ECONOMETRÍA ESPACIAL

Tradicionalmente en la ciencia regional, los estudios econométricos convencionales han analizado las unidades espaciales (p.e. municipios, estados) como “islas independientes” asumiendo implícitamente que los valores de una unidad y otra son independientes de su posición geográfica. Este bien podría no ser el caso, podríamos pensar en que la relación entre las características de dos unidades que comparten una frontera sea mayor que la de dos unidades que se encuentren distanciadas. El comprobar esta hipótesis es precisamente de lo que se ocupa la econometría espacial. Así, parte de la idea de que las observaciones tienen un cierto ordenamiento espacial y que éste da lugar a la existencia de diversas relaciones entre ellas. Este ordenamiento es complejo, al menos más complejo que por ejemplo el estudio de series de tiempo debido a su naturaleza multidimensional y bidireccional.[1]

Para comenzar es importante tener claro el concepto de Matriz de Pesos Espaciales (MPE). Esta matriz es la que se define para medir la proximidad o relación espacial de las observaciones, pues representa la fuerza de interacción potencial entre las distintas localizaciones de éstas. La MPE es generalmente denotada por W y es de orden (nxn). La manera más general en que puede ser definida es una matriz de contigüidad binaria, generada a partir de información topológica, en donde el valor de w_{ij} será 1 si las unidades i y j comparten alguna frontera, y tomará el valor de 0 en caso contrario. En algunos casos, para mayor facilidad de interpretación, la MPE es estandarizada por

[1] En series de tiempo la relación es unidireccional: y se relaciona con y-1 o y-2 pero no a la inversa; además, la única dimensión de estudio es el tiempo.

renglones, esto es, W es transformada de modo que la suma de los elementos de un renglón es 1.[2]

Debido a que la matriz de simple contigüidad no puede diferenciar la fuerza de las relaciones espaciales entre ubicaciones adyacentes, se han propuesto relaciones más complejas[3]: Cliff y Ord(1981) proponen una matriz asimétrica en donde los valores vienen dados por una combinación de la distancia entre las unidades y un cálculo de extensión relativa de las fronteras que éstas comparten: $w_{ij} = (d_{ij})^{-a}(\beta_{ij})^{b}$ donde d_{ij} es la distancia entre la unidad i y j, β_{ij} es la proporción del límite interior de la unidad i que está en contacto con la unidad j, a y b son parámetros.

Dacey(1968) toma el área relativa de las unidades espaciales: $w_{ij} = d_{ij}\alpha_i\beta_{ij}$ donde d_{ij} es un factor de contigüidad binario, α_i es la porción que representa la unidad i dentro del área total de estudio, β_{ij} es el mismo factor de Cliff y Ord.

Bodson y Peeters(1975) asignan los valores con un peso de accesibilidad general al combinar la influencia de varios canales de comunicación entre las unidades espaciales en una función lógica. $w_{ij} =\Sigma_j k_j\{a/[1+b*\exp(-c_j d_{ij})]\}$ en donde k_j es la importancia relativa de las vías de comunicación, d_{ij} es la distancia entre las unidades i y j; a, b y c son parámetros que deben ser estimados.

Podemos ver así, que uno de los problemas metodológicos más difíciles en el análisis espacial es la especificación correcta de la MPE. Dado que siempre existirá cierto grado de arbitrariedad - la mayoría de las aplicaciones de ciencia regional definen la MPE

---

[2] A parte de facilitar la interpretación de coeficientes, no hay razón matemática o estadística para estandarizar, por esto, la estandarización no debe realizarse de manera automática. Considere los siguientes ejemplos: cuando la matriz estandarizada resulta asimétrica, las estimaciones y procedimientos de los test son más complejos; además, si, como veremos más adelante, la matriz está basada en alguna función inversa de distancia, el estandarizar puede provocar la pérdida de la interpretación económica de esto (Gómez, 1999)

[3] Ejemplos encontrados en Bao, Shumming.

como alguna combinación de la matriz de contigüidad simple y la relación de distancia- los autores recomiendan comparar distintos tipos de matriz a fin de encontrar el más apropiado.[4]

Siguiendo a Anselin (1990) el trabajo de análisis con datos espaciales se divide en 2 enfoques: *Data driven approach* en donde se asume la "aleatoriedad" y tanto el patrón, la estructura y la interacción espaciales son derivados de los datos únicamente, sin ser limitados por alguna noción teórica; y el *Model driven approach* que comienza con una especificación teórica que es confrontada con los datos. La mayoría de los métodos bajo esta categoría se ocupan de la estimación y diagnóstico de las especificaciones de los modelos espaciales.[5]

- DATA DRIVEN APPROACH

Cae en la categoría de *explanatory data analysis* (EDA) popularizado por Tukey en 1977. Todas las técnicas comparten el supuesto de una distribución aleatoria del patrón espacial y que éste, la estructura y forma de dependencia espaciales se derivan únicamente de los datos. Para encontrar esta asociación espacial, definida como autocorrelación espacial, se proponen estadísticos globales y estadísticos locales.

El estadístico global más común es la I de Moran definida por:

$$I(d) = \Sigma\, \Sigma\, w_{ij}(x_i - \mu_x)\,(x_j - \mu_x) \,/\, (S^2 \Sigma\, \Sigma\, w_{ij})$$

donde $S^2 = (1/n)\ \Sigma\ (x_i - \mu_x)^2$, x_i denota el valor observado en la unidad i, μ_x es el promedio de$\{x_i\}$sobre las n ubicaciones, y w_{ij} es la medida de peso espacial definida

---

[4] Cabe mencionar que para incorporar el efecto espacial en la econometría se podrían utilizar también operadores de desfase espacial (rezagos), sin embargo éstos implicarían definir una dirección específica del rezago de las observaciones y de un orden de contgüidad. (Gómez, 1999)
[5] Bao, Shuming

anteriormente. La media teórica de la I de Moran es -1/(n-1). I tendrá signo positivo cuando el valor observado de las ubicaciones dentro de una distancia determinada (d)[6] tiendan a ser similares, es negativo cuando son desiguales y se aproxima a cero cuando hay un ordenamiento aleatorio e independiente en el espacio (Goodchild, 1986). Otro estadístico es la C de Geary que está basada en la suma ponderada del cuadrado de la diferencia entre observaciones:

$$C(d) = (n-1) / (2 \Sigma \Sigma w_{ij}) \{\Sigma \Sigma w_{ij}(x_i - x_j)^2 / (1/n) \Sigma (x_i - \mu_x)^2\}$$

donde x_i y x_j están en valores estandarizados y w_{ij} son los elementos de la matriz de pesos espaciales estandarizada por renglones. A mayor valor de C (>>1) indica que los valores de las unidades dentro de la distancia (d) tienden a ser diferentes, mientras que un valor pequeño (C<<1) indica similitud.

Por último, mencionamos el estadístico definido por Getis en 1992:

$$G(d) = \Sigma \Sigma w_{ij}(d)x_i x_j / \Sigma \Sigma x_i x_j$$

de igual manera que I y C, una G aproximadamente normalizada se puede calcular (Z(G)) de la cual, valores positivos altos indican que el patrón espacial esta dominado por *clusters* de valores altos y viceversa.

Los estadísticos globales prueban la hipótesis nula de no existencia de autocorrelación espacial, y están basados en el supuesto de *estabilidad estructural* sobre el espacio. Este supuesto podría bien ser no realista, especialmente con un número grande de observaciones: la idea de patrones diferentes para regiones distantes entre sí no es muy difícil de concebir. Para capturar patrones locales de espacialidad existen los estadísticos denominados locales, como el *pocket plot* de Cressie (1991), el *Moran scatterplot* de

[6] No sin cierto grado de arbitrariedad.

Anselin(1993) el estadístico *G* de Getis y Ord(1992) y el *Indicador local de asociación espacial (LISA por sus siglas en inglés)* de Anselin(1994).

Además de su utilidad como estimadores de autocorrelación espacial, la I de Moran y la G de Getis son usados para ilustrar la concentración espacial de valores altos o bajos. Es importante recordar que siempre se debe tener cierta precaución en la interpretación de los resultados, pues todos los estadísticos dependen de la definición de la matriz de pesos espacial y el número de unidades espaciales de la muestra.

- MODEL DRIVEN APPROACH

Los dos principales problemas de los procesos espaciales estocásticos son: la *dependencia o autocorrelación espacial* y la *heterogeneidad o estructura espacial.* La causa de la primera es, en general, los errores de medición. Estos errores a su vez son generados por: a) al analizar las unidades de estudio nos podemos dar cuenta que la delimitación es por lo general aleatoria: las divisiones políticas de los municipios por ejemplo; b) con la cada vez más fuerte presencia de tecnología y automatización de la agregación de datos el riesgo de no detectar errores aumenta; y c) la combinación de procesos espaciales como la difusión, el intercambio, la interacción y efectos de *spill over*.[7] La heterogeneidad por su parte es causada por la falta de "estabilidad" espacial del comportamiento o de las relaciones que están siendo estudiadas,[8] debido al tamaño de la unidad de observación, sus dotaciones iniciales de recursos o los procesos de concentración que éstas conllevan generando entonces un mosaico irregular de objetos de análisis. Otro aspecto de la heterogeneidad espacial es la heterocedasticidad que ésta provoca, pero de esto, nos ocuparemos más adelante.

[7] Gómez (1999)

[8] Por ejemplo, las formas funcionales pueden variar para diferentes áreas económicas (Gómez, 1999))

Existen 2 *tipos de modelos espaciales autrorregresivos* que son empleados en la estadística espacial: Conditional Autorregresive (CAR) y Simultaneous Autorregresive (SAR) en donde la diferencia entre ellas es la especificación de la función de covarianza inversa. Para CAR esta especificación es: $(I - \rho C)/\sigma^2$ donde C es una matriz binaria de contigüidad simple; para SAR es $(I - \rho W)^t(I - \rho W)/\sigma^2$ donde W es una matriz de pesos espacial estandarizada por renglones.[9]

La forma general del modelo de proceso espacial es:

$$y = \rho W_1 y + X\beta + \varepsilon$$

$$\varepsilon = \lambda W_2 \varepsilon + \mu$$

donde W_1 y W_2 son matrices de pesos espaciales y $\mu \sim N(0, \Omega)$.

Existen 3 *especificaciones* de los modelos espaciales autorregresivos:

1.- Autorregresión espacial simple: la variable independiente es únicamente un rezago espacial de la variable dependiente.

$$(y - \alpha) = \rho W(y - \alpha) + \varepsilon$$

o reescrito: $$y = \rho Wy + \varepsilon$$

en donde ρ es el coeficiente de autocorrelación espacial (y su signo determina la relación de los efectos entre vecinos) y $\varepsilon \sim N(0,\Omega)$. Como el error está correlacionado con la variable explicativa Wy, los estimadores de MCO serán sesgados e ineficientes.

2.- Modelo espacial autorregresivo con dependencia substancial (de acuerdo a la terminología de Anselin 1993): además del rezago espacial, existen otras variables explicativas exógenas.

$$y = \rho Wy + X\beta + \varepsilon$$

[9] CAR es normalmente usado únicamente para relaciones de primer orden, mientras SAR es usado para dependencia de primer y segundo orden.

donde X es un vector de variables explicativas que se asumen no correlacionadas con el término de error. El mismo problema que en el caso anterior se presenta, y MCO producirá estimadores sesgados e ineficientes..

3.- Modelo espacial autorregresivo con dependencia espacial del error: este modelo tiene sólo variables explicativas exógenas[10] pero ε sigue un proceso autorregresivo espacial.

$$y = X\beta + \varepsilon$$

$$\varepsilon = \lambda W \varepsilon + \mu$$

de aquí, $$(y - X\beta) = \lambda W(y - X\beta) + \varepsilon$$

donde λ es el coeficiente escalar del error espacial (y su significancia representará que un choque aleatorio afectará no sólo a la región en donde se provocó, sino que se transmitirá a todo el sistema) y $\mu \sim N(0, \Omega)$. De aquí, los estimadores MCO serán insesgados, pero ineficientes.

Existe una cuarta especificación propuesta por Rey y Montouri (1999) llamada *Spatial Cross-regressive.* Lo que hace es introducir una variable de rezago espacial exógena definida como (Wy_o) y que contiene los valores iniciales de la variable relevante como pesos en la MPE. Por ejemplo, Lim (2003) utiliza esta especificación para la estimación del grado de convergencia regional a través del crecimiento del ingreso quedando el modelo: $$g = \alpha_i + \beta y_o + \tau W y_o + \varepsilon$$

Este modelo es estimado por MCO. En esta especificación, la variable dependiente está afectada tanto por el valor inicial de la observación como por los valores iniciales de las demás regiones.

[10] Del modelo general , $\rho \equiv 0$ y el error original tiene la matriz de covarianzas no-esférica: $E[\varepsilon \varepsilon'] = (I - \lambda W)^{-1} \sigma^2 I (I - \lambda W)^{-1}$

Comprobar la existencia de dependencia espacial y heterogeneidad en modelos espaciales usualmente depende de cómo la estructura espacial o la MPE sea definida. Dado que no hay una definición única, estas pruebas son más complejas que los modelos econométricos "estándar". Existen 2 tipos de pruebas para la dependencia espacial: probar la dependencia del error y probar para dependencia sustancial. El enfoque común es de aplicar la I de Moran y la prueba del Multiplicador de Lagrange (LM) a los residuos de la regresión de MCO. Esperaríamos que estos residuos tengan autocorrelación en caso de que la relación espacial no haya sido tomada en cuenta específica y correctamente en el modelo (Gómez de Antonio, 1999).

La I de Moran[11] satisface una distribución normal asintótica en muestras grandes, se calcula:

$$I = e'We/e'e \sim N(\mu_I,\sigma^2_I)$$

donde e es el vector de residuos de la regresión original. LM puede calcularse para ambas pruebas, sigue una distribución $\chi^2(1)$ asintótica y la hipótesis nula es que no existe dependencia espacial. De acuerdo a Anselin(1993) el tipo de dependencia puede ser definido comparando ambos LM, el de mayor significancia tiende a ser la alternativa correcta. El cálculo viene dado por:

$$LM(err) = \{e'We/\sigma^2\}^2 / tr[W'W + W^2]$$

tr representa el "matrix trace operator", σ^2 es un estimado de Maximum likelihood para la varianza del error (p.e. σ^2=e'e/n).

[11] Para probar la dependencia del error únicamente.

$$LM(lag) = \{e'Wy/\sigma^2\}^2 / \{(WXb)'MWXb/\sigma^2 + tr[W'W + W^2]\}$$

aquí Wy es el rezago espacial, b es el vector de estimadores MCO para los parámetros β, M es una matriz de proyección, $M=I-X(X'X)^{-1}X'$.

Ante la existencia de dependencia espacial, los estimadores MCO serán sesgados e inconsistentes. Estimar por EGLS tampoco llevará a consistencia por la naturaleza multidimensional de la dependencia espacial. Anselin(1988) y Anselin y Bera(1998) sugieren que la inferencia en espacialidad tendrá características deseables si es hecha de acuerdo al enfoque de Maximum Likelihood.[12] El uso de Variables Instrumentales, Método General de Momentos o Mínimos Cuadrados de 2 Etapas han sido sugeridos por algunos otros autores.

Es importante hacer notar lo siguiente, puede suceder que ante la prueba para dependencia espacial, uno, ambos o ninguno de los estadísticos sean significativos. La estrategia para elegir una correcta especificación del modelo espacial está normalmente basada en la descrita por Anselin y Florax (1995) o por Florax *et al*(2003). De acuerdo con Anselin y Florax para escoger el modelo más apropiado se comparan los estadísticos LM_λ y $LM\rho$: si LM_λ es más significativo que $LM\rho$ y además, LM_λ es robusto mientras el otro no, el modelo debe ser especificado con dependencia en el error; en el caso de que ambas condiciones se cumplan para $LM\rho$ la correcta especificación es de dependencia sustancial[13].

[12] Este último enfoque, es el que se siguió en este proyecto. De acuerdo con Anselin (1988), si las condiciones de regularidad para una función likelihood se cumplen, los estimadores tendrán consistencia, serán asintóticamente eficientes y seguirán una distribución normal. Además casi siempre son insesgados. (En Bao, Shumming)

[13] Lim (2003)

Florax *et al*(2003) define 6 pasos para encontrar la especificación correcta:

1.- Estimar el modelo por MCO

2.- Probar dependencia espacial por LM_λ o LM_ρ

3.- Si no se comprueba la existencia de dependencia espacial, se deben tomar los estimadores de MCO.

4.- Si ambos estadísticos prueban dependencia, la especificación correcta será la del estadístico más significativo.

5.- Si LM_λ es significativo mientras que LM_ρ no lo es, la especificación correcta es de un modelo de dependencia en el error.

6.- Si LM_ρ es significativo mientras que LM_λ no lo es, la especificación correcta es de un modelo de dependencia sustancial.[14]

Una vez comprobada la dependencia espacial, las propiedades distribucionales de los tests paramétricos tradicionales para heterocedasticidad no serán válidas. El último paso entonces para concluir el análisis es el de hacer un test de Chow para comprobar la no existencia de heterogeneidad. Las hipótesis del test son las siguientes:

$$H_o : y = X\beta + \varepsilon$$

$$H_1 : y = \begin{matrix} X_i & 0 \\ 0 & X_j \end{matrix} \quad \begin{matrix} \beta_i \\ \beta_j \end{matrix} + \varepsilon$$

donde $X_i(n\text{x}k)$ y $X_j(n\text{x}k)$ son subconjuntos de observaciones de la variable explicativa, β_i y β_j coeficientes de regresión relativos. Cuando el error ε sigue un proceso espacial autorregresivo, el estadístico de prueba de Chow se define como:

$$C = \{e_R'(I - \lambda W)'(I - \lambda W)e_R - e_U'(I - \lambda W)'(I - \lambda W)e_U\} / \sigma^2 \sim \chi^2(k)$$

[14] De igual manera que el R^2, el *log likelihood* siempre aumenta cuando aumentamos el número de variables en el modelo, Lim(2003) utiliza al AIC (Criterio de información de Akaike) para corregir esto. La regla es que a medida que el AIC sea menor, el modelo ajusta mejor; y es calculado como:
AIC= -2L + 2K (donde L es el *log likelihood* maximizado y K es el numro de variables).

donde λ es el estimado de ML para el parámetro espacial, e_R (e_U) son los residuos para una regresión restringida (no restringida), y σ^2 es el estimado de la varianza del error para ya sea el modelo restringido, el no restringido o ambos. Esta prueba sigue una distribución $\chi^2(k)$ bajo la hipótesis nula de no existencia de heterogeneidad espacial (β_i = β_i).[15]

[15] Para prevenir este problema, también se puede estimar los estadísticos LM robustos y "ahorrarnos" este último paso. (Bao, Shumming)