

MODELIZACIÓN ESPACIAL DINÁMICA CON DATOS DE PANEL: APLICACIÓN AL CASO DE LOS PRECIOS DE LA VIVIENDA EN INGLATERRA

Bernard FINGLETON

Universidad de Cambridge

Alain PIROTTE

Universidad de Paris II Panthéon-Assas y TEPP (CNRS)

Resumen

Los modelos espaciales dinámicos incorporan efectos de «memoria», lo que implica persistencia temporal en los datos e interacción contemporánea entre las diferentes localizaciones. La variación del precio de la vivienda en Inglaterra es ideal para este tipo de modelos, los cuales utilizaremos para estimar y predecir los precios de la vivienda en 353 distritos ingleses, incluyendo dos variantes con y sin interacción espacial autorregresiva entre los errores. El modelo preferido muestra una dependencia significativa tanto en los errores como en el tiempo, pero entre los precios de los distritos solo se obtiene una interdependencia contemporánea marginalmente significativa. Bajo el modelo preferido, tanto la demanda como la oferta de vivienda es significativa y presenta el signo apropiado.

Palabras clave: panel dinámico espacial, precios de la vivienda, predicción, interacción espacial.

Abstract

Dynamic spatial panel models incorporate 'memory' effects, so that the dependent variable depends on its level in the previous period, together with contemporaneous interaction involving observations in different locations. English house price variation is ideally suited to this type of estimator, which we use to estimate and forecast house prices across 353 English districts, using variants both with and without autoregressive spatial interaction among the errors. The preferred model shows significant error dependence, and significant temporal dependence, but only marginally significant contemporaneous price interdependence across districts. Under the preferred model both demand for, and supply of, housing is significant and appropriately signed.

Key words: dynamic spatial panel, house prices, prediction, spatial interaction.

JEL classification: C33, C53, R31, R32.

I. INTRODUCCIÓN

EN los últimos años se han producido avances considerables en los modelos que trabajan con datos de panel mediante extensiones de la metodología estándar, basadas no solo en la incorporación de efectos dinámicos, sino también en las interrelaciones espaciales entre los individuos que forman la muestra analizada, ya se traten estos de regiones, ciudades o países; en particular, cabe mencionar los trabajos de LeSage y Pace (2009), Baltagi (2013), Elhorst (2014) y Pesaran (2015). La diferencia respecto a los modelos convencionales para datos de panel radica en que éstos buscan explicar la variación de la variable dependiente; por ejemplo, los precios medios de la vivienda observados a lo largo del tiempo en distintas regiones inglesas, a partir de una serie de factores relativos a la oferta y la demanda de vivienda. Los modelos dinámicos presuponen, además, que los precios de la vivienda tienen «memoria», es decir, que los precios en el período t dependen, aparte de otras variables, de los precios en el período $t - 1$ (es de-

cir, el año precedente si trabajamos con series de datos anuales). La interacción espacial puede manifestarse de dos formas. En primer lugar, con la existencia de interrelación contemporánea entre el valor de la variable dependiente en el distrito j (por ejemplo) y el valor de la variable en el distrito k , de modo que el precio observado en el período t en j está relacionado con el precio en el período t en k . Estas interacciones espaciales son endógenas, lo que significa que el precio en j determina el precio en k , pero, a su vez, el precio en k determina también el precio en j . Además, la interrelación espacial es multidireccional, es decir, el precio en j es influido no solo por el precio en el distrito k , sino por el precio en otros distritos. En segundo lugar, podría existir una interrelación espacial autorregresiva, afectando al término de las perturbaciones del modelo (denotado por SAR-RE).

El resto del artículo está estructurado como sigue: en la sección segunda describimos cuál debe ser la estructura formal del modelo para poder estimar tanto los efectos dinámicos como los efec-

tos contemporáneos de la interacción espacial, afectando a la variable endógena y a las perturbaciones; para ello, nos basamos en Baltagi *et al.* (2014a). La sección tercera expone la metodología de estimación del modelo. La idea es simplemente ampliar al caso espacial el enfoque del método generalizado de los momentos (GMM, por sus siglas en inglés) habitual de los modelos dinámicos que trabajan con datos de panel (Arellano y Bond, 1991), para obtener estimaciones consistentes de los parámetros. Al estimador utilizado se le denomina GMM-SL-SAR-RE. En la sección cuarta se formula un estimador alternativo, denotado como GMM-SL-RE, que ignora cualquier proceso de autocorrelación espacial (SAR-RE) en las perturbaciones. La sección quinta introduce un modelo de precios de la vivienda, en los distritos ingleses, para el que resulta apropiado el uso de estos estimadores. La sección sexta presenta los datos del estudio, mientras que la sección séptima contiene los resultados obtenidos con ambos algoritmos de estimación. La sección octava se ocupa de la predicción utilizando los resultados anteriores. Por último, la sección novena se dedica a las conclusiones del estudio.

II. EL MODELO DINÁMICO ESPACIAL BASADO EN DATOS DE PANEL

En esta sección resumimos el modelo utilizado para tratar con los datos. En lo que sigue, se asume que los datos se generan a partir de la ecuación

$$\mathbf{y}_t = \gamma \mathbf{y}_{t-1} + \rho_1 \mathbf{W}_N \mathbf{y}_t + \mathbf{x}_t \beta + \varepsilon_t \quad [1]$$

donde \mathbf{y}_t es un vector de observaciones de dimensión $(N \times 1)$ sobre la variable dependiente en el período t , siendo N el número de distritos bajo estudio (individuos). Además, \mathbf{x}_t es una matriz de observaciones $(N \times k)$, en el período t , para k regresores. \mathbf{W}_N es una matriz de interacción espacial $(N \times N)$ que define las interacciones que tienen lugar, simultáneamente, entre cada par de los N distritos. En la formulación más simple, el elemento (i, j) de \mathbf{W}_N , que denotamos como $w_{ij,N}$, toma valor 1 si existe interacción recíproca entre el par de distritos i y j , y cero en caso contrario. La diagonal principal de \mathbf{W}_N tiene todos sus elementos iguales a cero, ya que, por definición, un distrito no puede interactuar consigo mismo. No es necesario que los elementos de \mathbf{W}_N tomen solo valores 1 ó 0, en una dicotomía simplista que usualmente denota proximidad espacial, sino que pueden ser el resultado de un tratamiento más elaborado como se

describirá más adelante. El vector $(N \times 1)$ de $\mathbf{W}_N \mathbf{y}_t$, al que suele identificarse como el retardo espacial, está formado por las sumas ponderadas del vector observado \mathbf{y}_t , de modo que para el distrito j , el elemento j de $\mathbf{W}_N \mathbf{y}_t$ equivale a la suma de los valores en \mathbf{y}_t observados en los distritos que se encuentran próximos a j . Los términos γ y ρ_1 son parámetros a estimar, y reflejan la memoria en los datos de panel y la intensidad de la interacción espacial autorregresiva. El término β es un vector de dimensión $(k \times 1)$ formado por los k parámetros relativos a los k regresores en \mathbf{x}_t .

El vector de perturbación ε_t es de orden $(N \times 1)$ y representa los factores no observables que también influyen en \mathbf{y}_t . Si aceptamos que puede existir dependencia espacial en las perturbaciones, una estructura autorregresiva suele ser una de las opciones habituales, de modo que

$$\varepsilon_t = \rho_2 \mathbf{M}_N \varepsilon_t + \mathbf{u}_t. \quad [2]$$

En la ecuación [2], \mathbf{M}_N es una matriz similar a \mathbf{W}_N en el sentido de que define la interacción entre las perturbaciones correspondientes a cada distrito; con frecuencia, suele trabajarse bajo la hipótesis de que $\mathbf{M}_N = \mathbf{W}_N$. Por su parte, ρ_2 es un parámetro que determina la intensidad de la interacción espacial autorregresiva en las perturbaciones. Dicho de otra forma, $\varepsilon_t = (\mathbf{I}_N - \rho_2 \mathbf{M}_N)^{-1} \mathbf{u}_t$ donde \mathbf{I}_N es la matriz identidad de orden N . El término \mathbf{u}_t está compuesto de dos efectos, $\mathbf{u}_t = \boldsymbol{\mu} + \mathbf{v}_t$. Uno, denotado como $\boldsymbol{\mu}$ y de dimensión $(N \times 1)$, es un efecto invariable en el tiempo que refleja la heterogeneidad no observable entre los distritos; el segundo es el vector \mathbf{v}_t , de dimensión $(N \times 1)$, que representa un componente aleatorio no observado. Expresado de forma más compacta:

$$\varepsilon = [\mathbf{I}_T \otimes (\mathbf{I}_N - \rho_2 \mathbf{M}_N)^{-1}] \mathbf{u} \quad [3]$$

donde \mathbf{I}_T es una matriz identidad de orden T . Como es habitual, los efectos aleatorios se asumen independiente e idénticamente distribuidos, con media nula y varianza constante, $\boldsymbol{\mu} \sim iid(0, \sigma_\mu^2)$ y $\mathbf{v} \sim iid(0, \sigma_v^2)$, además de ser independientes entre sí. Por consiguiente, a partir de [2], el término de perturbación para el distrito i , ε_{it} , es igual a la suma de un efecto específico del distrito μ_i , un efecto idiosincrásico que varía en el tiempo v_{it} , y una combinación de los términos de las perturbaciones en otros distritos, definida por \mathbf{M}_N y ρ_2 .

El proceso definido por las ecuaciones [1] y [2] puede escribirse de forma simplificada

mediante una ecuación recursiva. Si denotamos $\mathbf{G}_N = (\mathbf{I}_N - \rho_1 \mathbf{W}_N)$ y $\mathbf{B}_N = (\mathbf{I}_N - \rho_2 \mathbf{M}_N)$, se llega a la siguiente expresión:

$$\mathbf{y}_t = \mathbf{G}_N^{-1} [\gamma \mathbf{y}_{t-1} + \mathbf{x}_t \beta + \mathbf{B}_N^{-1} \mathbf{u}_t]. \quad [4]$$

III. UN ESTIMADOR ESPACIAL GMM

El método de estimación de los parámetros de la ecuación [1] debe tener en cuenta el problema que de la endogeneidad de algunas de las variables, están correlacionadas con las perturbaciones. Por ejemplo, \mathbf{y}_{t-1} está correlacionado con ε_t porque ε_t está a su vez correlacionado con ε_{t-1} en virtud del hecho de que ambos ε_t y ε_{t-1} contienen el mismo efecto asociado al distrito, invariable en el tiempo μ . Además, $\mathbf{W}_N \mathbf{y}_t$ está correlacionado con ε_t porque $\mathbf{W}_N \mathbf{y}_t$ está correlacionado con \mathbf{y}_t el cual depende de ε_t . Dicho de otro modo, $\mathbf{W}_N \mathbf{y}_t$ también es endógeno.

Este resultado es bien conocido en la literatura (Anselin, 1988). Puede demostrarse indirectamente utilizando un modelo espacial autoregresivo de corte transversal simple, tal como

$$\mathbf{y} = \rho \mathbf{W}_N \mathbf{y} + \mathbf{e}. \quad [5]$$

Si tratamos de estimar ρ por mínimos cuadrados ordinarios (MCO), la estimación de ρ así obtenida será inconsistente debido a la endogeneidad de $\mathbf{W}_N \mathbf{y}$. El estimador por MCO de ρ , digamos r , es $r = (\mathbf{y}'_L \mathbf{y}_L)^{-1} \mathbf{y}'_L \mathbf{y}$, donde $\mathbf{y}_L = \mathbf{W}_N \mathbf{y}$. Reemplazando [5] en la estimación de r , se obtiene que $r = (\mathbf{y}'_L \mathbf{y}_L)^{-1} \mathbf{y}'_L (\rho \mathbf{y}_L + \mathbf{e}) = \rho + (\mathbf{y}'_L \mathbf{y}_L)^{-1} \mathbf{y}'_L \mathbf{e}$. El segundo término de esta expresión no tiende a cero con el tamaño de la muestra, es decir, asintóticamente el $\text{plim} \left(\frac{1}{N} \right) (\mathbf{y}'_L \mathbf{e}) \neq 0$.

Para evitar la endogeneidad causada por el hecho de que \mathbf{y}_{t-1} está correlacionado con las perturbaciones ε_t , debido a la presencia de μ , tomamos primeras diferencias en los datos, de modo que $\Delta \varepsilon_t = \varepsilon_t - \varepsilon_{t-1} = \mathbf{B}_N^{-1} (\mu + v_t) - \mathbf{B}_N^{-1} (\mu + v_{t-1}) = \mathbf{B}_N^{-1} \Delta v_t$. En consecuencia, la ecuación de trabajo, eliminado μ , queda

$$\Delta \mathbf{y}_t = \gamma \Delta \mathbf{y}_{t-1} + \rho_1 \mathbf{W}_N \Delta \mathbf{y}_t + \Delta \mathbf{x}_t \beta + \Delta \varepsilon_t. \quad [6]$$

El problema de la endogeneidad del retardo espacial sigue estando presente, ya que si $\mathbf{W}_N \mathbf{y}_t$ está correlacionado con ε_t y $\mathbf{W}_N \mathbf{y}_{t-1}$ lo está con ε_{t-1} , $\mathbf{W}_N \Delta \mathbf{y}_t$ también lo estará con $\Delta \varepsilon_t$. El enfoque que presentamos adapta la metodología de Arellano y Bond (1991) para obtener un esti-

mador GMM basado en varias condiciones sobre los momentos. Las ecuaciones correspondientes son complejas y no avanzamos en esta línea. Por simplificar, supongamos que los niveles de las variables son independientes de Δv_t , además introducimos retardos en los niveles para asegurar el mantenimiento del supuesto de independencia (el número de retardos depende de la endogeneidad o exogeneidad de las variables). El supuesto de independencia permite que los niveles de las variables sean utilizados como instrumentos. Asumimos que las variables representadas por \mathbf{x}_t son estrictamente exógenas (esta hipótesis puede relajarse en la práctica), lo que significa que son independientes de v_t para periodos pasados, presentes y futuros. En concreto, si Δv_t toma valores para $t = 3, \dots, T$, esas variables exógenas, \mathbf{x}_t , son independientes de Δv_t para cualquier periodo t . En el caso de \mathbf{y}_t la situación es diferente porque \mathbf{y}_t depende de v_t , con lo que \mathbf{y}_{t-2} es independiente de Δv_t ; sin embargo, \mathbf{y}_{t-1} no lo es, ya que $\Delta v_t = v_t - v_{t-1}$. Baltagi *et al.* (2014a) introducen instrumentos espaciales como el retardo espacial de la endógena $\mathbf{W}_N \mathbf{y}_t$ o los retardos espaciales de las variables exógenas $\mathbf{W}_N \mathbf{x}_t$. Para asegurar la independencia con Δv_t , solo puede incorporarse en el conjunto de instrumentos $\mathbf{W}_N \mathbf{y}_{t-2}$ con $t = 3, \dots, T$; por el contrario, los instrumentos con $\mathbf{W}_N \mathbf{x}_t$ van de $t = 3, \dots, T$ puesto que la hipótesis de exogeneidad de \mathbf{x}_t hace innecesario cualquier *lagging* para mantener la independencia con Δv_t , $t = 3, \dots, T$. De hecho, los instrumentos $\mathbf{W}_N \mathbf{x}_t$ pueden utilizarse sea cual sea el valor que tome t . Denotamos todo el conjunto de instrumentos, tanto espaciales como no espaciales, mediante \mathbf{Z}^* .

La utilización de estos instrumentos en la primera etapa del proceso genera estimaciones consistentes de las perturbaciones. Con respecto al estimador GMM-SL-SAR-RE, la estimación consistente de las perturbaciones pueden utilizarse, en una segunda etapa, para estimar los parámetros ρ_2 , σ_μ^2 y σ_v^2 . La base para este procedimiento de estimación se concreta en el enfoque de Kapoor *et al.* (2007) y presupone la estimación por GMM de ρ_2 , σ_v^2 y $\sigma_1^2 = \sigma_v^2 + T \sigma_\mu^2$; en nuestro caso, ha sido necesario adaptar el planteamiento de Kapoor *et al.* (2007) puesto que ellos obtienen estimaciones consistentes de las perturbaciones, utilizando MCO, al asumir una perturbación de tipo ruido blanco. Partiendo de una estimación consistente de las perturbaciones ε , es posible generar unas ecuaciones de momentos muestrales cuya resolución conduce a las estimaciones de los parámetros faltantes ρ_2 , σ_v^2 y $\sigma_1^2 = \sigma_v^2 + T \sigma_\mu^2$. No profundizamos más en esta línea,

por la propia complejidad de las ecuaciones de los momentos; creemos que es suficiente con hacer un bosquejo general de la situación.

Asumiendo que se ha resuelto la estimación de ρ_2 , σ_μ^2 y σ_v^2 , la siguiente etapa implica utilizar un estimador GMM espacial consistente para poder obtener $\hat{\delta} = (\hat{\gamma}, \hat{\rho}_1, \hat{\beta})$, combinando la matriz de instrumentos \mathbf{Z} con las estimaciones obtenidas en la etapa 2, tal y como se detalla en Baltagi *et al.* (2014a).

IV. UN ESTIMADOR GMM ALTERNATIVO

Una alternativa respecto al estimador antes descrito asume la no existencia de autocorrelación en el término de error compuesto, de modo que $\rho_2 = 0$ con lo cual

$$\mathbf{y}_t = \mathbf{G}_N^{-1} [\gamma \mathbf{y}_{t-1} + \beta \mathbf{x}_t + \tilde{\varepsilon}_t] \quad [7]$$

donde $\tilde{\varepsilon}_t = \mu + v_t$ junto con $\mu \sim iid(0, \sigma_\mu^2)$ y $v_t \sim iid(0, \sigma_v^2)$. La estimación de este caso sigue un procedimiento similar al del GMM-SL-SAR-RE, utilizando las mismas condiciones de ortogonalidad citadas en Arellano y Bond (1991) además de las condiciones de ortogonalidad descritas previamente. Igual que antes, los instrumentos en niveles se utilizan a la hora de estimar los datos en diferencias. No obstante, por convención $\rho_2 = 0$, por lo que no es necesario estimar este parámetro. La segunda etapa en el GMM-SL-RE utiliza los residuos de los datos en diferencias para producir estimaciones más precisas de ρ_1 , γ y β .

V. APLICACIÓN A LOS PRECIOS DE LA VIVIENDA EN INGLATERRA

La muestra objeto de análisis está compuesta por 353 Autoridades Unitarias y Distritos con Autoridad Local (UALAD, por sus siglas en inglés) de Inglaterra, excluidas las Islas Scilly. El gráfico 1 muestra la distribución geográfica de los precios medios de la vivienda en 2007; destaca la tremenda heterogeneidad existente entre los 353 distritos, con precios mucho más altos en la región de Londres y el sureste que en el resto del país. Este patrón se repite para otros años. El gráfico 2 ofrece un mayor detalle de la estructura de precios en el sureste de Inglaterra; está claro que algunos distritos del centro de Londres, como Kensington-Chelsea y Westminster, tenían unos precios excesivamente altos, con una media superior a £1M y £0.5M respectivamente. En el polo opuesto, los precios medios en las zonas más remo-

tas del país se situaron por debajo de las £250.000. El gráfico 3 muestra el efecto de la crisis de 2008, que puso fin a las constantes alzas registradas por los precios de la vivienda durante la década anterior.

A continuación desarrollamos un modelo de precios medios de la vivienda por distrito, como aplicación de la especificación descrita al comienzo de este trabajo. Para llegar a ese modelo, utilizamos una hipótesis sencilla relativa al mecanismo de formación de precios de mercado en la que la oferta y la demanda de vivienda constituyen los dos factores, junto a ciertas covariables adicionales más unos efectos aleatorios individuales no observables. Bajo esta premisa, podemos escribir:

$$\begin{aligned} q_{it} &= f(y_{it}, x_{1it}, \omega_{it}) \\ q_{it} &= \text{cantidad demandada en el distrito } i \text{ en el} \\ &\text{período } t \\ \tilde{q}_{it} &= f(y_{it}, x_{2it}, \varsigma_{it}) \\ \tilde{q}_{it} &= \text{cantidad ofertada en el distrito } i \text{ en el} \\ &\text{período } t \\ y_{it} &= \text{precio} \\ x_{1it}, x_{2it} &= \text{covariables} \\ \omega_{it}, \varsigma_{it} &= \text{efectos (aleatorios) no modelizados.} \end{aligned}$$

Para avanzar en esta línea, vamos a suponer una relación lineal como en la ecuación [8], en la que la cantidad demandada q_{it} en el distrito i en el período t está relacionada negativamente con el precio y_{it} y con el precio en el período precedente y_{it-1} ; supondremos también que la demanda está relacionada positivamente con el retardo espacial del precio, denotado por $\sum_{j=1}^N w_{ij} y_{jt}$, $i \neq j$, $t = 1, 2, \dots, T$, lo que implica que precios altos en distritos próximos, de acuerdo a la matriz \mathbf{W}_N , generan un aumento de la demanda en el distrito considerado. Nos referimos a este hecho como «demanda desplazada», indicando que unos precios altos en los distritos próximos provocarán, posiblemente, la relocalización de la demanda hacia el distrito i

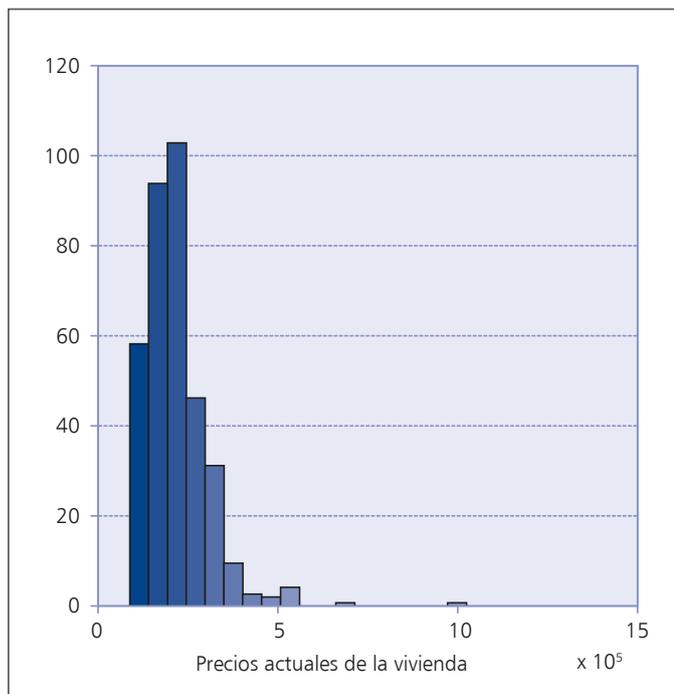
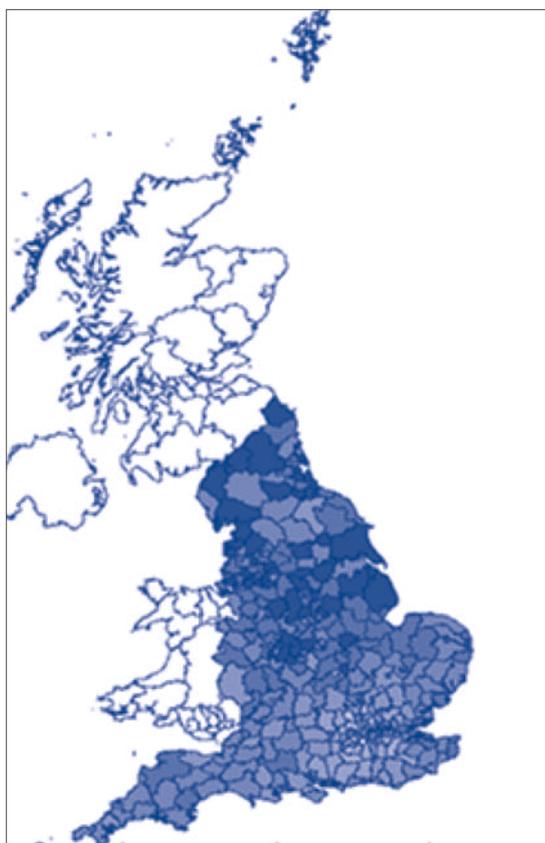
$$q_{it} = a_1 x_{1it} - a_2 y_{it} - a_3 y_{it-1} + a_4 \sum_{j=1}^N w_{ij} y_{jt} + \omega_{it} \quad [8]$$

Igualmente, la cantidad ofertada se plantea como una función lineal de acuerdo a

$$\tilde{q}_{it} = b_1 x_{2it} + b_2 y_{it} + b_3 y_{it-1} - b_4 \sum_{j=1}^N w_{ij} y_{jt} + \varsigma_{it} \quad [9]$$

El signo negativo del retardo espacial de los precios en [9] refleja la hipótesis de que unos precios altos en los distritos vecinos [1] provocarán la dis-

GRÁFICO 1
PRECIOS MEDIOS DE LA VIVIENDA EN 2007



minución de la oferta en el distrito i . El resultado es un nuevo impulso hacia la relocalización de la oferta desde i ya que, presumiblemente, los rendimientos superiores existentes en los distritos vecinos atraerán inversores y promotores inmobiliarios, en detrimento del distrito i . Normalizando la función de oferta respecto a y_{it} se obtiene:

$$y_{it} = \frac{1}{b_2} \tilde{q}_{it} - \frac{b_3}{b_2} y_{it-1} - \frac{b_1}{b_2} x_{2it} + \frac{b_4}{b_2} \sum_{j=1}^N w_{ij} y_{jt} - \frac{s_{it}}{b_2} \quad [10]$$

reemplazando $q_{it} = \tilde{q}_{it}$ y simplificando obtenemos

$$y_{it} = \gamma y_{it-1} + \rho_1 \sum_{k=1}^N w_{ik} y_{kt} + \beta_1 x_{1it} + \beta_2 x_{2it} + \varepsilon_{it} \quad [11]$$

$$i = 1, \dots, N; t = 1, \dots, T,$$

donde y_{it} indica el precio de la vivienda en el distrito i en el período t , x_{1it} es la demanda en el distrito i en el período t , x_{2it} es la oferta en ese mismo distrito y período, y β_1 y β_2 son parámetros. Además, $w_{ij,N}$

es el elemento de fila i y columna j de la matriz $W_{N,N}$ de orden $(N \times N)$, basada en las frecuencias de desplazamientos diarios de los habitantes desde cada distrito hasta su lugar de trabajo (esto es, la especificación de la matriz se basa en los datos de *commuting*). Los elementos de la diagonal principal de esta matriz son cero. Por otro lado, γ y ρ_1 son parámetros.

Debe recordarse que existen ciertas condiciones que definen el espacio de parámetros admisible para el modelo que estamos elaborando, lo que significa que las estimaciones que obtengamos deben pertenecer a ciertos intervalos predefinidos. Solo en ese caso obtendremos un modelo estable y estacionario en el que, por ejemplo, las predicciones no serán explosivas ni se encontraran demasiado alejadas de lo que razonablemente cabría esperar. Los detalles de esas condiciones de estabilidad son ciertamente complejos, por lo que remitimos al lector a Baltagi *et al.* (2014a). En términos generales,

GRÁFICO 2
PRECIOS MEDIOS DE LA VIVIENDA EN 2007, SURESTE DE INGLATERRA

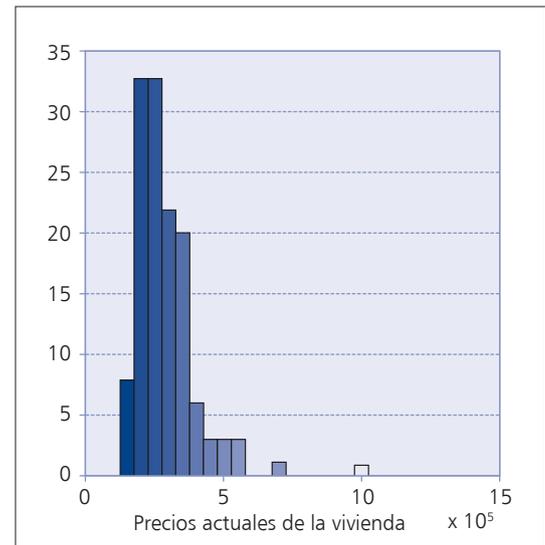
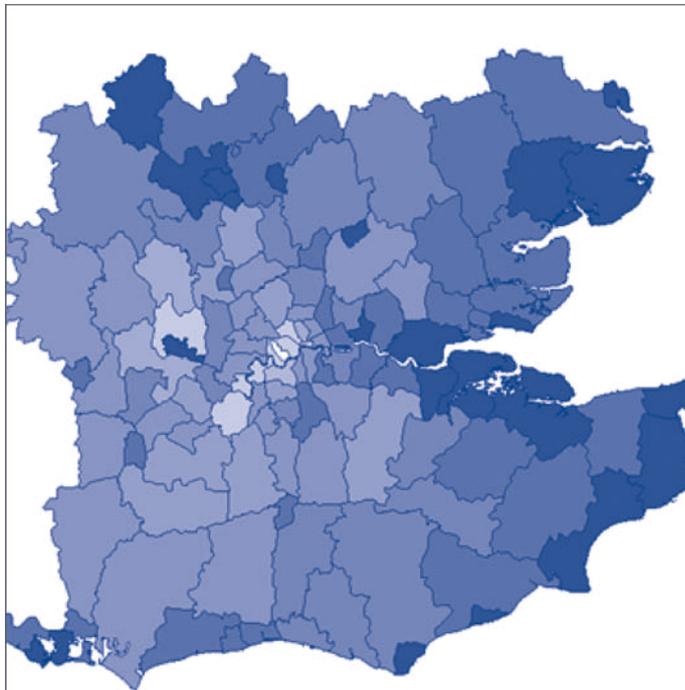
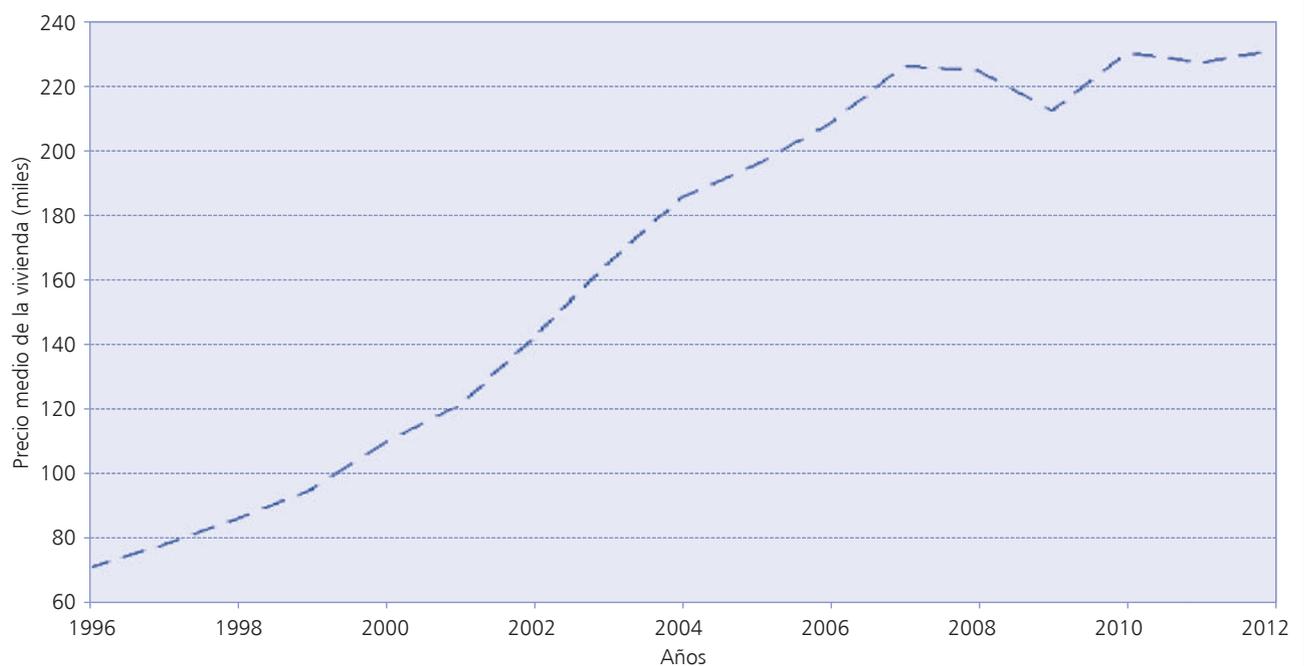


GRÁFICO 3
PRECIOS MEDIOS DE LA VIVIENDA



estas condiciones dependen de las propiedades matemáticas de la matriz W_N , resumidas en el vector de valores propios. La información relevante se concreta en los valores propios menor y mayor (2) respectivamente, de esa matriz denotados por e_{\min} y e_{\max}

$$\begin{aligned} |\gamma| < 1, e_{\min}^{-1} < \rho_1 < e_{\max}^{-1}, |\gamma| < \rho_1 e_{\max} \text{ si } \rho_1 > 0 \\ |\gamma| < 1 - \rho_1 e_{\min} \text{ si } \rho_1 < 0. \end{aligned} \quad [12]$$

El parámetro autorregresivo temporal, γ , debe ser menor que 1 en valor absoluto ya que, de otro modo, tendríamos un sistema explosivo con aumentos de precios *ad infinitum* o bien se produciría una serie de precios alternantes, con valores positivos y negativos, tendiendo hacia valores cada vez más extremos. De modo semejante, existen límites superior e inferior para los valores admisibles de ρ_1 . Si cada fila de W_N suma 1, entonces $e_{\max} = 1$; se evitarán resultados catastróficos si $\rho_1 < 1$. Con respecto al límite inferior, por simplicidad podríamos pensar en utilizar -1 como aproximación si bien, estrictamente hablando, ese límite dependerá de la estructura de W_N . En concreto, si se obtiene una estimación negativa de ρ_1 , esta debería estar por encima del límite inferior, e_{\min}^{-1} . Por regla general $\rho_1 > 0$, lo que significa que los precios de la vivienda en distritos vecinos tienden a ser similares. Un valor negativo indicaría que predominan las diferencias, sin excluir otros factores (3).

Una posibilidad, en el contexto del estimador GMM-SL-SAR-RE, y excluida del caso GMM-SL-RE, es que existan perturbaciones espaciales autorregresivas

$$\varepsilon_{it} = \rho_2 \sum_{k=1}^N m_{ik} \varepsilon_{kt} + U_{it} \quad [13]$$

donde

$$U_{it} = \mu_i + v_{it} \quad [14]$$

siendo M_N una matriz ($N \times N$) de pesos espaciales, $\mu_i \sim iid(0, \sigma_\mu^2)$ y $v_{it} \sim iid(0, \sigma_v^2)$. En este caso, suponemos que las perturbaciones son espacialmente dependientes debido a un error compuesto. La dependencia especial en las perturbaciones puede deberse, por ejemplo, a la existencia de variables, también espacialmente dependientes omitidas en el modelo. Como se ha dicho, la estructura de M_N es flexible y puede reflejar cualquier supuesto sobre la interdependencia de las perturbaciones; en este caso, vamos a tratarla como una matriz de contigüidades estandarizada por fila, de modo que $m_{ij} = m_{ij}^* / \sum_{j=1}^N m_{ij}^*$ donde $m_{ij}^* = 1$ si los distritos i y j comparten frontera común, y 0 en caso contrario. Debido a la estandarización por filas, e_{\max} para M_N

es igual a 1, con lo que el rango de parámetros que asegura un modelo estable es $e_{\min}^{-1} < \rho_2 < e_{\max}^{-1}$, de ahí que deba cumplirse $\rho_2 < 1$.

VI. DATOS

Todos los datos utilizados en la estimación se refieren al período 2002–2007. También están disponibles los datos para 2001, pero como este fue el año del censo utilizado para construir W_N , evitamos incluirlo para garantizar la absoluta exogeneidad de W_N respecto de los datos objeto de estimación. La crisis económica mundial posterior a 2007 tuvo un gran impacto en muchos mercados, entre ellos el de la vivienda. Por esta razón, evitamos tomar datos referidos al período posterior a la crisis a efectos de estimación. Los datos de precios proceden de la Oficina Nacional de Estadística (ONS) y del Registro de la Propiedad del Reino Unido, lo que permite generar precios medios en cada distrito (como se observa en el gráfico 1) en base en los precios acordados en las transacciones sobre inmuebles. La oferta de vivienda se mide a través del número de inmuebles en cada distrito (4).

La demanda de vivienda depende del salario medio y el empleo. Los datos de salarios se han obtenido de NOMIS, la base de datos de estadísticas laborales de la ONS. Más concretamente, utilizamos la media del salario bruto semanal \bar{w} (remuneración, en libras a la semana, pagada en el lugar de trabajo y sin distinción de sexo) que resulta de la *Encuesta Anual del Tiempo de Trabajo y Salarios (ASHE)*. El empleo total en cada distrito y año E_{it} se ha obtenido de la *Encuesta Anual a Empresas (ABS)* también disponible en NOMIS.

La demanda agregada local intradistrito se mide por $I_{it} = \bar{w}_{it} E_{it}$. La demanda procedente de distritos próximos, situados a una distancia apta para el desplazamiento diario, se mide por los ingresos en esos distritos ponderados por la proporción de trabajadores que se desplazan diariamente desde cada uno de ellos. La frecuencia de desplazamientos diarios al lugar de trabajo procede del Censo de 2001 de la Oficina Nacional de Estadística británica. El cuadro n.º 1 muestra una parte de los datos en forma de tabla de doble entrada cuyas filas representan el distrito de origen y las columnas el distrito de destino.

Si estandarizamos los datos sobre frecuencias de desplazamientos diarios, de modo que la suma de cada fila sea igual a la unidad, se obtienen las pon-

deraciones necesarias para calcular los ingresos en distritos próximos situados a una distancia de *commuting*. Denotando la tabla formada por las proporciones de desplazamientos como \tilde{C} , la demanda en el distrito i en el periodo t viene dada por $x_{it} = \sum_{k=1}^N \tilde{C}_{ik} l_{kt}$.

La matriz W_N refleja las proporciones de desplazamientos diarios de los residentes en cada distrito, omitiendo los movimientos con origen y destino en el propio distrito de modo que los elementos de la diagonal principal son todos cero. Como es habitual, la matriz se ha estandarizado para que la suma de cada fila sea 1. Esta matriz permite obtener el retardo espacial de la serie de precios por distritos, $W_N y_t = \sum_{k=1}^N w_{jk} y_{kt}$. Si el distrito j presenta una alta proporción de desplazamientos diarios de trabajadores que viajan al distrito k , asumimos que, si bien el distrito k es una alternativa viable como lugar de residencia, los trabajadores prefieren vivir en el distrito j , posiblemente porque los precios más bajos allí compensan el coste de desplazarse a diario al distrito k porque, por ejemplo el distrito k podría estar situado en el centro de una ciudad con viviendas caras. Asumimos que lo que ocurre con los precios en k está directamente relacionado con lo que ocurre en j . Por tanto, si los precios suben en el distrito k , la demanda de vivienda aumenta en j y los precios, consecuentemente, también subirán.

Con respecto a los errores autorregresivos, la matriz de contigüidades M_N de orden $(N \times N)$ y filas estandarizadas, está basada en los distritos del gráfico 1 con los que se comparte frontera.

VII. RESULTADOS DE LA ESTIMACIÓN

Utilizando el algoritmo GMM-SL-SAR-RE, el cuadro n.º 2 muestra que las estimaciones de los coeficientes de demanda y oferta, $\hat{\beta}_1$ y $\hat{\beta}_2$, son significativas y tienen el signo esperado, ya que los precios tienden a aumentar cuando aumenta la

demanda y a disminuir cuando se incrementa la oferta. El estimador del parámetro espacial autorregresivo $\hat{\rho}_1$ es escasamente relevante (solo es significativo, para una alternativa bilateral, a un nivel de significación del 10 por 100), lo que indica que los precios en un distrito están relacionados positivamente, aunque de forma débil, con los precios en otros distritos próximos. No obstante y de acuerdo al estimador $\hat{\gamma}$, los precios guardan una relación positiva y muy significativa con los precios del año precedente. Por lo que respecta a los efectos no observados, es evidente que existe una considerable heterogeneidad asociada al componente no sistemático de los errores compuestos, según se refleja en las magnitudes relativas de las estimaciones de las varianzas $\hat{\sigma}_\mu^2$ y $\hat{\sigma}_v^2$. Adviértase también que la estimación del parámetro espacial autorregresivo $\hat{\rho}_2$ toma un valor alto, positivo y significativo (5), subrayando la similitud entre las perturbaciones de los distritos vecinos. Por último, queremos notar que los valores de las estimaciones obtenidas se encuentran dentro de los límites definidos por las condiciones de estabilidad.

Las estimaciones del cuadro n.º 2 solo cuentan una parte de la historia ya que, para interpretar los efectos de la demanda y la oferta, se deberían evaluar las derivadas de y respecto a x que, en presencia de un retardo espacial y/o en el tiempo, no coinciden con las estimaciones de β_1 y β_2 . Este aspecto es relevante y así ha sido subrayado, por ejemplo, por LeSage y Pace (2009) y Elhorst (2014). En nuestro caso, existen efectos tanto a corto como a largo plazo. Los efectos a corto plazo se corresponden con las derivadas matriciales

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_{1k}} & \dots & \frac{\partial y_1}{\partial x_{Nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_N}{\partial x_{1k}} & \dots & \frac{\partial y_N}{\partial x_{Nk}} \end{bmatrix}_t = (\mathbf{I}_N - \rho_1 W_N)^{-1} \begin{bmatrix} \beta_k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \beta_k \end{bmatrix}, [15]$$

CUADRO N.º 1

MUESTRA PARCIAL DE LA MATRIZ DE ORDEN 353 X 353 DE DESPLAZAMIENTOS DIARIOS DE TRABAJADORES

	LUTON	MID BEDFORDSHIRE	BEDFORD	SOUTH BEDFORDSHIRE	BRACKNELL FOREST
Luton.....	54.399	1.071	865	6.876	28
Mid Bedfordshire.....	3.961	29.925	5.675	1.662	30
Bedford.....	1.888	4.458	50.165	599	17
South Bedfordshire.....	9.446	987	716	27.556	43
Bracknell Forest.....	12	0	8	6	30.840

CUADRO N.º 2

ESTIMACIÓN DEL MODELO GMM-SL-SAR-RE

PARÁMETRO	ESTIMA	ERROR TÍP.	RATIO Z	VALOR P*
γ	0,8394	0,0539	15,57	0,000001
ρ_1	0,09183	0,05484	1,674	0,094030
β_1	0,0007972	0,0002079	3,834	0,000126
β_2	-0,003278	0,0009265	-3,538	0,000403
ρ_2	0,7885	0,158005	4,630*	0,000004
σ_u^2	9,5537			
σ_v^2	2,2314			

Notas: + Ratio Z asociado a la distribución N(0,1). *Equivalente a menos la media de 1.000 replicaciones bootstrap (0,0567509) dividida por la desviación típica de la distribución bootstrap (0,158005).

mientras que los efectos a largo plazo son

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_{1k}} & \dots & \frac{\partial y_1}{\partial x_{Nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_N}{\partial x_{1k}} & \dots & \frac{\partial y_N}{\partial x_{Nk}} \end{bmatrix} = \left((1-\gamma)I_N - \rho_1 W_N \right)^{-1} \begin{bmatrix} \beta_k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \beta_k \end{bmatrix} \quad [16]$$

El efecto a corto plazo refleja el impacto sobre y_t en el período t de una variación de una unidad en x_{kt} , siendo $k = 1,2$, en cada uno de los N distritos en el período t , y comprende los efectos tanto directos como indirectos. En cuanto al efecto a largo plazo, las derivadas miden el impacto total sobre y en el período T (con T tendiendo a infinito) de una variación de una unidad en x_{kt} , siendo $k = 1,2$ en cada uno de los N distritos mantenida hasta el período T . Dado el tamaño de estas matrices de derivadas, es conveniente resumirlas, lo que puede hacerse tomando medias: de la diagonal principal de la matriz de derivadas parciales para los efectos directos, y de los elementos de fuera de la diagonal para los efectos indirectos. La suma de las dos medias anteriores constituye el efecto total. Los resultados se muestran en el cuadro n.º 3.

Examinando el cuadro n.º 3 resulta evidente que, a corto plazo, el impacto directo inmediato de una variación unitaria (6) en la demanda es un aumento de los precios de la vivienda (7) de £0,798. Obsérvese que este impacto es ligeramente mayor que el valor estimado de β_1 , debido a los efectos retroactivos (o de *feedback*) de la interacción instantánea, por lo cual las variaciones de precios producidas en el entorno del distrito también afectarán, de forma inmediata, a los precios en el distrito. Por comparación, el efecto indirecto de una

CUADRO N.º 3

EFFECTOS A CORTO Y LARGO PLAZO

	CORTO PLAZO	LARGO PLAZO
x_1	0,000797	0,00079
Directo.....	0,000798	0,00522
Indirecto.....	0,000079	0,00637
Total.....	0,000878	0,01159
x_2	-0,003278	-0,00328
Directo.....	-0,003281	-0,02148
Indirecto.....	-0,000328	-0,02619
Total.....	-0,003609	-0,04768

variación unitaria en la demanda, en el corto plazo, es menor en una relación de 10 a 1.

Comentarios similares pueden hacerse para la ratio entre los efectos directos e indirectos de una variación unitaria en la oferta (8). Como es lógico, esos efectos son negativos y de distinta magnitud a causa de las diferentes unidades de medida utilizadas. A largo plazo, los efectos de desbordamiento son mucho más relevantes, explicando una proporción mayor del efecto total que los efectos directos, tanto en el caso de la oferta como de la demanda. En el caso de la demanda, el efecto directo a largo plazo es más de 6,5 veces mayor que el efecto a corto plazo. El efecto indirecto, que explica cerca del 55 por 100 del efecto total a largo plazo para la demanda, es casi 80 veces mayor que el efecto indirecto a corto plazo. En el caso de la oferta, las ratios entre los efectos directos e indirectos, tanto a corto como a largo plazo, son idénticas a los de la demanda.

El cuadro n.º 4 proporciona las estimaciones producidas por el modelo alternativo GMM-SL-RE,

utilizando el mismo conjunto de instrumentos. El rechazo de la hipótesis nula de que $\rho_2 = 0$ indica que esta especificación es errónea. No obstante, para fundamentar con más detalle la necesidad de incorporar al modelo el proceso de error autorregresivo, en la siguiente sección exploramos la capacidad predictiva comparada de ambos conjuntos de estimadores.

VIII. PREDICCIÓN

Tanto en el caso de utilizar el estimador GMM-SL-SAR-RE como el alternativo GMM-SL-RE, las ecuaciones de predicción pueden darnos los mismos efectos totales que las derivadas. Siguiendo a Chamberlain (1984), Sevestre y Trognon (1996) y Baltagi *et al.* (2014b) el predictor lineal es

$$E[\mathbf{y}_t] = \mathbf{G}_N^{-1} (\gamma E[\mathbf{y}_{t-1}] + \mathbf{x}_t \boldsymbol{\beta} + \mathbf{B}_N^{-1} E[\boldsymbol{\mu}]) \quad [17]$$

que puede entenderse como idéntico a la ecuación [4], aunque tomando esperanzas matemáticas condicionales $E[\cdot]$. Aplicamos la ecuación [17] tanto si consideramos un incremento unitario, $\delta_{it} = \mathbf{1}_N$, o no; $\mathbf{1}_N$ es un vector de 1 de dimensión $(N \times 1)$. A continuación, examinamos la variación en valor esperado de los precios $E[\mathbf{y}_t]$. Para el efecto total a corto plazo de una variación unitaria en la demanda, calculamos la diferencia entre $\hat{\mathbf{y}}_{ta}$ y $\hat{\mathbf{y}}_{tb}$, donde

$$\hat{\mathbf{y}}_{ta} = \hat{\mathbf{G}}_N^{-1} (\hat{\gamma} \hat{\mathbf{y}}_{t-1,a} + (\mathbf{x}_{1t} + \delta_{it}) \hat{\boldsymbol{\beta}}_1 + \bar{\mathbf{x}}_{2t}^{-1} \hat{\boldsymbol{\beta}}_2 + \mathbf{B}_N^{-1} \hat{\boldsymbol{\mu}}) \quad [18]$$

y

$$\hat{\mathbf{y}}_{tb} = \hat{\mathbf{G}}_N^{-1} (\hat{\gamma} \hat{\mathbf{y}}_{t-1,b} + \mathbf{x}_{1t} \hat{\boldsymbol{\beta}}_1 + \mathbf{x}_{2t} \hat{\boldsymbol{\beta}}_2 + \mathbf{B}_N^{-1} \hat{\boldsymbol{\mu}}). \quad [19]$$

Un tratamiento similar puede aplicarse a la oferta, de modo que los efectos a corto plazo son

$$\begin{aligned} \frac{\partial \mathbf{y}}{\partial \mathbf{x}_k} &= \frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{y}}_{iTa} - \hat{\mathbf{y}}_{iTb}) = \\ &= \text{media} (\hat{\mathbf{G}}_N^{-1} \hat{\boldsymbol{\beta}}_k \mathbf{1}_N), \quad k = 1, 2. \end{aligned} \quad [20]$$

El efecto a largo plazo total se corresponde con el impacto sobre \mathbf{y} en el periodo T (con T tendiendo a infinito) de una variación de una unidad en \mathbf{x}_1 (o \mathbf{x}_2) en cada uno de los N distritos, que se mantiene en el tiempo hasta T ; en consecuencia, para \mathbf{x}_1 y dado el incremento $\delta_{1t} = \mathbf{1}_N$, $t = 1, \dots, T$, se podría obtener la diferencia entre [18] y [19] a través de sucesivas iteraciones para $t = 1, \dots, T$. De forma similar, se puede obtener el efecto total a largo plazo de una variación en \mathbf{x}_2 como consecuencia de un incremento $\delta_{2t} = \mathbf{1}_N$, para $t = 1, \dots, T$. Conforme T aumenta, la diferencia entre $\hat{\mathbf{y}}_{Ta}$ y $\hat{\mathbf{y}}_{Tb}$ se va aproximando a la derivada. De hecho, es fácil de demostrar numéricamente que

$$\begin{aligned} \frac{\partial \mathbf{y}}{\partial \mathbf{x}_k} &= \frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{y}}_{iT_a} - \hat{\mathbf{y}}_{iT_b}) = \\ &= \text{media} ((-\hat{\gamma} \mathbf{I}_N + \hat{\mathbf{G}}_N)^{-1} \hat{\boldsymbol{\beta}}_k \mathbf{1}_N), \quad k = 1, 2. \end{aligned} \quad [21]$$

Estas ecuaciones de predicción permiten que el análisis se amplíe más allá de las estimaciones de los efectos a corto y a largo plazo proporcionadas por las derivadas, como demuestra Fingleton (2015). Por ejemplo, una posibilidad es investigar qué impacto tendría, en el cálculo de los efectos a largo plazo, la introducción de variaciones diferentes de la unidad permitiendo, incluso que estas variaciones puedan cambiar en el tiempo. También podrían aplicarse a varios regresores simultáneamente. De igual forma, se podrían obtener predicciones condicionales de los precios de la vivienda en base a hipótesis alternativas sobre la trayectoria de la oferta y la demanda en los diferentes distritos.

En el presente artículo empleamos una estrategia de validación cruzada para discutir la validez de los estimadores rivales «comparando su capacidad predictiva sobre datos que no han sido utilizados en la estimación del modelo» (Anselin, 1988). En primer lugar, obtenemos predicciones con los dos estimadores rivales, un periodo adelante; es decir, tomamos las estimaciones obtenidas utilizando

CUADRO N.º 4

ESTIMACIÓN DEL MODELO GMM-SL-RE

PARÁMETRO	ESTIMACIÓN	ERROR TÍP.	RATIO Z	VALOR P
γ	0,4378	0,01572	27,85	<0,000001
ρ_1	0,5092	0,01790	28,45	<0,000001
β_1	0,0005937	0,000076	7,75	<0,000001
β_2	-0,003917	0,000619	-6,33	<0,000001

datos correspondientes al período 2002 – 2007 y predecimos a partir de ellas los precios en 2008.

Nótese que en el caso de predecir con la ecuación [17] también necesitamos estimaciones de los efectos de los distritos, invariantes en el tiempo μ . El enfoque adoptado aquí es tomar la media de los residuos a lo largo del tiempo, de modo que

$$y_t = \gamma y_{t-1} + \rho_1 W_N y_t + x_t \beta + \varepsilon_t \quad [22]$$

$$\varepsilon_t = y_t - (\gamma y_{t-1} + \rho_1 W_N y_t + x_t \beta)$$

Dado $\varepsilon_t = B_N^{-1} u_t$ donde $u_t = \mu + v_t$

$$B_N^{-1} u_t = y_t - \gamma y_{t-1} - \rho_1 W_N y_t - x_t \beta \quad [23]$$

siendo

$$\hat{\mu}^{(t)} = \hat{B}_N^{-1} [\hat{G}_N y_t - \hat{\gamma} y_{t-1} - \hat{x}_t \hat{\beta}] - v_t.$$

Para hacer operativos estos resultados, asumimos que $v_t \sim N(0, \hat{\sigma}_v^2)$; a continuación, con esa distribución generamos valores aleatorios y calculamos la media, a lo largo del tiempo, de los términos $\hat{\mu}^{(t)}$ a fin de obtener una estimación del valor constante $\hat{\mu}$. En el caso del estimador GMM-SL-RE, $\rho_2 = 0$ de modo que $B_N^{-1} = I_N$.

Las predicciones producidas por los dos estimadores se ilustran en los gráficos 4 y 5. En ellos se puede apreciar la semejanza existente entre los

GRÁFICO 4
PREDICIONES DE PRECIOS DE LA VIVIENDA PARA UN PERÍODO ADELANTE, UTILIZANDO LA ESTIMACIÓN GMM-SL-SAR-RE

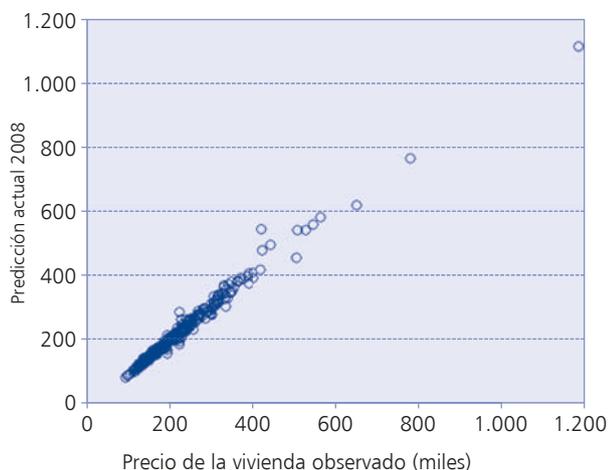
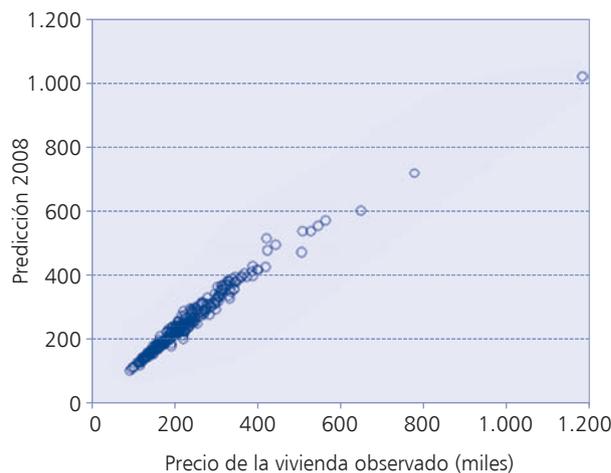


GRÁFICO 5
PREDICIONES DE PRECIOS DE LA VIVIENDA PARA UN PERÍODO ADELANTE, UTILIZANDO LA ESTIMACIÓN GMM-SL-RE



precios observados para 2008 y los precios resultantes de ambas predicciones. Igualmente parece claro que los precios deducidos del algoritmo GMM-SL-SAR-RE ofrecen un mejor ajuste (9), debido al error de especificación existente en el modelo GMM-SL-RE lo cual se debe a la omisión de ρ_2 . La diferencia entre las dos estimaciones se resume en los errores medios cuadráticos (RMSE) para un período adelante, que vienen dados por la expresión

$$\sqrt{\sum_{i=1}^N (y_{iT+1} - \hat{y}_{iT+1})^2 / N}, \text{ e iguales a } 24,804 \text{ y } 33,59,$$

respectivamente.

IX. CONCLUSIONES

En este artículo se explora la eficacia de dos estimadores alternativos que proporcionan estimaciones contradictorias de los determinantes de la variación de los precios de la vivienda en los distritos ingleses. Ambos muestran la importancia del retardo temporal, o efecto «memoria», muy significativo, si bien la estimación del parámetro $\hat{\gamma}$ obtenida utilizando el estimador GMM-SL-SAR-RE es de una cuantía mucho mayor que la que se obtiene a través del GMM-SL-RE. Tanto bajo el enfoque GMM-SL-SAR-RE como bajo GMM-SL-RE, se evidencia un efecto positivo y significativo de la demanda y un efecto negativo y significativo de la oferta. Las diferencias entre ambas estimaciones se deben a la omisión del parámetro del retardo

espacial ρ_1 , que es escasamente significativo en el caso del estimador GMM-SL-SAR-RE, pero elevado y fuertemente significativo en el caso del GMM-SL-RE. Esta diferencia refleja el hecho de que, en el primer caso, obtenemos un estimador estadísticamente significativo y elevado de $\hat{\rho}_2$ lo que modera el efecto derivado del retardo espacial al capturar los desbordamientos espaciales en las perturbaciones que de otro modo serían detectadas por el retardo espacial. Así, bajo el enfoque GMM-SL-SAR-RE, obtenemos que los precios de la vivienda solo interactúan de forma débil en el espacio (interacción reflejada por los flujos de desplazamientos diarios de los trabajadores a su lugar de trabajo, o *commuting*); consecuentemente, la sugerencia resultante del estimador GMM-SL-RE de que existe una fuerte interacción espacial entre los precios de los distintos distritos es engañosa. Aunque este modelo genera predicciones bastante precisas para un período adelante, dichas predicciones no son tan precisas como las provenientes de nuestro estimador preferido; el evidente error de especificación que supone asumir que $\rho_2 = 0$ conduce a una interpretación inexacta no solo del papel que juegan los desbordamientos espaciales, sino también de los efectos de la oferta y la demanda.

NOTAS

(1) Por simplificar, asumimos aquí que la misma matriz W_N interviene tanto en la ecuación de oferta como en la de demanda.

(2) En general, los límites para γ , ρ_1 y ρ_2 derivados de esos valores propios, son reales; caso de existir valores propios complejos, los límites para γ y ρ_1 se verían afectados en el mismo sentido.

(3) Un coeficiente negativo puede estar asociado a un efecto competición; por ejemplo, las ciudades grandes tienden a estar rodeadas por otras más pequeñas, mientras que los árboles de gran tamaño tienen a crecer rodeados de otros de porte más reducido.

(4) Debemos señalar que se han utilizado estimaciones para algunos distritos para 2010, y en todos ellos en 2012. Las estimaciones se basan en predicciones utilizando un modelo autorregresivo con un retardo, ajustado para datos del período 2001-2011 o 2001-2009.

(5) Como resultado del muestreo aleatorio con reemplazamiento usando las perturbaciones estimadas en la primera etapa. Replicando de este procedimiento de muestreo, 1.000 veces en nuestro caso, produce una distribución de referencia bajo la hipótesis nula de aleatoriedad espacial en los residuos.

(6) Obsérvese que no hemos tomado logaritmos, con lo que las estimaciones de los parámetros no pueden interpretarse como elasticidades.

(7) Medido en miles de £.

(8) En este caso, la ratio entre efectos directos e indirectos tanto para la demanda como para la oferta es idéntica, tanto a largo como a corto plazo. Esta es una propiedad matemática del modelo. Elhorst (2014, p. 106) plantea otras especificaciones alternativas en las que no se mantiene esa igualdad. Sin embargo, nuestra especificación está fundada sobre la hipótesis enunciada en la sección anterior sobre los precios de la vivienda lo que nos remite al planteamiento inicial.

(9) Elhorst (2005) llega a un resultado similar en el sentido de que tener en cuenta la autocorrelación espacial en el término de error proporciona un mejor comportamiento predictivo.

BIBLIOGRAFÍA

- ANSELIN, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- ARELLANO, M., y S. BOND (1991), «Some tests of specification for panel data: Monte Carlo evidence and an application to employment», *Review of Economic Studies*, 58: 277-297.
- BALTAGI, B. H. (2013), *Econometric Analysis of Panel Data*, Wiley and Sons, Chichester.
- BALTAGI, B. H.; FINGLETON, B., y A. PIROTTE (2014a), «Spatial lag models with nested random effects: An instrumental variable procedure with an application to English house prices», *Journal of Urban Economics*, 80: 76-86.
- (2014b), «Estimating and forecasting with spatial dynamic panel model», *Oxford Bulletin of Economics and Statistics*, 76: 112-136.
- CHAMBERLAIN, G. (1984), «Panel data», en GRILICHES, Z., and INTRILIGATOR, M., (eds.), *The Handbook of Econometrics*, Chapter 22, North-Holland, Amsterdam, pp. 1247-1318.
- ELHORST, J. P. (2005), «Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels», *Geographical Analysis*, 37: 85-106.
- (2014), *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, Springer, Heidelberg.
- FINGLETON, B. (2015), «Spatial panels, UK house prices and the global crisis», paper presented at the *Fourteenth International Workshop of Spatial Econometrics and Statistics*, París.
- KAPOOR, M.; KELEJIAN, H. H., y I. R. PRUCHA (2007), «Panel data models with spatially correlated error components», *Journal of Econometrics*, 140: 97-130.
- LESAGE, J., y R. K. PACE (2009), *Introduction to Spatial Econometrics*, CRC Press, Londres.
- PESARAN, M. H. (2015), *Time Series and Panel Data Econometrics*, Oxford University Press, Oxford.
- SEVESTRE, P., y A. TROGNON (1996), «Dynamic linear models», en MÁTYÁS, L. and SEVESTRE, P. (eds.): *The Econometrics of Panel Data: a Handbook of Theory with Applications*, Chapter 7, Kluwer Academic Publishers, Dordrecht, pp. 121-144.