

Modelos para datos espaciales con estructura transversal o de panel. Una revisión*

JEAN PAELINCK ^a, JESÚS MUR ^b, F. JAVIER TRÍVEZ ^b

^a *Emeritus Professor, Erasmus University Rotterdam, Burgemeester Oudlaan, 50, 3062 PA Rotterdam, Países Bajos. E-mail: j.paelinck@upcmail.nl*

^b *Universidad de Zaragoza, Facultad de CC.EE., Campus Rio Ebro, c/ María de Luna s/n, 50018 Zaragoza, España. E-mail: jmur@unizar.es, fjtrivez@unizar.es*

RESUMEN

En este trabajo analizamos la situación actual de la Econometría espacial, como disciplina de naturaleza econométrica especializada en el manejo de datos y relaciones de tipo espacial. Se trata de un campo de investigación que ha evolucionado muy rápidamente en las últimas décadas produciendo una gran variedad de técnicas e instrumentos diferentes. El pequeño análisis bibliográfico que realizamos en la primera parte del trabajo pone de manifiesto esta tensión. A continuación, la discusión sobre métodos se estructura en dos grandes apartados dedicados, respectivamente, a los modelos puros de corte transversal y a los modelos de datos panel. Esta revisión nos ha permitido identificar ciertas lagunas que deberían recibir mayor atención por parte de los usuarios, aunque la valoración global es positiva.

Palabras clave: Econometría espacial, modelos transversales, modelos de datos panel.

Models for Spatial Data with Panel or Cross-Sectional Structure. A Review

ABSTRACT

This paper analyzes the current situation of spatial econometrics; this is an econometric discipline specialized in data and relationships of a spatial type. The research in this field has evolved very quickly in recent decades producing a variety of different techniques and instruments. This tension is clearly evident in the literature review we included in the first part of the paper. Then we discuss the methods which are divided into two main sections devoted respectively to pure cross sectional models and panel data models. The conclusion we draw from this review is very positive; however, we point out some gaps that, we think, merit further attention from users

Keywords: Spatial Econometrics, Cross-Sectional Models, Panel Data Models.

Clasificación JEL: C21, C18, C33, R15

* *Los autores quieren expresar su agradecimiento por la ayuda financiera recibida del proyecto ECO2012-36032-C03-01 del Ministerio de Economía y Competitividad del Gobierno de España y del Departamento de Industria e Innovación del Gobierno de Aragón y del FON-DO SOCIAL EUROPEO a través del proyecto del grupo GAEC, Ref. S60.*

1. INTRODUCCIÓN

Los últimos años han sido muy activos en el pequeño mundo de la Econometría espacial, en los que se han sucedido varias celebraciones: los cincuenta años de la aparición del primer número de *Journal of Regional Science* (2012); treinta y cinco años de los textos de Paelinck y Klaassen (1979) y de Cliff y Ord (1981), los veinticinco años del texto de Anselin (1988) o el primer lustro de la Asociación de Econometría Espacial (SEA).

La disciplina acumula suficiente experiencia como para poder mirar hacia atrás con perspectiva, como Anselin (2010), con el objetivo de corregir errores y tomar impulso. Esta rama de la Econometría ha madurado muy rápidamente, bajo grandes presiones, produciendo un amplio abanico de nuevas y poderosas técnicas para el análisis cuantitativo de datos espaciales. No debería calificarse más como *‘una disciplina joven con un brillante futuro’*. En la sección 2 se ofrecen algunos datos que confirman su fortaleza actual y las buenas perspectivas de desarrollo. El ratio de publicaciones en los últimos años ha sido elevado y muy constante, lo que no ha evitado ciertos desequilibrios que deberán ser resueltos a medio plazo.

Remitimos a los lectores a autores más autorizados que nosotros para hacer un estado de la cuestión, como Anselin *et al.* (2004), Anselin y Rey (2010), Arbia (2006), Elhorst (2014), Getis *et al.* (2003), Griffith (2003), Griffith y Paelinck (2011), Haining (2003) o LeSage y Pace (2009), por citar los más populares. Nuestro objetivo es más modesto y se limita a ofrecer una panorámica, seguramente sesgada, del tipo de técnicas que se han extendido en el ámbito del análisis de datos espaciales. Esta contribución debe entenderse como una actualización de la que realizamos hace unos años (Paelinck *et al.*, 2004)

Nos preocupan, en particular, tres cuestiones que, a nuestro entender, todavía no han sido resueltas satisfactoriamente. La primera se refiere a la selección de la matriz de contactos \mathbf{W} . No es ninguna exageración afirmar que la especificación del modelo se inicia con la construcción de esta matriz que, después, juega un papel crucial en las diferentes etapas del trabajo econométrico. El problema, como es bien sabido, es la gran incertidumbre que rodea la decisión de qué matriz utilizar y por qué. En la literatura encontramos muchas sugerencias sobre cómo proceder pero con escasa fundamentación teórica. Como indican Corrado y Fingleton (2012), carecemos de rutinas formalizadas para abordar la cuestión que, esencialmente, es un problema de decisión.

Todo lo que tiene que ver con el tópico de causalidad es otro gran déficit de la literatura sobre Econometría espacial, lo que contrasta notablemente con su relevancia en el ámbito más general. Gibbons y Overman (2012) se muestran muy críticos con la situación. Su argumento es que no hay nada singular en esta disciplina que explique la omisión del tópico de causalidad en un ámbito espacial, y nosotros nos sumamos a esta crítica.

En tercer lugar, y no por ello menos importante, debe situarse el problema de la identificación en los modelos espaciales convencionales. Este punto se ha discutido en el pasado, pero habitualmente en conexión con la matriz de contactos. Existe consenso general en que esta matriz es necesaria, entre otras cosas, para asegurar la identificación de los modelos, lo que significa que la identificación descansa sobre bases bastante cuestionables (sobre la matriz de contactos, \mathbf{W} , en definitiva). Manski (1993) reabrió este debate que no puede declararse, ni mucho menos, resuelto. Al contrario, la preocupación se ha extendido como se comprueba en los trabajos de Pinkse y Slade (2010) o de Gibbons y Overman (2012).

Sobre estos aspectos volveremos más adelante. El trabajo se organiza en seis secciones. En la Sección 2 repasamos las publicaciones, vinculadas a problemas de econometría espacial, que han aparecido en las revistas más relevantes para el análisis económico en general y el análisis espacial en particular. En la Sección 3 centramos la atención en los modelos de corte transversal y en la Sección 4 en los modelos de datos panel, en su versión estática y dinámica. En ambos casos, comentamos brevemente los principios y criterios de análisis que subyacen en los dos tipos de modelos; discutimos también las deficiencias que, creemos, deberán corregirse a medio plazo. En la Sección 5 presentamos una pequeña propuesta que combina estructuras espaciales y temporales. El trabajo finaliza con una sección de conclusiones y perspectivas futuras.

2. VISIBILIDAD INTERNACIONAL Y DESINTERÉS NACIONAL

Hace unos pocos años, Arbia (2011) publicó una pequeña investigación bibliográfica dedicada a conocer donde se habían publicado los trabajos de Econometría espacial en el periodo 2007-2011, que coincide con el de actividad de la Asociación de Econometría Espacial (SEA). Arbia encontró 45 artículos publicados en las 7 revistas líder en el ámbito de la Econometría y 146 más en las 9 revistas más citadas en temas de análisis espacial. Si extendemos el estudio para el periodo 2012 a 2014, vemos que esta tendencia se ha mantenido ligeramente acentuada (Tabla 1).

Está claro que 8 años es muy poco tiempo para definir tendencias, pero los datos sobre trabajos publicados en revistas de impacto confirman la vitalidad del trabajo que se desarrolla en esta disciplina. Para ser honestos, debe reconocerse que prácticamente la mitad de los trabajos publicados en revistas de Econometría están relacionados con modelos de datos panel, y una proporción elevada tratan con problemas de no estacionariedad y raíces unitarias en los que el espacio tiene alguna incidencia. En cierto sentido, puede decirse que la investigación sobre los modelos panel ha contribuido a hacer más visible la propia Econometría espacial. La colección de trabajos publicados en revistas de corte

regional es más heterogéneo, reflejo de los problemas típicos que preocupan en este ámbito (especificación/selección de modelos, aproximadamente el 10% de las publicaciones; estimación, 15%; contraste de modelos espaciales, 30%; miscelánea y aplicaciones diversas, 45%).

La situación en el caso de las revistas españolas es muy diferente. En las 5 revistas que pueden considerarse como de referencia en este ámbito se publicaron solo 43 trabajos relacionados con la Econometría espacial en el primer periodo y 28 en el segundo, y casi todos ellos concentrados en dos revistas: *Investigaciones Regionales*, publicación de cabecera de la Asociación Española de Ciencia Regional, y *Revista de Estadística Española* que le ha dedicado un par de monográficos recientemente (también *Estudios de Economía Aplicada* le dedicó un monográfico, en 2004 vol 22-3, al tópico de la economía espacial que está fuera de nuestra ventana de análisis). Esta debilidad se extiende también al tipo de papeles que han publicado estas revistas, casi exclusivamente aplicaciones con muy poca investigación de tipo metodológico.

Tabla 1¹

Artículos sobre Econometría espacial publicados en revistas de referencia (2007-2014).
Media anual

<i>Revistas econométricas</i>			<i>Revistas de ciencia regional</i>			<i>Revistas españolas</i>		
<i>Journal</i>	<i>2007-11</i>	<i>2012-14</i>	<i>Journal</i>	<i>2007-11</i>	<i>2012-14</i>	<i>Journal</i>	<i>2007-11</i>	<i>2012-14</i>
Econométrica	0.2	0.3	GA	4.8	4.3	EEA	0.8	0.3
JoE	4.0	4.7	JGS	1.8	4.7	IR	4.6	5.3
JAE	1.8	1.3	RS	1.2	2.0	EE	2.8	3.6
ER	0.2	3.0	PRS	6.8	3.0	REA	-	-
JTSE	0.2	-	JRS	1.2	4.7	SERIES	0.4	-
ET	1.0	1.7	ARS	3.2	6.7			
EJ	1.6	1.7	SEA	2.0	7.7			
			IRSR	4.0	2.0			
			RSUE	4.2	7.0			
TOTAL	9.0	12.7		29.2	42.1		8.6	9.2

JoE: *Journal of Econometrics*; JAE: *Journal of Applied Econometrics*; ER: *Econometric Reviews*; JTSE: *Journal of Time Series Econometrics*; ET: *Econometric Theory*; EJ: *Econometrics Journal*.

GA: *Geographical Analysis*; JGS: *Journal of Geographical Systems*; RS: *Regional Studies*; PRS: *Papers in Regional Science*; JRS: *Journal of Regional Science*; ARS: *Annals of Regional Science*; SEA: *Spatial Economic Analysis*; IRSR: *International Regional Science Review*; RSUE: *Regional Science and Urban Economics*.

EEA: *Estudios de Economía Aplicada*; IR: *Investigaciones Regionales*; EE: *Revista de Estadística Española*; REA: *Revista de Economía Aplicada*; SERIES: *Fusión de Investigaciones Económicas y Spanish Economic Review*.

Fuente: Elaboración propia.

¹ Los datos del periodo 2012-2014 incluyen papeles que, en octubre de 2014, aparecían como 'Aceptados, pendientes de publicación', en la página web de la revista.

3. MODELOS DE CORTE TRANSVERSAL

En el análisis regional es frecuente trabajar con ecuaciones donde se manejan datos que son estrictamente de naturaleza espacial como, por ejemplo, en modelos de crecimiento regional en los cuales el crecimiento de una región se pone en conexión con la dinámica previa de la propia región y con los resultados obtenidos por las regiones vecinas, con las que aquélla interactúa. Los modelos de precios hedónicos, en el caso de la vivienda, relacionan precios de venta con las características de la propia vivienda, con las características de las viviendas que la rodean, con las ‘*amenities*’ existentes en el vecindario y con los precios de venta de las viviendas situadas en las inmediaciones. Lo mismo puede decirse de los modelos de emisiones de contaminantes en los que resulta esencial conocer la disposición espacial del conjunto de la cadena productiva así como la de los mercados de referencia. En otros casos lo que destaca es la fuerte disparidad de los datos que se analizan como, por ejemplo, en los modelos de economía urbana o en los análisis de agrupamientos espaciales, donde las unidades de observación son tremendamente dispares entre sí.

Estos ejemplos muestran con claridad que si uno quiere modelizar datos espaciales deberá poner en cuestión el supuesto simplificador de que una observación es independiente de las restantes, e intercambiable con ellas. Al contrario, lo habitual es la interdependencia en forma de efectos de desbordamiento (acciones de terceros agentes, localizados en un punto del espacio, cuyas decisiones tienen efectos más allá de su propia región) o de mecanismos de interacción explícita entre los agentes, como en los modelos de competencia entre municipios por la prestación de servicios sociales. También la heterogeneidad es pauta común en este tipo de datos, lo cual puede distorsionar severamente el análisis. Todo ello tiene graves consecuencias que debilitan seriamente la calidad de las estimaciones y la inferencia obtenida en modelos no espaciales. No parece necesario insistir en el hecho de que, si uno está manejando datos de naturaleza espacial, los instrumentos usados tienen que ser consistentes con el entorno.

En general podemos distinguir tres tipos diferentes de efectos de interacción: entre los términos endógenos, proveniente de los términos exógenos o asociados al término de error no observable. En terminología econométrica, como indica en Anselin (1988), se trata de modelos con dependencia espacial sustantiva (*Spatial Lag Model* o SLM), con efectos de desbordamiento en los regresores (*Spatial Lag in X-Model* o SLX) o con dependencia en el error (*Spatial Error Model* o SEM).

El enfoque habitual consiste en iniciar la discusión con un modelo simple, no espacial y lineal, y contrastar si hay suficiente evidencia en su contra. En caso afirmativo, el analista extiende la especificación incorporando los efectos de interacción espacial necesarios. Este planteamiento es el denominado específico-a-general (abreviado como *Stge*). El enfoque opuesto es el general-a-espe-

cífico (denominado *Gets*) en el que la discusión se inicia con la especificación más general posible, donde deberían encontrarse anidados todos los modelos de interés incluyendo los distintos tipos de efectos espaciales que se considere. La misión del investigador será tratar de simplificar ese modelo general de anidación (*GNS*, en adelante) descartando hipótesis débiles y redundantes.

Un modelo no espacial típico es de la forma:

$$\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1)$$

donde \mathbf{y} es un vector columna, ($n \times 1$), con una observación para cada unidad de la muestra, \mathbf{x} es una matriz ($n \times k$) de observaciones de los regresores, $\boldsymbol{\beta}$ es un vector ($k \times 1$) de parámetros desconocidos y $\boldsymbol{\varepsilon}$ un vector ($n \times 1$) de perturbaciones aleatorias, habitualmente supuestas de tipo ruido blanco. El estimador MCO es la opción obvia para el modelo no espacial de (1).

Los efectos de interacción espacial endógena se producen porque, al tomar decisiones, los agentes localizados en la unidad i tienen en consideración las decisiones de los agentes localizados en otros puntos del espacio, y viceversa. Esta situación es típica, por ejemplo, en modelos de interacción estratégica para corporaciones locales donde se toman decisiones sobre impuestos, gasto público, etc., asumiendo que las demás corporaciones locales van a reaccionar a esas mismas decisiones (Brueckner 2003).

Los efectos de interacción también pueden ser exógenos y provenir de cierto regresor o conjunto de regresores. El grado de conservación de un inmueble es un factor determinante de su propio precio pero también afecta a los precios de los inmuebles circundantes. En los modelos de crecimiento y de localización espacial resulta habitual referirse a las externalidades generadas por ciertas infraestructuras y factores de localización esenciales, tales como nudos de comunicaciones o instituciones académicas, cuyo impacto trasciende las fronteras regionales (Ertur y Koch 2007).

El tercer tipo de interacción tiene lugar a través de los términos de error de la ecuación y no responden, necesariamente, a una hipótesis concreta; pueden estar asociados a desajustes entre las unidades administrativas (artificiales) y las áreas de influencia (en las que realmente se resuelven las relaciones), a la omisión de alguna variable relevante para la especificación, con cierta estructura espacial o, a la propagación de un shock no observable a través de una red urbana (Fingleton y López-Bazo, 2006).

Un modelo general con los tres tipos de efectos puede ser como el siguiente:

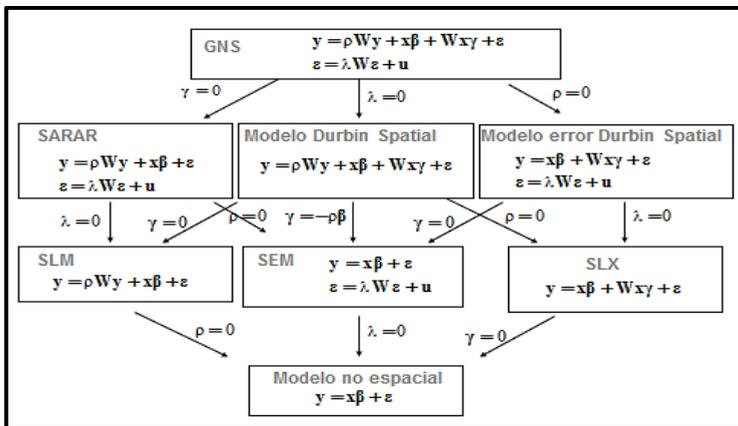
$$\left. \begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \mathbf{x}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda \mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u} \end{aligned} \right\} \quad (2)$$

donde $\mathbf{W}\mathbf{y}$ es el denominado *retardo espacial* de la variable endógena, que canaliza los mecanismos de interacción endógena que tienen lugar en los modelos;

los efectos de interacción exógena se concretan en el término Wx y los de interacción en el error en $W\varepsilon$. ρ y λ son los coeficientes de autocorrelación espacial asociados a los respectivos mecanismos de interacción, lo mismo que γ , vector de parámetros desconocidos de orden $(k \times 1)$. Por último, W es la matriz de pesos (también denominada de contigüidad, de contactos, ...) que habitualmente se supone no negativa y con valores cero en la diagonal principal. Esta matriz describe la distribución sobre el espacio de las unidades de observación (a su vez, se supone que la localización de estas unidades permanece fija).

En la Figura 1 detalla un catálogo de modelos potencialmente útiles en una aplicación estándar de base espacial. Estos modelos están relacionados mediante una estructura de anidación tipo *Gets* (la estructura *Sgte* sería la inversa)

Figura 1
Algunos modelos espaciales de interés. Estructura de anidación *Gets*



Fuente: Elaboración propia

La estructura de anidación (se lea en sentido descendente o ascendente) descansa en una batería de contrastes de especificación desarrollados, en su mayor parte, sobre el principio del Multiplicador de Lagrange (ver Anselin *et al.*, 1996, y Florax y de Graaff, 2004, para los detalles).

Estos modelos pueden estimarse utilizando diferentes algoritmos. De hecho, posiblemente sea este aspecto el que más contribuciones ha recibido en los últimos años. El método tradicional de estimación es, por supuesto, máxima verosimilitud, ML (Ord, 1975), aunque no deben olvidarse los estimadores cuasi-máximo verosímiles, QML (Lee 2004), los estimadores de variables instrumentales, IV (Liu y Lee, 2013), el método generalizado de los momentos, GMM (Kelejian y Prucha 1999, Fingleton y Le Gallo, 2008), los métodos bayesianos impulsados por LeSage y Pace (2009) ni los métodos semiparamétricos (Basile *et al.*, 2014) y no paramétricos (Robinson, 2011).

Para completar esta panorámica sobre los modelos econométricos de corte transversal, queremos volver a los tres puntos que mencionábamos en la sección de introducción, esto es causalidad, construcción de \mathbf{W} e identificación.

Uno de los aspectos más llamativos de la literatura sobre Econometría espacial es la práctica ausencia de todo lo que se refiere a causalidad. El problema no es menor porque, por ejemplo, el análisis de impactos (Lesage y Pace, 2009) que cuenta con una voluminosa literatura de aplicación, descansa sobre supuestos explícitos de causalidad entre variables. La literatura sobre metodología econométrica le ha dedicado mucha atención (Granger, 1980; Sims, 1980; Heckman, 1999 por mencionar solo unas pocas referencias). Hoover (2004) indica que más de la mitad de los artículos existentes en los archivos JSTOR, usan palabras pertenecientes a la '*familia causal*' (*causa, causas, causal, causalmente, causalidad, causación*). Sin embargo este debate no ha prendido en la literatura sobre Econometría espacial.

¿Qué tiene de especial este campo? Nada en nuestra opinión. En un corte transversal se carece de la perspectiva temporal de los datos, lo que dificulta el análisis al no poder utilizar el principio de precedencia temporal, pero el objetivo fundamental del investigador sigue siendo el mismo: (i) identificar relaciones causales significativas entre variables y (ii) identificar parámetros causales. A nuestro juicio, esta omisión tiene consecuencias fatales porque, al no estar clara la dirección de causalidad, buena parte de las especificaciones econométricas son débiles. Hay algunas excepciones en las que la dirección de causalidad es diáfana como el clima y los recursos naturales en modelos de asentamientos urbanos, pero lo habitual es lo contrario (por ejemplo, las amenidades en modelos de precios hedónicos de la vivienda). En un modelo tipo SLM como el siguiente:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3)$$

Bajo la cláusula '*todo lo demás constante*', el análisis de impactos nos dice que si cambiamos el valor de la variable x localizada en un punto concreto del espacio, en una unidad por ejemplo, observaremos cambios en la variable y en el resto de puntos del espacio. Algunos de ellos se producirán en el mismo lugar donde la variable x se ha alterado (este es el *efecto directo*) mientras que otros se distribuirán sobre el espacio (es el *efecto indirecto*). La suma de los dos es el denominado *efecto total*.

Sin embargo, la ecuación de (3) *solo* es una ecuación de correlación, útil para detectar regularidades sobre el espacio mientras que el análisis de multiplicadores precisa de la noción de causalidad. Partridge *et al.* (2012) exponen el problema con claridad:

“One common complaint is that standard spatial econometrics do not do a good job of differentiating when outcomes in nearby areas are spatially corre-

lated (perhaps due to common explanatory factors) versus spatial causality when outcomes in area A affect outcomes in area B - i.e., the notion of "spatial dependence" is too often applied in sloppy fashion. Both of these imply very different causal mechanisms." (p.168)

Gibbons y Overman (2012) abogan por desarrollar una metodología alternativa basada en lo que ellos denominan enfoque ‘*experimentalista*’ en oposición al dominante enfoque ‘*estructuralista*’ (en el que la teoría es la principal fuente de identificación del modelo). La sugerencia es interesante aunque difícilmente aplicable en el caso de la economía espacial en el que rara vez puede contarse con hechos contra-factuales. Herrera *et al.* (2014a) devuelven la cuestión a sus aspectos puramente estadísticos desarrollando un nuevo enfoque basado en el concepto tradicional de *contenido informativo incremental*, acuñado por Granger (1980). Por lo que acabamos de comentar, la situación en este ámbito concreto es muy pobre y debería corregirse. A efectos prácticos, es importante que los usuarios sean conscientes de esta debilidad y no se abuse de hipótesis simplificadoras y heroicas, en muchos casos.

El problema de construir la matriz de contactos ha estado siempre presente en la literatura sobre Econometría espacial. En el caso temporal, hay una solución obvia al problema de cómo estructurar la dependencia en base a los principios de causación y de precedencia temporal. Sin embargo, la situación es más compleja en el caso espacial donde las relaciones son, potencialmente, multidireccionales. Esta es la finalidad de la matriz **W**, captar esas relaciones.

Como se ha dicho, existe una abundante literatura dedicada a esa matriz (Harris *et al.*, 2011, y Corrado y Fingleton, 2012, para revisiones recientes) que se desarrolla en torno a tres preguntas clave: ¿cuál es la función de esa matriz?, ¿cómo se construye esa matriz? y, finalmente, en realidad ¿necesitamos esa matriz?

La respuesta a la primera cuestión es simple: los efectos de desbordamiento y de interacción espacial son fenómenos clave en la especificación de un modelo de corte transversal; basta con recordar el mantra enunciado por Tobler (1970): la primera ley de la geografía es que “*everything is related to everything else, but near things are more related than distant things*”. Sin embargo, estos efectos de desbordamiento no son observables por lo que necesitamos aproximarlos utilizando instrumentos y la matriz **W** juega un papel esencial aquí.

La respuesta a la segunda cuestión es bastante más compleja porque, de hecho, hay muy pocas restricciones que deban considerarse cuando se construye esta matriz (ceros en la matriz diagonal, no puede ser idempotente y debe estar asintóticamente acotada, Kapoor *et al.*, 2007). Bajo estas premisas, la decisión final corresponde al usuario lo que introduce un alto grado de arbitrariedad en la solución. Generalizando, pueden distinguirse dos grandes enfoques al problema de construir **W**:

- (i) Especificar la matriz de forma exógena.
- (ii) Especificar la matriz de forma endógena.

El exógeno es el enfoque más popular entre los usuarios y está basado en juicios previos, subjetivos en gran medida, relativos a la estructura geográfica de los datos. Ejemplos de este enfoque son el criterio de vecindad, los k-vecinos, los umbrales de distancia o los diferentes kernels en los que se utiliza la distancia como argumento (Harris *et al.*, 2011). El segundo enfoque toma en consideración tanto la topología del espacio como la naturaleza de los datos para construir la matriz a partir de ellos. Bodson y Peeters (1975) pueden considerarse los pioneros en este enfoque al que después se han sumado muchos otros como Getis y Aldstadt (2004), Aldstadt y Getis (2006) con el procedimiento AMOEBa, los filtros espaciales de Tiefelsdorf y Griffith (2007), el enfoque de máxima entropía de Fernández *et al.* (2009) o el criterio CCC de Mur y Paelinck (2011).

En los últimos años han surgido propuestas más ambiciosas con las que se pretende inferir la matriz de pesos directamente a partir de los datos utilizados en el análisis. Es una apuesta compleja por la gran cantidad de parámetros que es necesario evaluar. Bhattacharjee y Jensen-Butler (2013) y Beenstock y Feldsensen (2012) solucionan el problema en un contexto de datos panel mientras que Benjanuvatra (2012) permanece en un contexto puramente transversal. Otras propuestas para construir la matriz \mathbf{W} de forma no exógena pueden encontrarse en Autant-Bernard y LeSage (2011), Basile *et al.* (2012), Frenken *et al.* (2010), Maggioni *et al.* (2007), Mora y Moreno, (2010), o Ponds *et al.* (2007).

Añadimos un tercer enfoque, complementario con los anteriores, y que entronca con una observación muy atinada de Bavaud (1998, p.153): “*there is no such thing as ‘true’, ‘universal’ spatial weights, optimal in all situations*”; *in fact, the weighting matrix “must reflect the properties of the particular phenomena, properties which are bound to differ from field to field.”* En esta línea se contempla la cuestión de \mathbf{W} como un problema de selección de modelos. De hecho, matrices de contactos diferentes equivalen a conjuntos de regresores diferentes, sea en modelos estáticos o en dinámicos, y consecuentemente también a modelos diferentes. En este sentido deben entenderse las propuestas surgidas en torno al test J , como criterio para comparar modelos no anidados (Kelejian, 2008, Burridge y Fingleton, 2010, Burridge, 2012), la probabilidad a posteriori de Lesage (1997) o el test de Herrera *et al.* (2014b) basado en la entropía simbólica.

Una crítica habitual a la que se enfrentan los usuarios de Econometría espacial es que sus modelos no están identificados. De otro modo, el investigador asume tal cantidad de aprioris sobre el Proceso Generador de Datos, PGD, que las conclusiones no son otra cosa que auto-predicciones.

La polémica sobre la identificación tiene dos vertientes. La primera está relacionada, una vez más, con la matriz de contactos. Como proponen LeSage y Pace (2009), supongamos un proceso espacial autoregresivo no restringido, con solo tres observaciones:

$$\left. \begin{aligned} y_1 &= \alpha_{12}y_2 + \alpha_{13}y_3 + x_1\beta + u_1 \\ y_2 &= \alpha_{21}y_1 + \alpha_{23}y_3 + x_2\beta + u_2 \\ y_3 &= \alpha_{31}y_1 + \alpha_{32}y_2 + x_3\beta + u_3 \end{aligned} \right\} \quad (4)$$

y_i es la variable endógena en la unidad i , x_i y u_i son las variables exógenas y el término de error en la misma unidad; β y $\{\alpha_{ij}; i, j = 1, 2, 3\}$ son parámetros desconocidos donde $\alpha_{ii} = 0$. Hay seis parámetros y solo tres observaciones lo que significa que el sistema está sub-identificado. LeSage y Pace (2009, p.8) concluyen: “*The solution to the over-parameterization problem that arises when we allow each dependence relation to have relation-specific parameters is to impose structure on the spatial dependence.*” Esto es, necesitamos aportar una matriz de pesos al problema para poder re-escribirlo como:

$$\begin{bmatrix} \alpha_{12}y_2 + \alpha_{13}y_3 \\ \alpha_{21}y_1 + \alpha_{23}y_3 \\ \alpha_{31}y_1 + \alpha_{32}y_2 \end{bmatrix} = \begin{bmatrix} 0 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 0 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = Ay \Leftrightarrow \begin{bmatrix} 0 & \varpi_{12} & \varpi_{13} \\ \varpi_{21} & 0 & \varpi_{23} \\ \varpi_{31} & \varpi_{32} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \rho Wy \quad (5)$$

En esta expresión solo interviene un parámetro, ρ , aunque viene acompañado de muchas restricciones: $\rho = \alpha_{ij} / \varpi_{ij} = \alpha_{ik} / \varpi_{ik}$, que sirven para reducir el grado de incertidumbre del sistema (obviamente, se supone que son ciertas). Más importante, el sistema de (5) está identificado. Sin embargo hay que reconocer que esta solución solo es parcialmente satisfactoria, dada la fuerte carga que suponen las nuevas restricciones.

La segunda vertiente del problema de identificación tiene que ver con el ‘problema de la reflexión’ de Manski (1993), introducido en la literatura sobre interacción social pero con claros paralelismos en modelos espaciales. Como el propio Manski (1993) define el problema:

“*[It] arises when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in some group influences the behaviour of the individuals that comprise the group. The term reflection is appropriate because the problem is similar to that of interpreting the almost simultaneous movements of a person and his reflection in a mirror. Does the mirror image cause the person's movements or reflect them? An observer who does not understand something of optics and human behaviour would not be able to tell.*” (p. 532).

Analíticamente, supongamos un modelo de interacción como el siguiente:

$$\left. \begin{aligned} y_r &= \rho E[y_r|z_r] + x_r'\beta + E[x_r|z_r]'\gamma + u_r \\ u_r &= \varphi E[u_r|z_r] + v_r \end{aligned} \right\} \quad (6)$$

y es la variable que mide el resultado del individuo r , z son atributos que caracterizan al grupo de referencia del individuo r (aula de docencia, grupo étnico, vecinos, etc.) x miden las características observables del individuo r y u los atributos individuales no observables (habilidades personales, clima social de una región, etc.). Es habitual asumir que los atributos observables y no observables son ortogonales.

El parámetro ρ captura los *efectos endógenos*; esto es, hasta qué punto el resultado de una persona depende de la media de su grupo de referencia, definida por $E[y_r|z_r]$. El vector γ mide la porción de los resultados individuales que puede atribuirse a la media de las características exógenas de su grupo de referencia; son los *efectos exógenos contextuales*. β mide el impacto de las características exógenas del individuo sobre sus propios resultados. Finalmente φ captura los *efectos correlados no observables*, debido a los cuales los individuos del mismo grupo tienden a comportarse de forma similar porque también comparten características individuales, institucionales o medioambientales similares. Tras reagrupar términos y asumiendo ausencia de efectos correlados no observables, el problema de falta de identificación surge claramente en la forma reducida:

$$y_r = x_r'\beta + E[x_r|z_r]'(\rho\beta + \gamma)\frac{1}{1-\rho} + \frac{\rho}{1-\rho}E[u_r|z_r] + u_r \quad (7)$$

Únicamente el vector β está identificado. Para el resto de parámetros obtenemos diversas combinaciones que no permiten separar los efectos contextuales, γ , de los endógenos, ρ . Gibbons y Overman (2012) lo exponen crudamente:

“This is Manski’s ‘reflection problem’: only the overall effect of neighbors’ characteristics is identified, not whether they work through exogenous or endogenous neighborhood effects. These issues are intuitive: How can you distinguish between something unobserved and spatially correlated driving spatial correlation in y from the situation where y is spatially correlated because of direct interaction between outcomes? Further, how can you tell whether an individual is affected by the behavior of their group, or by the characteristics of their group when group behavior depends on the characteristics of the group? In many circumstances you cannot, without imposing further restrictions.” (p. 178).

4. MODELOS DE DATOS PANEL

La literatura sobre Econometría espacial de la última década se ha caracterizado, sobre todo, por el gran interés en el manejo de ecuaciones basadas en paneles, lo cual está relacionado con la mayor disponibilidad de bases de datos donde se combinan observaciones en el tiempo y en la sección transversal. Los datos panel abren muchas más posibilidades de modelización para el investigador en comparación con el ámbito, a menudo estrecho, de una simple ecuación de corte transversal. Además, el volumen de información muestral aumenta proporcionalmente con la serie temporal, contribuyendo a superar una de las restricciones tradicionales de los modelos espaciales: el déficit de información.

El paso de un modelo espacial como el de (2), en el que hay n observaciones en el corte transversal a un modelo espacio-temporal en el que a las n observaciones transversales añadimos T periodos temporales no plantea mayor problema, sin más que añadir un subíndice t dedicado al tiempo. Por ejemplo:

$$\left. \begin{aligned} \mathbf{y}_t &= \rho \mathbf{W} \mathbf{y}_t + \mathbf{x}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{x}_t \boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_t &= \lambda \mathbf{W} \boldsymbol{\varepsilon}_t + \mathbf{u}_t \end{aligned} \right\} \quad (8)$$

El modelo puede estimarse por cualquiera de los métodos habituales en un ámbito puramente transversal, supuesto que los elementos fundamentales de la especificación (parámetros, matrices de pesos) se mantienen constantes de un corte a otro y que estos cortes son independientes entre sí. Bajo el supuesto de independencia, la función de densidad conjunta se factoriza en el producto de las funciones de densidad marginal de cada uno de los T cortes, lo que aumenta la potencia de la estimación ML al haber más información muestral en el algoritmo.

La principal objeción contra la agrupación indiscriminada de los datos, el *pooling*, es que el modelo resultante no tiene en cuenta la heterogeneidad espacial ni la temporal. Está claro que las unidades espaciales pueden ser muy diferentes entre sí, entre otras razones, porque existen aspectos diferenciales no observables y persistentes en el tiempo (instituciones sociales, factores culturales, etc.) que interfieren en el funcionamiento de los modelos. Lo mismo puede ocurrir con los diferentes cortes, afectados por shocks no observables que afectan por igual a todos los individuos pero que cambian de periodo en periodo. Si no se tienen en cuenta estos efectos no observables, el riesgo de obtener estimaciones sesgadas e inconsistentes se incrementa (la literatura panel ha discutido mucho sobre ello; ver, por ejemplo, Hsiao 2003). Parece lógico, en consecuencia, incorporar estos dos tipos de efectos, extendiendo la especificación de (8) como:

$$\left. \begin{aligned} \mathbf{y}_t &= \rho \mathbf{W} \mathbf{y}_t + \mathbf{x}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{x}_t \boldsymbol{\gamma} + \boldsymbol{\eta} + \xi_t \boldsymbol{\tau} + \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_t &= \lambda \mathbf{W} \boldsymbol{\varepsilon}_t + \mathbf{u}_t \end{aligned} \right\} \quad (9)$$

siendo $\boldsymbol{\eta} = [\eta_1 \ \eta_2 \ \dots \ \eta_n]'$ el vector de efectos individuales no observables, $\boldsymbol{\tau}$ un vector ($n \times 1$) de unos y ξ_t el efecto temporal no observable correspondiente al periodo t .

Los efectos no observados pueden ser tratados como fijos o aleatorios lo que suele desembocar en una polémica árida que la literatura tradicional resuelve con el test de Hausman. Esta discusión, en lo que se refiere al tratamiento de los efectos individuales, ha tenido mayor interés en la literatura sobre Econometría espacial donde identificamos dos posturas encontradas. Anselin *et al.* (2006) se manifiestan abiertamente en contra del uso de efectos fijos dado que “(the) spatial models rely on asymptotics in the cross-sectional dimension (...), this would preclude the fixed effects model from being extended with a spatial lag or spatial error term”. Su recomendación son los efectos aleatorios. Sin embargo, debe recordarse que, para que este enfoque resulte aceptable, las unidades de observación deben ser representativas de una población. Consecuentemente, el número de unidades debe poder crecer indefinidamente y los efectos aleatorios no observados deben ser ortogonales con los efectos observados de los individuos. Estas tres circunstancias no son tan fáciles de verificar en una aplicación concreta.

Elhorst (2003) no ve ningún problema en el uso de efectos fijos dado que “the inconsistency of ρ is not transmitted to the estimators of the slope coefficients in the demeaned equation (...). This implies that the large sample properties of the fixed effects model when $n \rightarrow \infty$ do apply for the demeaned equation. (...). Hence, the fixed effects model is compelling, even when n is large and T is small”. Esta posición es compartida por otros autores como Beck (2001) para quien las unidades espaciales, en general, son entidades fijas con entidad propia, no el resultado de un muestreo previo, por lo que la inferencia debe ser de tipo condicional. Beenstock y Felsenstein (2007) abundan en el mismo argumento.

A pesar de estas consideraciones, la literatura de aplicación está claramente dominada por el enfoque de efectos aleatorios (Baltagi, 2005, Baltagi *et al.*, 2007, por citar algunos ejemplos) que, a menudo, se contempla como una opción de compromiso entre la rigidez del modelo *pool* y la exhaustividad del modelo de efectos fijos. Es decir, mantiene la simplicidad de la especificación *pool* pero aporta más información al usuario. Otra consideración importante es que el modelo de efectos aleatorios evita la pérdida de grados de libertad en casos en los cuales la dimensión transversal crece rápidamente mientras que la temporal permanece finita. Además, los efectos fijos individuales solo pueden

estimarse consistentemente si se dispone de un T suficientemente grande, lo cual supone una limitación grave si el objetivo del análisis pasa por interpretar esos efectos. Por el contrario, si el objetivo es obtener buenas estimaciones del resto de coeficientes, la inconsistencia del estimador de los efectos fijos (en caso de T pequeño) en ningún caso se transmite a los de las pendientes. Finalmente, el enfoque de efectos aleatorios evita los problemas creados por la presencia en la especificación de regresores invariantes o muy estables en el tiempo, cuyo impacto no puede tratarse adecuadamente en un enfoque de efectos fijos tradicional.

Como indica Elhorst (2014), la cuestión de si resulta preferible un modelo de efectos aleatorios o de efectos fijos sigue abierta.

Aparte de estas precisiones sobre la interpretación de los efectos no observables en un modelo panel con estructura espacial, la discusión metodológica posterior no difiere en lo sustancial de la correspondiente a un modelo de corte transversal tradicional. Una vez que el usuario ha decidido qué enfoque resulta preferible (efectos fijos o aleatorios) se trata de poner en marcha una estrategia de construcción del modelo tipo *Gets* o *Stge*, similar a la mostrada en la Figura 1. Obviamente, los contrastes de especificación correspondientes a este caso tienen que ser los adecuados al contexto panel en el que se encuadra la discusión. La batería de contrastes disponible en estos momentos es amplia, desarrollada en la última década en base a los trabajos, entre otros, de Baltagi (2005), Baltagi *et al.* (2007), Elhorst (2003, 2008, 2012b), Kapoor *et al.* (2007) o Lee y Yu (2010a y b, 2012a y b).

La conversión de un modelo estático, como el de (8), en un modelo dinámico con estructura espacial aumenta significativamente la potencia del análisis al combinar ambas dimensiones, temporal y espacial, en un solo instrumento. Una ecuación típica podría ser como la siguiente:

$$\left. \begin{aligned} (a) \quad & \mathbf{y}_t = \alpha \mathbf{y}_{t-1} + \rho_0 \mathbf{W} \mathbf{y}_t + \rho_1 \mathbf{W} \mathbf{y}_{t-1} + \mathbf{x}_t \boldsymbol{\beta} + \mathbf{W} \mathbf{x}_t \gamma_0 + \mathbf{W} \mathbf{x}_{t-1} \gamma_1 + \mathbf{v}_t \\ (b) \quad & \mathbf{v}_t = \phi \mathbf{v}_{t-1} + \lambda \mathbf{W} \mathbf{v}_t + \boldsymbol{\eta} + \xi_t \boldsymbol{\tau} + \boldsymbol{\varepsilon}_t \\ (c) \quad & \boldsymbol{\eta} = \kappa \mathbf{W} \boldsymbol{\eta} + \boldsymbol{\zeta} \end{aligned} \right\} \quad (10)$$

Obviamente, la estructura dinámica y espacial de (10) es ajustable a las necesidades del investigador. Elhorst (2014) sugiere un término de error, \mathbf{v}_t , compuesto de correlación serial, $\phi \mathbf{v}_{t-1}$, correlación espacial contemporánea, $\lambda \mathbf{W} \mathbf{v}_t$, junto a efectos individuales y temporales no observables, $\boldsymbol{\eta}$ y $\xi_t \boldsymbol{\tau}$, y un términos de error idiosincrático, $\boldsymbol{\varepsilon}_t$. Los efectos individuales y temporales pueden tratarse como efectos fijos o aleatorios dependiendo de cómo se inter-

prete la especificación. En el último caso, es razonable suponer que también los efectos individuales mantengan cierta estructura espacial.

Un aspecto interesante de (10) son las condiciones de deben verificar los parámetros para asegurar estacionariedad en la especificación. En la ecuación (10c), las condiciones sobre K son estrictamente de naturaleza espacial y son las usuales:

$$\frac{1}{\omega_n^-} < \kappa < \frac{1}{\omega_n^+} \quad (11)$$

siendo ω_n^- y ω_n^+ los valores propios de \mathbf{W} más grandes, negativo y positivo respectivamente, para la muestra de tamaño n . En el caso de la ecuación (10b), Elhorst (2008) demuestra que los valores propios de la matriz $\phi(\mathbf{I} - \lambda\mathbf{W})^{-1}$ tienen que estar dentro del círculo unidad lo que equivale a:

$$\left. \begin{array}{l} |\phi| < 1 - \lambda\omega_n^+ \quad \text{si } \lambda \geq 0 \\ |\phi| < 1 - \lambda\omega_n^- \quad \text{si } \lambda < 0 \end{array} \right\} \quad (12)$$

Existe un efecto de retroalimentación entre los parámetros de dependencia temporal y espacial, en el que también interviene la configuración espacial de la muestra a través de los valores propios de la matriz de pesos \mathbf{W} . Finalmente Parent y LeSage (2010) obtienen las condiciones de estacionariedad para la ecuación (10a). Nuevamente los valores propios, en este caso de la matriz $(\alpha\mathbf{I} + \rho_1\mathbf{W})(\mathbf{I} - \rho_0\mathbf{W})^{-1}$, tienen que estar dentro del círculo unidad lo que equivale a:

$$\left. \begin{array}{l} \alpha > 1 - (\rho_0 + \rho_1)\lambda\omega_n^+ \quad \text{si } \rho_0 + \rho_1 \geq 0 \\ \alpha > 1 - (\rho_0 + \rho_1)\lambda\omega_n^- \quad \text{si } \rho_0 + \rho_1 < 0 \end{array} \right\} \quad (13)$$

El resultado de estacionariedad involucra a los coeficientes de correlación serial con los de correlación espacial. En concreto, el primero tiene que ajustarse a la suma de los dos coeficientes de correlación espacial, filtrada por los valores propios de la matriz de pesos. Las condiciones de (13) son menos estrictas que las avanzadas tanto por Yu *et al.* (2008), $|\alpha| + |\rho_0| + |\rho_1| < 1$, como por Lee y Yu (2010b), $\alpha + \rho_0 + \rho_1 < 1$. De hecho, ninguno de los dos toma en consideración el impacto de la matriz de pesos. Más detalles sobre la cuestión de la estacionariedad en este tipo de modelos pueden encontrarse en Elhorst (2012a, 2014).

La estimación de modelos panel, dinámicos o estáticos, con estructura espacial empieza a ser una cuestión relativamente estándar a pesar de la complejidad inherente al problema. En general, predominan los enfoques basados en la función de verosimilitud del modelo (Anselin *et al.*, 2006, Baltagi, 2005, Elhorst 2003, 2010b) o en su versión quasi-máximo verosímil más flexible (Lee, 2004, Lee y Yu, 2010b). El método LSDVC, mínimos cuadrados de efectos fijos corregidos, de Kiviet (1985) ha sido adaptado al caso de paneles estáticos por Lee y Yu (2010a) y Baltagi (2005). Kapoor *et al.* (2007) introducen el método GMM para tratar modelos estáticos con estructura espacial en el error, que después se ha generalizado también para modelos dinámicos (Kukenova y Monteiro, 2009, Baltagi *et al.*, 2014). Baltagi y Liu (2011) desarrollan un algoritmo basado en variables instrumentales con buenas propiedades para el caso de modelos estáticos. Finalmente, Parent y Lesage (2010) introducen el algoritmo de estimación bayesiana MCMC, adaptable tanto a modelos dinámicos como a modelos estáticos.

Es decir, existen suficientes métodos de estimación, con buenas propiedades, para atender las diferentes necesidades que puedan presentarse cuando se están utilizando datos panel con estructura espacial. Una restricción importante que el investigador debe considerar es que, en todos ellos, se está asumiendo implícitamente la constancia de la matriz de pesos \mathbf{W} a lo largo del tiempo. Sin embargo, es evidente que esta matriz puede evolucionar debido, por ejemplo, a cambios en el entorno. Esta faceta está poco explorada en la literatura, aunque Lee y Yu (2012a) y Qua y Lee (2015) demuestran que es posible endogeneizar el comportamiento de esa matriz, en función de otras variables del entorno, y obtener estimadores consistentes y asintóticamente normales de los parámetros de un modelo (estático, dinámico) utilizando un enfoque quasi-máximo verosímil.

Por último, en lo que respecta a software, existe una buena oferta de código para trabajar con modelos panel (Elhorst, 2012a, para más detalles). Los usuarios de MATLABTM cuentan con fuentes muy útiles en los sitios http://oak.cats.ohiou.edu/*lacombe/research.html, de Donald Lacombe, www.regroningen.nl, de Paul Elhorst, y <http://www.spatial-econometrics.com/> de James Lesage. La oferta en R, por ahora, se limita al manejo de modelos estáticos con el paquete *splm* (Millo y Piras, 2013) y los recursos en STATATM se concentran en los paquetes *xsmle*, para modelos estáticos, y *spregdpd* para modelos dinámicos.

5. EPÍLOGO 2004-2014

En los últimos diez años, desde nuestra revisión anterior sobre la situación de la econometría espacial, se han producido un buen número de aportaciones en el campo. La complejidad es una de las ideas que ha tomado más fuerza. Un

caso especial de complejidad es la no convexidad, que puede considerarse como un reto de singular importancia para los practicantes de la Econometría donde, por ejemplo, la existencia de regímenes múltiples es una situación normal. En el análisis espacial, la no convexidad es habitual en dinámicas micro-regionales y de tipo local.

Los modelos espacio-temporales han sido un instrumento clásico de investigación en Física teórica y aplicada, y resulta natural pensar en su adaptación en el campo de la Econometría espacial. En aquél caso, las estructuras espacio-temporales se han representado utilizando ecuaciones diferenciales parciales (EDP en adelante). Si consideremos solo el tiempo, t , y el espacio, x , interactuando en una función $g(x,t)$, una EDP para puede escribirse como:

$$h(x; t; g; g_x; g_t; g_{xx}; g_{tt}; g_{xt}; \dots) = 0 \quad (14)$$

Siendo $h(-)$ función de las variables tiempo y espacio, de una función desconocida, g , y de un cierto número finito de sus derivadas parciales. Un miembro de esta familia genérica es la ecuación de ondas, especificada como:

$$\ddot{f}(x; t) = \alpha^2 f''(x, t) \quad (15)$$

donde el doble punto, \ddot{f} , significa la segunda derivada en el tiempo (esto es, aceleración) y la doble prima, f'' , su segunda derivada en el espacio (esto es, curvatura), α^2 es un parámetro estrictamente positivo. La expresión de (15) es una ecuación de interacción local. Sin embargo, en Economía espacial (lo mismo que en Física cuántica), la dimensión local no es la norma: Domina la interacción global para lo que necesitamos modelos de interacción espacial más generales como:

$$\ddot{f}(x; t) = \alpha^2 \int_{-1}^{+1} w(x, \xi) f''(\xi, t) d\xi \quad (16)$$

siendo $w(x, \xi)$ la denominada '*función de descuento espacial*'. La interacción de esta función con una variable expresa el concepto de potencial en el intervalo $[-1; +1]$, por lo que la ecuación de (16) debería denominarse ecuación diferencial parcial potencializada, EDPP.

En un contexto finito, la ecuación EDPP de (16) puede re-escribirse como:

$$\Delta_t^2 f(x; t) = \alpha^2 \sum_{-1}^{+1} w(x; \xi) \Delta_t^2(\xi; t) \quad (17)$$

El resultado del sumatorio depende, obviamente, del proceso de interacción espacial seleccionado (ver, por ejemplo, Kaashoek y Paelinck, 1996). Este modelo se ha aplicado en diferentes ocasiones, por ejemplo, al caso de la región Rhône-Alpes (Coutrot *et al.*, 2009), con el objetivo de analizar el desarrollo de industrias basadas en el conocimiento a través del empleo en actividades rela-

cionadas con este tipo de industrias. Se disponía de 3 observaciones temporales para 39 ciudades de la región y los primeros resultados indicaban la existencia de, al menos, dos regímenes (Griffith y Paelinck, 2011) para lo que es necesario una ecuación como:

$$\Delta_t^2 \ln(n_{t0i}) = \lambda [a \Delta \ln(n_{0i}) + b \Delta \ln(n_{1i}) + c \Delta \ln(n_{3i})] + (1 - \lambda) [a^* \Delta \ln(n_{0i}) + b^* \Delta \ln(n_{1i}) + c^* \Delta \ln(n_{3i})] \quad (18)$$

λ es una variable binaria que identifica el cambio de régimen en un caso típico de no convexidad en el análisis espacial. Los principales resultados de la estimación de ese modelo aparecen en la Tabla (2), en la columna I los correspondientes al modelo restringido y en la columna II los del modelo ampliado a dos regímenes.

Tabla 2
No convexidad en Rhône-Alpes. Modelo de dos regímenes

Parámetros	(I)	(II)
a	-0.0041	0.0122
b	-0.0049	0.0072
c	0.0002	-0.0015
a*	-	-0.0059
b*	-	-0.0003
c*	-	0.0008
(Pseudo-)R ²	0.5156	0.9990

Fuente: Elaboración propia.

Los resultados son interesantes desde diferentes puntos de vista. En primer lugar, los regímenes tienen signos opuestos. En segundo lugar, la estimación es, casi, una interpolación por lo que todos los coeficientes deberían ser altamente significativos. Finalmente, estos resultados proceden de una especificación en diferencias de segundo orden.

Este ejemplo sirve para subrayar nuestra apuesta por el desarrollo de nuevos instrumentos, como el modelo en diferenciales parciales, adaptados a la casuística particular del análisis económico espacial en el que lo micro tiene especial relevancia.

6. CONCLUSIONES Y PERSPECTIVAS DE FUTURO

Está claro que el pequeño mundo de la Econometría espacial ha evolucionado muy rápidamente en las últimas décadas. Ha pasado de constituir una disciplina marginal que concitaba el interés de un reducido grupo de usuarios, a posicionarse en el núcleo central de la investigación en economía cuantitativa. Como usuarios, lo más satisfactorio que nosotros podemos decir es que, en la actualidad, ya no tenemos que explicar ‘qué es eso de la matriz *W* y para qué

vale?. Es decir, la disciplina se ha normalizado con todas sus limitaciones y restricciones.

Evidentemente quedan muchas cuestiones pendientes. En este trabajo hemos aludido a tres que nos preocupar especialmente como son el análisis de causalidad, el problema de la identificación de los modelos espaciales y la construcción de la matriz \mathbf{W} . Estamos seguros que existen otras muchas igualmente interesantes (cointegración y estructura espacial, difusión de la volatilidad sobre el espacio, topología y redes espaciales, etc.). En cualquier caso, uno de las urgencias más importantes en estos momentos, en los que el uso de los modelos espaciales se ha generalizado, sea insistir en el buen uso de todos estos instrumentos.

REFERENCIAS

- ALDSTADT J. y GETIS, A. (2006). "Using AMOEBA to Create a Spatial Weight Matrix and Identify Spatial Clusters". *Geographical Analysis*, 38(4), pp. 327-343.
- ANSELIN, L. (1988). *Spatial econometrics: methods and models*. Dordrecht: Kluwer.
- ANSELIN, L. (2010). "Thirty Years of Spatial Econometrics". *Papers in Regional Science*, (89)1, pp. 3-25.
- Anselin, L.; Bera, A.K.; Florax, R. y Yoon. M.J. (1996). "Simple diagnostic tests for spatial dependence". *Regional Science and Urban Economics*, 26(1), pp. 77-104.
- ANSELIN, L.; FLORAX, R. y REY, S. (eds.) (2004). *Advances in Spatial Econometrics. Methodology, Tools and Applications*. Berlin: Springer.
- ANSELIN, L.; LE GALLO, J. y JAYET, H. (2006). "Spatial Panel Econometrics". En Mátyás, L. y Sevestre, P. (eds): *The Econometrics of Panel Data* (pp 625-660). Berlin: Springer.
- ANSELIN, L. y REY, S. (eds.) (2010). *Perspectives on Spatial Data Analysis*. Berlin: Springer-Verlag.
- ARBIA, G. (2006). *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Berlin: Springer-Verlag.
- ARBIA, G. (2011). "A Lustrum of SEA: Recent Research Trends Following the Creation of the Spatial Econometrics Association (2007-2011)". *Spatial Economic Analysis*, 6 (4), pp. 377-395.
- AUTANT-BERNARD, C. y LESAGE, J. (2011). "Quantifying Knowledge Spillovers Using Spatial Econometric Models". *Journal of Regional Science*, 51(3), pp. 471-496.
- BALTAGI, B. H. (2005). *Econometric analysis of panel data* (3ed.) Chichester: Wiley.
- BALTAGI, B. H.; SONG, S. H.; JUNG, B. C. y KOH, W. (2007). "Testing for serial correlation, spatial autocorrelation and random effects using panel data". *Journal of Econometrics* 140(1), pp. 5-51.
- BALTAGI, B. H. y LIU, L. (2011). "Instrumental variable estimation of a spatial autoregressive panel model with random effects". *Economic Letters*, 111, pp. 135-137.

- BALTAGI, B. H.; FINGLETON, B. y PIROTTE, A. (2014). "Estimating and Forecasting with a Dynamic Spatial Panel Data Model". *Oxford Bulletin of Economics and Statistics*. 76(1), pp. 112-138.
- BASILE, R.; CAPELLO, R. y CARAGLIU, A. (2012). "Technological Interdependence and Regional Growth in Europe". *Papers in Regional Science*, 91(4), pp. 697-722.
- BASILE, R.; DURBÁN, M.; MÍNGUEZ, R.; MONTERO, J. M. y MUR, J. (2014): "Modeling regional economic dynamics: spatial dependence, spatial heterogeneity and nonlinearities". *Journal of Economic Dynamics & Control*. (forthcoming).
- BAVAUD, F. (1998). "Models for Spatial Weights: A Systematic Look". *Geographical Analysis*, 30(2), pp. 153-171.
- BECK, N. (2001). "Time-series-cross-section data: What have we learned in the past few years?". *Annals of Review Political Science*, 4, pp. 271-293.
- BEENSTOCK, M. y FELSENSTEIN, D. (2007). "Spatial vector autoregressions". *Spatial Economic Analysis*, 2(2), pp. 167-196.
- BEENSTOCK, M. y FELSENSTEIN, D. (2012). "Nonparametric estimation of the spatial connectivity matrix using spatial panel data". *Geographical Analysis*, 44, pp. 386-397.
- BHATTACHARJEE, A. y JENSEN-BUTLER, C. (2013). "Estimation of the Spatial Weights Matrix under Structural Constraints". *Regional Science and Urban Economics*, 43, pp. 617-634.
- BENJANUVATRA, S. (2012). *Essays on Spatial Econometrics*. PhD Thesis: University of York.
- BODSON, P. y PEETERS, D. (1975). "Estimation of the Coefficients of a Linear Regression in the Presence of Spatial Autocorrelation: An Application to a Belgian Labour-demand Function". *Environment and Planning*, 7(4), pp. 455-472.
- BRUECKNER, J. K. (2003). "Strategic interaction among local governments: An overview of empirical studies". *International Regional Science Review*, 26(2), pp.:175-188.
- BURRIDGE, P. (2012). "Improving the J test in the SARAR model by likelihood estimation". *Spatial Economic Analysis*, 7(1), pp. 75-107.
- BURRIDGE, P. y FINGLETON, B. (2010). "Bootstrap inference in spatial econometrics: the J-test". *Spatial Economic Analysis I*, 5, pp. 93-119.
- CLIFF, A. y ORD, K. (1981). *Spatial Processes: Models & Applications*. London: Pion.
- CORRADO, L. y FINGLETON, B. (2012). "Where is the economics in spatial econometrics?". *Journal of Regional Science*, 52(2), pp. 210-239.
- COUTROT, B.; PAELINCK, J.; SALLEZ, A. y SUTTER, R. (2009). "On Potentialized Partial Finite Difference equations: Analyzing the Complexity of Knowledge-Based Spatial Economic Activities". *Région et Développement*, 29, pp. 201-228.
- ELHORST, J. P. (2003). "Specification and estimation of spatial panel data models". *International Regional Science Review*, 26(3), pp. 244-268.
- ELHORST, J. P. (2008). "Serial and spatial autocorrelation". *Economic Letters*, 100(3), pp. 422-424.
- ELHORST, J.P. (2010). "Dynamic panels with endogenous interaction effects when T is small". *Regional Science and Urban Economics*, 40(5), pp. 272-282.
- ELHORST, J. P. (2012a). "Matlab software for spatial panels". *International Regional Science Review*, 37(3), pp. 389-405.

- ELHORST, J. P. (2012b). "Dynamic spatial panels: models, methods, and inferences". *Journal of Geographical Systems*, 14(1), pp. 5-28.
- ELHORST, J. P. (2014). *Spatial Econometrics. From Cross-Sectional Data to Spatial Panels*. Berlin: Springer.
- ERTUR, C. y KOCH, W. (2007). "Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence". *Journal of Applied Econometrics*, 22(6), pp. 1033-1062.
- FERNÁNDEZ, E.; MAYOR, M. y RODRÍGUEZ, J. (2009). "Estimating Spatial Autoregressive Models by GME-GCE Techniques". *International Regional Science Review*, 32(2), pp. 148-172.
- FINGLETON, B. y LE GALLO, J. (2008). "Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties". *Papers in Regional Science*, 87, pp. 319-339.
- FINGLETON, B. y LÓPEZ-BAZO, E. (2006). "Empirical growth models with spatial effects". *Papers in Regional Science*, 85, pp. 177-198.
- FLORAX, R. y DE GRAAFF, Th. (2004). "The performance of diagnostic tests for spatial dependence in linear regression models: A meta-analysis of simulation studies". En Anselin, L.; Florax, R. y Rey, S. (ed.): *Advances in Spatial Econometrics*. Berlin: Springer, pp. 29-67.
- FRENKEN, K., PONDS, R. y VAN OORT, F. (2010). "The Citation Impact of Research Collaboration in Science-based Industries: A spatial-institutional Analysis". *Papers in Regional Science*, 89(2), pp. 351-371.
- GETIS, A. y ALDSTADT, J. (2004). "Constructing the Spatial Weight Matrix Using a Local Statistic". *Geographical Analysis*, 36(2), pp. 90-104.
- GETIS, A.; MUR, J.; ZOLLER H. (eds.) (2003). *Spatial Econometrics and Spatial Statistics*. Basingstoke: Palgrave Macmillan.
- GIBBONS, S. y OVERMAN, H. G. (2012). "Mostly pointless spatial econometrics?". *Journal of Regional Science*, 52(2), pp.172-191.
- GRANGER, C. (1980). "Testing for Causality: A Personal Viewpoint". *Journal of Economic Dynamics and Control*, 2(1), pp. 329-352.
- GRIFFITH, A. (2003). *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization*. Berlin: Springer-Verlag.
- GRIFFITH, D. y PAELINCK, J. (2009). "Specifying a joint space- and time-lag using a bivariate Poisson distribution". *Journal of Geographical Systems*, 11(1), pp. 23-36.
- GRIFFITH, A. y PAELINCK, J. (2011). *Non-Standard Spatial Statistics and Spatial Econometrics*. Berlin: Springer-Verlag.
- HAINING, R. (2003). *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University Press.
- HARRIS, R.; MOFFAT, J. y KRAVTSOVA, V. (2011). "In Search of W". *Spatial Economic Analysis*, 6(3), pp. 249-270.
- HECKMAN, J. (1999). *Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective*. New York: NBER, Working paper n. 7333.
- HERRERA, M.; MUR, J. y RUIZ, M. (2014a). "Detecting Causal Relationships Between Spatial Processes". *Papers in Regional Science* (forthcoming).

- HERRERA, M.; MUR, J. y RUIZ, M. (2014b). *Selecting the Most Adequate Spatial Weighting Matrix: A Study on Criteria*. Trabajo presentado en el VII Seminario Jean Paelinck, Zaragoza, Noviembre
- HOOVER, K. (2004). "Lost causes". *Journal of the History of Economic Thought*, 26(2), pp. 149-164.
- HSIAO, Ch. (2003). *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- KAASHOEK, J. y PAELINCK, J. (1996). "Studying the Dynamics of Pre-geographical Space by Means of Space- and Time-Potentialised Partial Differential Equations". *Geographical Systems*, 3, pp. 259-277.
- KAPOOR, M.; KELEJIAN, H. y PRUCHA, I. (2007). "Panel Data Models with Spatially Correlated Error Components". *Journal of Econometrics*, 140(2), pp. 97-130.
- KELEJIAN, H. H. (2008). "A spatial J-test for model specification against a single or a set of non-nested alternatives". *Letters in Spatial and Resource Sciences*, 1(1), pp. 3-11.
- KELEJIAN, H. H. y PRUCHA, I. R. (1999). "A generalized moments estimator for the autoregressive parameter in a spatial model". *International Economic Review*, 40(2), pp. 509-533.
- KIVIET, J. (1985). "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models". *Journal of Econometrics*, 68(1), pp. 53-78.
- KUKENOVA, M. y MONTEIRO, J. A. (2009). Spatial dynamic panel model and system GMM: A monte carlo investigation. <http://ideas.repec.org/p/pram/prapa/11569.html>
- LEE, L. F. (2004). "Asymptotic distribution of quasi-maximum likelihood estimators for spatial autoregressive models". *Econometrica*, 72(6), pp. 1899-1925.
- LEE, L. F. y YU, J. (2010a). "Estimation of spatial autoregressive panel data models with fixed effects". *Journal of Econometrics*, 154(2), pp. 165-185.
- LEE, L. F. y YU, J. (2010b). "Some recent developments in spatial panel data models". *Regional Science and Urban Economics*, 40, pp. 255-271.
- LEE, L. F. y YU, J. (2012a). "QML estimation of spatial dynamic panel data models with time varying spatial weights matrices". *Spatial Economic Analysis*, 7(1), pp. 31-74.
- LEE, L. F. y YU, J. (2012b). "Spatial panels: Random components versus fixed effects". *International Economic Review*, 53, pp. 1369-1388.
- LESAGE, J.P. (1997). "Bayesian estimation of spatial autoregressive models". *International Regional Science Review*, 20, pp. 113-129.
- LESAGE, J. P. y PACE, R. K. (2009). *Introduction to spatial econometrics*. Boca Raton: CRC Press, Taylor & Francis Group.
- LIU, X. y LEE, L. F. (2013). "Two-stage least squares estimation of spatial autoregressive models with endogenous regressors and many instruments". *Econometric Review* 32(5-6), pp. 734-753.
- MAGGIONI, M.; NOSVELLI, M. y UBERTI, T. E. (2007). "Space Versus Networks in the Geography of Innovation: A European Analysis". *Papers in Regional Science*, 86(3), pp. 471-493.
- MANSKI, C. (1993). "Identification of Endogenous Social Effects: The Reflection Problem". *The Review of Economic Studies*, 60(3), pp. 531-542.
- MILLO, G. y PIRAS, G. (2012). "Splm: spatial panel data models in R". *Journal of Statistical Software*, 47(1), pp. 1-38.

- MORA, T. y MORENO, R. (2010). "Specialisation Changes in European Regions: The Role Played by Externalities Across Regions". *Journal of Geographical Systems*, 12(3), pp. 311-334.
- MUR, J. y PAELINCK, J. (2011). "Deriving the W-matrix via p-median Complete Correlation Analysis of Residuals". *Annals of Regional Science*, 47(2), pp. 253-267.
- MUR, J. y PAELINCK, J. (2013). "Theoretical Spatial Economics and Spatial Econometrics: Time-and-Space Non-Convexities Galore". *Estadística Española*, 180, pp. 2-24.
- ORD, K. (1975). "Estimation methods for models of spatial interaction". *Journal of the American Statistical Association*, 70, pp. 120-126.
- PAELINCK, J. y KLAASSEN, L. (1979). *Spatial Econometrics*. Farnborough: Saxon House.
- Paelinck, J.; Mur, J. y Trávez, J. (2004). "Econometría Espacial: más luces que sombras". *Estudios de Economía Aplicada*, 22(3), pp. 383-404.
- PARENT, O. y LESAGE, J. P. (2010). "A spatial dynamic panel model with random effects applied to commuting times". *Transportation Research Part B* 44, pp. 633-645.
- PARTRIDGE, M.; BOARNET, M.; BRAKMAN, S. y OTTAVIANO, G. (2012) "Introduction: Whither Spatial Econometrics". *Journal of Regional Science*, 52(2), pp. 167-171.
- PINKSE, J. y SLADE, M. (2010). "The Future of Spatial Econometrics". *Journal of Regional Science*, 50, pp. 103-117.
- PONDS, R.; VAN OORT, F. y FRENKEN, K. (2007). "The Geographical and Institutional Proximity of Research Collaboration". *Papers in Regional Science*, 86(3), pp. 423-443.
- QUA, X. y LEE, L. (2015). "Estimating a spatial autoregressive model with an endogenous spatial weight matrix". *Journal of Econometrics*, 184(1), pp. 209-232.
- ROBINSON, P. (2011). "Asymptotic theory for nonparametric regression with spatial data". *Journal of Econometrics*, 165(1), pp. 5-19.
- SIMS, C. (1980). "Macroeconomics and Reality". *Econometrica*, 48(1), pp. 1-48.
- TIEFELSDORF, M. y GRIFFITH, D. (2007). "Semiparametric Filtering of Spatial Autocorrelation: The Eigenvector Approach". *Environment and Planning A*, 39(5), pp. 1193-1221.
- TOBLER, W. (1970). "A Computer Movie Simulating Urban Growth in the Detroit Region". *Economic Geography*, 46(2), pp. 234-240.
- YU, J.; DE JONG, R. y LEE, L. (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), pp. 118-134.