Autor(es): **Angel Alañón Pardo**

Título: **El uso práctico de las técnicas de econometría espacial: La productividad del trabajo industrial**

Resumen:

En este trabajo se describen, con un carácter eminentemente práctico, algunas de las técnicas de Estadística y de Econometría Espacial diseñadas para el trabajo empírico con datos y con procesos espaciales. En primer lugar se justifica la necesidad de su uso. Luego se describen las técnicas espaciales más generales, cuya utilización se ejemplificará mediante el análisis y la posterior contrastación de un sencillo modelo econométrico explicativo de la productividad del trabajo en el sector industrial.

Angel Alañón Pardo

Universidad Complutense de Madrid

Facultad de Ciencias Económicas y Empresariales

Departamento de Economía Aplicada I

E-28223 Madrid

-

-

INDICE

1.Introducción

El principal objetivo de este trabajo es difundir el conocimiento y el uso de las técnicas más utilizadas en econometría y en estadística espacial.

Estas técnicas están especialmente diseñadas para el trabajo empírico con datos y con procesos espaciales. Su importancia radica en que muchas veces su uso es indispensable para obtener resultados válidos. Sin embargo, el conocimiento y la utilización de estas técnicas, muy difundidas en gran parte de las Ciencias Sociales y de las de la Naturaleza, todavía no están muy generalizados entre los economistas (Anselin y Florax, 1995).

Para intentar paliar esta carencia, en el presente trabajo vamos a proceder a la descripción de algunas de estás técnicas, ejemplificando su uso con datos de la productividad del trabajo industrial de las regiones europeas.

A este respecto hay que destacar que de la gran cantidad de trabajos publicados sobre convergencia y crecimiento regional en la Unión Europea, a la hora de tener en cuenta el espacio casi ninguno va más allá de un análisis descriptivo que no utiliza técnicas específicas del análisis espacial.

Para ello, comenzaremos justificando el uso de técnicas específicas para el trabajo con datos espaciales, por las características propias de éstos y de los fenómenos que se trata de estudiar, fundamentalmente por la existencia de autocorrelación espacial.

Seguidamente, realizaremos un análisis exploratorio de la distribución espacial de la productividad. El principal objetivo de este apartado será determinar, al menos de un modo tentativo, si la productividad regional del trabajo en la industria europea se distribuye de manera aleatoria o si presenta algún tipo de asociación espacial

Después se expondrá un modelo sencillo explicativo de la productividad del trabajo, que se contrastará econométricamente tratando de recoger las pautas de asociación, en caso de que éstas existan, apuntadas en el apartado anterior.

Para terminar, veremos las conclusiones que se pueden derivar de los resultados del modelo, delimitando su alcance y sus limitaciones, y proponiendo nuevas líneas y posibilidades de investigación.

2. Las Técnicas de Análisis Espacial

En este apartado veremos, en primer lugar, las características distintivas de los datos y procesos espaciales, que obligan a utilizar técnicas de análisis específicamente desarrolladas o adaptadas para ellos. Después trataremos la medición de los conceptos definidos en el punto anterior, e introduciremos las técnicas de análisis espacial que vamos a utilizar a lo largo del trabajo.

2.1 La Especificidad de los Fenómenos y de los Datos Espaciales

El trabajo con datos y con fenómenos o procesos espaciales tiene algunas características específicas que impiden el uso de algunas de las técnicas econométricas convencionales, entendiendo como tales las que se basan en los desarrollos conseguidos para series temporales.

Conscientes de este hecho Paelinck y Klaassen (1979) destacaron cinco características que obligaban a la utilización de lo que desde entonces se conoce como Econometría Espacial:

1. - El papel de la interdependencia espacial en los modelos espaciales
2. – La asimetría en las relaciones espaciales.
3. – La importancia de factores explicativos radicados en otros lugares,
4. – La diferenciación entre interacción ex ante y ex post.
5. – La modelización explícita del espacio.

Detrás de estas cinco características subyace la existencia de lo que se denomina autocorrelación o dependencia espacial. La autocorrelación espacial se puede definir como la coincidencia de valores similares con una situación espacial similar (Anselín y Bera, 1996).

La autocorrelación espacial tiene dos causas fundamentales. Por un lado la existencia de procesos espaciales, como se deja ver en las características mencionadas anteriormente. Y, por el otro, por errores de diversa naturaleza: imperfección de los datos, desajuste entre el área y el fenómeno que se trata de explicar etc (Anselin, 1988).

Aparte de la existencia de autocorrelación espacial, hay que destacar el alto grado de heterogeneidad que suelen presentar los datos espaciales.

Se puede argumentar que tanto la autocorrelación como la heterogeneidad no son exclusivos del trabajo con datos espaciales, ya que también están presentes en el trabajo con series temporales. Sin embargo mientras que la autocorrelación en series temporales sólo puede seguir una única dirección, con datos espaciales puede ser multidirecciónal. Los datos espaciales presentan una mayor heterogeneidad que los temporales -piénsese por ejemplo en las diferencias de densidad de recursos entre municipios o entre regiones.

Además, con datos espaciales, esta heterogeneidad, que se manifiesta en el análisis de regresión en forma de heteroscedasticidad, se puede confundir con la autocorrelación espacial, lo que complica en gran medida el análisis (Anselin y Bera, op. cit.).

Si se trabaja con un modelo econométrico en el que existe autocorrelación espacial, y ésta no se elimina o modeliza correctamente entonces ni el ajuste, ni la inferencia, ni muchos contrastes de hipótesis serán fiables. Dependiendo de la naturaleza de dicha autocorrelación las estimaciones del modelo podrán ser sesgadas, inconsistentes o ineficientes.

2.2. Dependencia Espacial: Expresión, Detección y Técnicas para su Tratamiento

a) La expresión de la dependencia espacial

Como veíamos en el punto anterior el carácter multidireccional es uno de los factores distintivos del análisis espacial frente al temporal. La forma más común de recoger esta posible interacción multidireccional es mediante el uso de las denominadas matrices de ponderación o de ordenación espacial.

Se trata de matrices cuadradas, representadas generalmente por W, en las que cada elemento, $w_{ij,}$ recoge la interacción entre las observaciones que representa.

Los elementos de la matriz pueden ser de diversa naturaleza. Así nos podemos encontrar con las matrices de distancias, en las que w_{ij} suele representar la inversa de la distancia entre las observaciones i y j –ya que la tesis que se trata de corroborar en la mayoría de los procesos espaciales es que la interacción es función inversa de la distancia. En las llamadas matrices de contactos cada elemento está representado por un 1 o por un 0, dependiendo de si se supone que hay o no proximidad entre cada par de observaciones. Normalmente se suele considerar que existe proximidad si las observaciones comparten una frontera común o sí la distancia que las separa está dentro del umbral fuera del cual se supone que el proceso espacial desaparece o pierde fuerza.

Aunque las matrices mencionadas en el párrafo anterior son las más utilizadas, también es corriente el uso de matrices construidas en función de la longitud de la frontera que comparten las observaciones, del área etc.

Por razones fundamentalmente operativas, estas matrices se suelen estandarizar por filas, esto es, cada elemento se suele dividir por el sumatorio de los elementos de su fila, con el objeto de que el nuevo sumatorio sea la unidad.

b) La detección de la autocorrelación espacial

En este apartado vamos a estudiar algunas de los procedimientos que se utilizan para detectar la autocorrelación espacial tanto en el análisis descriptivo como en el análisis de regresión.

Antes de describir esos procemientos vamos a introducir los conceptos de autocorrelación espacial positiva y de autocorrelación espacial negativa.

La definición de autocorrelación espacial dada en el apartado 2.1 como coincidencia de valores similares en localizaciones próximas responde al caso más corriente de autocorrelación espacial: la autocorrelación espacial positiva.

Si cogemos el mapa de cualquier país de mundo podremos encontrar regiones que se caractericen por tener municipios grandes o pequeños, ya sea en población, en extensión o en riqueza, muy próximos entre si.

Por oposición, en el caso de la autocorrelación espacial negativa la asociación entre la localización de valores similares es menor que si estos estuvieran gobernados por un proceso aleatorio. En Odland (1988) se propone como ejemplo extremo de autocorrelación espacial negativa un tablero de ajedrez donde el color de las casillas indica el distinto valor de las observaciones.

No obstante, hay que destacar que, normalmente, de lo que se trata es de estudiar si hay o no autocorrelación espacial positiva.

b1) La detección en el análisis exploratorio

Los estadísticos mas utilizados para la detección de la dependencia espacial son la I de Moran, la c de Geary y el Scatterplot de Moran .

La I de Moran se expresa formalmente como:

I= N / Soå $_{i}$ å $_{j}$ w_{ij} (x_i - m) (x_j - m) / å $_{i}$ (x_i - m)^2

N es el número de observaciones, w_{ij} el elemento de la matriz de ordenación espacial correspondiente al par de observaciones i y j, x_i y x_j son las observaciones para las localizaciones i y jj, y So es una constante de escala (So =å $_{i}$ å $_{j}$ w_{ij}). Cuando se trabaja con matrices estandarizadas, dado que cada fila suma la unidad, So equivale a N, y el estadístico se convierte en el ratio de un producto cruzado espacial partido por una varianza:

I = å $_{i}$ å $_{j}$ w_{ij} (x_i - m) (x_j - m) / å $_{i}$ (x_i - m)^2

Si el valor del coeficiente de la I es mayor que su valor esperado, la media, estaremos ante un caso de autocorrelación espacial positiva, siendo negativa en el caso contrario.

La c de Geary es similar a la I de Moran:

c = (N-I) / 2 So [å $_{i}$ å $_{j}$ w_{ij} (x_i - x_j) ^{2} / å $_{i}$ (x_i - m)^2]

El valor esperado para este estadístico es la unidad. Valores por debajo de la unidad son indicativos de autocorrelación espacial positiva, mientras que los que están por encima indican autocorrelación espacial negativa.

En ambos estadísticos la inferencia se fundamenta en la z estandarizada, esto es restándole su media y dividiendo el resultado por la desviación típica. Generalmente, en función de la naturaleza de los datos y del fenómeno que se trate de estudiar, asumiremos que esta z sigue una distribución normal, pero también se puede hacer el supuesto de que sigue una distribución aleatoria.

Cuando el estadístico es significativo su probabilidad asociada será muy baja. En el caso de la I de Moran los estadísticos positivos indican autocorrelación espacial positiva y los negativos negativa. Sin embargo en la c de Geary los estadísticos positivos indican autocorrelación negativa y los negativos positiva.

Basándonos en la I de Moran, mediante el Scatterplot de Moran podemos identificar tanto a las agrupaciones de observaciones que tienen un valor similar (autocorrelación espacial positiva) como a las de valores heterogéneos (autocorrelación espacial negativa). Esta herramienta, que nos permite ver el grado de estabilidad que tiene el proceso espacial, requiere matrices de ordenación espacial estandarizadas.

b2) La detección en el análisis de regresión

Como se indicó en el apartado 2.1 la utilización de técnicas de estimación convencionales, como puedan ser los mínimos cuadrados ordinarios, en presencia de autocorrelación espacial puede falsear el significado de las medidas de ajuste, de la inferencia y de las asunciones que normalmente se suponen en sus estimadores (insesgadez, mínima varianza y consistencia).

En este apartado vamos a ver las formas que puede adoptar la autocorrelación espacial en el análisis de regresión y los test que se han diseñado para detectarla.

La autocorrelación espacial se puede dar en la variable dependiente, en el término de error, o en ambos a la vez.

Sea una ecuación cualquiera:

(1) y = bx + x ,

E (x) = 0 ; E ($x_{ix\ j}$) = 0 i ¹ j; E (x_i^2) = s^2

Si la variable dependiente está autocorrelacionada espacialmente nos encontraremos con que el modelo correcto podría ser:

2. y = r Wy + bx + x

Donde el valor de y en cada punto del espacio se relaciona con los demás valores de y mediante el término autorregresivo espacial r Wy.

La consecuencia de no incluir ese término autorregresivo espacial es que nuestras estimaciones serán sesgadas e inconsistentes.

Siguiendo la terminología anglosajona estos modelos se suelen denominar tipo "lag", por el carácter autorregresivo espacial de la variable dependiente.

También puede ocurrir que sean los términos de error los que estén autocorrelacionados espacialmente:

(3) y = bx + Î

a) Î = lWÎ + x

b) Î = lWx + x

En (3a) nos encontramos con que el término de error sigue un proceso autorregresivo espacial mientras que en (3b) se trata de una media móvil espacial.

Las consecuencias de no considerar cualquiera de estas modelizaciones del término de error son que nuestros estimadores ya no serán eficientes, por lo que la inferencia y las medidas de ajuste tradicionales serán equívocas.

Estos modelos se suelen denominar tipo "error autorregresivo" o "error media móvil".

Y, por último, también nos podemos encontrar con que tanto la variable dependiente como el término de error estén autocorrelacionados espacialmente:

(4) y = r Wy + bx + Î

a) Î = lWÎ + x

b) Î = lWx + x

Evidentemente, si el proceso que estamos estudiando sigue el tipo de autocorrelación señalado en (4) las consecuencias de modelizarlo correctamente serán más serias que las de estimar un modelo convencional en el caso de (2) o de (3).

Estos dos últimos modelos son más difíciles de modelizar que los anteriores y, de hecho, según nuestro conocimiento, no hay ningún programa econometríco que lo haga. No obstante, el que vamos a utilizar en este trabajo, SpaceStat 1.80, incluye un test para la detección del (4b), que también se conoce como SARMA, acrónimo inglés para proceso autorregresivo espacial y media móvil.

A continuación vamos a analizar algunos de los test que sirven para detectar la autocorrelación espacial en el análisis de regresión: dos generales, que no especifican la causa de la autocorrelación; dos para la autocorrelación de la variable dependiente, dos para la del término de error cuando este sigue un proceso autorregresivo, y uno para el modelo SARMA.

Los test generales son aquellos que indican la presencia de autocorrelación espacial pero que no hacen referencia a la naturaleza de ésta. Los más utilizados son el test de Moran y el de Kelejian-Robinson.

El primero es igual al que vimos en el apartado 2.1.b2 y se aplica sobre los residuos de la regresión por mínimos cuadrados. Requiere que los términos de error sigan una distribución normal, y, recoge algunos errores de especificación por lo que a veces no es muy fiable (Anselin y Rey, 1991).

El test de Kelejian-Robinson tiene la ventaja, frente al de Moran, de no necesitar que los términos de error sigan una distribución normal. Además, tampoco requiere que el modelo sea lineal. Sin embargo, no esta recomendado para muestras pequeñas y en los experimentos llevados a cabo por Anselin y Florax (1995) no consiguió buenos resultados. Se distribuye siguiendo una x^2 con tantos grados de libertad como variables explicativas (incluyendo al término constante).

El resto de los test se basan en el multiplicador de Lagrange (LM) y requieren que el término de error siga una distribución normal.

Los dos primeros detectan la autocorrelación espacial en la variable dependiente y se distribuyen como una x^2 con un grado de libertad. La hipótesis nula en ambos es que en la ecuación (2) r = 0, esto es, que no hay autocorrelación espacial en la variable dependiente. Sin embargo mientras que en el primero se considera también que no hay autocorrelación espacial en los términos de error, l = 0, el segundo es robusto ante la presencia de ese otro tipo de autocorrelación.

Y lo mismo ocurre con los test diseñados para detectar la autocorrelación espacial en los términos de error. Se distribuyen también como una x^2 con un grado de libertad, y la hipótesis nula en ambos es que en la ecuación (3) l = 0. De igual forma, el primer test considera que no hay autocorrelación en la variable dependiente, r = 0, y el segundo es robusto ante este tipo de autocorrelación.

Por último hay que mencionar el test para el modelo SARMA, autocorrelación espacial en la variable dependiente y en las perturbaciones, siguiendo éstas un proceso de media móvil. En este test la hipótesis nula en (4b) es que l = r = 0, esto es, que no hay autocorrelación espacial. Este estadístico se distribuye como una x^2 con dos grados de libertad.

c) Las técnicas para el tratamiento de la dependencia espacial

Una vez que detectamos la presencia de dependencia espacial tenemos dos opciones o eliminarla, para así poder aplicar las técnicas de estimación convencionales, o incoporarla a nuestro modelo. En este apartado vamos a ver las principales formas de modelizar la autocorrelación espacial, en concreto el modelo tipo "lag" y el modelo "error autorregresivo".

Dadas sus características, simultaneidad para el modelo "lag", y errores no esféricos en el "error", éstos no se pueden estimar mediante los mínimos cuadrados ordinarios, sino que hay que acudir a otras técnicas, como pueden ser la estimación por máxima verosimilitud o las variables instrumentales.

Si nuestro modelo cumple la hipótesis de normalidad, en este trabajo utilizaremos la técnica de estimación por el procedimiento de máxima verosimilitud.

El ajuste en los modelos estimados por este procedimiento que incoporan un componente espacial, ya sea tipo "lag" o tipo "error" no se interpretan como en los mínimos cuadrados ordinarios. Así, aunque el programa utilizado arroja un R^2, éste ni tiene el mismo significado que en mínimos cuadrados ordinarios, ni es comparable entre modelos.

Los medidas de ajuste que sí son susceptibles de comparación entre modelos, incluso con los mínimos cuadrados ordinarios, son el logaritmo de máxima verosimilitud (LIK), y los criterios de información de Akaike (AIC), y de Schwartz (SC). El ajuste será mejor cuanto mayor LIK tenga, y menores AIC y SC. A la hora de utilizar estos criterios hay que tener en cuenta que los dos últimos tienden a favorecer a los modelos "error" frente a los "lag" (Anselin, 1992).

La significación de los coeficientes no se hace por medio de la "t" de Student, en su lugar se utiliza una "z" estandarizada que tiene su misma interpretación.

En ambos modelos se utiliza el ratio de verosimilitud (LR) para comprobar la validez del parámetro autorregresivo.

Cuando se trabaja con muestras pequeñas para ver si esta bien especificado el modelo se verifica que los resultados del test de Wald, (W) del ratio de máxima verosimilitud (LR) y del multiplicador de lagrange (LM), sigan el orden siguiente:

$$W \geq LR \geq LM$$

En el "error" también se verifica que cumpla hipótesis de "common factor" para comprobar si el modelo está bien especificado.

Por último hay que verificar que no queda autocorrelación espacial sin modelizar mediante test basados en el multiplicador de Lagrange.

3. Exposición y contrastación del modelo

El objetivo final de este apartado no es conseguir un modelo que explique totalmente la productividad del trabajo en la industria de las regiones europeas, sino averiguar si existe un componente espacial y calibrar su importancia.

En primer lugar expondremos un modelo explicativo de la productividad del trabajo bastante sencillo. Después someteremos a las variables que lo componen a un análisis exploratorio para ver si se observan pautas de asociación espacial.

Luego contrastaremos el modelo por la técnica de los mínimos cuadrados ordinarios, con test de autocorrelación espacial, para, si se confirman los resultados del análisis exploratorio, modelizar explícitamente esa autocorrelación o dependencia espacial.

3.1. Exposición del modelo

Nuestro modelo inicial tienen como variable dependiente a la productividad del trabajo industrial (YL), y una sola variable explicativa (TM) que es el tamaño medio de los establecimientos productivos, según muestra la ecuación (5).

(5) YL = f (TM)

Así, suponemos que cuanto mayor sea el establecimiento productivo más fácil será que acceda a economías internas de escala, que incorpore innovaciones productivas o que realice actividades de Investigación y Desarrollo, factores que suelen favorecer el crecimiento de la productividad. Asimismo también parece existir cierta correlación entre tamaño y actividad exportadora, y entre esta última y la productividad.

Estos planteamientos no pretenden ser originales ya que podríamos encontrar las mismas ideas en autores como Kaldor o Kalecki, aunque el referente directo en que nos hemos inspirado es el trabajo de Bueno Lastra (1990).

3.2. Análisis Exploratorio

En este apartado se va a comprobar si existen indicios de que la productividad del trabajo en la industria europea está autocorrelacionada espacialmente.

La medida de la productividad del trabajo industrial (YL) que utilizaremos aquí será el cociente entre Valor Añadido Bruto Industrial a precios de mercado en ecus y el número de ocupados en la industria. El indicador que vamos a utilizar para la variable TM es el tamaño medio de los establecimientos industriales que tienen más de 20 trabajadores. El año de referencia será 1989, por ser del que se dispone de más datos. Las regiones estudiadas serán las NUTs II de Alemania, Francia, Holanda, Italia y Luxemburgo. Estos datos provienen de Eurostat.

Los estadísticos utilizados para detectar la existencia de autocorrelación espacial seran los ya mencionados anteriormente: la c de Geary y la I de Moran. En el caso de que haya indicios de asociación espacial presentaremos los resultados del Scattetplot de Moran para ilustrarla de alguna forma. Previamente se someterá a YL a un test de normalidad para ver si podemos incluir este supuesto en el cómputo de estos estadísticos.

Las matrices de ordenación espacial que se vamos a utilizar son M1, M2, M3, SM1 SM2, y SM3. M1 considera que hay interacción entre las regiones que comparten frontera. M2 es una matriz de orden superior que considera que una región tendrá interacción no sólo con sus regiones fronterizas, sino también con las que compartan frontera con estas últimas. M3 se construye con la misma lógica que M2 suponiendo que existe interacción también con las colindantes a las que se añadieron en M2. SM1, SM2 y SM3 son, respectivamente, M1, M2 y M3 estandarizadas.

Como podemos ver en c.1. los resultados del test de Wald nos permiten no rechazar la hipótesis de que YL sigue una distribución normal, y calcular así los estadísticos de Moran y de Geary siguiendo esta asunción.

c.1. Test de normalidad de Wald		
Variable	**Test**	**Prob.**
YL	1,970687	0,37331088

Los resultados de los test de Moran y de Geary , c.2, son altamente significativos, indicando la existencia de autocorrelación espacial positiva, según todas las matrices utilizadas.

Parece también, que la productividad del trabajo industrial experimenta un proceso autorregresivo decreciente. Según los resultados de la I de Moran, con matrices con y sin estandarizar, y de la c Geary con matrices estandarizadas la asociación espacial crece de M1 y SM1 a M2 y SM2 para decrecer depués en M3 y SM3. Con la c de Geary y las matrices sin estandarizar el proceso sí sería autorregresivo decreciente estricto ya que el valor más significativo es el de M1 y el menor el de M3.

Los resultados del Moran Scatterplot, c.3, confirman la asociación espacial positiva de la productividad del trabajo en las regiones comunitarias. Sin embargo, según éstos la asociación positiva es más fuerte (valor de la I), y más extendida (número de observaciones a las que afecta) con la matriz de primer orden , SM1.

c.2 Asociación Espacial de YL según I de Moran y c de Geary					
Matriz	**I**	**Media**	**Desv. Est**	**Z . val.**	**Prob.**
M1	0,3458173	-0,015	0,083610	4,320060	0,000016
M2	0,2289519	-0,015	0,046892	5,210640	0,000000
M3	0,1430578	-0,015	0,032429	4,885805	0,000001
SM1	0,4207043	-0,015	0,090731	4,806400	0,000002
SM2	0,3129794	-0,015	0,050395	6,515769	0,000000
SM3	0,1932247	-0,015	0,034593	6,030315	0,000000
Matriz	**c**	**Media**	**Desv. Est**	**Z . val.**	**Prob.**
M1	0,3935491	1,000	0,112523	-5,389562	0,000000
M2	0,6139006	1,000	0,078250	-4,934169	0,000001
M3	0,7541527	1,000	0,058328	-4,214937	0,000025
SM1	0,4533712	1,000	0,102150	-5,351241	0,000000
SM2	0,6187243	1,000	0,058927	-6,470285	0,000000
SM3	0,7637366	1,000	0,041098	-5,748802	0,000000

Por tanto una vez detectada la presencia de autocorrelación espacial positiva en la variable dependiente, hay que comprobar si ésta afecta también al modelo de regresión por mínimos cuadrados ordinarios realizando test de autocorrelación sobre los residuos

c.3 MORAN SCATTER PLOT YL MATRICES SM1, SM2 y SM3			
SM1 MORAN'S I: 0.420704			
POSITIVE SPATIAL ASSOCIATION		**NEGATIVE SPATIAL ASSCIATION**	
Large x – Large Wx: **29 observations**	**Small x – Small Wx:** **23 observations**	**Large x – Small Wx:** **7 observations**	**Small x – Large Wx:** **7 observations**
1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 18, 23, 25, 26, 27, 28, 29, 30, 43, 53, 54, 56, 57, 58, 59, 61, 62, 63, 64	12, 13, 14, 15, 16, 17, 21, 22, 35, 36, 37, 38, 39, 40, 41, 42, 45, 46, 47, 48, 49, 50, 52	19, 31, 34, 44, 51, 55, 60	6, 20, 24, 32, 33, 65, 66
SM2 MORAN'S I: 0.312979			
POSITIVE SPATIAL ASSOCIATION		**NEGATIVE SPATIAL ASSCIATION**	
Large x – Large Wx: **32 observations**	**Small x – Small Wx:** **18 observations**	**Large x – Small Wx:** **4 observations**	**Small x – Large Wx:** **12 observations**
1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 18, 19, 23, 25, 26, 27, 28, 29, 43, 44, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64	15, 16, 20, 22, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 45, 48, 49, 50	30, 31, 34, 51	6, 12, 13, 14, 17, 21, 24, 46, 47, 52, 65, 66
SM3 MORAN'S I: 0.193225			
POSITIVE SPATIAL ASSOCIATION		**NEGATIVE SPATIAL ASSCIATION**	
Large x – Large Wx: **29 observations**	**Small x – Small Wx:** **13 observations**	**Large x – Small Wx:** **7 observations**	**Small x – Large Wx:** **17 observations**
1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 18, 19, 23, 26, 27, 28, 29, 43, 44, 54, 55, 56, 57, 59, 60, 61, 62, 63, 64	22, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 50, 52	25, 30, 31, 34, 51, 53,58	6, 12, 13, 14, 15, 16, 17, 20, 21, 24, 45, 46, 47, 48, 49, 65, 66

3.3. Diagnósticos espaciales y modelización espacial

La estimación del modelo (5) por el procedimiento de los mínimos cuadrados proporciona los siguientes resultados:

ORDINARY LEAST SQUARES ESTIMATIONDEPENDENT VARIABLE YL OBS 66 VARS 2 DF 64 R2 0.1498 R2-adj 0.1365 LIK -212.689 AIC 429.378 SC 433.758

RSS 2433.05 F-test 11.2730 Prob 0.00132776

SIG-SQ 38.0165 (6.16575) SIG-SQ(ML) 36.8645 (6.07161)

VARIABLE COEFF S.D. t-value Prob

CONSTANT 18.6431 2.03625 9.155584 0.000000

TM 0.0694427 0.0206827 3.357524 0.001328

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 5.172640

TEST ON NORMALITY OF ERRORS

TEST DF VALUE PROB

Jarque-Bera 2 0.269227 0.874054

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST DF VALUE PROB

Breusch-Pagan test 1 0.100374 0.751381

SPECIFICATION ROBUST TEST

TEST DF VALUE PROB

White 2 0.816402 0.664845

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHTS MATRIX M1 (not row-standardized)

TEST MI/DF VALUE PROB

Moran's I (error) 0.376885 **4.802709 0.000002**

Lagrange Multiplier (error) 1 **18.607550 0.000016**

Robust LM (error) 1 **15.315312 0.000091**

Kelejian-Robinson (error) 2 **28.141204 0.000001**

Lagrange Multiplier (lag) 1 3.397711 0.065287

Robust LM (lag) 1 0.105474 0.745357

Lagrange Multiplier (SARMA) 2 **18.713024 0.000086**

FOR WEIGHTS MATRIX M2 (not row-standardized)

TEST MI/DF VALUE PROB

Moran's **I** (error) 0.249350 **6.038760 0.000000**

Lagrange Multiplier (error) 1 **22.321004 0.000002**

Robust LM (error) 1 **18.050208 0.000022**

Kelejian-Robinson (error) 2 **36.833201 0.000000**

Lagrange Multiplier (lag) 1 **5.356336 0.020647**

Robust LM (lag) 1 1.085539 0.297462

Lagrange Multiplier (SARMA) 2 **23.406544 0.000008**

FOR WEIGHTS MATRIX M3 (not row-standardized)

TEST MI/DF VALUE PROB

Moran's **I** (error) 0.159914 **5.992533 0.000000**

Lagrange Multiplier (error) 1 **16.085135 0.000061**

Robust LM (error) 1 **11.150594 0.000840**

Kelejian-Robinson (error) 2 **25.759631 0.000003**

Lagrange Multiplier (lag) 1 **8.643457 0.003282**

Robust LM (lag) 1 3.708916 0.054123

Lagrange Multiplier (SARMA) 2 **19.794051 0.000050**

FOR WEIGHTS MATRIX SM1 (row-standardized weights)

TEST MI/DF VALUE PROB

Moran's **I** (error) 0.429983 **5.052086 0.000000**

Lagrange Multiplier (error) 1 **21.079920 0.000004**

Robust LM (error) 1 0.217617 0.640861

Kelejian-Robinson (error) 2 **28.141204 0.000001**

Lagrange Multiplier (lag) 1 **21.198304 0.000004**

Robust LM (lag) 1 0.336002 0.562146

Lagrange Multiplier (SARMA) 2 **21.415921 0.000022**

FOR WEIGHTS MATRIX SM2 (row-standardized weights)

TEST MI/DF VALUE PROB

Moran's **I** (error) 0.324381 **7.057762 0.000000**

Lagrange Multiplier (error) 1 **34.659602 0.000000**

Robust LM (error) 1 **4.847518 0.027686**

Kelejian-Robinson (error) 2 **36.833201 0.000000**

Lagrange Multiplier (lag) 1 **29.889300 0.000000**

Robust LM (lag) 1 0.077216 0.781106

Lagrange Multiplier (SARMA) 2 **34.736818 0.000000**

FOR WEIGHTS MATRIX SM3 (row-standardized weights)

TEST MI/DF VALUE PROB

Moran's **I** (error) 0.203477 **6.833154 0.000000**

Lagrange Multiplier (error) 1 **24.664459 0.000001**

Robust LM (error) 1 **4.957623 0.025976**

Kelejian-Robinson (error) 2 **25.759631 0.000003**

Lagrange Multiplier (lag) 1 **19.723064 0.000009**

Robust LM (lag) 1 0.016228 0.898634

Lagrange Multiplier (SARMA) 2 **24.680687 0.000004**

Como podemos ver el ajuste del modelo es bastante bajo, la variable explicativa es signficativa, y, con las evidencias disponibles, podemos aceptar las hipótesis de homoscedasticad y de normalidad, y confiar en que no existen problemas de multicolinealidad. Sin embargo, tal y como se intuía tras el análisis exploratorio, los test de autocorrelación espacial son bastante signficativos por lo que tanto la significación como el ajuste de nuestro modelo están en cuestión.

Por ello vamos a incorporar al modelo esta autocorrelación espacial, ya sea como una retardo de la variable dependiente, tipo "lag", o como términos de error autorregresivos, tipo "error".

Dado que todas las matrices de ordenación espacial presentan estadísticos de autocorrelación espacial bastante significativos para cada una de ellas se han probado ambos modelos. Como se indica en el anexo 2 sólo parece que tres modelos eliminen la autocorrelación espacial de los residuos: dos tipo "error", uno con SM1 (Mo1) y otro con SM2 (Mo2), aunque el primero casi permite el rechazo, y un tipo "lag" con SM1 (Mo3).

Como se puede comprobar en el anexo 3 en los tres modelos tanto el término constante como la variable tamaño medio y el término autorregresivo son significativos, Ninguno de los tres presenta problemas de heteroscedasticiad, y los tipo "error" Mo1 y Mo2, verifican la hipótesis de common factor.

Los tres modelos consiguen un mejor ajuste (LIK, AIC y SC) que el de mínimos cuadrados ordinarios. El modelo con peor ajuste es Mo1. Mo2 y Mo3 alcanzan un ajuste similar, ya que aunque el de Mo2 sea algo mayor siguiendo los criterios de información no hay que olvidar que, estos tienden a favorecer a los modelos "error" frente a los "lag".

Sin embargo en ninguno de los modelos se verifica estrictamente el orden de magnitud esperado para el test de Wald, el ratio de verosimilitud, y el multiplicador de Lagrange ($W \geq LR \geq LM$):

Mo1: $(5.22)^2 \geq 19.27 \leq 21.07$; Mo2: $(6.42)^2 \geq 19.52 \leq 34.65$;

Mo3: $(5.83)^2 \geq 20.77 \leq 21.19$

Como se puede apreciar, en todos ellos LR < LM, sobre todo en Mo2. Mo1 y Mo3 se aproximan más, y dado que éste último es el que muestra menos discrepancia y que consigue un mejor ajuste nos quedamos con este modelo, que quedaría como sigue:

6. $YL = 5.14 + 0.58SM1YL + 0.05TM$

(2.63) (0.10) (0.01)

Elasticidad (YL/ WYL) = 0,58; Elasticidad (YL/TM)= 0,21.

4. Conclusiones

Los resultados del análisis de regresión confirman tanto los obtenidos por el análisis exploratorio, como los de los diagnostícos especiales del modelo de mínimos cuadrados ordinarios: la productividad del trabajo industrial en la regiones europeas está autocorrelacionada espacialmente. La interpretación económica de este hecho puede ser la existencia de algún tipo de asociación espacial interregional en la productividad del trabajo industrial. Esta asociación, que vincula la productividad de una región con la de las adyacentes, se puede medir mediante la elasticidad del componente autorregresivo de la ecuación (6).

Las razones que provocan esta asociación espacial pueden ser varias: la integración de la industria europea, la competencia interna y externa, externalidades de diversa naturaleza etc. Sin embargo, con un modelo tan sencillo como el que se ha expuesto no es posible determinar a que obedece.

A este respecto proponemos varias mejoras que pueden incrementar nuestro conocimiento sobre el tema:

1-Inclusión de más variables explicativas y el perfeccionamiento de las existentes para mejorar el ajuste de modelo y explicar mejor la productividad.

2-Añadir más observaciones y probar con otros años de referencia.

3-Hacer estudios sectoriales para ahondar en las causas y en la naturaleza de esta asociación.

ANEXO 1: NUTs II INCLUIDAS EN EL ESTUDIO Y CONTIGÜIDAD SEGÚN M1

(1)ILE DE FRANCE: 2,3,4,5,7 (2)CHAMPAGNE-ARDENNE: 1,3,7,9,11

(3)PICARDIE: 1,2,4,8 (4) HAUTE-NORMANDIE : 1,3,5,6

(5)CENTRE: 1,4,6,7,12,14,17,19 (6)BASSE-NORMANDIE: 4,5,12,13

(7)BOURGOGNE:1,2,5,11,18,19 (8)NORD-PAS-DE-CALAIS: 3

(9)LORRAINE:2,10,11,43,65 (10)ALSACE: 9,11,56

(11)FRANCHE-COMTE: 2,7,9,10,18 (12)PAYS DE LA LOIRE:5,6,13,14

(13)BRETAGNE: 6,12 (14)POITOU-CHARENTES: 5,12,15,17

(15)AQUITAINE: 14,16,17 (16)MIDI-PYRENEES: 15,17,19,20

(17)LIMOUSIN: 5,14,15,16,19 (18)RHONE-ALPES: 7,11,19,20,21,23,24

(19)AUVERGNE: 5,7,16,17,18,20 (20)LANGUEDOC-ROUSSILLON: 16,18,19,21

(21)PROVENCE-ALPES-COTE D`AZUR: (22)CORSE: 21,31,42

18,20,22,23,25 (23)PIEMONTE: 18,21,24,25,26,30

(24)VALLE D`AOSTA: 18, 23 (25)LIGURIA: 21,23,30,31

(26)LOMBARDIA: 23,27,28,30 (27)TRENTINO-ALTO ADIGE: 26,28

(28)VENETO: 26,27,29,30 (29)FRIULI-VENEZIA GIULIA: 28

(30)EMILIA-ROMAGNA:23,25,26,28,31,33 (31)TOSCANA:23,25,26,28,31,33

(32)UMBRIA: 31,33,34 (33)MARCHE: 30,31,32,34,35

(34)LAZIO: 31,32,33,35,36,37,42 (35)ABRUZZO: 33,34,36

(36)MOLISE: 34,35,37,38 (37)CAMPANIA: 34,36,38,39

(38)PUGLIA: 36,37,39 (39)BASILICATA: 37,38,40

(40)CALABRIA: 39,41 (41)SICILIA: 40,42

(42)SARDEGNA: 22,34,41 (43)LUXEMBOURG (GRAND-DUCHE): 9,64,65

(44)GRONINGEN: 45,46,62 (45)FRIESLAND: 44,46,47

(46)DRENTHE: 44,45,47,62 (47)OVERIJSSEL: 45,46,48,49,62,63

(48)GELDERLAND:47,49,50,52,54,55,63 (49)FLEVOLAND: 47,48,50,51

(50)UTRECHT: 48,49,50,51 (51)NOORD-HOLLAND: 49,50,52

(52)ZUID-HOLLAND: 48,50,51,54 (53)ZEELAND: 54

(54)NOORD-BRABANT: 48,52,53,55,63 (55)LIMBURG (NL): 48,54,63

(56)BADEN-WUERTTEMBERG: 10,57,61,64 (57)BAYERN: 56,61

(58)BERLIN: 62 (59)BREMEN: 62

(60)HAMBURG: 62,66 (61)HESSEN: 56,57,62,63,64

(62)NIEDERSACHSEN: 44,46,47,58,59,63,66 (63)NORDRHEIN-WESTFALEN:47,48,54,55,61,62,64

(64)RHEINLAND-PFALZ: 43,56,61,63,65 (65)SAARLAND: 9,43,64

(66)SCHLESWIG-HOLSTEIN: 60,62

ANEXO 2: TEST DE AUTOCORRELACIÓN ESPACIAL MODELOS "LAG" Y "ERROR"

	MOD. LAG		**MOD. ERROR**			**MOD. LAG**		**MOD. ERROR**	
	ERROR DEP.		**LAG DEP.**			**ERROR DEP.**		**LAG DEP.**	
M1	**Val.**	**Prob.**	**Val**	**Prob.**	**SM1**	**Val**	**Prob.**	**Val**	**Prob.**
m1	12.31	0.000449	0.08	0.764631	**M1**	0.12	0.724609	0.91	0.337925
m2	13.13	0.000290	0.66	0.416336	**M2**	0.00	0.332450	0.00	0.990448
m3	6.35	0.011698	0.56	0.453247	**M3**	0.93	0.951442	0.04	0.839783
sm1	11.87	0.000568	3.44	0.063432	**Sm1**	0.46	0.493510	0.00	0.930966

sm2	22.05	0.000003	8.49	0.003568	**Sm2**	1.74	0.186842	3.41	0.064740
sm3	11.68	0.000632	4.12	0.042310	**Sm3**	0.09	0.758066	0.17	0.679850
M2	**Val.**	**Prob.**	**Val**	**Prob.**	**SM2**	**Val.**	**Prob.**	**Val**	**Prob.**
m1	11.97	0.000541	0.00	0.980134	**M1**	4.83	0.027861	0.00	0.933548
m2	10.44	0.001233	0.01	0.896340	**M2**	2.71	0.099624	0.00	0.995159
m3	4.05	0.044137	0.57	0.449902	**M3**	0.23	0.630531	0.05	0.817178
sm1	12.01	0.000527	6.74	0.009402	**Sm1**	2.38	0.122847	2.63	0.104675
sm2	17.59	0.000027	9.75	0.001784	**Sm2**	1.96	0.161199	2.79	0.094682
sm3	8.04	0.004563	3.86	0.049228	**Sm3**	0.13	0.714000	0.00	0.971865
M3	**Val.**	**Prob.**	**Val**	**Prob.**	**SM3**	**Val.**	**Prob.**	**Val**	**Prob.**
m1	9.09	0.002567	0.02	0.870272	**M1**	11.01	0.000902	0.05	0.817797
m2	8.25	0.004065	0.00	0.971024	**M2**	11.01	0.000904	0.15	0.694679
m3	1.90	0.167712	0.01	0.900146	**M3**	3.81	0.050850	0.26	0.606223
sm1	8.93	0.002804	8.99	0.002701	**Sm1**	7.36	0.006667	7.28	0.006938
sm2	14.10	0.000173	15.39	0.000087	**Sm2**	11.45	0.000712	11.54	0.000679
sm3	4.33	0.037281	6.04	0.013963	**Sm3**	3.61	0.057364	3.58	0.058174

ANEXO 3: ESTIMACIÓN Mo1, Mo2 y Mo3

Mo1: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION; MATRIX: SM1

DEPENDENT VARIABLE YL OBS 66 VARS 2 DF 64

R2 0.1480 Sq. Corr. 0.1498 R2(Buse) 0.1362

LIK -203.050 AIC 410.100 SC 414.480

SIG-SQ 24.9300 (4.99299)

VARIABLE COEFF S.D. z-value Prob

CONSTANT 18.9755 2.41647 7.852557 0.000000

TM 0.0690445 0.0214 3.226373 0.001254

LAMBDA 0.570166 0.109084 5.226856 0.000000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST DF VALUE PROB

Breusch-Pagan test 1 1.059663 0.303292

Spatial B-P test 1 1.059949 0.303227

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHTS MATRIX SM1 (row-standardized weights)

TEST DF VALUE PROB

Likelihood Ratio Test 1 19.277856 0.000011

TEST ON COMMON FACTOR HYPOTHESIS

TEST DF VALUE PROB

Likelihood Ratio Test 1 1.527171 0.216538

Wald Test 1 1.495952 0.221295

Mo2 :SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION; MATRIX: SM2

DEPENDENT VARIABLE YL OBS 66 VARS 2 DF 64

R2 0.1360 Sq. Corr. 0.1498 R2(Buse) 0.1511

LIK -202.929 AIC 409.858 SC 414.237

SIG-SQ 25.3791 (5.03777)

VARIABLE COEFF S.D. z-value Prob

CONSTANT 18.8453 2.81857 6.686097 0.000000

TM 0.0661708 0.0193035 3.427920 0.000608

LAMBDA 0.727884 0.113307 6.423973 0.000000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST DF VALUE PROB

Breusch-Pagan test 1 1.820605 0.177241

Spatial B-P test 1 1.820616 0.177240

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHTS MATRIX SM2 (row-standardized weights)

TEST DF VALUE PROB

Likelihood Ratio Test 1 19.520064 0.000010

TEST ON COMMON FACTOR HYPOTHESIS

TEST DF VALUE PROB

Likelihood Ratio Test 1 0.771458 0.379766

Wald Test 1 0.735529 0.391097

Mo3: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION; MATRIX: SM1

DEPENDENT VARIABLE YL OBS 66 VARS 3 DF 63

R2 0.3006 Sq. Corr. 0.4592

LIK -202.300 AIC 410.600 SC 417.169

SIG-SQ 24.1736 (4.91666)

VARIABLE COEFF S.D. z-value Prob

W_YL 0.589409 0.100928 5.839923 0.000000

CONSTANT 5.14736 2.63815 1.951122 0.051042

TM 0.0572872 0.0172986 3.311666 0.000927

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST DF VALUE PROB

Breusch-Pagan test 1 1.241603 0.265162

Spatial B-P test 1 1.241926 0.265100

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHTS MATRIX SM1 (row-standardized weights)

TEST DF VALUE PROB

Likelihood Ratio Test 1 20.778262 0.000005

-

-

-

ANEXO 4: ESTIMACIÓN MÍNIMOS CUADRADOS PARA ESPAÑA

ORDINARY LEAST SQUARES ESTIMATION

DEPENDENT VARIABLE YL OBS 17 VARS 2 DF 15

R2 0.4766 R2-adj 0.4417

LIK -47.5882 AIC 99.1765 SC 100.843

RSS 268.796 F-test 13.6561 Prob 0.00215927

SIG-SQ 17.9197 (4.23317) SIG-SQ(ML) 15.8115 (3.97637)

VARIABLE COEFF S.D. t-value Prob

CONSTANT 9.98732 2.26732 4.404896 0.000512

TM 0.57011 0.154275 3.695416 0.002159

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 4.177357

TEST ON NORMALITY OF ERRORS

TEST DF VALUE PROB

Jarque-Bera 2 1.194912 0.550209

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST DF VALUE PROB

Breusch-Pagan test 1 1.945649 0.163057

BIBILIOGRAFÍA

Anselin, Luc. (1988): *Spatial Econometrics: Methods and models*. Kluwer Academic, Dordrecht.

Anselin, Luc y Rey, Sergio (1991): "Properties of tests for spatial dependence in linear regression models", *Geographical Analysis,* 23, 112-31.

Anselin, Luc (1992): *SpaceStat Tutorial: A workbook for using SpaceStat in the analisys of spatial data*, Regional Research Institute, West Virginia University, Morgantown, WV.

Anselin, Luc (1995): *SpaceStat version 1.80. User's guide*, Regional Research Institute, West Virginia University, Morgantown, WV.

Anselin, Luc y Florax, Raymond (1995): "Small sample properties of test for spatial dependence in regression models: Some further results" en Anselin, Luc y Florax, Raymond, eds., *New Directions in Spatial Econometrics,* pp. 21-74. Springer-Verlag, Berlín.

Anselin, Luc y Bera, Anil K. (1996): "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics", *Research Paper 9617,* Regional Research Institure, West Virginia University, Morgantown, WV.

Aznar Grasa, Antonio, Mur Lacambra, Jesús y Trivez Bielsa, F. Javier (1986): "Métodos econométricos en el análisis regional", Ponencia presentada en la XXII Reunión de Estudios Regionales, Pamplona.

Bueno Lastra, Juan (1990): *Los desequilibrios regionales. Teoría y realidad española*, Pirámide, Madrid.

Eurostat (varios años): *Regiones: Anuario Estadístico*, Eurostat, Luxemburgo.

Eurostat (varios años): *Structure and activity of Industry, Data by regions,* Eurostat, Luxemburgo.

Gómez de Antonio, Miguel (1998): "Econometría espacial: algunos aspectos generales", *Documento de trabajo 9901*, Facultad de Ciencias Económicas y Empresariales, Universidad Complutense de Madrid.

Haining, Robert (1990): *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press, Cambridge.

López-Bazo, Enrique, Mora Corral, Antonio J., Surinach Caralt, Jordi y Vayá Valcarce, Esther (1996): "¿Se ven afectados los resultados sobre convergencia regional cuando se consideran los efectos espaciales?", Ponencia presentada en la XXII Reunión de Estudios Regionales, Pamplona.

López-Bazo, Enrique, Mora Corral, Antonio J., Surinach Caralt, Jordi y Vayá, Valcarce, Esther (1997): "Convergencia regional en la Unión Europea ante el nuevo entorno económico", *Información Comercial Española*, 762, 25-41.

Odland, John (1988): *Spatial Autocorrelation*, Sage Pub.., Londres.

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