un sub-índice t , que va de $t = 1, \dots, T$, apilando las observaciones del corte transversal periodo a periodo. Sin modificar ningún supuesto en el término de error, este tipo de extensión espacio-temporal da lugar a los modelos *pooled*. Por ejemplo, el modelo *pooled* de Cliff-Ord puede escribirse como:

$$\begin{aligned} \mathbf{y}_t &= \rho\mathbf{W}\mathbf{y}_t + \alpha\mathbf{1} + \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{X}_t\boldsymbol{\theta} + \mathbf{u}_t, \\ \mathbf{u}_t &= \lambda\mathbf{W}\mathbf{u}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T, \end{aligned} \quad (11)$$

³Este modelo es muy utilizado en análisis de efectos pares o contextuales. Por ejemplo, la relación entre rendimiento académico (\mathbf{r}) y nivel socio-económico (\mathbf{nes}) individual y contextual puede especificarse como $\mathbf{r} = \alpha\mathbf{1} + \mathbf{nes}\boldsymbol{\beta}_1 + \overline{\mathbf{nes}}\boldsymbol{\beta}_2 + \mathbf{u}$, en donde $\overline{\mathbf{nes}} = \mathbf{W} \times \mathbf{nes}$ es el efecto contextual o rezago espacial del nivel socio-económico, siendo \mathbf{W} construida en función de pertenencia a la misma escuela.

donde las unidades de corte transversal ($i = 1, \dots, n$) están apiladas por periodo temporal t y el término ε_t es un vector de dimensión n de errores aleatorios *i.i.d.*

En econometría espacial se trabaja habitualmente con paneles micro, es decir, con un número amplio de unidades de corte transversal y un número reducido de periodos temporales. Por otra parte, las técnicas espaciales requieren que los datos de panel se encuentren balanceados en el sentido que las mismas unidades espaciales deben ser observadas en cada uno de los periodos. Esto plantea un problema adicional ya que los datos de panel tienden a sufrir el fenómeno de desgaste muestral, es decir, pérdida información entre diferentes periodos, ocasionando paneles desbalanceados.

Las críticas a los modelos *pooled* es que omiten posibles fuentes de heterogeneidad. Por ejemplo, no consideran particularidades de cada región que pueden ser específicas e invariantes en el tiempo. Características como religión, normas jurídicas o estructura laboral no muestran variabilidad temporal y su omisión aumenta el riesgo de sesgo en las estimaciones. Una posible solución es incorporar en los modelos un término $\boldsymbol{\mu}$, de dimensión n , que capture el efecto específico individual de las regiones. Otra fuente de heterogeneidad puede provenir de particularidades específicas entre periodos temporales, como un año de crisis o recesión. Estos eventos causarán diferencias entre periodos temporales que posiblemente sesgarán las estimaciones. Como solución a esto puede añadirse un término $\boldsymbol{\phi}_t$ que controle por efectos temporales específicos.

Al incorporar los efectos específicos, individuales y temporales, el modelo en 11 es re-expresado como:

$$\begin{aligned} \mathbf{y}_t &= \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \mathbf{u}_t, \\ \mathbf{u}_t &= \lambda \mathbf{W} \mathbf{u}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T. \end{aligned} \tag{12}$$

Los efectos específicos pueden ser considerados como efectos fijos o aleatorios. En el caso un modelo de efectos fijos, una variable ficticia es introducida por cada unidad espacial (sin considerar a la unidad base) y otra variable variable ficticia es utilizada para capturar el efecto temporal específico.

En el caso de los modelos con efectos aleatorios, debe especificarse la distribución de estos efectos, como por ejemplo: $\boldsymbol{\mu} \sim \mathcal{N}(0, \sigma_{\boldsymbol{\mu}}^2 \mathbf{I}_n)$, $\boldsymbol{\phi}_t \sim \mathcal{N}(0, \sigma_{\boldsymbol{\phi}}^2)$. Entonces, la especificación en el caso de considerar un *SDM* con efectos aleatorios será:

$$\begin{aligned} \mathbf{y}_t &= \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta} + \underbrace{\boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t}_{\mathbf{u}_t}, \\ \boldsymbol{\varepsilon}_t &\sim \mathcal{N}[0, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_n], \quad \boldsymbol{\mu} \sim \mathcal{N}[0, \sigma_{\boldsymbol{\mu}}^2 \mathbf{I}_n], \quad \boldsymbol{\phi}_t \sim \mathcal{N}(0, \sigma_{\boldsymbol{\phi}}^2). \end{aligned}$$

En el caso de considerar un *SDEM* con efectos aleatorios, la especificación vendrá dada

por:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{X}_t\boldsymbol{\theta} + \underbrace{\boldsymbol{\mu} + \boldsymbol{\phi}_t\boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t}_{\mathbf{u}_t}, \\ \boldsymbol{\varepsilon}_t &= \lambda\mathbf{W}\boldsymbol{\varepsilon}_t + \boldsymbol{\eta}_t, \\ \boldsymbol{\eta}_t &\sim \mathcal{N}[0, \sigma_{\boldsymbol{\eta}}^2\mathbf{I}_n], \quad \boldsymbol{\mu} \sim \mathcal{N}[0, \sigma_{\boldsymbol{\mu}}^2\mathbf{I}_n], \quad \boldsymbol{\phi}_t \sim \mathcal{N}(0, \sigma_{\boldsymbol{\phi}}^2). \end{aligned}$$

Estas versiones aleatorias suponen que el término remanente de error $\boldsymbol{\varepsilon}_t$ posee estructura espacial. Otras especificaciones modifican este supuesto tal que el mismo proceso espacial aplica tanto a los efectos individuales como al error remanente (ver Kapoor et al., 2007).

Respecto a los estadísticos para detectar dependencia espacial existen diversos desarrollos. La presencia de estructura espacial residual y sustantiva en modelos *pooled* puede ser detectada usando las versiones de los tests *LM* desarrolladas por Anselin et al. (2008) adaptadas a la estructura espacio-temporal. También existen versiones de los tests *LM* para modelos no-espaciales de efectos fijos (Debarsy y Ertur, 2010) y aleatorios (Baltagi et al., 2003).

Alternativamente, pueden utilizarse otros tests generales que permiten detectar la presencia de dependencia espacial. Recuérdese que, bajo un corte transversal, la dependencia espacial ha sido reducida a la estimación de un parámetro debido a la insuficiente información para estimar la correlación entre pares de observaciones. Con datos de panel, esta limitación es reducida dependiendo de la longitud de la dimensión temporal y es posible obtener estimaciones de la correlación entre las observaciones del corte transversal. Un estadístico que detecta dependencia de corte transversal es el CD de Pesaran (2004). Este estadístico puede ser aplicado a una amplia variedad de paneles, incluyendo dinámicos y no-estacionarios, ya sea bajo efectos fijos o aleatorios.

La literatura rápidamente ha ampliado el análisis de paneles estáticos hacia paneles dinámicos. La incorporación de dinámica temporal genera un modelo básico, que en su forma vectorial puede escribirse, para un periodo t :

$$\mathbf{y}_t = \tau\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t\boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (13)$$

con $\boldsymbol{\varepsilon}_t \sim i.i.d(0, \sigma_{\boldsymbol{\varepsilon}}^2)$, con \mathbf{y}_t e \mathbf{y}_{t-1} de orden $n \times 1$ y efectos específicos individuales $\boldsymbol{\mu}' = [\mu_1, \mu_2, \dots, \mu_n]$ y temporales $\boldsymbol{\phi}_t$, con $t = 1, \dots, T$.

La ecuación (13) permite capturar la persistencia temporal del proceso, incluyendo la heterogeneidad entre las regiones y entre periodos. Igual que en los modelos estáticos, los efectos específicos pueden ser tratados como efectos fijos o aleatorios.

Tal como sucedía en los modelos estáticos, los efectos espaciales pueden introducirse en la ecuación (13) en diferentes formas. Anselin et al. (2008) propone un conjunto de especificaciones:

- Modelo simultáneo espacio-temporal:

$$\mathbf{y}_t = \tau \mathbf{y}_{t-1} + \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (14)$$

- Modelo recursivo espacial puro:

$$\mathbf{y}_t = \gamma \mathbf{W} \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (15)$$

- Modelo recursivo espacio-temporal:

$$\mathbf{y}_t = \tau \mathbf{y}_{t-1} + \gamma \mathbf{W} \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (16)$$

- Modelo dinámico espacio-temporal:

$$\mathbf{y}_t = \tau \mathbf{y}_{t-1} + \gamma \mathbf{W} \mathbf{y}_{t-1} + \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (17)$$

El modelo dinámico espacio-temporal incorpora todas las alternativas de rezagos espacio-temporales de la variable dependiente. Las características de este modelo fueron analizadas por Yu et al. (2008). El modelo recursivo espacio-temporal, ecuación (16), ha sido aplicado en cuestiones de predicciones espacio-temporales (por ejemplo, véase Giacomini y Granger, 2004).

Los modelos previamente mencionados son tipos de modelos espacio-temporales *SLM* (*dynSLM*) ya que solo incorporan elementos dinámicos espacio-temporales endógenos. Esta clasificación puede ampliarse hacia modelos espacio-temporales de Durbin (*dynSDM*) en cuatro versiones:

$$\mathbf{y}_t = \tau \mathbf{y}_{t-1} + \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (18)$$

$$\mathbf{y}_t = \gamma \mathbf{W} \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (19)$$

$$\mathbf{y}_t = \tau \mathbf{y}_{t-1} + \gamma \mathbf{W} \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (20)$$

$$\mathbf{y}_t = \tau \mathbf{y}_{t-1} + \gamma \mathbf{W} \mathbf{y}_{t-1} + \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta} + \boldsymbol{\mu} + \boldsymbol{\phi}_t \boldsymbol{\nu}_n + \boldsymbol{\varepsilon}_t, \quad (21)$$

en donde se han introducido efectos espaciales locales, $\mathbf{W} \mathbf{X}_t$, además de los efectos ya presentes en los modelos previos.

Revisiones relativamente recientes de estos modelos pueden encontrarse en Anselin et al. (2008), Lee y Yu (2010), Elhorst (2010b) y Lee y Yu (2014).

3.3. Cuestiones inferenciales y asintóticas

La estimación de los modelos espaciales no puede realizarse por métodos simples. La inclusión de variables $\mathbf{W} \mathbf{y}$ o $\mathbf{W} \mathbf{u}$ provocan que el estimador *MCO* sea inconsistente y/o ineficiente,

dependiendo del caso (Anselin, 1988, pp. 57-59). Como consecuencia, la literatura ha desarrollado tres métodos de estimación para modelos de efectos espaciales. El método más utilizado desde los inicios de la disciplina fue la estimación por máxima verosimilitud, *ML*. Otra alternativa de gran desarrollo en los 90's está basada en variables instrumentales o método de momentos generalizados (*IV/GMM*), y también se ha desarrollado la estimación bayesiana (LeSage y Pace, 2009, cap. 5).

La estimación de estos modelos presenta una serie de dificultades relacionadas a los parámetros espaciales y al comportamiento de las matrices \mathbf{W} , $(\mathbf{I} - \rho\mathbf{W})^{-1}$ y/o $(\mathbf{I} - \lambda\mathbf{W})^{-1}$ presentes en todos los métodos de estimación. Una de ellas es la necesidad de imponer alguna restricción sobre el rango de valores que puede alcanzar ρ , lo mismo se aplica para λ . Sin restricciones sobre este parámetro, el espacio paramétrico da lugar a discontinuidades y no asegura la invertibilidad de la matriz $(\mathbf{I} - \rho\mathbf{W})$. Para evitar este problema y también facilitar la estimación de ρ , Ord (1981) propuso restringir los posibles valores del parámetro en un rango $1/r_{min} < \rho < 1/r_{max}$, donde r_{min} y r_{max} son los autovalores, mínimo y máximo, de la matriz \mathbf{W} . Bajo una matriz espacial estandarizada por filas automáticamente se impone una restricción sobre el máximo autovalor tal que ρ se encontrará restringido entre los valores $1/r_{min} < \rho < 1$. Algunos autores como Kelejian y Robinson (1995) han sugerido limitar el espacio paramétrico de ρ al intervalo $(-1, 1)$ en similitud al tratamiento que se realiza en series temporales, sin embargo este rango de valores puede ser demasiado restrictivo para el caso espacial.

En el caso de los modelos dinámicos espacio-temporales también deben cumplirse con ciertas restricciones de estacionariedad. Como Elhorst (2014b) explica, la estimación de un modelo dinámico que incorpore los términos \mathbf{y}_{t-1} , $\mathbf{W}\mathbf{y}_{t-1}$, $\mathbf{W}\mathbf{y}_t$ será estable si se cumplen las condiciones: $|\tau| < 1$ y $|\tau| + |\gamma| + |\rho| < 1$. En situaciones en donde esta condición no se cumpla, el modelo estará espacialmente cointegrado y una transformación del modelo será necesaria para su estimación (Yu et al., 2012).

En el caso de la estimación por *ML*, una dificultad adicional se encuentra en la maximización de la función de log-verosimilitud que involucra al determinante de una matriz de dimensión $n \times n$, $\ln |\mathbf{I} - \rho\mathbf{W}|$. Este determinante debe ser evaluado repetidamente para cada valor estimado de ρ y el procedimiento puede ser problemático para grandes bases de datos. En la era de *big data*, donde puede tenerse millones de observaciones la precisión de este determinante es clave. Una solución propuesta por Ord (1975) es estimar dicho término mediante la siguiente transformación:

$$\ln |\mathbf{I} - \rho\mathbf{W}| = \sum_{i=1}^n \ln (1 - \rho r_i), \quad (22)$$

siendo r_i los autovalores de la matriz \mathbf{W} , que reduce la complejidad pero puede ser problemática en tiempo computacional para grandes bases de datos.

Respecto a la estimación por *IV/GMM*, los principales avances de este método fueron realizados por Kelejian y Prucha (1998, 1999), con generalizaciones realizadas por Kelejian y Prucha (2010), Arraiz et al. (2010) y Drukker et al. (2013). Las condiciones de momentos no incluyen la estimación del determinante, ecuación (22), estableciendo ventajas computacionales claras y notorias si nos encontramos con grandes bases de datos. Otra virtud de este método es la posibilidad de instrumentalizar otras variables explicativas endógenas, además de la endogeneidad espacial. Este es un punto en donde *ML* o estimación bayesiana no presentan ninguna solución.

Una potencial desventaja de la estimación por *IV/GMM* es que no impone ninguna restricción sobre el espacio paramétrico en el proceso de estimación. Por lo tanto puede suceder que el valor estimado del coeficiente espacial (ρ o λ) se encuentre fuera del rango de estabilidad dando lugar a estimaciones no-interpretables.

En el caso de la estimación de los paneles espaciales dinámicos (estacionarios), hay diferentes alternativas que extienden los procedimientos de máxima verosimilitud condicional y no-condicional, así como existen procedimientos alternativos por *IV/GMM*. Elhorst (2010a) muestra diferentes procedimientos basados en *MV* y Yu et al. (2008) desarrollan un estimador de cuasi-máxima verosimilitud que muestra propiedades comparables a los estimadores *MV*.

Respecto a las propiedades asintóticas de los estimadores, en econometría estándar suponer que $n \rightarrow \infty$ no incluye ningún tipo de información adicional. Pero en el espacio, el incremento del tamaño muestral necesita especificar de qué forma es expandido el tamaño muestral en el espacio geográfico. Elhorst (2014b, pp. 55) comenta que existen dos tipos de asintótica comúnmente utilizadas en el contexto espacial. Una de ellas es la estructura “infill”, donde las observaciones espaciales se encuentran delimitadas en el espacio geográfico a medida que $n \rightarrow \infty$. Otra estructura es “increasing domain”, donde los límites geográficos de las regiones se expanden cuando $n \rightarrow \infty$. Por otra parte, hay dos tipos de diseño muestral en el espacio: (a) diseño estocástico donde las unidades espaciales son aleatoriamente seleccionadas; y (b) diseño fijo donde las unidades espaciales se encuentran relacionadas a un espacio no aleatorio, posiblemente irregular.

La literatura de econometría espacial se enfoca principalmente en la estructura asintótica “increasing domain” bajo un diseño muestral fijo esto es debido a que mayoritariamente se utilizan datos areales. Por lo tanto, una posición con mayor respaldo ha sido la estimación por efectos fijos, dado que el espacio es usualmente muestreado en forma exhaustiva y la inferencia realizada puede considerarse como poblacional. De todas formas, hay argumentos para realizar estimaciones por efectos aleatorios ya que no todos los estudios utilizan de forma exhaustiva el espacio geográfico y, por otra parte, los datos de panel para un intervalo temporal pueden considerarse como una muestra representativa de una superpoblación (concepto relacionado a un proceso generador de datos espacio-temporal, ver Graubard y Korn, 2002, y referencias allí citadas para mayor detalle).

Las propiedades de los estimadores dependen crucialmente de determinadas condiciones de regularidad que debe cumplir la matriz de pesos espaciales. Un requisito mencionado por Kelejian y Prucha (1998) establece que la sumatoria de las columnas y filas de las matrices \mathbf{W} , $(I - \rho\mathbf{W})^{-1}$, antes de normalizar \mathbf{W} , debe estar uniformemente limitada en valor absoluto cuando n tiende a infinito. Además, Lee (2004) establece que la sumatoria de filas y columnas de \mathbf{W} , antes de estandarizar, no debe divergir a infinito a una velocidad igual o mayor que la tasa a la que crece el tamaño muestral n . El cumplimiento de al menos una de las anteriores condiciones permiten que la correlación entre dos unidades espaciales tienda a cero cuando la distancia que las separa tiende a infinito y de esta forma limitan la correlación a un rango de valores manejables. Por ejemplo, si se trabaja con el criterio contigüidad la primera condición es satisfecha automáticamente. En cambio, si el criterio utilizado es la inversa de la distancia la primera condición no es satisfecha pero si satisface la condición que establece Lee (2004).

Las anteriores condiciones establecen situaciones límites a la conectividad de la matriz de pesos. Pero aún cuando se cumplan, las propiedades finitas de los estimadores pueden verse afectadas. Smith (2009) muestra que matrices espaciales fuertemente conectadas (medidas por el número medio de vecinos), con una densidad media de vínculos de 30% o superior, generan un sesgo descendente en los parámetros espaciales bajo *ML*. En todos los casos, este sesgo implica que la autocorrelación espacial positiva puede no ser detectada debido a la presencia de matrices densamente conectadas. Por lo general, la conectividad de la matriz por criterios geográficos son poco densas, no superando el 10% de los vínculos posibles. El problema puede ser importante cuando la matriz es construida por criterios de redes sociales o no-espacial.

4. Interpretación de los modelos con rezago espacial endógeno

Un punto de interés, válido para los modelos de corte transversal y de datos de panel, es la interpretación de los resultados obtenidos. Por ejemplo, en los inicios de la especialidad el modelo *SLM* fue adoptado como un modelo central debido a que brinda una estimación de la interdependencia espacial. La interpretación de los resultados obtenidos era directa, en el sentido que un valor significativo del parámetro ρ era evidencia de la importancia de los efectos desbordamiento o contagio. Sin embargo, LeSage y Pace (2009) destacan la errónea interpretación de estos resultados y muestran la necesidad de considerar derivadas parciales del impacto de un cambio en una unidad de las variables explicativas.

Para ver más claramente el punto destacado por LeSage y Pace (2009), debe tenerse presente que los efectos desbordamiento en un modelo *SDM*, por ejemplo, se deben a la presencia de $\mathbf{W}\mathbf{y}$ (efecto global) y $\mathbf{W}\mathbf{X}$ (efecto local):

$$\mathbf{y}_t = \rho\mathbf{W}\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{X}_t\boldsymbol{\theta} + \mathbf{u}_t. \quad (23)$$

Supongamos que se desea analizar el impacto del incremento unitario de la k –ésima variable explicativa, este incremento es considerado igual para todas las localizaciones. Bajo un modelo de regresión no espacial, el efecto total sobre la variable dependiente será igual a la estimación del coeficiente β_k , *ceteris paribus*, sin importar la localización. Pero ante un modelo como el presentado en (23), el efecto total depende de las unidades vecinas de cada localización y de la magnitud de los coeficientes que acompañan a las variables espaciales.

El efecto total puede ser descompuesto en un efecto directo y otro indirecto. El efecto directo es el impacto del cambio de la variable explicativa sobre la variable dependiente en cada localidad, este efecto tenderá a ser similar al obtenido por un modelo de regresión no espacial si ρ y θ_k son cercanos a 0. El efecto indirecto se debe a la dinámica espacial generada por la presencia de ρ y θ_k que afectará a todas las unidades el modelo.

Bajo un modelo de Durbin pueden distinguirse dos tipos de efectos indirectos producto de la interdependencia entre las unidades. Uno es de tipo global, afectando a todas las unidades: y_i aumenta inicialmente β_{ik} unidades, y esto genera un nuevo incremento producto del cambio en las y'_j s de los S vecinos, capturado por el término endógeno $\rho W y$. Es decir, para cada i –ésima localidad existirá un nuevo impacto igual a $\rho \sum_{j \neq i}^S w_{ij} \beta_{jk}$ unidades, retro-alimentando el cambio en la variable dependiente y que solo cesará al converger (suponiendo $|\rho| < 1$). Este efecto es identificado como un *efecto espacial global* para distinguirlo del *efecto espacial local* que surge del incremento unitario en la k –ésima variable explicativa de los S vecinos de la localidad i : $\sum_{j \neq i}^S w_{ij} \theta_{jk}$. Este efecto espacial refuerza al efecto inicial β_k y es de tipo local en el sentido que no posee un efecto dinamizador como el generado por la presencia de ρ .

En términos matriciales, y de forma más compacta, el *efecto marginal total* en el modelo *SDM* puede ser representado por la matriz de derivadas parciales de y respecto al cambio en una unidad de la k –ésima variable explicativa en X , dado un periodo t :

$$\begin{aligned}
\left[\frac{\partial E(\mathbf{y})}{\partial x_{1k}} \quad \dots \quad \frac{\partial E(\mathbf{y})}{\partial x_{nk}} \right]_t &= \begin{bmatrix} \frac{\partial E(y_1)}{\partial x_{1k}} & \dots & \frac{\partial E(y_1)}{\partial x_{nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial E(y_n)}{\partial x_{1k}} & \dots & \frac{\partial E(y_n)}{\partial x_{nk}} \end{bmatrix}_t, \\
&= (\mathbf{I}_n - \rho \mathbf{W})_t^{-1} \begin{bmatrix} \beta_k & w_{12} \theta_k & \dots & w_{1n} \theta_k \\ w_{21} \theta_k & \beta_k & \dots & w_{2n} \theta_k \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} \theta_k & w_{n2} \theta_k & \dots & \beta_k \end{bmatrix}_t, \\
&= \mathbf{S}_t [\beta_k \mathbf{I}_n + \theta_k \mathbf{W}]_t, \tag{24}
\end{aligned}$$

donde w_{ij} es el elemento (i, j) –ésimo de W , β_k es el k –ésimo elemento del vector β , θ_k es el elemento k –ésimo del vector θ y \mathbf{S}_t es igual a $(\mathbf{I}_n - \rho \mathbf{W})_t^{-1}$. Obsérvese que para

la estimación de los efectos marginales se utiliza el valor esperado de la forma reducida del modelo, $E(\mathbf{y}_t) = (\mathbf{I} - \rho \mathbf{W})_t^{-1} [\mathbf{X}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{X}_t \boldsymbol{\theta}]$.

En la expresión (24) puede observarse la presencia de un efecto directo y otro indirecto. El *efecto directo* es representado por los elementos de la diagonal principal de la ecuación (24). Este cambio será diferente para cada región por lo que no existe un único efecto directo. Por su parte, el *efecto indirecto (spatial spillover)*, proviene de los elementos fuera de la diagonal principal de (24) y, nuevamente, será diferente para cada región. En este caso se tiene un sumatorio de efectos para cada localidad.

Para el caso de un *SLM*, $\mathbf{y}_t = \rho \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$, los *efectos marginales totales* se simplifican ya que $\theta_k = 0$ siendo la expresión final igual a $\mathbf{S}_t [\beta_k \mathbf{I}_n]_t$. El *efecto directo* será igual al que surge en el modelo *SDM*, pero el *efecto indirecto* se torna más simple siendo igual a la suma de los elementos fuera de la diagonal principal de $\beta_k \mathbf{S}_t$, conteniendo solo efectos globales.

Ya sea que se estime un modelo *SLM* o *SDM*, siempre se tendrá un efecto marginal que no es único para todas las localidades (en contraposición a lo que sucede bajo un modelo no-espacial). Debido a esta característica, LeSage y Pace (2009) proponen una medida resumen para cada uno de los efectos considerando valores promedio. Por ejemplo para un modelo *SLM*, para cada periodo temporal se estiman los siguientes efectos:

$$\begin{aligned} \overline{M}_{Total} &= \frac{\beta_k}{n} \mathbf{1}'_n \mathbf{S} \mathbf{1}_n = \frac{\beta_k}{n} \sum_i \sum_j S_{i,j} \\ \overline{M}_{Directo} &= \frac{\beta_k}{n} \text{tr}(\mathbf{S}) = \frac{\beta_k}{n} \sum_i S_{i,i} \\ \overline{M}_{Indirecto} &= \frac{\beta_k}{n} \mathbf{1}'_n \mathbf{S} \mathbf{1}_n - \frac{\beta_k}{n} \text{tr}(\mathbf{S}) = \frac{\beta_k}{n} \sum_i \sum_{j \neq i} S_{i,j} \end{aligned}$$

Los efectos indirectos estimados deberían ser utilizados para conocer si los estos efectos espaciales son significativos o no. El problema es que los mismos surgen de una combinación de coeficientes y puede suceder que algunos coeficientes sean significativos otros sin poder conocer la significancia conjunta involucrada en los efectos indirectos. LeSage y Pace (2009) proponen simular la distribución de los efectos directos e indirectos usando la matriz de varianza-covarianza implicada en la estimación por máxima verosimilitud. Si hemos estimado un *SLM*, el procedimiento consiste en extraer una combinación r de parámetros de esta matriz de varianza-covarianza:

$$\left[\alpha_r \quad \boldsymbol{\beta}'_r \quad \rho_r \quad \sigma_r^2 \right]' = \left[\hat{\alpha} \quad \hat{\boldsymbol{\beta}}' \quad \hat{\rho} \quad \hat{\sigma}^2 \right]' + P' \boldsymbol{\epsilon},$$

donde P simboliza matriz triangular superior de la descomposición de Cholesky de la matriz de varianza-covarianza de los parámetros de máxima verosimilitud, y $\boldsymbol{\epsilon}$ es un vector de números aleatorios de dimensión $(K + 3)$, provenientes de una distribución normal estándar. Si R

extracciones parámetros son realizadas de esta forma, entonces los efectos promedios totales, directos e indirectos para cada variable explicativa pueden ser aproximados como:

$$\bar{\delta}_k = \frac{1}{R} \sum_{r=1}^R \delta_{kd},$$

$$t - valor_k = \bar{\delta}_k / \left[\frac{1}{R} \sum_{r=1}^R (\delta_{kd} - \bar{\delta}_k)^2 \right],$$

donde $\bar{\delta}_k$ es el efecto promedio de obtenido en R simulaciones, generalmente $R = 1000$, y dado el $t - valor$ de los efectos, puede obtenerse la significancia de cada efecto.

Los efectos directos e indirectos desarrollados para corte transversal o paneles estáticos pueden extenderse hacia paneles dinámicos. La ventaja del modelo dinámico es que pueden obtenerse efectos directos e indirectos de corto y largo plazo. Por ejemplo, bajo un modelo *dynSDM3* (ecuación 20), los efectos de corto plazo se obtienen ignorando τ y γ . La matriz de derivadas parciales de y con respecto a un cambio en una unidad de la variable explicativa $k - \acute{e}sima$ en las n regiones, para un punto temporal dado puede obtenerse como:

$$\left[\frac{\partial E(\mathbf{y})}{\partial x_{1k}} \quad \dots \quad \frac{\partial E(\mathbf{y})}{\partial x_{nk}} \right]_t = (\mathbf{I}_n - \rho \mathbf{W})^{-1} [\beta_k \mathbf{I}_n + \theta_k \mathbf{W}]. \quad (25)$$

Para obtener los efectos de largo plazo se puede suponer se alcanza un estado estacionario ($\mathbf{y}_t = \mathbf{y}_{t-1} = \mathbf{y}^*$) tal que:

$$\left[\frac{\partial E(\mathbf{y})}{\partial x_{1k}} \quad \dots \quad \frac{\partial E(\mathbf{y})}{\partial x_{nk}} \right]_t = [(1 - \tau) \mathbf{I}_n - (\rho + \gamma) \mathbf{W}]^{-1} [\beta_k \mathbf{I}_n + \theta_k \mathbf{W}]. \quad (26)$$

Estos resultados permiten determinar los efectos directos en el corto y largo plazo, y los efectos indirectos del corto y largo plazo (*spatial spillovers*). Desde estos modelos pueden obtenerse trayectorias secuenciales periodo a periodo como las presentadas por Debarsy et al. (2012).

Es importante recordar que estos efectos deben ser cuidadosamente interpretados y no están libres de controversias ya que su obtención tiene un fuerte componente causal. La relación sobre la adecuada interpretación de los resultados y qué se entiende por causalidad en el espacio no es un tema abierto. Por ejemplo, algunos autores sugieren que el parámetro espacial ρ del *SLM* (o del *SDM*) es un parámetro causal (Gibbons y Overman, 2012). La posición más adecuada no es posiblemente esa, más bien el parámetro ρ únicamente captura la dinámica espacial de la propia variable endógena y no debería interpretarse como causal. Este posición proviene de las similitudes espaciales a lo que sucede en un proceso AR(1) en el tiempo, $y_t = \rho y_{t-1} + \varepsilon_t$, en donde ρ nunca se interpreta como un parámetro causal en series de tiempo. Para una discusión sobre este punto véase Herrera et al. (2014).

5. Aplicaciones empíricas

Las aplicaciones empíricas de esta sección están basadas en dos modelos económicos bien conocidos. El primero es un modelo de crecimiento económico estimado por Ertur y Koch (2007) usando datos de corte transversal de 91 países en el periodo 1960–1995. El otro ejemplo es un modelo de consumo de tabaco estimado por Baltagi y Li (2004) usando un panel de datos de 50 estados norteamericanos desde 1963 hasta 1992.

5.1. Modelo de crecimiento (Ertur y Koch, 2007)

Ertur y Koch (2007) investigan la interdependencia tecnológica en el crecimiento económico entre países. Basados en un modelo neoclásico, los autores consideran una función de producción Cobb-Douglas que refleja rendimientos constantes a escala respecto al trabajo y capital físico para el país i en el periodo temporal t :

$$Y_i(t) = A_i(t)K_i^\alpha(t)L_i^{1-\alpha}(t), \quad (27)$$

donde $Y_i(t)$ es el producto, $K_i(t)$ es el nivel de capital físico, $L_i(t)$ es el nivel de trabajo, y $A_i(t)$ es el nivel agregado de tecnología especificado como:

$$A_i(t) = \Omega(t)k_i^\phi(t) \prod_{j \neq i}^n A_j^{\delta\omega_{ij}}(t). \quad (28)$$

El nivel agregado de tecnología $A_i(t)$ de un país i depende de tres elementos. Primero, una cierta proporción de progreso tecnológico exógeno e idéntico en todos los países, $\Omega(t) = \Omega(0)e^{\mu t}$, donde μ es la tasa constante de crecimiento. Segundo, el nivel tecnológico de cada país crece con el nivel agregado de capital físico per trabajador $k_i^\phi(t) = (K_i(t)/L_i(t))^\phi$ con $\phi \in [0; 1]$ capturando la fortaleza de las externalidades domésticas de la acumulación del capital físico. Finalmente, el tercer término captura el efecto externo de conocimiento incorporado en el capital localizado en diferentes países cuyo impacto trasciende los límites nacional con una intensidad decreciente, $\delta \in [0; 1]$. Los términos ω_{ij} representan la conectividad entre el país i y sus vecinos j .

Siguiendo el desarrollo de Solow, los autores suponen que en cada país i , una fracción constante del producto s_i es ahorrada y que el mano de obra crece exógenamente a la tasa l_i . Además, se supone una tasa anual de depreciación del capital físico constante e idéntica para todos los países. La evolución del producto por trabajador en el país i es gobernado por la conocida ecuación dinámica de Solow que permite obtener el ingreso por trabajador de estado estacionario de los países (Ertur y Koch, 2007, pp. 1038, ecuación 9). Luego de algunas

manipulaciones se obtiene el ingreso real por trabajador del país i en estado estacionario:

$$\begin{aligned} \ln y_i^*(t) &= \frac{1}{1-\alpha-\phi} \ln \Omega(t) + \frac{\alpha+\phi}{1-\alpha-\phi} \ln s_i - \frac{\alpha+\phi}{1-\alpha-\phi} \ln(l_i+g+\delta) \\ &\quad - \frac{\alpha\gamma}{1-\alpha-\phi} \sum_{j \neq i}^n w_{ij} \ln s_j + \frac{\alpha\gamma}{1-\alpha-\phi} \sum_{j \neq i}^n w_{ij} \ln(l_j+g+\delta) \\ &\quad + \frac{\gamma(1-\alpha)}{1-\alpha-\phi} \sum_{j \neq i}^n w_{ij} \ln y_j^*(t). \end{aligned} \quad (29)$$

Este es un modelo de Solow espacialmente aumentado y coincide con el predictor obtenido por Solow más un efecto espacial. Considerando $t = 0$ por simplicidad, la anterior ecuación permite encontrar el ingreso real por trabajador a lo largo de la trayectoria de crecimiento balanceado:

$$\begin{aligned} \ln \left[\frac{Y_i}{L_i} \right] &= \beta_0 + \beta_1 \ln s_i + \beta_2 \ln(l_i+g+\delta) \\ &\quad + \theta_1 \sum_{j \neq i}^n w_{ij} \ln s_j + \rho \sum_{j \neq i}^n w_{ij} \ln \left[\frac{Y_j}{L_j} \right] + \varepsilon_i, \end{aligned} \quad (30)$$

donde $1/(1-\alpha-\phi) \ln \Omega(0) = \beta_0 + \varepsilon_i$, $i = 1, \dots, n$, siendo β_0 una constante y ε_i un término de error con $\ln \Omega(0)$ reflejando la tecnología, dotación de recursos, etc. país-específico. Se supone además que $g + \delta = 0,05$. Finalmente, se tienen las siguientes restricciones teóricas entre coeficientes: $\beta_1 = -\beta_2 = \frac{\alpha+\phi}{1-\alpha-\phi}$ y $\theta_2 = -\theta_1 = \frac{\alpha\gamma}{1-\alpha-\phi}$.

La ecuación (30) coincide con el modelo *SDM* y es estimada usando información del ingreso real, inversión y crecimiento poblacional de una muestra de 91 países durante 1960 – 1995. Respecto a la matriz de pesos espaciales, se consideran dos matrices con criterios de distancia geográfica: la primera utiliza la función inversa de la distancia al cuadrado (principal matriz que refleja una aproximación gravitacional) y la segunda se basa en la función exponencial negativa de la distancia al cuadrado (utilizada como prueba de robustez en la especificación).

Ambas matrices, \mathbf{W}_{inv} y \mathbf{W}_{exp} son estandarizadas por filas tal que cada elemento es transformado como: $\omega_{hij} = \omega_{hij}^* / \sum_{j=1}^n \omega_{hij}^*$; $h = inv, exp$, en donde

$$\omega_{inv,ij}^* = \begin{cases} 0 & \text{si } i = j \\ d_{ij}^{-2} & \text{en otro caso} \end{cases}, \quad \omega_{exp,ij}^* = \begin{cases} 0 & \text{si } i = j \\ e^{-2d_{ij}} & \text{en otro caso} \end{cases}, \quad (31)$$

siendo d_{ij} la distancia del gran-círculo entre capitales de los países i y j .

Los autores analizan varias especificaciones siguiendo restricciones teóricas sobre los parámetros, incluyendo modelos espaciales alternativos. En el Cuadro 1 se presentan los resultados del modelo de Solow tradicional estimado por *MCO* y los resultados del modelo *SDM* estimado por *ML* bajo las dos matrices alternativas. El parte superior de Cuadro 1 muestra las estima-

ciones no restringidas y la parte inferior muestra los resultados bajo las restricciones teóricas $\beta_1 = -\beta_2$ y $\theta_2 = -\theta_1$. Para el modelo de Solow, los resultados del test I de Moran (Cliff y Ord, 1981) rechaza fuertemente la hipótesis nula de no autocorrelación espacial. Este resultado refleja el error de especificación del modelo de Solow tradicional.

Cuadro 1: Estimaciones alternativas

<i>Modelos no restringidos</i>	<i>Solow</i>	<i>SDM/W_{inv}</i>	<i>SDM/W_{exp}</i>
constante	4,65**	0,48	1,81
ln(<i>s</i>)	1,28***	0,82***	0,80***
ln(<i>l</i> + 0,05)	-2,71***	-1,50***	-1,58***
W × ln(<i>s</i>)		-0,28	-0,24
W × ln(<i>l</i> + 0,05)		0,24	0,52
W × ln(<i>y</i>)		0,71***	0,62***
I Moran (W_{inv})	0,41***		
I Moran (W_{exp})	0,40***		
<i>Modelos restringidos</i>			
constante	8,38***	2,32***	3,11***
ln(<i>s</i>) - ln(<i>l</i> + 0,05)	1,38***	0,85***	0,82***
W × [ln(<i>s</i>) - ln(<i>l</i> + 0,05)]		-0,22	-0,19
W × ln(<i>y</i>)		0,71***	0,62***
I Moran (W_{inv})	0,40***		
I Moran (W_{exp})	0,38***		
<i>Test de restricciones</i>			
α implicado	4,43**	2,07	1,88
α implicado	0,58***	0,28**	0,30**
ϕ implicado		0,18*	0,15
γ implicado		0,56***	0,51***
<i>AIC</i>		97,54	105,37
<i>Entropía</i>		30,83	34,74
<i>Bayes (prob. post.)</i>		0,98	0,02

Note: *, ** y *** denota significancia al 10 %, 5 % y 1 % respectivamente.

Las otras columnas del Cuadro 1 muestran el modelo de Solow aumentado. En los modelos no restringidos los coeficientes tienen el signo predicho por la teoría y el coeficiente de autocorrelación espacial, ρ , es positivo y significativo. Las restricciones teóricas $\beta_1 = -\beta_2$ y $\theta_2 = -\theta_1$ no son rechazadas por los test LR (2,07 y 1,88, respectivamente) por lo que los modelos pueden estimarse bajo las restricciones teóricas.

Las estimaciones de los modelos restringidos confirman el error de especificación del modelo tradicional de Solow mediante los resultados del test I de Moran. Los modelos de Solow aumentado muestran un α implicado cercano a un tercio con valores significativos. Además, el valor estimado de γ es cercano a 0,5 y significativo indicando la importancia de la interdependencia tecnológica entre países y la importancia del vecindario en la determinación del ingreso real (ρ

estimado significativo).

Finalmente, los criterios de información AIC y $Entropía$ muestran preferencia por el modelo bajo \mathbf{W}_{inv} y similar evidencia es mostrada por la probabilidad posterior de la estimación bayesiana. El desarrollo teórico del modelo de crecimiento guía la especificación econométrica pero no resuelve la incertidumbre sobre (1) el criterio de conectividad que define qué países son vecinos, y (2) el supuesto de pesos espaciales que define cómo los vecinos se afectan entre sí. Los autores definen un criterio de conectividad basado en la distancia y plantean cierta incertidumbre sobre la fortaleza de la interacción. Ertur y Koch usaron el criterio AIC para elegir la primera matriz y, en este ejemplo, se han agregado dos alternativas que confirman esa elección.

5.2. Consumo de tabaco en datos de panel

Baltagi y Li (2004) explican la demanda de cigarrillos en 46⁴ estados norteamericanos en un periodo de 30 años, desde 1963 a 1992. La variable dependiente, $\log(c)$, es el logaritmo de la venta per cápita de cigarrillos en términos reales, medida en paquetes de cigarrillos por personas de 14 o más años. Las variables explicativas son: $\log(p)$, el logaritmo del precio de venta promedio del paquete de tabaco y $\log(ing)$, logaritmo del ingreso disponible per cápita. La matriz de pesos espaciales es especificada como una matriz de contigüidad binaria, estandarizada por fila. Esta base de datos ha sido utilizada por Debarsy et al. (2012), Elhorst (2014a) y Kelejian y Piras (2014).

La motivación inicial del trabajo fue el efecto contrabando, en donde compradores de cigarrillos de estados limítrofes compraban cigarrillos en los estados vecinos si el precio era ventajoso. Esto podría ser para consumo propio o para revender en su estado de origen.

En el Cuadro 2 se presentan los resultados de modelos estáticos y dinámicos. De los dos modelos estáticos presentados, el modelo *pooled* permite una primera aproximación para detectar la presencia de elementos faltantes como efectos espaciales, sustantivos y residuales. Esta información es obtenida por las versiones robustas de los tests LM , LM_{error}^r y LM_{lag}^r , que detectan la omisión de un efecto espacial en el error o un efecto espacial en la variable dependiente, respectivamente. Ambos tests rechazan la hipótesis nula de no autocorrelación residual y sustantiva, permitiendo avanzar en la especificación espacial. Adicionalmente, se testea en el modelo *pooled* la significancia conjunto de efectos fijos individuales y temporales encontrando evidencia a favor de la inclusión de ambos efectos.

Considerando los resultados del modelo *pooled*, la selección especificación del modelo puede seguir la secuencia de abajo hacia arriba de la Figura 1. En nuestro caso aplicaremos la estrategia de utilizar el modelo estático más complejo posible, SDM , y a partir de allí reducir elementos no significativos.

⁴Los estados omitidos son Alaska, Colorado, Hawaii, North Carolina, y Oregon.

Cuadro 2: Estimaciones alternativas

<i>Modelos</i>	<i>Estáticos</i>		<i>Dinámicos</i>	
	<i>Pooled</i>	<i>SDM</i>	<i>dynSDM1</i>	<i>dynSDM4</i>
$\log(c)_{-1}$			0,87***	0,87***
$\mathbf{W} \times \log(c)_{-1}$				0,02
$\mathbf{W} \times \log(c)$		0,23***	0,05**	0,03
$\log(p)$	-0,86***	-1,00***	-0,26***	-0,26***
$\log(y)$	0,27***	0,60***	0,10***	0,10***
$\mathbf{W} \times \log(p)$		0,06	0,16***	0,16***
$\mathbf{W} \times \log(y)$		-0,29***	-0,02	-0,02
LM_{error}^r	144,84***			
LM_{lag}^r	58,26***			
$Wald(\boldsymbol{\theta} = 0)$		16,06***		
$LR(fact. com.)$		8,62**		

Nota: *, ** y *** denota significancia al 10 %, 5 % y 1 % respectivamente. Los modelos *SDM* incluyen efectos fijos.

Los resultados del modelo estático *SDM* se presentan en la tercera columna del Cuadro 2. La hipótesis nula de $\boldsymbol{\theta} = 0$ permite ver si el modelo *SDM* puede ser simplificado a un *SLM*, la evidencia del test de *Wald* permite rechazar la hipótesis a favor de la inclusión de rezagos espaciales exógenos. La posible reducción del modelo *SDM* a un *SEM* es realizada por el test *LR* de factores comunes: $H_0 : \boldsymbol{\theta} + \rho\boldsymbol{\beta} = 0$, con 2 grados de libertad. Finalmente, la significancia del parámetro $\hat{\rho}$ muestra que no es posible reducir el modelo *SDM* en un *SLX*.

Una vez detectado el modelo estático final, puede procederse a interpretar los resultados. Los coeficientes de las dos variables explicativas tienen el signo esperado con elasticidad unitaria negativa para el precio y de 0.6 positiva para el ingreso indicando que un mayor ingreso aumenta la demanda de cigarrillos. Comparados estos coeficientes con los obtenidos por el modelo *pooled* parece claro que existe una corrección, sin embargo, esta comparación no es válida debido a la presencia del parámetro ρ . Es necesario computar los efectos marginales para proveer de una interpretación adecuada. En el Cuadro 3 se muestran los efectos marginales del modelo estático *SDM* basados un 1000 repeticiones. Los resultados son interpretados como de largo plazo, una vez que la dinámica espacial del consumo converge. El efecto de principal interés es el impacto marginal indirecto del precio y como puede observarse el resultado es una elasticidad-precio -0,21 y significativo al 5 %, evidencia que no es consistente con la hipótesis inicial de Baltagi y Lee. Esto puede ser explicado por la ausencia de elementos dinámicos temporales.

Para conocer si existe dinámica temporal, se han estimado dos versiones dinámicas del modelo *SDM*: *dynSDM1* (ecuación 18) y *dynSDM4* (ecuación 21), ambos considerando efectos fijos individuales y temporales. Como se muestra en el Cuadro 3, ambos modelos detectan una alta persistencia en el consumo de cigarrillos con un coeficiente del $\ln(c)_{-1}$ de 0.87. Sin embargo, el modelo *dynSDM4* no detecta persistencia espacio-recursiva siendo $\mathbf{W} \times \ln(c)_{-1}$

no significativo. Por lo tanto, el modelo más adecuado es el *dynSDM1*.

Cuadro 3: Efectos directos e indirectos de modelos *SDM* con efectos fijos

<i>Modelos</i>	<i>SDM</i>	<i>dynSDM1</i>
<i>Corto plazo</i>		
efecto directo $\log(p)$		-0,26***
efecto directo $\log(y)$		0,10***
efecto indirecto $\log(p)$		0,15***
efecto indirecto $\log(y)$		-0,02
<i>Largo plazo</i>		
efecto directo $\log(p)$	-1,01***	-1,94***
efecto directo $\log(y)$	0,59***	0,78***
efecto indirecto $\log(p)$	-0,21**	0,59
efecto indirecto $\log(y)$	-0,19**	0,25

Nota: *, ** y *** denota significancia al 10 %, 5 % y 1 % respectivamente.

Bajo este modelo final, *dynSDM1*, nuevamente se obtienen los efectos marginales de las variables explicativas que se presentan en el Cuadro 3. Debido a la presencia de dinámica temporal es posible obtener efectos marginales de corto y largo plazo. Centrando el interés en el efecto indirecto del precio, puede deducirse que efecto un efecto contrabando en el corto plazo, elasticidad-precio 0,15 y significativa: el incremento del precio en estados vecinos aumenta el consumo en cada región. Este efecto en el largo plazo continúa siendo positivo pero no es significativo. En conclusión, existe evidencia de spillovers espaciales del precio del cigarrillo a corto plazo.

6. Comentarios Finales

El desarrollo de la econometría espacial ha sido muy importante en los últimos años acompañando la mayor disponibilidad de información georreferenciada. El capítulo presenta los desarrollos más consolidados dentro del área para datos de corte transversal y datos de panel destacando tópicos relativos a la especificación e interpretación de los modelos espaciales.

Los datos espaciales presentan características particulares. Los datos pueden ser continuos o discretos, pueden provenir de agregaciones u observaciones puntuales en el espacio, la localización puede ser regular o irregular, y dichas localizaciones pueden ser consideradas provenientes de un espacio continuo o discreto. La agregación altera la variabilidad e impacta en la inferencia a realizar. También el soporte de la información geográfica es importante cuando se combinan diferentes tipos de datos espaciales. Estas características deben ser consideradas y las conclusiones inferenciales deberían realizarse con suma cautela teniendo en cuenta la escala de análisis, más aún si estas conclusiones son orientadas a acciones políticas.

Un tema central en esta disciplina es la construcción de la matriz de pesos espaciales y los problemas inferenciales que puede generar una mala elección. La normalidad asintótica y consistencia de los estimadores está bien establecida bajo el supuesto de una matriz espacial estrictamente exógena. Sin embargo, cuando los pesos espaciales involucran variables endógenas los métodos no son válidos. Algunos recientes desarrollos flexibilizan este supuesto, por ejemplo, Kelejian y Piras (2014) proponen una matriz de pesos espaciales endógenos instrumentalizada. Cheng y Lee (2017) utilizan esta estrategia para desarrollar un test de endogeneidad tipo Hausman para los pesos espaciales. Qu et al. (2017) y Qu y Lee (2017) desarrollan enfoques de estimación permitiendo correlación entre el término error y las variables utilizadas en la construcción de los pesos espaciales. Esta endogeneidad es solucionada utilizando el método de función de control. Estos modelos son actualmente de frontera y existen varios avances que flexibilizan la matriz de contactos permitiendo que la misma cambie entre periodos temporales.

Los ejemplos empíricos presentados en el capítulo pretenden mostrar dos vías alternativas habituales en econometría espacial aplicada. El trabajo de Ertur y Koch (2007) es un buen ejemplo de la aplicación del enfoque *theory-driven* y cómo la estimación final íntegramente vincula la interpretación a la teoría de crecimiento económico. A pesar de ser un modelo derivado desde la teoría, la especificación de la matriz es un a priori de los investigadores y por ello la necesidad de comparar matrices alternativas. Esta incertidumbre está presente en todo modelo espacial y existen diferentes de alternativas que ayudan a la elección de la matriz más adecuada (véase Herrera et al., 2017). El segundo caso aplicado es un ejemplo del enfoque *data-driven*, explorando especificaciones espaciales alternativas y con conclusiones obtenidas del modelo final. En este caso no hay un desarrollo teórico que fundamente el modelo finalmente estimado, son los datos los que permiten adecuar la mejor especificación. La interpretación de los resultados son realizados según la vigente visión de efectos marginales simulados. La interpretación de estos efectos marginales posee un fuerte significado causal que no deja de ser controversial y es un tema que merece un mayor debate a futuro.

Finalmente, cabe mencionar las diferentes alternativas computacionales. Bajo Matlab existe una amplia cantidad de códigos que cubren todo el rango de modelos presentados en el capítulo. El programa *R* tiene diferentes paquetes que permiten estimar modelos de corte transversal y paneles espaciales estáticos. GeoDa es un excelente programa para iniciarse en el análisis espacial, permitiendo la estimación de modelos de corte transversal y paneles pooled bajo *ML*. GeoDaSpace es útil si se desea estimar modelos de corte transversal por *IV/GMM*. También existe un conjunto de opciones desarrolladas bajo Python. Finalmente, la versión 15 de Stata trae incorporado un módulo para análisis espacial y hay una nutrida variedad de comandos desarrollados por usuarios que cubren todo el espectro de los modelos espaciales aquí analizados.

Referencias

- Anselin, L. (1988). *Spatial econometrics: Methods and models*, Volume 4. Kluwer Academic Publishers, Dordrecht.
- Anselin, L. (2005). Exploring spatial data with GeoDa: A workbook. Urbana-Champaign: Center for Spatially Integrated Social Science.
- Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1), 3-25.
- Anselin, L., A. Bera, R. Florax, y M. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, 26(1), 77-104.
- Anselin, L., J. Le Gallo, y H. Jayet (2008). Spatial panel econometrics. En L. Matyas y P. Sevestre (Eds.), *The econometrics of panel data: fundamentals and recent developments in theory and practice*, vol. 46, 625-660. Springer.
- Arraiz, I., D. Drukker, H. Kelejian, e I. Prucha (2010). A spatial Cliff-Ord-type model with heteroskedastic innovations: Small and large sample results. *Journal of Regional Science*, 50(2), 592-614.
- Baltagi, B. y D. Li (2004). Prediction in the panel data model with spatial correlation. En L. Anselin, R. Florax, y S. Rey (Eds.), *Advances in spatial econometrics*, 283-295. Springer.
- Baltagi, B., S. Song, y W. Koh (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117(1), 123-150.
- Bartels, C. y R. Ketellapper (1979). *Exploratory and explanatory statistical analysis of spatial data*. Martinus Nijho, Boston, MA.
- Bennett, R. (1979). *Spatial time series*. Pion, London.
- Bivand, R. (2010). Exploratory spatial data analysis. En M. Fischer y G (Eds.), *Handbook of Applied Spatial Analysis*, 219-254. Springer, Hiedelberg.
- Brock, W. y S. Durlauf (1995). Discrete choices with social interactions I: Theory. Santa Fe, Santa Fe Institute WP 95-10.
- Brueckner, J. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review*, 26(2), 175-188.
- Cheng, W. y L.-f. Lee (2017). Testing endogeneity of spatial and social networks. *Regional Science and Urban Economics*, 64, 81-97.
- Cliff, A. y J. Ord (1981). *Spatial Processes: Models and Applications*. Pion: London.
- Cressie, N. (1993). *Statistics for spatial data*. New York, Wiley.
- Debarys, N. y C. Ertur (2010). Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, 40(6), 453-470.
- Debarys, N., C. Ertur, y J. P. LeSage (2012). Interpreting dynamic space-time panel data models. *Statistical Methodology*, 9(1), 158-171.
- Drukker, D. , P. Egger, e I. Prucha (2013). On two-step estimation of a spatial autoregressive

- model with autoregressive disturbances and endogenous regressors. *Econometric Reviews*, 32(5-6), 686-733.
- Elhorst, J. (2010a). Dynamic panels with endogenous interaction effects when T is small. *Regional Science and Urban Economics*, 40(5), 272-282.
- Elhorst, J. (2010b). Spatial panel data models. En M. Fischer y A. Getis (Eds.), *Handbook of Applied Spatial Analysis. Software Tools, Methods and Applications*, 377-407. Springer-Verlag Berlin Heidelberg.
- Elhorst, J. (2014a). Matlab software for spatial panels. *International Regional Science Review*, 37(3), 389-405.
- Elhorst, J. (2014b). *Spatial Econometrics. From Cross-sectional data to Spatial Panels*. SpringerBriefs in Regional Science, Springer.
- Ertur, C. y W. Koch (2007). Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics*, 22(6), 1033-1062.
- Giacomini, R. y C. Granger (2004). Aggregation of space-time processes. *Journal of Econometrics*, 118(1), 7-26.
- Gibbons, S. y H. Overman (2012). Mostly pointless spatial econometrics?. *Journal of Regional Science*, 52(2), 172-191.
- Gotway, C. y L. Young (2002). Combining incompatible spatial data. *Journal of the American Statistical Association*, 97(458), 632-648.
- Graubard, B. y E. Korn (2002). Inference for superpopulation parameters using sample surveys. *Statistical Science*, 17(1), 73-96.
- Griffith, D. A. (1996). Some guidelines for specifying the geographic weights matrix contained in spatial statistical models. En S. Arlinghaus (Ed.), *Practical Handbook of Spatial Statistics*, p. 65-82. CRC Press Book: Boca Raton, FL.
- Halleck Vega, S. y J. Elhorst (2015). The SLX model. *Journal of Regional Science*, 55(3), 339-363.
- Hendry, D. (1979). Predictive failure and econometric modelling in macroeconomics: The transactions demand for money. En P. Ormerod (Ed.), *Economic Modelling: Current Issues and Problems in Macroeconomic Modelling in the UK and the US*, 217-242. Heinemann Education Books, London.
- Herrera, M. (2015). Econometría espacial usando Stata. Breve guía aplicada para datos de corte transversal. Documentos de Trabajo del IELDE (13).
- Herrera, M., J. Mur, y M. Ruiz (2016). Detecting causal relationships between spatial processes. *Papers in Regional Science*, 95(3), 577-594.
- Herrera, M., J. Mur, y M. Ruiz (2017). A comparison study on criteria to select the most adequate weighting matrix. XIth World Conference of the Spatial Econometrics Association, Singapore.

- Hordijk, L. (1979). Problems in estimating econometric relations in space. *Papers in Regional Science*, 42(1), 99-115.
- Kapoor, M., H. Kelejian, e I. Prucha (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140, 97-130.
- Kelejian, H. y D. Robinson (1995). Spatial correlation: a suggested alternative to the autoregressive model. En L. Anselin y R. Florax (Eds.), *New directions in Spatial Econometrics*, 75-95. Springer.
- Kelejian, H. y G. Piras (2014). Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics*, 46, 140-149.
- Kelejian, H. e I. Prucha (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1), 99-121.
- Kelejian, H. e I. Prucha (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2), 509-533.
- Kelejian, H. e I. Prucha (2010). Specication and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1), 53-67.
- Krugman, P. (1991). Increasing returns and Economic Geography. *Journal of Political Economy*, 99(3), 483-499.
- Lee, L.-f. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews*, 22(4), 307-335.
- Lee, L.-f. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6), 1899- 1925.
- Lee, L.-f. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2), 333-374.
- Lee, L.-f. y J. Yu (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2), 165-185.
- Lee, L.-f. y J. Yu (2014). Spatial panel models. En B. Baltagi (Ed.), *The Oxford Handbook of Panel Data*, 363-401. Oxford University Press.
- LeSage, J. y R. Pace (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC press.
- Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3), 531-542.
- Openshaw, S. (1984). *The modifiable areal unit problem*. Geo Books, Norwich. Printed by Headley Brothers Ltd. Kent.
- Openshaw, S. y P. Taylor (1979). A million or so correlation coecients: three experiments on the modifiable areal unit problem. En N. Wrigley y P. Taylor (Eds.), *Statistical Methods in the Spatial Sciences*, 127-144. Pion.

- Ord, J. (1981). Towards a theory of spatial statistics: a comment. *Geographical Analysis*, 13(1), 86-91.
- Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70(349), 120-126.
- Pace, R. y S. Zhu (2012). Separable spatial modeling of spillovers and disturbances. *Journal of Geographical Systems*, 14(1), 75-90.
- Paelinck, J. y L. Klaassen (1979). *Spatial Econometrics*. Saxon House, Farnborough.
- Pesaran, M. (2004). General diagnostic tests for cross section dependence in panels. Cambridge Working Papers in Economics No. 0435. University of Cambridge.
- Qu, X. y L.-F. Lee (2017). Estimation of an endogenous SAR model with spatial weights constructed by bilateral variables. XIth World Conference of the Spatial Econometrics Association, Singapore.
- Qu, X., L.-f. Lee, y J. Yu (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics*, 197(2), 173-201.
- Smith, T. (2009). Estimation bias in spatial models with strongly connected weight matrices. *Geographical Analysis*, 41(3), 307-332.
- Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46, 234-240.
- Waller, L. y C. Gotway (2004). *Applied Spatial Statistics for Public Health Data*, v. 368. John Wiley & Sons.
- Yu, J., R. de Jong, y L.-f. Lee (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146 (1), 118-134.
- Yu, J., R. de Jong, y L.-f. Lee (2012). Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration. *Journal of Econometrics*, 167(1), 16-37.