

Modelos de regresión espacio-temporales en la estimación de la renta municipal: el caso de la Región de Murcia¹

CHASCO, C. (*) y F.A. LÓPEZ, (**)

(*) *Departamento de Economía Aplicada. Universidad Autónoma de Madrid. Campus de Cantoblanco. Carretera de Colmenar Viejo Km. 15.500 28049 MADRID Tlfno.: 491 397 4266*

(**) *Departamento de Métodos Cuantitativos e Informáticos. Universidad Politécnica de Cartagena. e-mail:*

E-mail: () coro.chasco@uam.es;(**) Fernando.Lopez@upct.es*

RESUMEN

En los últimos años ha surgido un interés creciente por las especificaciones y estimaciones de relaciones econométricas basadas en Datos de Panel. Este interés se ha trasladado al campo de la Econometría Espacial, con la introducción de la componente temporal en modelos que surgieron con un enfoque puramente espacial. Una especificación que incorpora esta componente temporal son los modelos SUR espaciales. Estas estructuras sólo consideran relaciones de vecindad contemporáneas entre unidades espaciales para cada instante de tiempo, sin tener en consideración relaciones espacio-temporales. El objetivo de este artículo es plantear una generalización de los modelos SUR espaciales mediante la extensión de la matriz diagonal de coeficientes de dependencia espacial, a una matriz diagonal inferior de coeficientes de dependencia espacio temporal. De este modo, se introduce en el modelo de regresión tanto dependencia espacial contemporánea como no contemporánea. Los modelos desarrollados en este artículo se aplican a la estimación de la renta bruta disponible municipal en la Región de Murcia, contrastando su superioridad frente a aquellas especificaciones que únicamente corrigen los efectos espaciales contemporáneos.

Palabras Clave: Econometría espacial, Modelos SUR, Renta bruta disponible.

Estimation of Municipal Income with Space-Time Regression Models: a Case for the Region of Murcia

ABSTRACT

Recently, there is an increasing interest in panel data specifications. This approach has also been extended to spatial econometrics, in which space-time models both introduce time dimension and correct for spatial effects. Spatial SUR models only consider contemporary neighborhood relationships between

¹ Los autores agradecen los comentarios de dos evaluadores anónimos que han permitido mejorar sensiblemente este artículo. Una versión preliminar del mismo se presentó en la XVIII Reunión Anual de Asepelt-España en León, 2004.

spatial units in each period of time. The aim of this paper is the proposal of an alternative specification for spatial SUR models, which includes not only spatial lags but also space-time autocorrelation effects (or non-contemporary spatial dependence) through an extended W weight matrix. We also demonstrate that this is frequently a better specification with higher predictive capacity. An application to the estimation of municipal household income in the Region of Murcia reveals the superiority of this new specification against those ones that only consider purely contemporary spatial effects.

Key words: Spatial Econometric, SUR Models, Household Income, Region of Murcia.

Clasificación JEL: C33; C51; C21, C53, D14, O18.

1. INTRODUCCIÓN

En la actualidad, existe un interés creciente por las especificaciones y estimaciones de relaciones econométricas basadas en datos de panel. Este interés puede explicarse por el hecho de que los modelos de datos de panel permiten al investigador estudiar fenómenos de comportamiento más complejo y, de este modo, extienden las posibilidades que presentan los modelos puramente espaciales o temporales: son generalmente modelos más informativos porque contienen una mayor variación y, por tanto, introducen menor colinealidad entre las variables. Efectivamente, la especificación de datos de panel, al incorporar un mayor número de grados de libertad en el modelo, suele incrementar la eficiencia de las estimaciones e incluye efectos que no pueden ser considerados en los modelos de corte puramente transversal o temporal (Hsiao 1986; Baltagi 2001). Otro de los factores que indudablemente han ayudado a fomentar el interés por este tipo de modelos es la creciente disponibilidad de bases de datos que ofrecen la información en la doble dimensión espacial y temporal.

Pero los modelos habituales de datos de panel suelen ignorar los efectos espaciales de dependencia y heterogeneidad espacial y son escasas las contribuciones que ofrecen alguna aportación metodológica en esta línea. Anselin (1988, cap. 10), en su conocido libro sobre econometría espacial, presenta varios modelos que bajo el nombre de SUR espaciales incluyen estructuras autorregresivas espaciales, bien en los errores o bien mediante la inclusión en el modelo de la variable endógena espacialmente retardada como explicativa. Más recientemente, también este autor (Anselin 2001) presenta una breve taxonomía de modelos de datos de panel en los que especifica una triple dependencia: espacial, temporal y espacio-temporal. La integración entre los modelos tradicionales de datos de panel y la econometría espacial es ya una realidad, gracias a los trabajos de Arellano (2003) y Anselin *et al.* (2004). Otros autores presentan sus propuestas tanto en trabajos teóricos (Baltagi *et al.* 2003 y Pace *et al.* 2000) como empíricos (Elhorst 2001 y 2003, Baltagi y Li 2003 y Mobley 2003).

Una de las dificultades que se plantea cuando se pretende extrapolar los modelos clásicos de la econometría espacial al ámbito de los modelos de datos de panel -con el fin de introducir estructuras de dependencia espacio-temporal- es el relacionado con el concepto de vecindad. El problema surge a la hora de definir un concepto de vecindad global y tiene su origen en la dificultad de conciliar las dos escalas: espacial y temporal, debido a las distintas métricas que se utilizan en cada dimensión.

La matriz de contactos, W , propia de la econometría espacial y a través de la cual se introduce la estructura de vecindades en un modelo espacial, puede ser también la clave para la introducción de una estructura de vecindades de tipo espacio-temporal. En el caso de los modelos descritos por Anselin, la extensión de la matriz de pesos espaciales W se reduce a una matriz bloque-diagonal en la que únicamente se contemplan las relaciones de vecindad contemporánea entre regiones, sin permitir otras relaciones de interacción espacial muy frecuentes en la realidad de tipo espacio-temporal. Otros autores como Pace *et al.* (2000) han definido la matriz W diferenciando entre vecinos espaciales y vecinos temporales. En concreto, construyen una matriz W como suma de ambas contribuciones² y cuantifican mediante parámetros distintos los efectos espaciales, los temporales y los espacio-temporales.

² $W = \phi_S S + \phi_T T + \phi_{ST} ST + \phi_{TS} TS$, donde S determina una estructura de vecindades en el espacio y T en el tiempo.

En este artículo, nuestra propuesta es lograr una generalización de los modelos SUR espaciales trabajando con una matriz W diagonal inferior por bloques. Esta extensión de la matriz de contactos hace posible un concepto de vecindad global que tiene en cuenta relaciones de vecindad espacial contemporáneas y no contemporáneas o espacio-temporales. Los modelos obtenidos con esta nueva matriz W introducirán como factores explicativos no sólo retardos espaciales contemporáneos, sino también retardos espaciales no contemporáneos.

Este concepto general de vecindad enlaza con la discusión sobre la necesidad de introducir retardos espaciales y retardos espacio-temporales en los modelos de regresión. Aunque lo habitual en econometría espacial es considerar que la presencia de dependencia espacial se produce de forma instantánea³ ($Y_t = \rho W Y_t$), también es cierto que debería contemplarse la posibilidad de que el efecto de dependencia espacial se produjera con algún retardo temporal ($Y_t = \rho W Y_{t-k}$) o de ambas posibilidades a la vez.

Con el fin de contrastar la bondad de los diferentes modelos teóricos descritos en este artículo, se realiza un ejercicio de estimación de la renta bruta disponible por habitante en los municipios de la Región de Murcia. La renta es una variable económica cuya distribución en el ámbito provincial presenta fuertes rasgos de autocorrelación espacial (ver Chasco, 2003; cap. 5). El método que con más frecuencia se ha utilizado para su estimación en ámbitos territoriales inferiores, como el municipal, ha sido el método indirecto basado en un modelo econométrico. Básicamente esta metodología consiste en la realización de una estimación econométrica en un nivel geográfico superior, el provincial, que posteriormente se extrapola a un ámbito geográfico inferior, el municipal.

En la realización de estas estimaciones, nunca se había prestado atención a los problemas que los efectos espaciales previsiblemente presentan en la estimación mínimo cuadrática del modelo econométrico provincial. Por eso, con la intención de solventar este problema, últimamente han aparecido diversos estudios en los que, haciendo uso de técnicas de econometría espacial, se han realizado nuevas estimaciones de la renta municipal (Alañón 2001, Chasco 2003). Con la introducción de estos nuevos modelos de regresión se eliminan los problemas de autocorrelación y heterogeneidad espacial que efectivamente se presentan en la estimación de la renta provincial. En este artículo, extendemos las especificaciones presentadas hasta la fecha, incluyendo en el modelo la dimensión espacio-temporal. Dentro del contexto socioeconómico, estos modelos que contemplan de forma simultánea las dos estructuras autorregresivas, contemporáneas y no contemporáneas, son más enriquecedores, puesto que cuando se produce un shock en cierta localización (ej. un crecimiento de la renta), muy probablemente éste se difundirá sobre localizaciones vecinas durante un cierto periodo de tiempo. Los resultados obtenidos confirman la superioridad de estas estructuras frente a las ya conocidas.

La estructura de este artículo es la siguiente. En la Sección 2 se exponen las características más importantes de los modelos SUR, presentando nuevas especificaciones que incluyen, bajo ciertos supuestos, no sólo retardos espaciales, sino también retardos espacio-temporales. La Sección 3 está dedicada a presentar algunos detalles del proceso de inferencia de los modelos propuestos. En la Sección 4 se realiza un ejercicio de predicción-extrapolación de la renta bruta

³ Por W se denota una matriz cuadrada no estocástica cuyos elementos w_{ij} reflejan la intensidad de la interdependencia existente entre cada par de regiones i y j . Se supone que W es una matriz de constantes conocidas y que todos los elementos de la diagonal son ceros. El producto WY se entiende como un retardo espacial de la variable Y .

disponible de los hogares a los municipios de la Región de Murcia, basado en un modelo SUR espacial con variable endógena espacial y espacio-temporalmente retardada. Finalmente, la Sección 5 resume las principales conclusiones de este trabajo.

2. MODELOS DE REGRESIÓN SUR ESPACIO TEMPORALES

A continuación se describen diferentes especificaciones de los modelos de regresión SUR espacio-temporales. En la sección 2.1, se presenta una breve introducción del modelo SUR espacial como base para los desarrollos posteriores. En la sección 2.2, se describe el modelo SUR espacial con estructuras de autocorrelación espacial en su especificación. Y en la sección 2.3, se presenta la innovación de este artículo: un modelo SUR espacial con retardos espacio-temporales.

2.1. Formulación general del modelo SUR espacial.

Como punto de partida de este artículo, se considerará la especificación de un modelo de regresión que expresa la totalidad de potenciales dependencias espacio-temporales así como formas de heterogeneidad:

$$y_{it} = X_{it} \beta_{it} + \varepsilon_{it} \quad ; \quad E[\varepsilon_{it}] = 0 \quad ; \quad E[\varepsilon_{it} \varepsilon_{js}] \neq 0 \quad ; \quad i = 1, \dots, N \quad ; \quad t = 1, \dots, T \quad (1)$$

donde por y_{it} se denota la observación de la variable endógena en la unidad espacial i y el instante t , X_{it} es un vector fila de k variables explicativas en la unidad i e instante t , β_{it} es un vector columna de k parámetros espacio-temporales y con ε_{it} denotamos el término de error. La varianza residual $E[\varepsilon_{it} \varepsilon_{js}] \neq 0$ expresa todas las posibles especificaciones de dependencia espacio-temporal y estructuras de heterogeneidad espacial.

Este modelo (1) es claramente inviable debido a la ausencia de grados de libertad necesarios para estimar el total de los parámetros β_{it} . Por eso, será necesario imponer a este parámetro general β_{it} ciertas restricciones de no variabilidad en alguna de sus dimensiones. Así, cuando el vector de coeficientes varía sobre el espacio pero es constante en el tiempo (β_i) se obtiene el siguiente modelo:

$$y_{it} = X_{it} \beta_i + \varepsilon_{it} \quad ; \quad E[\varepsilon_{it}] = 0 \quad ; \quad E[\varepsilon_{it} \varepsilon_{js}] = \sigma_{ij} \quad ; \quad i = 1, \dots, N \quad ; \quad t = 1, \dots, T \quad (2)$$

El modelo (2) es conocido como modelo SUR (“Seemingly Unrelated Regression” o modelo de ecuaciones aparentemente no relacionadas) y fue originalmente propuesto por Zellner (1962) con la idea inicial de especificar un sistema de ecuaciones, definido en un contexto espacio temporal, en el que se recogiera el fenómeno de dependencia espacial contemporánea entre los términos de error.

En segundo lugar, cuando en el modelo (1) se impone sobre el vector de coeficientes (β_{it}) la restricción de no variabilidad espacial, se obtiene una nueva especificación SUR:

$$y_{it} = X_{it} \beta_t + \varepsilon_{it} \quad ; \quad E[\varepsilon_{it}] = 0 \quad ; \quad E[\varepsilon_{it} \varepsilon_{is}] = \sigma_{ts} \quad ; \quad i = 1, \dots, N \quad ; \quad t = 1, \dots, T \quad (3)$$

El modelo (3) es conocido como modelo SUR espacial (en adelante, SSUR). Como puede observarse, en este modelo los términos del error están correlacionados en el tiempo. En forma matricial, la ecuación para cada periodo

de tiempo se expresa como:

$$Y_t = X_t \beta_t + \varepsilon_t \quad (4)$$

donde Y_t y ε_t son N por 1 vectores y X_t es una matriz N por K_t de variables explicativas. El número de variables independientes, K_t puede ser diferente para cada ecuación (periodo de tiempo). Este modelo será operativo en el caso de disponer de más observaciones para la dimensión espacial que para la temporal ($N > T$), que es lo habitual en Ciencia Regional.

La estimación de este modelo se realiza de forma simultánea para todos los instantes de tiempo t . Las ecuaciones se apilan para cada periodo de tiempo como se muestra en la siguiente expresión:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & X_T \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_T \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix} \quad (5)$$

o de forma compacta: $Y = X\beta + \varepsilon$ (6)

donde Y es un vector NT por 1 de variables dependientes, X es una matriz diagonal por bloques de dimensión NT por K ($K = \sum K_t$ número total de variables independientes), β es un vector K por 1 de coeficientes y ε es un vector NT por 1 de términos de la perturbación aleatoria. Esta especificación genera una matriz de varianzas y covarianzas del término de errores Ω de la forma:

$$E[\varepsilon \varepsilon'] = \Omega = \Sigma \otimes I \quad (7)$$

donde $\Sigma = \{\sigma_{ts}\}$ es una matriz T por T , y \otimes es el producto de Kronecker.

2.2 Especificación de los modelos SSUR-ERR y SSUR-LAG.

En el modelo SSUR, es posible considerar la existencia de autocorrelación espacial dentro de cada ecuación de tiempo. Los efectos espaciales pueden incorporarse a este modelo, bien incluyendo dependencia espacial de tipo residual o bien mediante la introducción de retardos espaciales de la variable endógena. Se obtienen así dos nuevas especificaciones:

2.2.1. Modelo SSUR-ERR

En el primer caso, la dependencia espacial residual se especifica mediante un esquema autorregresivo de primer orden en cada ecuación:

$$\begin{aligned} Y_t &= X_t \beta_t + u_t \\ u_t &= \lambda_t W u_t + \varepsilon_t \end{aligned} \quad (8)$$

con: $E[\varepsilon_t \varepsilon_s'] = \sigma_{ts} I$

Por tanto, para cada instante de tiempo t , los errores siguen un proceso autorregresivo espacial con diferente parámetro⁴ λ_t .

⁴ En nuestro caso, consideramos la misma matriz de ponderaciones W para cada instante de tiempo, pero puede generalizarse sin dificultad para el caso de diferentes matrices W_t para cada periodo.

En forma compacta

$$\begin{aligned} Y &= X \beta + u \\ u &= (\Lambda \otimes W) u + \varepsilon \end{aligned} \tag{9}$$

donde $\Lambda = \text{diag}\{\lambda_t\}$ es una matriz cuadrada de orden T diagonal de coeficiente de dependencia espacial y donde u es un vector apilado NT por 1.

Aquí, la matriz de varianzas y covarianzas es la siguiente:

$$E[u u'] = \Omega = B(\Sigma \otimes I)B' \tag{10}$$

siendo B es una matriz bloque-diagonal de orden NT por NT en la que cada elemento de la diagonal principal se expresa como $B_t = (I - \lambda_t W)^{-1}$. Este modelo se conoce como SUR espacial con autocorrelación espacial residual (SSUR-ERR).

2.2.2. Modelo SSUR-LAG

El segundo tipo de dependencia espacial que podría especificarse en el modelo SSUR consiste en introducir retardos espaciales de la variable endógena en cada periodo de tiempo como una variable explicativa más, tal y como se muestra en la siguiente expresión:

$$\begin{aligned} Y_t &= \rho_t W Y_t + X_t \beta_t + \varepsilon_t \\ E[\varepsilon_t \varepsilon_s'] &= \sigma_{\varepsilon}^2 I \end{aligned} \tag{11}$$

el sistema expresado en forma compacta se escribe como:

$$Y = (\Gamma \otimes W) Y + X\beta + \varepsilon \tag{12}$$

con:
$$E[\varepsilon \varepsilon'] = \Sigma \otimes I \tag{13}$$

donde Γ es la matriz de coeficientes de dependencia espacial T por T diagonal siendo ρ_t el valor diagonal y la matriz I es una matriz identidad de orden NT. Este modelo se denomina SUR espacial con variables dependientes espacialmente retardadas (SSUR-LAG).

2.3. Un modelo SUR espacial con retardos espacio-temporales

Este apartado está dedicado a generalizar los modelos SSUR planteados en la sección anterior con el fin de que incorporen, no sólo retardos espaciales, sino también retardos espacio temporales. Se persigue con estas nuevas estructuras, adaptar la realidad del proceso de difusión espacial de muchos fenómenos, ya que cuando se produce un shock una determinada región, la regiones vecinas se ven afectadas, no sólo de forma inmediata ($Y_t = \rho W Y_t$), sino también transcurrido cierto periodo de tiempo ($Y_t = \rho W Y_{t-k}$). Antes de introducir los modelos formalmente, haremos una pequeña justificación desde el punto de vista interpretativo.

En econometría espacial, la idea de que la dependencia espacial se introduce en los modelos de regresión de forma instantánea ($Y_t = \rho W Y_t$) es aceptada de forma natural. No obstante, en el caso de modelos con un único corte transversal, la inclusión simultánea de ambos tipos de retardos, espaciales y espacio temporales, ha sido ya propuesta por algunos autores. Así, Upton y Fingleton (1985, pág 369) consideran problemática la idea de un efecto espacial instantáneo, sobre todo cuando el tiempo t se interpreta en su sentido más estricto

y sugieren la inclusión de un retardo temporal en el efecto espacial, proponiendo modelos del tipo $Y_t = \rho WY_{t-1}$. Más recientemente, Elhorst (2001, 2003), considerando también esta cuestión, plantea modelos para un único corte transversal en los que introduce como factores exógenos una amplia variedad de retardos espaciales no instantáneos, tanto de la variable endógena como de las exógenas.

No cabe duda de que el debate sobre la instantaneidad de la dependencia espacial se encuentra relacionado con el origen que da lugar al efecto de autocorrelación espacial en un proceso estocástico. Dos son las causas principales que producen dependencia espacial en una variable territorial: de un lado, la existencia de errores de medida y de otro, los fenómenos de interacción espacial (Anselin 1988, pág. 11). La primera de estas causas (los errores de medida) podría considerarse, sin problemas, como un efecto “instantáneo”, aun interpretando el instante t en su sentido más estricto. Pero en los fenómenos socioeconómicos de interacción espacial (como, por ejemplo, la distribución de la renta disponible por habitante) la instantaneidad en la dependencia espacial tiene menos sentido, por lo que sería razonable incluir un retardo temporal en el efecto espacial. Sin duda, los fenómenos de difusión espacial en un proceso y los efectos de desbordamiento o “spillover” suelen requerir de un periodo de tiempo para su constatación.

En el caso de los procesos económicos resulta difícil atribuir claramente el resultado de la dependencia espacial a una causa u otra y, en ocasiones, se producirán contribuciones de ambos tipos (dependencia contemporánea y no contemporánea), sin que sea posible cuantificar el grado en que cada uno se produce. Por tanto, los modelos que únicamente plantean dependencia espacial instantánea no son capaces de identificar, por sí mismos, los mecanismos subyacentes en un proceso económico. Puesto que el objetivo debe ser compatibilizar el modelo econométrico con la teoría económica, creemos que deberían plantearse especificaciones que recojan, no sólo las estructuras espaciales contemporáneas, sino también las no contemporáneas o retardadas en el tiempo.

Aceptada la idea de introducir dependencia espacial retardada, lo más adecuado será incluir dichas estructuras, no tanto en modelos para un único corte transversal, como hasta ahora se ha venido haciendo, sino mediante modelos con paneles de datos que recojan de forma simultánea la dependencia espacial contemporánea y/o inter-temporal. Es con esta idea con la que proponemos dos nuevos tipos de modelos SSUR: uno de ellos incorporaría la dependencia espacio-temporal en los términos del error (SSUR-ERR*) y el otro, en forma de un retardo espacial de la variable endógena como explicativa (SSUR-LAG*).

2.3.1. Modelo SSUR-ERR*

Se trata de un modelo SSUR que introduce retardos espacio-temporales en el término del error atendiendo a la siguiente especificación:

$$\begin{aligned}
 Y_t &= X_t \beta_t + u_t \\
 u_t &= \sum_{k=1}^t \lambda_{tk} W u_k + \varepsilon_t \\
 E[\varepsilon_t \varepsilon_s'] &= \sigma_{\varepsilon}^2 I
 \end{aligned}
 \tag{14}$$

El modelo (14) se puede expresar en forma compacta si definimos una matriz de coeficientes de dependencia espacio temporales agrupados en una matriz diagonal inferior:

$$\Lambda_* = \begin{bmatrix} \lambda_{11} & 0 & \dots & 0 \\ \lambda_{21} & \lambda_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \lambda_{T1} & \lambda_{T2} & \dots & \lambda_{TT} \end{bmatrix} \quad (15)$$

De este modo, la ecuación (14) puede ahora expresarse en forma compacta como:

$$\begin{aligned} Y &= X\beta + u \\ u &= (\Lambda_* \otimes W) u + \varepsilon \\ E[\varepsilon \varepsilon'] &= \Sigma \otimes I \end{aligned} \quad (16)$$

En este caso la matriz de varianzas y covarianzas del término de error para la especificación espacio temporal es de la forma:

$$E[u u'] = \Omega = B_* (\Sigma \otimes I) B_*' \quad (17)$$

siendo $B_* = (I - \Lambda_* \otimes W)^{-1}$

2.3.2. Modelo SSUR-LAG*

El segundo de los modelos introduce el efecto de dependencia espacial mediante dos tipos de retardos, espaciales y espacio temporales, como factores exógenos de un modelo SSUR:

$$\begin{aligned} Y_t &= \sum_{k=1}^t \rho_{tk} W Y_k + X_t \beta_t + \varepsilon_t \quad ; \quad \text{con } t = 1, \dots, T \\ E[\varepsilon \varepsilon'] &= \Sigma \otimes I \end{aligned} \quad (18)$$

El modelo (18) también se puede escribir de forma compacta:

$$Y = (\Gamma_* \otimes W) Y + X\beta + \varepsilon \quad (19)$$

donde por Γ_* se entiende una matriz diagonal inferior equivalente a la presentada para Λ_* en la expresión (15).

En las expresiones (14) y (18), la selección de los parámetros de dependencia espacio- temporales de las matrices Λ_* y Γ_* juega un papel determinante. Centraremos este trabajo en la comparación de los dos modelos siguientes:

- (i) $\lambda_{ts} = 0 \quad \forall s \neq t; \lambda_{tt} \neq 0$ (resp.: $\rho_{ts} = 0 \quad \forall s \neq t; \rho_{tt} \neq 0$) $(s, t = 1, \dots, T)$ (20)
- (ii) $\lambda_{ts} = 0 \quad \forall s \neq t-k, s \neq t; \lambda_{tt} \neq 0; \lambda_{t-t-k} \neq 0$ (resp.: $\rho_{ts} = 0 \quad \forall s \neq t-k, s \neq t; \rho_{tt} \neq 0; \rho_{t-t-k} \neq 0$) para algún k entero positivo no nulo comprendido entre 1 y T . $(s, t = 1, \dots, T)$ (21)

La primera de las especificaciones de la matriz de coeficientes de dependencia espacio temporal Λ_* (respectivamente: Γ_*) sólo contempla el caso de dependencia espacial instantánea. En este caso no hay interacción espacio-temporal, siendo Λ_* (resp.: Γ_*) una matriz diagonal, por lo que $\Lambda_* = \Lambda$ (resp.: $\Gamma_* = \Gamma$); es decir, para cada periodo de tiempo t , se estima un único coeficiente autorregresivo espacial contemporáneo. Bajo las restricciones impuestas en (20) este primer modelo coincidiría con los modelos SSUR-ERR (8) o SSUR-LAG (11).

En la segunda especificación (21), se plantea un efecto de autocorrelación espacial en dos instantes de tiempo (salvo para el primer periodo o periodo inicial). Por un lado, una dependencia espacial contemporánea tomando $\lambda_{tt} \neq 0$

(resp.: $\rho_{tt} \neq 0$) y, por otro, una dependencia espacial no contemporánea para un retardo k , al considerar $\lambda_{t,t-k} \neq 0$ (resp.: $\rho_{t,t-k} \neq 0$). La especificación en este segundo caso, para una única variable espacial retardada en el momento pasado $t-k$ es la siguiente:

a) Modelo SSUR-ERR*:

$$\begin{aligned} \text{Para } t = 1: & \quad Y_1 = X_1 \beta_1 + u_1 \quad ; \quad u_1 = \lambda_{11} W u_1 + \varepsilon_1 \\ \text{Para } t > 1: & \quad Y_t = X_t \beta_t + u_t \quad ; \quad u_t = (\lambda_{t,t-k} + \lambda_{tt}) W u_{t-k} + \lambda_{tt} W_t u_t + \varepsilon_t \end{aligned} \quad (22)$$

b) Modelo SSUR-LAG*:

$$\begin{aligned} \text{Para } t = 1: & \quad Y_1 = \rho_{11} W Y_1 + X_1 \beta_1 + \varepsilon_1 \\ \text{Para } t > 1: & \quad Y_t = (\rho_{t,t-k} + \rho_{tt}) W Y_{t-k} + \rho_{tt} W Y_t + X_t \beta_t + \varepsilon_t \end{aligned} \quad (23)$$

Como puede observarse, en las expresiones (22) y (23) se especifican las dos causas de la dependencia espacial, instantánea o contemporánea y retardada o no contemporánea⁵.

3. ESTIMACIÓN DE LOS MODELOS SUR CON RETARDOS ESPACIO TEMPORALES.

La estimación de los modelos propuestos en la sección anterior no puede realizarse por el método de mínimos cuadrados ordinarios (MCO) debido a las consecuencias que sobre estos estimadores tiene la presencia del efecto de dependencia espacial (Anselin 1988, cap. 6). Aunque el proceso de inferencia podría llevarse a cabo mediante varios métodos, la alternativa más frecuente es la estimación por máxima verosimilitud (MV), que es la que nosotros proponemos en este artículo. En concreto, los estimadores MV se obtienen a partir de la maximización del logaritmo de la función de verosimilitud asociada al modelo espacial especificado, tal y como se presenta a continuación para los nuevos modelos SSUR-ERR* y SSUR-LAG*.

En primer lugar, para el modelo SSUR-ERR* el logaritmo de la función de verosimilitud del proceso generado por la expresión (16) se obtiene como:

$$\ln L = - (NT/2) \ln \pi - (1/2) \ln |\Omega| - (1/2) \varepsilon' \Omega^{-1} \varepsilon \quad (24)$$

con $\Omega = [(I - \Lambda^* \otimes W)' (\Sigma \otimes I)^{-1} (I - \Lambda^* \otimes W)]^{-1}$ y teniendo en cuenta que:

$$- (1/2) \ln |\Omega| = - (1/2) [N \ln |\Sigma| - 2 \ln |I - \Lambda^* \otimes W|] = - (N/2) \ln |\Sigma| + \ln |I - \Lambda^* \otimes W| \quad (25)$$

⁵ Como ya se indicó anteriormente, el objetivo que se persigue con este tipo de especificaciones es adaptar el proceso de difusión espacial propio de muchos fenómenos socioeconómicos distinguiendo en la expresión habitual de dependencia espacial no sólo el aspecto contemporáneo de la misma (representado por coeficiente ρ_{tt}), sino también el componente no contemporáneo o retardado que expresa este fenómeno de difusión (y que se representa a través del coeficiente $\rho_{t,t-k}$). Con esto lo que se pretende es demostrar la posible existencia en el fenómeno de dependencia espacial de dos componentes, contemporáneo y no contemporáneo que permita obtener un mejor conocimiento de este fenómeno de la dependencia espacial. No se aborda en este ejercicio la influencia que un retardo puramente temporal (Y_{t-k}) pueda tener también sobre la variable endógena no tanto porque este elemento no sea realmente importante –que lo es– cuanto por las dificultades de identificación que plantean los modelos llamados modelos espacio-temporales recursivos (Anselin *et al.* 2004, pag. 19). La combinación de retardos espaciales (contemporáneos y no contemporáneos) con retardos temporales será objeto de análisis en futuras publicaciones.

Sustituyendo esta expresión en (24) se obtiene:

$$\ln L = - (NT/2) \ln \pi - (N/2) \ln |\Sigma| + \ln |I - \Lambda_* \otimes W| - (1/2\sigma^2)(Y - X\beta)'(I - \Lambda_* \otimes W)'(\Sigma \otimes I)(I - \Lambda_* \otimes W)(Y - X\beta) \quad (26)$$

Para maximizar esta función, se consideran los estimadores MCG, que se obtienen mediante las expresiones:

$$\begin{aligned} \tilde{\beta} &= \left[X'(I - \Lambda_* \otimes W)'(I \otimes \Sigma)^{-1}(I - \Lambda_* \otimes W)X \right]^{-1} X'(I - \Lambda_* \otimes W)'(I \otimes \Sigma)(I - \Lambda_* \otimes W)Y \\ \tilde{\sigma}^2 &= \frac{1}{NT} (Y - X\tilde{\beta})'(I - \Lambda_* \otimes W)'(I \otimes \Sigma)(I - \Lambda_* \otimes W)(Y - X\tilde{\beta}) \end{aligned} \quad (27)$$

Al introducir estos estimadores en (26) se obtiene la función de verosimilitud concentrada:

$$\ln L = - (NT/2) (\ln \pi + 1) - (NT/2) \ln \tilde{\sigma}^2 + \ln |I - \Lambda_* \otimes W| \quad (28)$$

Esta función que depende de los parámetros especificados en Λ_* y debe maximizarse numéricamente.

En segundo lugar, para obtener los estimadores MV del modelo SSUR-LAG*, bastará considerar la transformación:

$$\Omega^{-1/2} (A_*Y - X\beta) = v \quad (29)$$

donde $A_* = I - \Gamma_* \otimes W$, $\Omega = \Sigma \otimes I$ para un SSUR-LAG* y v es un vector de perturbaciones aleatorias independientes con una distribución normal estándar. Aunque el término de error v presenta una distribución conjunta que se comporta adecuadamente, no puede ser observado y la función de verosimilitud tiene que estar basada en Y . Por este motivo, es necesario introducir el concepto de Jacobiano que permite derivar a partir de v la distribución conjunta de Y de la siguiente manera:

$$J = \det(\partial v / \partial Y) = |\Omega|^{-1/2} |A_*| = |\Sigma|^{-n/2} |A_*| \quad (30)$$

entonces, el correspondiente logaritmo de la función de verosimilitud será el siguiente:

$$L = - (NT/2) \ln \pi - (N/2) \ln |\Sigma| + \ln |A_*| - (1/2)v'v \quad (31)$$

con
$$v'v = (A_*Y - X\beta)'(\Sigma \otimes I)^{-1} (A_*Y - X\beta) \quad (32)$$

Esta función plantea importantes problemas para su maximización ya que no pueden desarrollarse procedimientos análogos al caso SSUR-LAG debido a la no diagonalidad de la matriz Γ_*

Ahora bien, en el caso mas simple del modelo que proponemos SSUR-LAG* (expresión 23), el proceso de inferencia puede también realizarse como si se tratara de la estimación por MV de un modelo SSUR en el que se han incluido como exógenas retardos espacio temporales de la variable endógena. Esto es consecuencia propiedad de invarianza de los estimadores máximo verosímiles. Ya que si en la expresión (23) tomamos

$$\alpha_1 = \rho_{11} \quad ; \quad \alpha_k = \rho_{t,t-k} + \rho_{tt} \quad (33)$$

En el modelo (23) puede incluirse el término WY_{t-k} como una variable exógena acompañada del correspondiente coeficiente α_k . Una vez obtenidos los estimadores máximo verosímiles $\hat{\alpha}_k$ se deduce que

$$\hat{\rho}_{kk} = \hat{\alpha}_k \quad ; \quad \hat{\rho}_{t,t-k} = \hat{\alpha}_k - \hat{\alpha}_t \quad (34)$$

4. ESTIMACIÓN DE LA RENTA BRUTA DISPONIBLE EN LOS MUNICIPIOS DE LA REGIÓN DE MURCIA.

En este apartado se realizará una estimación de la renta bruta disponible de los hogares (renta familiar disponible) por habitante (RFBDpc) a nivel provincial que supone un avance metodológico de la propuesta realizada en López *et al.* (2003), Chasco (2003) y Alañón (2001). En España, existe un interés creciente por la información estadística microterritorial y, por tanto, por la obtención de magnitudes socioeconómicas para las que no se dispone de información oficial. El caso más habitual serían las estimaciones de renta bruta disponible, que es una variable de gran interés para la elaboración de políticas e investigación de mercados locales. La metodología que se utiliza se basa en un proceso de estimación que intenta conciliar las consecuencias adversas del problema de la unidad espacial modificable (MAUP), que se plantea en todo ejercicio de regresión ecológica, y la presencia de los efectos espaciales en los dos ámbitos de la aplicación: el ámbito agregado de las provincias, donde se estimará el modelo ecológico tipo SSUR-LAG*, y el ámbito desagregado de los municipios, sobre el que se extrapolarán los resultados del modelo anterior.

Del proceso completo de predicción-extrapolación de la RFBDpc, consistente en 6 etapas (ver Chasco 2003, pp. 185 y ss.), en este artículo nos centraremos sobre todo en las fases de selección de variables, modelización espacial y predicción-extrapolación de dicha magnitud a los municipios de la Región de Murcia. Las demás cuestiones implicadas en el proceso de predicción (fundamentación teórica, análisis exploratorio espacial y análisis de los resultados) serán tratadas de forma más general.

4.1. Fundamento y selección de las variables explicativas de la RFBDpc

En el caso que se propone, la renta familiar bruta disponible (RFBD) es una magnitud económica, conocida también como renta familiar disponible, que es definida por el Instituto Nacional de Estadística (INE) en la Contabilidad Regional de España⁶, como la suma total de los ingresos procedentes del trabajo, más las rentas de capital, prestaciones sociales y transferencias, menos los impuestos directos pagados por las familias y las cuotas pagadas a la seguridad social. Es decir, se trata del nivel de renta de que disponen las economías domésticas para gastar y ahorrar (o la suma de todos los ingresos efectivamente percibidos por las economías domésticas durante un período). Éste debería constituir el punto de partida de cualquier estimación de la RFBD, que debe diferenciarse de otras magnitudes económicas similares, como el PIB o el VAB.

Además, dado que el objetivo de este artículo consiste en la estimación de la

⁶ En el período de redacción del artículo, en la Contabilidad Regional de España sólo se disponía de las series de RFBD provincial en el período 1995-2001.

RFBD municipal, resulta de gran importancia tener en cuenta los principios de las teorías microeconómicas sobre la renta de los hogares. Conviene destacar que una forma de combatir las consecuencias adversas del MAUP (falta de adecuación entre las relaciones económicas en distintos ámbitos o escalas de agregación) consiste en realizar una selección de buenas variables explicativas de la RFBD, no sólo en el ámbito agregado o ecológico, que es donde tendrá lugar el proceso de estimación del modelo econométrico, sino también en el nivel desagregado microterritorial al que se extrapolarán estos resultados y que es el que se desea predecir. Además, la búsqueda de variables explicativas de la RFBDpc debe considerar otra importante restricción: la disponibilidad de dicha información no sólo para el ámbito agregado (provincial), sino también para el nivel desagregado de la predicción (municipios de la Región de Murcia), en el período temporal considerado (años 1996 y 2001).

Así, a partir de las teorías microeconómicas y los resultados de algunos análisis empíricos sobre la distribución de la renta personal⁷, así como de un análisis exhaustivo de gran cantidad de variables socioeconómicas, se propone una relación de 8 indicadores con buena capacidad explicativa de la RFBD provincial, disponibles además para el ámbito municipal de la Región de Murcia en los años 1996 y 2001. Se han considerado estos dos períodos temporales por constituir los más recientes para los que se dispone de cierta información municipal de relevancia procedente del Padrón de 1996 y el Censo de 2001. Se trata de un retardo temporal de 5 años para el que deseamos contrastar el posible efecto de difusión espacial en la distribución de la RFBD provincial.

Las 8 variables seleccionadas como buenas explicativas de la RFBD por habitante (medida en €) tanto en el ámbito de las provincias españolas como de los municipios de la Región de Murcia son: líneas telefónicas de uso doméstico por habitante, líneas RDSI y ADSL por habitante, tasa de demandas de empleo (por población de 16 y más años), distancia de los habitantes al municipio cabecera comercial más cercano, tasa de instrucción de 2º y 3º grado (por población de 16 y más años), tasa de directivos y empresarios no agrarios (por habitante), parque de turismos por habitante y precio del metro cuadrado de la vivienda. Dado que la consideración de todas estas variables en un mismo modelo de regresión produce problemas de multicolinealidad, se realizó un análisis factorial que redujo las 8 variables explicativas iniciales a 2 factores. Estos factores se calcularon por el método de componentes principales y rotación varimax para cada período temporal obteniéndose resultados muy similares: un 77% de la varianza total en 1996 y un 79% en 2001. La composición de los factores rotados es también muy similar en ambos períodos: en el factor 1 (F1) tienen gran peso las variables tasa de instrucción, tasa de empresarios, líneas RDSI, precio de la vivienda y distancia a la cabecera comercial, mientras que en el factor 2 (F2) predominan las líneas de teléfonos de uso doméstico, tasa de demandas empleo y parque de turismos. F1 es un indicador sintético de nivel educativo-categoría profesional y F2 un indicador de consumo-empleo.

⁷ Ver, por ejemplo, en Pena (1996) una relación de teorías microeconómicas explicativas de la distribución de la renta disponible personal (teoría de la habilidad, teoría estocástica, teoría de la elección individual, teoría del capital humano, teorías de la desigualdad educativa, teoría del ciclo vital, teorías de la redistribución de los ingresos por las administraciones públicas y teorías de justicia distributiva), así como los resultados obtenidos por los autores a partir del análisis empírico de las Encuestas de Presupuestos Familiares.

4.2. Especificación y estimación de un modelo SSUR-LAG* de la RFBDpc provincial en los años 1996 y 2001

En este apartado se presentan los resultados de la estimación de la RFBDpc provincial mediante el modelo SSUR-LAG* en los años 1996 y 2001, que es el que se utilizará para extrapolar dicha magnitud a los municipios de la Región de Murcia.

La especificación de todo modelo de econometría espacial requiere de una matriz de pesos espaciales, W , que exprese las relaciones de vecindad o interacción existente entre las unidades geográficas del espacio considerado. Son varias las especificaciones que han sido utilizadas en la literatura económica: desde las sencillas matrices binarias de interacciones espaciales, hasta matrices generales de distancias, basadas o no en puntos críticos de corte. En el contexto general de las regiones europeas, suelen utilizarse las matrices de distancias, sobre todo para evitar que la existencia de regiones/provincias insulares deje las correspondientes filas/columnas de la matriz con ceros, lo que podría afectar a los resultados de la inferencia. La distancia geográfica es considerada como una buena variable “proxy” cuando no se dispone (como es el caso de esta aplicación) de variables de vecindad económica como la red de infraestructuras o de transportes entre provincias/municipios.

Por otro lado, deben evitarse también los problemas de identificación reseñados por Manski (1993), según el cual los pesos de la matriz W deben ser exógenos al modelo. Por todo esto, en este trabajo se ha optado por considerar una especificación de la matriz $W = \{w_{ij}\}$ como función inversa de la distancia geográfica d_{ij} al cuadrado entre los centroides de las provincias i, j , que es una variable exógena:

$$w_{ij} = \frac{1}{d_{ik}^2} \quad ; \quad w_{ij}^* = \frac{w_{ij}}{\sum_{k \in N_i} w_{ij}} \quad (35)$$

siendo N_i el conjunto de índices de las provincias vecinas a la i -ésima. Esta matriz se ha estandarizado por filas⁸ de tal forma que es la matriz W^* con la que se desarrolla el modelo de regresión espacial. Por tanto, la influencia que ejerce una provincia j sobre otra i depende inversamente de la distancia relativa que las separa.

Una vez definida la matriz de pesos espaciales, se procede al proceso de especificación, estimación y contraste del modelo espacio-temporal. Para ello, hemos seguido la estrategia de modelización propuesta por Anselin (1988, pag. 203 y ss.) que, partiendo de una estimación por separado de los modelos espaciales en cada momento temporal, por el método de mínimos cuadrados ordinarios (MCO), contrasta la existencia de efectos espaciales en dichos modelos así como la diagonalidad de la matriz de varianzas y covarianzas interecuaciones Ω (expresión 7). En la Tabla 2 se presentan los principales resultados de los modelos. Así, en primer lugar se han estimado dos modelos de regresión provincial de la RFBDpc por mínimos cuadrados ordinarios (MCO), para los años 1996 y 2001, a los que se han aplicado contrastes sobre los efectos espaciales de heteroscedasticidad y autocorrelación espacial. Los resultados ponen claramente de manifiesto, en ambas ecuaciones, la significatividad individual de los coeficientes estimados (con signos positivos correctos), así

⁸ En adelante, se designará con un asterisco en el superíndice (W^*) a la matriz W estandarizada por filas, mientras que se reserva el asterisco en el subíndice (W_*) para la matriz W bloque-diagonal inferior.

como el hecho de que el factor de consumo-empleo tiene mayor influencia sobre la RFBDpc que el factor nivel de instrucción-categoría profesional, tanto en 1996 como 2001.

Asimismo, dado que la distribución de los residuos de ambas regresiones es una normal (tal y como demuestra el test de Jarque-Bera), es posible recurrir a los tests del multiplicador de Lagrange (LM). En ambos modelos (1996 y 2001), los contrastes de Breush-Pagan y White permiten aceptar la hipótesis nula de homoscedasticidad e indican la existencia de dependencia espacial en los residuos de la regresión (LMERR), así como la necesidad de especificar como explicativa la variable endógena espacialmente retardada (LM-LAG). En efecto, la elección de la especificación correcta (modelo del error espacial vs. modelo del retardo espacial) vendrá dada por los contrastes LM robustos, de los cuales únicamente el LMLAG, robusto a la presencia de dependencia espacial residual, resulta significativo. Estos resultados aconsejan la introducción de la variable RFBDpc espacialmente retardada como explicativa, en ambos modelos, dando lugar a lo que se conoce como modelo del retardo espacial ("spatial-lag model"). En el contexto de los modelos de estimación de la renta, esta especificación permite expresar los efectos "spillover" o de externalidades propios de esta variable, tan frecuentemente citados en la literatura económica.

La estimación de este modelo para los 2 períodos temporales considerados produce algunos cambios en la cuantía de los coeficientes estimados (que siguen siendo significativos), aunque no en el signo. Por su parte, el coeficiente autorregresivo que acompaña a la variable RFBDpc espacialmente retardada es, en ambos modelos, elevado, positivo y significativo, lo que pone de manifiesto la existencia de un efecto "contagio" (dependencia espacial), de tipo contemporáneo, en la distribución de la renta provincial. Sin embargo, el contraste ML de autocorrelación espacial en los residuos pone de manifiesto que, pese a la introducción explícita del retardo espacial contemporáneo en el modelo, los residuos no son todavía completamente aleatorios. Esta última evidencia podría deberse a la necesidad de introducir explícitamente el efecto de dependencia espacial no contemporánea, al menos para el último período (2001). Además, sería también deseable contrastar la posible diagonalidad de la matriz de varianzas y covarianzas de los residuos en 1996 y 2001 y, por tanto, la conveniencia de especificar los 2 modelos anteriores como un modelo SSUR.

Comenzamos por este último aspecto especificando un único modelo espacio-temporal de tipo SSUR-LAG (expresión 11) de la RFBDpc provincial en los períodos 1996 y 2001. Se trata de un modelo que corrige la dependencia espacial en los residuos MCO mediante la introducción, como explicativas del modelo, de dos variables espacialmente retardadas de la renta (una para cada período temporal), cuyos efectos sobre la variable endógena se miden a través de los coeficientes autorregresivos de dependencia contemporánea

ρ_t ($\hat{\rho}^{96,96}$ y $\hat{\rho}^{01,01}$). De nuevo, las variables explicativas son positivas y significativas, como también los coeficientes autorregresivos. En este caso, ya no es posible rechazar la hipótesis nula de ausencia de dependencia espacial residual (según el test LMERR), por lo que conseguimos el cumplimiento de esta hipótesis básica. Y, además, a través del test LM de diagonalidad de la matriz de varianzas y covarianzas (Ω) de la perturbación aleatoria en ambos períodos, es posible también certificar la superioridad del modelo SSUR-LAG frente a las especificaciones de corte-transversal para cada período temporal⁹.

⁹ Este contraste de diagonalidad de la matriz Ω es un contraste asintótico similar al propuesto por Breush-Pagan para contrastar la presencia de heteroscedasticidad (varianzas no constantes) en los residuos de un modelo y, como en este caso, está basado en el enfoque del multiplicador de

Por último, nos interesa también contrastar en el modelo SSUR-LAG de la expresión (11) la significación y aportación de un retardo espacial no contemporáneo para la explicación de la RFBDpc de 2001, de forma que del modelo (11) pasemos al modelo SSUR-LAG* de la expresión (23). En la Tabla 2, se trata de los resultados del último modelo en el que los coeficientes estimados son positivos y significativos (excepto el término independiente en la ecuación del año 2001). La influencia de las variables sobre la RFBDpc está más igualada en este modelo respecto del primero: aunque la variable consumo-empleo sigue siendo la más influyente en 1996, ya no lo es en el año 2001. En cuanto a los coeficientes autorregresivos de dependencia espacial, todos ellos resultan estadísticamente significativos. La interpretación de los mismos se deduce de la expresión (23): $\hat{\rho}^{96,96}=0,74$, $\hat{\rho}^{01,01}=0,34$ y $\hat{\rho}^{01,96}=0,30$. Es decir, el efecto conjunto (contemporáneo y no contemporáneo) de la RFBDpc espacialmente retardada es muy alto y similar en ambos años (0,74 en 1996 y 0,64 en 2001). En el primer período considerado, sólo se conoce este efecto conjunto de 0,74 según el cual cada aumento de la RFBDpc de las provincias vecinas a una dada en 1 unidad, en 1996, produce un incremento de la RFBDpc de dicha provincia en 0,74 unidades (aproximadamente). En 2001, este efecto conjunto de dependencia espacial (evaluado en 0,64), puede desdoblarse en un efecto de dependencia espacial contemporánea (0,34) y no contemporánea (0,30) que, como puede observarse, reparte en efecto conjunto de dependencia espacial prácticamente en partes iguales. Estos resultados confirman la hipótesis inicial de que el efecto de dependencia espacial en variables económicas, como la renta disponible, tiene un doble origen como efecto instantáneo y retardado (o de difusión espacio-temporal) un número “k” de períodos temporales. Por tanto, el modelo finalmente elegido para estimar la RFBDpc provincial en 2001 es el siguiente:

$$\begin{cases} \hat{Y}_{96} = 0,74 \cdot W^* Y_{96} + 1993,96 + 402,32 \cdot F196 + 464,38 \cdot F296 \\ \hat{Y}_{01} = (0,30 + 0,34) W Y_{96} + 0,34 W Y_{01} + 1505,84 + 516,08 \cdot F101 + 471,70 \cdot F201 \end{cases} \quad (36)$$

Desde el punto de vista empírico y de la teoría económica, una especificación como ésta resulta muy verosímil dado que el nivel de RFBDpc alcanzado por una provincia (o también un municipio), en 2001, depende del comportamiento de factores exógenos en dicha provincia (o municipio), en 2001, como F1 (indicador de nivel educativo-categoría profesional) y F2 (empleo-consumo), y del valor de RFBDpc en las provincias (municipios) vecinos, no sólo en el año 2001 (dependencia espacial contemporánea) sino también en un momento anterior, es decir, 1996. Este último elemento expresa el factor de difusión que se produce en una zona cuando el nivel alto/bajo de RFBDpc de una provincia o municipio tiene consecuencias futuras sobre el nivel alto/bajo de RFBDpc en las provincias/municipios vecinos. Este modelo es válido, no sólo para el ámbito de las provincias españolas en el que ha sido estimado, sino también para el ámbito de los municipios de la Región de Murcia, sobre el que realizará la extrapolación de la RFBDpc, como se verá a continuación.

Lagrange. Cuando las diferentes ecuaciones en un modelo SUR pertenecen a diferentes períodos temporales, un contraste de autocorrelación temporal de expresión desconocida es equivalente a un test sobre la diagonalidad de la matriz de varianzas y covarianzas inter-ecuaciones (ver cómo Anselin, 1988, pag. 141, explica este contraste para el caso de modelos SUR estándar).

4.3. Predicción-extrapolación y análisis de la RFBDpc de los municipios de la Región de Murcia en los años 1996 y 2001

El uso, con fines de predicción espacial en ámbitos desagregados (municipios), de modelos ecológicos (estimados en un agregado provincial) que incluyen la variable endógena espacialmente retardada como explicativa, implica una previa re-especificación de los mismos. Efectivamente, como lo que se desea es conocer el valor de la variable endógena municipal, en un año determinado, debe evitarse la presencia del valor retardado contemporáneo de dicha variable entre los regresores del modelo, de la siguiente manera:

$$\hat{Y} = (I - \hat{\rho} \cdot W)^{-1} X\hat{\beta}$$

$$\begin{cases} \hat{Y}_{i,96} = (I - 0.74 \cdot W^*)^{-1} (1993,96 + 402,32 \cdot F196_i + 464,38 \cdot F296_i) \\ \hat{Y}_{i,01} = (I - 0.34 \cdot W^*)^{-1} (0.65 \cdot W^* \hat{Y}_{i,96} + 1505,84 + 516,08 \cdot F101_i + 471,70 \cdot F201_i) \end{cases} \quad (37)$$

donde $i = 1, \dots, 45$ municipios y el término $(I - \hat{\rho}W)^{-1}$ es un multiplicador espacial global (Anselin, 2002) que une la variable endógena con todos los regresores. La matriz W se define como en la expresión (35), como función inversa de la distancia geográfica al cuadrado entre los centros geográficos de dos municipios, siendo W^* la matriz W estandarizada por filas.

Como puede observarse, se trata de un proceso recursivo en el que la extrapolación del valor de la RFBDpc municipal en el año 2001 (\hat{Y}_{01}) requiere de la previa predicción de la RFBDpc municipal en el año 1996 (\hat{Y}_{96}). Una vez obtenidos con este modelo los resultados de RFBD total para los 45 municipios de la Región de Murcia, éstos han sido ajustados para que la suma provincial coincida con los datos oficiales publicados por el INE. En la Tabla 1, se presentan los resultados obtenidos para la estimación de la RFBDpc de los municipios de la Región de Murcia en 2001. Además de las cifras obtenidas con el modelo SSUR-LAG*, se incluyen también los resultados correspondientes al resto de modelos de estimación de datos en 2001: modelo estándar por MCO, modelo del retardo espacial y modelo SSUR-LAG.

A modo de validación del modelo SSUR-LAG*, escogido en el ámbito provincial para la estimación municipal, se incluyen también en la Tabla 1 los porcentajes de desviación entre la RFBD total, estimada para el conjunto de municipios por cada método, y el dato oficial publicado por el INE para la Región de Murcia en 2001¹⁰. Como puede apreciarse, el modelo SSUR-LAG* con retardos espaciales contemporáneos y no contemporáneos es también en el ámbito municipal el mejor modelo en cuanto que produce un menor porcentaje de error en la estimación de la RFBD municipal. En efecto, el porcentaje de error registrado por este modelo es de -5% (en este caso por defecto), es decir: la suma de las cifras de RFBD estimadas por el modelo para los municipios de Murcia es un 5% inferior al dato ofrecido por la Contabilidad Regional del INE para esta Región.

¹⁰ Se calcula la ratio entre el dato de suma de RFBD total estimada para el conjunto de municipios de la Región de Murcia y el dato oficial publicado por la Contabilidad Regional del INE. Esta ratio se resta de 1 (predicción perfecta) y el resultado, multiplicado por cien, puede ser interpretado como el porcentaje de error por defecto (signo negativo) o por exceso (signo positivo)

Tabla 1

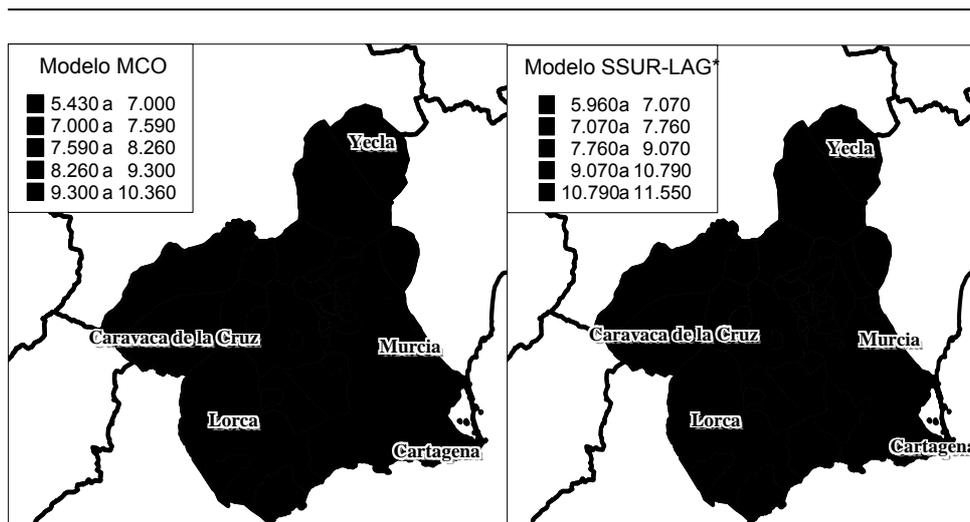
MUNICIPIO	Modelo básico (MCO)	Modelo retardo espacial	Modelo SSUR- LAG	Modelo SSUR- LAG*
Abanilla	7.794,76	7.629,81	7.626,61	7.276,11
Abarán	7.409,54	7.118,35	7.197,37	6.844,44
Águilas	7.848,58	7.964,80	7.969,71	7.907,97
Albudeite	5.438,71	6.094,46	6.375,17	5.968,57
Alcantarilla	8.267,89	8.124,34	8.200,47	7.947,38
Alcázares (Los)	9.781,04	10.502,24	10.251,49	10.793,37
Aledo	7.591,91	7.564,07	7.586,63	7.441,59
Alguazas	7.729,22	7.636,17	7.723,73	7.269,28
Alhama de Murcia	7.711,43	7.675,57	7.718,75	7.586,13
Archena	7.662,27	7.493,42	7.616,22	7.130,46
Beniel	8.023,39	8.323,90	8.329,37	8.138,80
Blanca	7.000,00	6.965,82	7.069,30	6.715,87
Bullas	7.259,37	6.961,62	6.991,51	6.620,77
Calasparra	7.112,75	6.969,43	6.999,52	6.631,03
Campos del Río	6.792,75	6.790,44	6.965,44	6.519,03
Caravaca de la Cruz	7.897,43	7.610,59	7.683,73	7.481,74
Cartagena	8.984,80	9.648,45	9.579,14	10.089,93
Cehegín	7.360,46	7.040,56	7.078,58	6.765,41
Ceutí	7.632,02	7.605,44	7.723,98	7.268,59
Cieza	8.010,84	7.574,45	7.647,05	7.331,70
Fortuna	7.209,58	7.377,48	7.425,51	7.079,14
Fuente Álamo de Murcia	8.376,69	8.437,04	8.362,17	8.344,14
Jumilla	7.718,22	7.859,68	7.835,85	7.636,25
Librilla	7.328,30	7.355,30	7.397,49	7.100,61
Lorca	8.134,99	7.960,26	7.974,40	7.888,87
Lorquí	8.721,58	8.183,82	8.220,40	7.764,01
Mazarrón	8.362,74	8.333,18	8.296,52	8.264,33
Molina de Segura	8.559,04	8.209,60	8.273,43	7.905,16
Moratalla	6.469,46	6.693,02	6.752,46	6.424,15
Mula	7.177,78	6.915,66	7.009,96	6.676,11
Murcia	9.422,13	9.111,12	9.138,97	9.075,90
Ojós	6.850,52	6.834,97	7.029,07	6.585,60
Pliego	6.860,94	6.714,42	6.777,66	6.385,55
Puerto Lumbreras	8.039,74	7.942,26	7.912,97	7.781,40
Ricote	6.564,39	6.673,19	6.840,34	6.412,85
San Javier	10.357,08	10.987,26	10.660,36	11.549,34
San Pedro del Pinatar	9.305,50	10.532,40	10.222,68	11.035,46
Santomera	8.691,53	8.517,09	8.525,95	8.303,88
Torre-Pacheco	8.740,02	9.639,40	9.451,18	10.136,37
Torres de Cotillas (Las)	7.994,03	7.771,70	7.855,64	7.481,45
Totana	7.644,33	7.604,54	7.628,36	7.506,45
Ulea	7.466,10	7.420,91	7.567,65	7.169,11
Unión (La)	7.826,57	9.129,43	9.062,88	9.557,91
Villanueva del Río Segura	6.688,13	6.854,25	7.035,38	6.531,63
Yecla	8.838,35	8.434,48	8.441,26	8.251,59

Desviación estimación/INE (% error) 0.78 (-22%) 0.83 (-17%) 0.83 (-17%) 0.95 (-5%)

Por otro lado, la estimación de la renta municipal mediante un modelo provincial MCO con un único corte transversal para 2001 y sin efectos espaciales, produce un porcentaje de error 17 puntos porcentuales inferior al caso anterior (-22%).

En la Figura 1, se han representado en un mapa los resultados obtenidos con el modelo básico MCO de corte-transversal sin efectos espaciales y el modelo SSUR-LAG* con dependencia espacial contemporánea y no contemporánea. Ambas estimaciones coinciden en estimar la distribución de la RFBDpc municipal en la Región de Murcia como un esquema de tipo centro-periferia. La estimación MCO centra los mayores valores de renta en torno a Murcia capital e incluye también, entre los municipios de mayor renta, a Caravaca de la Cruz y Yecla. Por su parte, el modelo SSUR-LAG* tiende a concentrar espacialmente las zonas de mayor/menor renta. Los municipios más ricos de la Región son los costeros del Mar Menor (San Javier, Los Alcázares, San Pedro del Pinatar), seguidos por Murcia, Cartagena y sus zonas de influencia. En este segundo modelo, la periferia estaría formada por los municipios del interior.

Figura 1



Nota: Los datos de RFBDpc se refieren a 2001; la unidad de medida es el euro per capita. Los intervalos han sido determinados por el método “natural breaks”, que agrupa las observaciones buscando la mínima varianza inter-grupos y la máxima inter-grupos.

5. CONCLUSIONES

La creciente disponibilidad de amplias bases de datos que contienen información referida a distintas escalas geográficas y periodos temporales, debería constituir una invitación al desarrollo de especificaciones econométricas capaces de capturar la dinámica de las relaciones existentes entre estos dos tipos de datos: espaciales y temporales. Por otro lado, la complejidad en su tratamiento está siendo atenuada, en gran parte, por el desarrollo de potentes herramientas informáticas que simplifican su manejo.

Este artículo se encuadra dentro del campo de la modelización econométrica basada en los datos de panel en el que, cada vez más, se consideran explícitamente los efectos espaciales propios de la econometría espacial. Sin embargo, es frecuente que los modelos de panel con efectos espaciales especifiquen el fenómeno de interacción espacial entre distintas observaciones (para cada período temporal) como un efecto instantáneo o contemporáneo, sin considerar las más que probables relaciones de interacción espacio-temporal o dependencia espacial no contemporánea.

En este contexto, nuestra propuesta consiste en la especificación simultánea, en un modelo econométrico espacio-temporal, los dos tipos de efectos de interacción espacial: dependencia contemporánea (instantánea) y no contemporánea (retardada). Para ello, se ha desarrollado una estructura matemática que considera no sólo el efecto de dependencia espacial instantánea o contemporánea para cada instante de tiempo (introduciendo la estructura $Y_t = \rho_t W Y_t$), sino también el efecto de dependencia espacial no contemporánea o retardada, que podría denominarse como dependencia espacio-temporal (mediante la estructura $Y_t = \rho_{t-k} W Y_{t-k}$).

Partiendo de la especificación de un modelo SUR espacial, que considera una matriz diagonal de coeficientes autorregresivos de dependencia espacial, se ha extendido dicho modelo mediante la consideración de una matriz diagonal-inferior de coeficientes autorregresivos espaciales. Esta modificación en la matriz de coeficientes de dependencia espacial lleva consigo la inclusión en el modelo de regresión no sólo de retardos espaciales contemporáneos (elementos de la diagonal principal), sino también de retardos espaciales no contemporáneos o espacio-temporales, lo que permite, de esta forma, captar simultáneamente ambos tipos de dependencia espacial.

El modelo desarrollado en el apartado 2.3 se ha utilizado para realizar la estimación de la renta familiar bruta disponible en los municipios de la Región de Murcia. La estimación de macromagnitudes socioeconómicas (como la RFBD) en niveles de desagregación territorial inferiores al nacional es, en sí mismo, un problema abierto pese a los muchos esfuerzos que se han realizado para obtener este tipo de información. Todo intento de mejorar el proceso de modelización econométrica en aras de obtener información desagregada microterritorial es una aportación de indudable interés económico y social. Por otro lado, la utilización de los modelos SSUR para la predicción-extrapolación de la renta municipal supone una novedad ya que, hasta la fecha, es la primera aplicación en este campo con modelos espacio-temporales o, al menos, con una especificación que contraste el efecto de difusión o dependencia espacio-temporal propio de un fenómeno como la distribución de la renta territorial. Efectivamente, como se contrasta en la aplicación desarrollada en el apartado 4, la presencia altamente significativa del coeficiente de dependencia espacial contemporánea y no contemporánea en el modelo de 2001 avala la hipótesis inicialmente planteada.

En cuanto a la estimación de la RFBDpc de los municipios de la Región de Murcia, los resultados obtenidos con el modelo SSUR-LAG* son los que menor porcentaje de error producen respecto de otros modelos espaciales y espacio-temporales más sencillos, amén de ser más consistentes con la estructura económica propia de los municipios de la Región: con un centro de riqueza concentrado en el eje Murcia-Mar Menor-Cartagena y una periferia localizada en los municipios del interior.

6. REFERENCIAS BIBLIOGRÁFICAS

- ALAÑÓN, A. (2001), “*La renta regional en España: análisis y estimación de sus determinantes*”. Tesis doctoral, Universidad Complutense de Madrid.
- ANSELIN, L. (1988) “*Spatial Econometrics: Methods and Models*”. Dordrecht: Kluwer
- ANSELIN, L. (2001) “*Spatial econometrics*”. En “A companion to theoretical econometrics”, ed. Baltagi, Oxford: Basil Blackwell; pp. 310-330.
- ANSELIN, L. (2002) “*Under the hood. Issues in the specification and interpretation of spatial regression models*”. *Agricultural Economics* 27; pp. 247–267.
- ANSELIN, L., J. LE GALLO y H. JAYET (2004), “*Spatial panel econometrics*”. In : Matyas L. et Sevestre P. (Eds.), “The Econometrics of Panel Data”, Kluwer Academic Publishers, 3ª edición (próxima publicación).
- ARELLANO, M. (2003). “Panel Data Econometrics”. Oxford University Press. Oxford, U.K.
- BALTAGI, B. H. (2001). “*Econometric Analysis of Panel Data*” (Segunda Edición). John Wiley & Sons, Chichester, United Kingdom.
- BALTAGI, B.H. y D. LI (2003) “*Prediction in the panel data model with spatial correlation*”. En: Anselin, L., R. Florax y S. Rey (Eds.), *New Advances in Spatial Econometrics*, Springer-Verlag; pp. 283-296.
- BALTAGI, B.H., S.H. SONG y W. KOH (2003) “*Testing panel data regression models with spatial error correlation*”. *Journal of Econometrics* (próxima publicación).
- CHASCO, C. (2003) “*Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales*”. Consejería de Economía e Innovación Tecnológica, Comunidad de Madrid.
- ELHORST, J.P. (2001) “*Dynamic models in space and time*”. *Geographical Analysis* 33; pp. 119-140.
- ELHORST, J.P. (2003). “*Specification and estimation of spatial panel data models*” *International Regional Science Review*, 26(3):244–268.
- HSIAO, C. (1986) “*Analysis of Panel Data*”, Cambridge University Press, Cambridge.
- LÓPEZ, F., J.A. GARCÍA y M. Ruiz (2003) “*Modelos de regresión espacio temporales en la estimación municipal de la renta. Estimación de la RFBDpc municipal en la Región de Murcia*”. Actas de la XVII Reunión Asepelt-España, Almería (CDROM).
- MANSKI C.F. (1993) “*Identification of endogenous social effects: the reflection problem*”. *Review of Economic Studies* 60; pp. 531-542.
- MOBLEY, L.R. (2003) “*Estimating hospital market pricing: An equilibrium approach using spatial econometrics*”. *Regional Science and Urban Economics* 33; pp. 489–516.

- PACE R.K., R. BARRY, O.W. GILLEY y C.F. SIRMANS (2000) “*A method for spatial-temporal forecasting with an application to real estate prices International*”. *Journal of Forecasting*, 16; pp. 229-246.
- PENA, B. dir. (1996) “*Distribución personal de la renta en España*” Ed. Pirámide, Madrid.
- UPTON, G. y B. FINGLETON (1985) “*Spatial data analysis by example*”. Volume 1. Wiley, Chichester.
- ZELLNER, A. (1962) “*An Efficient Method of Estimating Seemingly Unrelated Regressions and Test of Aggregation Bias*”. *Journal of the American Statistical Association*, 57, pp. 348-68.

Tabla 2 : Resultados de la estimación de los modelos de la renta disponible per capita

Model	Modelo básico (1996)	Modelo básico (2001)	Modelo sp-lag (1996)	Modelo sp-lag (2001)	Modelo SSUR-LAG	Modelo SSUR-LAG*
Estimation	MCO	MCO	MV	MV	MV	MV
$\hat{\beta}_0^{96}$	7930.03 (0.0000)	-	2551.29 (0.0068)	-	2393.63 (0.0033)	1993.96 (0.0078)
$\hat{\beta}_1^{96}$	593.551 (0.0000)	-	398.13 (0.0000)	-	420.309 (0.0000)	402.319 (0.0000)
$\hat{\beta}_2^{96}$	967.464 (0.0000)	-	601.766 (0.0000)	-	486.302 (0.0000)	464.376 (0.0000)
$\hat{\beta}_0^{01}$	-	10015.2 (0.0000)	-	2089.86 (0.0317)	2185.63 (0.0129)	1505.84* (0.1457)
$\hat{\beta}_1^{01}$	-	764.213 (0.0000)	-	513.507 (0.0000)	531.395 (0.0000)	516.082 (0.0000)
$\hat{\beta}_2^{01}$	-	1106.02 (0.0000)	-	599.101 (0.0000)	530.653 (0.0000)	471.696 (0.0000)
$\hat{\rho}^{96,96}$	-	-	0.6740 (0.0000)	-	0.6938 (0.0000)	0.7438 (0.0000)
$\hat{\rho}^{01,01}$	-	-	-	0.7854 (0.0000)	0.7759 (0.0000)	0.3354* (0.0703)
$\hat{\rho}^{01,96} + \hat{\rho}^{01,01}$	-	-	-	-	-	0.6422 (0.0022)
LIK	-389.61	-407.03	-378.48	-391.30	-746.633	-743.590
JB	0.5980** (0.7415)	0.6075 (0.7380)	-	-	-	-
Breush-Pagan	1.0602** (0.5885)	0.9528** (0.6210)	2.2835** (0.3193)	0.4462** (0.8000)	-	-
White	8.5520** (0.1283)	13.2635 (0.0210)	-	-	-	-
LMERR	11.2993 (0.0008)	10.2751 (0.0013)	4.6090 (0.0318)	6.1169 (0.0134)	0.8948** (0.3442)	2.0834** (0.1489)
LMERR robusto	0.2705** (0.6030)	0.5974** (0.4396)	-	-	-	-
LMLAG	20.6220 (0.0000)	28.0338 (0.0000)	-	-	-	-
LMLAG robusto	9.5932 (0.0020)	18.3561 (0.0000)	-	-	-	-
Test LM diagonalidad	-	-	-	-	46.2902 (0.0000)	49.3154 (0.0000)
Test Wald sobre la homogen. de los coeficientes	$\hat{\beta}_0$	-	-	-	9.8219 (0.0201)	-
	$\hat{\beta}_1$	-	-	-	0.4178** (0.5180)	-
	$\hat{\beta}_2$	-	-	-	4.1019 (0.0428)	-
	$\hat{\rho}^{t,t}$	-	-	-	0.9313** (0.3345)	-

Notas: *, ** es inferior al 95% y 90% de confianza, respectivamente. Entre paréntesis el nivel de significación de los estadísticos t-Student (en el caso MCO) o z (en el caso MV). *Modelo básico* es modelo no espacial. *MCO* es estimación por mínimos cuadrados ordinarios. *MV* es estimación máximo-verosímil. *Modelo sp-lag* es el modelo del retardo espacial ("spatial-lag"). *LIK* es logaritmo de verosimilitud. *JB* es test Jarque-Bera de normalidad de los residuos. *Breush-Pagan* es test de Breush-Pagan de homoscedasticidad. *White* es test de White de homoscedasticidad. *LMERR* es test del Multiplicador de Lagrange de autocorrelación espacial en los residuos. *LMERR robusto* es test del Multiplicador de Lagrange de autocorrelación espacial en los residuos robusto a la presencia de

dependencia espacial sustantiva en la variable endógena. *LMLAG* es test del Multiplicador de Lagrange de la variable endógena espacialmente retardada como adicional al modelo. *LMLAG robust* es test del Multiplicador de Lagrange de la variable endógena espacialmente retardada robusto a la existencia de dependencia espacial en los residuos. *Tests LM diagonalidad* es el test sobre la diagonalidad de la matriz de covarianzas de los residuos del modelo basado en el multiplicador de Lagrange.