AUTOCORRELACION ESPACIAL EN MODELOS DE USO DE SUELO

Francisco Martínez, Daniel Schwarz[1]
Universidad de Chile, Casilla 228-3, Santiago, Chile
Fax (562) 6718788, Email: fmartine@cec.uchile.cl

RESUMEN

El presente trabajo investiga el problema de la autocorrelación espacial, presente generalmente en los trabajos de economía urbana, la que en términos sencillos se podría definir como la falta de independencia de las observaciones de una variable localizada en diferentes puntos en el espacio. Esto se produce debido a que el valor de una variable no sólo depende de los atributos propios, sino que además se ve influenciada por el valor de la propia variable en otro punto del espacio. Un ejemplo clásico en la literatura es el precio de una vivienda que depende del valor del suelo en el entorno.

Específicamente se trabaja con el modelo de localización residencial de Santiago, MUSSA, el cual se estudia en busca de identificar posibles mejoras tratando la existencia de autocorrelación espacial. Para investigar esto se desarrolla un test que detecta que efectivamente la estructura de los errores de la función que se calibra para este modelo, llamada disposición a pagar, presenta problemas producto de errores correlacionados. Para solucionarlo se desarrolla un método de corrección que constituye una nueva sobre-estimación del modelo MUSSA que intenta explicar la porción no explicada de éste, vale decir los errores obtenidos en la calibración original.

A través de un método original de calibración para este tipo de modelos de elección discreta, se obtienen los parámetros de la corrección del modelo, los que sumados a los valores del modelo MUSSA arrojan un mejor ajuste que el del modelo original.

Finalmente, para comprobar que se ha eliminado el problema de errores autocorrelacionados, se aplica nuevamente el test al valor estimado de la disposición a pagar por localización residencial, estableciéndose que los errores del modelo corregido no presentan problemas producto de la autocorrelación espacial.

[1] Actualmente en AG Ltda.

Actas VIII Congreso Chileno de Ingeniería de Transporte, 1997

1. INTRODUCCION

Los estudios de economía urbana centran su preocupación en la localización de actividades y su interacción espacial. Cabe destacar que la información básica utilizada en estos estudios tiene una fuerte vinculación con el espacio, lo que conduce a que sea bastante *heterogénea* y esté *fuertemente relacionada con su entorno.*

Estas dos características son llamadas *efectos espaciales* y se conocen como **autocorrelación espacial** (o dependencia espacial) y **heterogeneidad espacial**, e inducen a que cualquier análisis de modelación requiera ser hecho bajo un marco especial (Anselin, 1988). La econometría espacial, que a comienzos de los años setenta se definió como los métodos que tratan satisfactoriamente las características espaciales de la información y de los modelos, es la ciencia que estudia estos fenómenos.

La **autocorrelación espacial** se puede definir, según Anselin, como la falta de independencia que se presenta en un set de observaciones en sección cruzada. Intuitivamente, esta dependencia espacial se puede parecer a la más familiar dependencia temporal encontrada en los textos de econometría estándar, en modelos rezagados y otros análisis de series temporales. Sin embargo el fenómeno espacial es más complejo, debido a la naturaleza multidireccional de la dependencia en el espacio en contrapartida con la clara dependencia unidireccional en series de tiempo. Esta razón impide la aplicación de métodos de estimación de econometría estándar usados para corregir los efectos de la correlación en series de tiempo y motiva el desarrollo de una marco metodológico diferente .

En el campo de la econometría práctica, los efectos espaciales de la información (principalmente autocorrelación espacial) han sido típicamente ignorados a la hora de estimar modelos en el ámbito de la economía urbana.

El objetivo de este trabajo es por una parte entender, identificar, cuantificar y corregir los sesgos producidos por la autocorrelación espacial de la información en el modelo de localización residencial. Se pretende desarrollar técnicas de econometría espacial que permitan identificar y eliminar este sesgo. Finalmente, se aplicarán estas técnicas y se compararán los resultados con los obtenidos anteriormente en el modelo de uso del suelo de Santiago MUSSA (U. de Chile - CIS, 1996).

El desarrollo del trabajo consta de cuatro partes. En la sección siguiente se presenta la discusión metodológica en el ámbito de la econometría espacial que permite por un lado detectar los sesgos producidos por la autocorrelación espacial mediante test de hipótesis y por otro eliminar o reducir este sesgo; también se describe en forma muy general el modelo de localización residencial de la ciudad de Santiago MUSSA. En la sección 3 se describe el método de corrección aplicado al modelo, abordando los tópicos referentes a la especificación funcional del modelo corregido y el método de calibración. En sección 4 se presentan los resultados de la corrección del modelo otorgando mayor énfasis a los avances logrados en cuanto a mejorar la calidad del modelo. Finalmente en la sección 5 se presentan las conclusiones y comentarios del presente trabajo.

2. ANTECEDENTES

2.1. Desarrollo de la econometría espacial

La detección de errores autocorrelacionados en secciones cruzadas, así como las consecuencias de los sesgos producidos por éstos es un tema ausente de los textos clásicos de econometría, probablemente debido a su énfasis predominante en fenómenos dinámicos basados en datos de series de tiempo. Por ejemplo, esta ausencia se observa en Maddala (1977) y Pindyck y Rubinfeld (1981), mencionándose en otros pero sin ahondar en el tema como Kmenta (1971) y Johnston (1984). Sólo desde mediados de los años ochenta se ha notado un desarrollo incipiente pero constante de métodos alternativos que consideren estos efectos y modelos que los incorporen (Anselin 1988, Anselin 1992), aún cuando existen trabajos que todavía los ignoran. Estos métodos, se encuentran principalmente en el marco de análisis de regresión simple en estudios de valor de suelo (Can 1992, Case 1992, Dubin 1992), estando aún ausentes en el contexto de los modelos de localización de actividades y en general de modelos de elección discreta. El grave problema de éstos últimos es que los sesgos producidos por la autocorrelación provocan que se violen los supuestos básicos de independencia, piedra angular de los modelos de elección discreta, ya que los errores no siguen una distribución i.i.d. (independientes e idénticamente distribuidos).

La autocorrelación espacial, definida como la falta de independencia que se presenta en un set de observaciones localizadas en sección cruzada, puede ser causada por una variedad de problemas de medición encontrados en trabajos aplicados. Ejemplos de estos son las delineaciones arbitrarias de unidades espaciales de observación (distritos censales, límites comunales, etc.), problemas de agregación espacial y más importante aún, la presencia de externalidades espaciales que corresponden a efectos georeferenciados, estos últimos por lo general no cuantificables directamente como por ejemplo la presencia de áreas verdes en zonas cercanas o contaminación ya sea acústica o ambiental. Además, e independiente del problema de medición, la organización espacial inherente y la estructura espacial de los fenómenos tienden a generar patrones complejos de interacción y dependencias que son de interés por sí solos. Los modelos resultantes de flujos, patrones, estructura y procesos espaciales, implícita o explícitamente incorporan elementos de dependencia espacial.

Para modelar en forma explícita los efectos de la dependencia espacial Anselin (1988) propone el siguiente modelo general para un set de observaciones en sección cruzada:

$$y = \rho W_1 y + X\beta + \varepsilon \tag{1}$$

$$\varepsilon = \lambda W_2 \varepsilon + \mu \tag{2}$$

$$\text{con} \quad \mu \rightarrow (0, \Omega) \tag{3}$$

donde y es la variable a modelar. El término $X\beta$ representa el efecto directo del vector de atributos locales X sobre y, con X la matriz de variables independientes y β el vector de parámetros asociados. Por otro lado $\rho W_1 y$ (con ρ un parámetro de escala) es el término autorregresivo (efecto espacial) en la variable dependiente y da cuenta del efecto espacial de las demás observaciones sobre ésta. $\lambda W_2 \varepsilon$ (λ es otro parámetro de escala) es el término autorregresivo en el error, con μ el término de error i.i.d. con media nula y matriz de varianzas - covarianzas Ω. W_1 y W_2 representan las matrices de peso espacial del proceso autorregresivo espacial en la variable dependiente (1) y en el término de error (2) respectivamente. Un análisis más profundo de las características de los modelos que incorporen los efectos espaciales se puede encontrar en Anselin (1988).

Las matrices de pesos espaciales representan el grado de influencia que ejerce una observación sobre otra, lo que se traduce en una matriz que presenta una estructura de influencias entre zonas. Por lo general en la mayoría de las aplicaciones en economía urbana las matrices de peso espacial están basadas en alguna combinación de distancia y contigüidad simple, lo que resulta bastante intuitivo, ya que en primer lugar uno tendería a pensar que una comuna vecina ejerce una influencia importante sobre la otra (Santiago - Providencia o Providencia - Las Condes) y asimismo que esta influencia es decreciente con la distancia (Vitacura - Conchalí). Trabajos como el de Anselin (1988) o Stetzer (1982) exploran en extenso la especificación de las matrices espaciales.

2.2. Métodos de detección de autocorrelación espacial

Un desarrollo simultáneo a los estudios empíricos y a los métodos de estimación de modelos han experimentado los métodos de testeo de autocorrelación espacial. En resumen, los desarrollos se centran en torno a cuatro tests, que corresponden al basado en el estadístico Moran I, el Wald (W), el de la razón de la verosimilitud (RV) y el del multiplicador de Lagrange (ML); estos tres últimos aprovechan las propiedades asintóticas de los estimadores máximo verosímiles (Anselin 1988).

La aplicación del test Moran I al modelo MUSSA presenta dos grandes problemas que nos hacen rechazar el uso de este test. Por un lado el test está diseñado para estimaciones de regresiones lineales con mínimos cuadrados ordinarios, lo que se contradice con la especificación no lineal estimada por máxima verosimilitud de MUSSA. Por otro lado se asume a priori una cierta distribución del error, lo que también resulta ser restrictivo.

Los test basados en estimaciones realizadas a través del método de máxima verosimilitud, vale decir el test Wald, el de la razón de la verosimilitud y el del multiplicador de Lagrange se pueden interpretar como diferentes formas de tratar un problema de variables omitidas (Anselin 1988). Estos comparan la distancia, en términos de los valores de la función de verosimilitud, entre un modelo completo y uno restringido, dado por la hipótesis nula de que no hay efectos espaciales ($\rho = 0$ o $\lambda = 0$ en las ecuaciones 1 y 2 respectivamente). La diferencia entre los tres test radica en la forma de evaluar la diferencia entre los modelos completos y restringidos. En general estos test si bien son menos restrictivos ya que son realizados a partir de estimaciones de parámetros con el método de máxima verosimilitud y no son específicos, pues no realizan el test de hipótesis sobre la estructura de los errores, sino sobre la mejora entre un modelo restringido y uno completo.

Un test específico para detectar los sesgos producidos por la autocorrelación espacial y no restrictivo, ni en la forma funcional del modelo, ni en su método de calibración ni en la distribución a priori del error, es el de Kelejian y Robinson (1992). La función de este test es justamente la de aceptar o rechazar la hipótesis nula de ausencia de autocorrelación (errores i.i.d.) a través de la construcción del estadístico S a partir de una regresión auxiliar entre los errores del fenómeno bajo estudio y las variables zonales que lo rigen. El estadístico S se construye a partir de una regresión auxiliar en donde la covarianza estimada de los errores ($\varepsilon_i\varepsilon_j$) del modelo se regresan en función del producto de variables zonales ($Z_{ik}Z_{jk}$), vale decir:

$$\varepsilon_i\varepsilon_j = \sum_k \gamma_k Z_{ik} Z_{jk} \tag{4}$$

donde γ_k representa el parámetro de la variable zonal k - ésima. Una vez obtenido el vector $\hat{\gamma}$ se calcula el estadístico S a partir de:

$$S = \frac{\hat{\gamma}' Z' Z \hat{\gamma}}{\hat{\sigma}_\varepsilon^4} \tag{5}$$

donde $\hat{\sigma}_\varepsilon^4$ es el cuadrado de la varianza de los errores. Este estadístico se distribuye chi - cuadrado con q grados de libertad, donde q es el número de variables de la regresión auxiliar, vale decir que se sugiere rechazar H_o si:

$$S > \chi_q^2(0.95) \tag{6}$$

donde $\chi_q^2(0.95)$ es el valor crítico de la distribución chi - cuadrado con q grados de libertad al 95 % de confianza.

2.3. Modelo de localización residencial MUSSA

En esta sección se presenta brevemente el modelo de localización al cual se intenta aplicar los métodos propuestos. Los fundamentos teóricos del modelo MUSSA se encuentran en el modelo de localización Bid - Choice (Martínez, 1992). La aplicación de este modelo al caso de Santiago, denominado MUSSA, se presenta en Martínez y Donoso (1995 a, b).

El modelo de localización residencial MUSSA se basa en la teoría de la elección discreta, cuyo modelo funcional es análogo a los modelos logit de elección discreta modal. La diferencia entre ambos radica simplemente en la función de utilidad indirecta usada, que en este caso se reemplaza por la función disposición a pagar (DP). La definición de la disposición a pagar, proviene de la teoría de los precios hedónicos desarrollada por Rosen (1974), que asume que la función disponibilidad a pagar por un bien se genera como una combinación de valores implícitos (hedónicos) de los atributos asociados al bien. En general la disposición a pagar del hogar **h** por una combinación vivienda-zona **v** es de la forma:

$$DP_{hv} = y_h + f\left(H_h, V_v, Z_v; \overline{P}; \beta_h; \omega_h^*\right) \tag{7}$$

en que $\mathbf{y_h}$ representa el ingreso del hogar tipo h, $\mathbf{H_h}$ es un vector de características socioeconómicas del hogar. $\mathbf{V_v}$ es un vector que describe la vivienda, $\mathbf{Z_v}$ describe la zona de localización. Por su parte $\overline{P}$ es el vector de precios fijos de otros bienes y servicios, β_h representa el vector de parámetros para un hogar h y ω_h^* un cierto nivel de utilidad fijo del hogar h. Además entre las variables del modelo se debe considerar la accesibilidad que es una variable del tipo **ZH**, ya que depende tanto de la zona como del hogar. A partir de la definición de la disposición a pagar se puede definir la probabilidad de que el hogar **h** se localice en la vivienda **v** como la probabilidad que ese hogar sea el mejor postor por esa vivienda, dado por:

$$P_{hv} = \frac{\exp(\mu DP_{hv})}{\sum_g \exp(\mu DP_{gv})} \tag{8}$$

bajo el supuesto que no existe especulación, es decir los consumidores revelan su verdadero valor por la vivienda; este supuesto no impone restricción en la modelación que se realiza en este estudio pero simplifica notación.

La especificación de la función DP en el estudio MUSSA es la siguiente:

$$\begin{aligned} DP_{hv} \; = \; & b_n \cdot HNPERS_h^{\lambda_0} + d_{nACC} \cdot HNPERS_h^{\lambda_{ACC}} \cdot ACC_{hv}^{\theta_{acc}} + \\ & c_{nHING} \cdot HING_h + \sum_{j=1}^{J} d_{nj} \cdot HNPERS^{\lambda_0} \cdot x_{hvj} \end{aligned} \tag{9}$$

donde $HNPERS_h$ representa el número de personas en el hogar **h**, ACC_{hv} representa la accesibilidad del hogar **h** por la vivienda-localizada **v**. $HING_h$ es el ingreso del hogar **h** y x_{hvj} es un vector de variables que describen el hogar, la vivienda y la zona, donde se incluyen : año de permanencia en el hogar, área de terreno de las casas, área construida de las casas, ingreso promedio de la zona, superficie de la zona destinada a usos industriales y a usos de comercio más servicios, y proporción de la superficie zonal destinada a usos de educación; ver Martínez y Donoso (1995b) o U. de Chile - CIS (1996).

3. METODO DE CORRECCION DEL MODELO MUSSA

3.1. Especificación modelo corregido

La forma de incorporar los efectos espaciales se inspira en la formulación del modelo general de Anselin (1988), ecuaciones (1), (2) y (3), con algunas variaciones que lo transforma en un caso particular de éste. Se requiere especificar una expresión de la función de disposición a pagar que trate la posible existencia de autocorrelación espacial, que llamaremos la función corregida. Se propone la siguiente especificación general:

$$\begin{gathered} DP_{hv}^c = X_{hv}\beta_h + \rho W DP_h^c + \varepsilon_{hv} \\ \text{con } \; \varepsilon_{hv} \rightarrow \text{i.i.d. Gumbel} \end{gathered} \tag{10}$$

en que la notación corresponde a la misma usada anteriormente, siendo ρ un parámetro de escala de la matriz de pesos espaciales W. DP_h^c es el vector de disposiciones a pagar corregida del hogar h por todas las viviendas y presenta un carácter recurrente, es decir DP^c_{hv} depende de las disposiciones a pagar del hogar en todas las opciones de localización espacial y por tipo de vivienda, aunque ellas están ponderadas por la matriz de pesos. La matriz de varianzas - covarianzas de los términos aleatorios ε_{hv}, que se denomina Ω, se supone una matriz diagonal, vale decir que está ausente el problema de autocorrelación espacial.

Debido a la complejidad de estimar una función del tipo (10), causada por la recurrencia que presenta la disposición a pagar, es que se usa una vector de variables proxy (Z) que estén fuertemente correlacionadas con la disposición a pagar DP_h^c. Esto conduce a una especificación del siguiente tipo:

$$DP_{hv}^c = X_{hv}\beta_h + \sum_{k=1}^{K}\sum_{j=1}^{N} Z_{jk} w_{vj}\delta_{hk} + \varepsilon_{hv}^{'} \tag{11}$$

en que Z_{jk} representa la variable proxy zonal **k** - ésima para la zona **j**. Un ejemplo de estas variables son : ingreso promedio zonal, superficie zonal por tipo de actividad (educación, industria, etc.), accesibilidad u otras si existiesen; vale decir son las variables usadas en el modelo MUSSA original. El término w_{ij}, que es un elemento de la matriz de pesos W, da cuenta del grado de correlación entre las zonas en que se encuentra la vivienda (**v**) y otra zona (**j**) ; es decir, pondera el grado de influencia que ejerce una zona sobre otra. δ_{hk} representa el parámetros del término espacial de la variable k - ésima ; N es el número de zonas y K el total de variables proxy.

De esta forma el modelo a calibrar propuesto, que corrige el sesgo producto de la autocorrelación espacial, tendría la siguiente forma matricial para un cierto hogar h:

$$\begin{bmatrix} DP_1 \\ \\ DP_v \end{bmatrix} = \underbrace{\begin{bmatrix} x_{11} & & x_{1k} \\ & & \\ x_{v1} & & x_{vk} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \\ \beta_k \end{bmatrix}}_{Efecto\ directo} + \underbrace{\begin{bmatrix} w_{11} & & w_{1N} \\ & \cdots & \\ w_{N1} & & w_{NN} \end{bmatrix} \begin{bmatrix} Z_{11} & & Z_{1m} \\ & \cdots & \\ Z_{N1} & & Z_{Nm} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \\ \delta_m \end{bmatrix}}_{Efecto\ espacial} \tag{12}$$

Cabe señalar que tanto las variables X como los términos de la matriz de pesos espaciales W pueden además contener parámetros exponenciales, no presentados en este caso.

3.2. Método de estimación y calibración del modelo corregido

La forma directa y tradicional de estimar las disposiciones a pagar corregidas sería establecer una nueva base de calibración en la que para cada vivienda se definiera un set de postores con sus respectivas variables, tal como se realizó en el trabajo de calibración de MUSSA (U. de Chile -

CIS, 1996), pero en este caso incluyendo los términos que permiten tratar el tema del efecto espacial.

Sin embargo, la notación matricial recién planteada permite apreciar que el número operaciones de cálculos requeridos en la calibración crecería enormemente. Basta ver a través de la ecuación (12), que de un orden k en el número de operaciones por cada postor se pasa a un orden k + m + N; si se considera k = 10 (número de variables en el modelo MUSSA), m = 5 (número de variables zonales más accesibilidad) y N=264, se pasa de 10 operaciones por postor a 279 operaciones, vale decir se realizan 28 veces más operaciones. Si pensamos que para el proceso de calibración de MUSSA se utiliza una muestra de 40 postores por vivienda, entonces el número de operaciones por vivienda sería 1120 veces mayor. Por lo tanto, se puede apreciar que el tiempo que tomaría realizar la calibración de un modelo de este tipo cae fuera de las posibilidades reales. Es por esto que se buscará una nueva forma de calibración, desarrollando una metodología para ello.

El modelo teórico que se estimará, expuesto en la ecuación (11), se puede expresar como:

$$DP_{hv}^{c} = DP_{hv}^{M} + DP_{hv}^{EE} + \varepsilon_{hv} \tag{13}$$

en que la disposición a pagar se descompone en dos términos. El primer término corresponde al efecto directo, o de la propia zona denotada por *v*, que se asume bien descrito por la componente determinística de la disposición a pagar calibrada en MUSSA ($DP^{M}{}_{hv}$). El segundo término DP_{hv}^{EE} corresponde a la matriz de términos de la disposición a pagar que da cuenta del efecto espacial. Así, la forma de estimar el modelo corregido es calibrando solamente el efecto espacial, manteniendo los parámetros del efecto directo calibrado en MUSSA.

Este método presenta la desventaja de una estimación secuencial, en que los efectos directos e indirectos se estiman en forma separada, lo que representa un método de sobreestimación. Sin embargo otorga el beneficio de poder realizar una calibración que es muy compleja. Para entender esta complejidad se debe recordar que la función a calibrar representa disposición a pagar por localización residencial, la que no se observa, ya que los datos observados son de la distribución de localización de la población.

Para abordar esta complejidad se aplica un método en dos etapas. En la primera se encuentra un término de corrección de la disposición a pagar calibrada en MUSSA, tal que sea capaz de reproducir la localización observada con mayor precisión utilizando la ecuación de localización (8). Como segunda etapa, se asume la hipótesis que este término se puede explicar por efectos espaciales. Esto permitirá además realizar el test de autocorrelación espacial sobre la expresiones finales DP^{c}.

Para abordar la primera etapa es necesario encontrar valores de la disposición a pagar que reproduzcan las observaciones de localización a un cierto nivel de agregación de la localización; esto es que sean capaces de reproducir los totales de hogares localizados, por vivienda y zona (*v*) y cluster de hogar (*h*). Estos valores, que representan el valor corregido de DP (DP^{c}), se pueden encontrar resolviendo un problema de punto fijo propio de la forma MNL Logit de la ecuación (8):

$$\overline{N}_{kv} = \frac{N_k \exp\left(\mu DP_{kv}^{c}\right)}{\sum_{k'} N_{k'} \exp\left(\mu DP_{k'v}^{c}\right)} \tag{14}$$

donde $\overline{N}_{kv}$ es el número de hogares del cluster k localizados en la vivienda-zona v y N_k es el número de hogares que pertenecen al cluster; N_{kv} y N_k son valores observables en un corte temporal. Luego se puede despejar DP^{c}_{hk} en función del vector completo DP^{c} obteniendo:

$$DP_{kv}^{c} = \frac{1}{\mu} \ln\left[\sum_{k'} N_{k'} \exp(\mu DP_{k'v}^{c})\right] \tag{15}$$

que tiene una forma de punto fijo con solución única y de convergencia rápida que se resuelve iterativamente. Cabe mencionar aquí que el criterio de corrección de la DP queda definido por la solución del punto fijo; además, el tratamiento agregado a nivel de cluster es consistente con el nivel de agregación que presenta el modelo MUSSA.

Teniendo estos valores se procede a la segunda etapa. Restando a los valores obtenidos el valor correspondiente de la disposición a pagar obtenida en MUSSA, se obtiene:

$$DP_{kv}^{EE} = DP_{kv}^{c} - DP_{kv}^{M} \tag{16}$$

que representa la corrección que se desea modelar.

$$\begin{aligned} DP_{hv}^{EE} = \overline{DP}_{kv}^{EE} + \varepsilon_{kv} = \sum_{tv=1}^{3} c_v + \sum_{\substack{j=1 \\ j\neq i}}^{264} \left(\delta_{ij} d_{ij}^{\alpha_h} + \beta_h p_{ij}^{\gamma_h}\right) (\lambda_{ZI_h} ZI_j^{\phi_{ZI_h}} + \lambda_{IND_h} IND_j^{\phi_{IND_h}} + \lambda_{EDU_h} EDU_j^{\phi_{EDU_h}} \\ + \sum_{TM=1}^{3} \lambda_{ACC_{hTM}} ACC_{hj}^{\phi_{ACC_{hTM}}} + \lambda_{CS_h} CS_j^{\phi_{CS_h}}) + \varepsilon_{kv} \end{aligned} \tag{17}$$

La variable accesibilidad del hogar (ACC) se desagregó por tasa de motorización para permitirle al modelo una mayor flexibilidad. Con igual propósito se incorporaron tres constantes que dicen relación con los seis tipos de viviendas del modelo MUSSA.

Para calibrar el modelo de corrección, ecuación (17), se usó el método de mínimos cuadrados; en ambos caso se pretende obtener:

$$\underset{k}{MIN}\ (DP_{kv}^{EE} - \overline{DP}_{kv}^{EE})^2 \tag{18}$$

Este problema se resuelve en forma independiente para cada una de los clusters, definidos por cinco categorías de ingreso de MUSSA; sin embargo, debido al bajo número de observaciones en las categorías cuatro y cinco se estimo un modelo conjunto para ambas categorías. El resultado del proceso de calibración permitirá encontrar los parámetros de la matriz de disposiciones a pagar que permite mejorar la estructura de probabilidades observadas y que además corrige el sesgo producto de la autocorrelación espacial, lo que será testeado una vez realizada la calibración.

Las forma funcional utilizada para modelar el efecto espacial, así como las variables usadas en este término y su descripción, se reportan en la siguiente ecuación y en el Cuadro 1:

Cuadro 1
Descripción de las variables zonales de la base de calibración

Variable	Parámetros asociados		DESCRIPCION
	Lineal	Exponencial	
	c_v		Constante por tipo de vivienda v = 1,2,3
d_{ij}		α_h	Distancia entre centroides de zonas i y j
δ_{ij}			Dummy de zonas cercanas : 1 si $d_{ij} < 15$; 0 si no
p_{ij}	β_h	γ_h	Proporción del borde de la zona i que es frontera con la zona j
ZI_j	λ_{ZI_h}	ϕ_{ZI_h}	Ingreso promedio zonal
IND_j	λ_{IND_h}	ϕ_{IND_h}	Superficie construida para uso "Industria" en la zona
CS_j	λ_{CS_h}	ϕ_{CS_h}	Superficie construida para uso "Servicios" y "Comercio" en la zona
EDU_j	λ_{EDU_h}	ϕ_{EDU_h}	Proporción de superficie construida para uso "Educación" en la zona
ACC_{hj}	$\lambda_{ACC_{hTM}}$	$\phi_{ACC_{hTM}}$	Accesibilidad del hogar **h** con tasa de motorización **TM** por la zona **j**

4. APLICACION AL MODELO MUSSA

En primer término se procedió a aplicar el test de autocorrelación espacial de Kelejian y Robinson (1992) a la disposición a pagar calibrada para el modelo MUSSA. El error de esta función corresponde a la corrección obtenida DP^{EE} a partir de la solución de punto fijo. El resultado de la aplicación de éste arrojó como valor del estadístico S = 5656 cuyo valor es mayor que 12,6 que corresponde al crítico de la distribución chi - cuadrado con 6 grados de libertad al 95 % de confianza. Como conclusión se puede señalar que el test indica que se debe rechazar la hipótesis nula de estar en presencia de errores independientes, o en otras palabras el test indica la presencia de errores correlacionados. Este resultado indica que al no existir un tratamiento específico de los efectos del espacio, el modelo de localización podría estar sesgado, lo que se ha demostrado a través del test. Corresponde por lo tanto desarrollar la corrección a éste para luego testear la especificación final, la cuál no debiera presentar problemas de autocorrelación espacial.

Se calibró el efecto espacial, especificado por la ecuación (17) obteniéndose los siguientes resultados:

Cuadro 2

Resultados del proceso de calibración : valor de los parámetros modelación efecto espacial

VARIABLES (test t)	PARAM	Categoría 1	Categoría 2	Categoría 3	Categoría 4 y 5
TIPO VIVIENDA 1 y 2	c_1	-0,244 (-18,8)	0,253 (26,8)	1,300 (131,3)	0,687 (19,6)
TIPO VIVIENDA 3 y 4	c_2	0,501 (30,9)	0,820 (79,7)	0,549 (58,7)	-0,051 (-1,7)
TIPO VIVIENDA 5 y 6	c_3	0,308 (16,8)	1,169 (95,2)	1,089 (102,2)	-0,146 (-3,9)
DISTANCIA ENTRE ZONAS	α_h	-0,717 (-63,9)	-1,560 (-150,3)	-0,857 (-49,7)	-3,902 (-104,1)
PERIMETRO COMUN	β_h	3,075 (156,6)	0,822 (77,0)	0,688 (116,0)	7,705 (28,1)
PERIMETRO COMUN	γ_h	1,200 (197,2)	2,433 (195,4)	7,404 (1330,7)	0,827 (34,0)
ZSUPCIND	ϕ_{IND}	0,043 (26,7)	0,061 (3,2)	0,627 (9,3)	1,646 (47,6)
ZSUPCIND	λ_{IND}	1,876 (262,4)	0,310 (19,1)	0,078 (3,7)	0,002 (7,6)
ZINGPROM	ϕ_{ZI}	0,245 (44,9)	-0,236 (-43,9)	-0,408 (-72,1)	-3,533 (-28,4)
ZINGPROM	λ_{ZI}	-0,553 (-76,2)	-0,099 (-10,6)	3,635 (504,9)	2,279 (25,5)
ZSUPCEDU / ZSUPTERR	ϕ_{EDU}	0,968 (122,0)	-0,369 (-62,8)	-0,090 (-18,1)	1,092 (22,0)

Continuación Cuadro 2

Resultados del proceso de calibración : valor de los parámetros modelación efecto espacial

ZSUPCEDU / ZSUPTERR	λ_{EDU}	0.031 (15.6)	0.397 (30.1)	-2.490 (-291.4)	0.042 (9.2)
ZSUPCCOM + ZSUPCSER	ϕ_{CS}	0.114 (20.4)	-0.659 (-50.7)	-3.251 (-566.6)	2.666 (46.9)
ZSUPCCOM + ZSUPCSER	λ_{CS}	-0.840 (-172.8)	1.457 (185.2)	0.453 (40.8)	0.000 (2.1)
ACC TM0	ϕ_{ACCTM0}	1.311 (231.5)	1.214 (52.7)	0.199 (22.7)	0.186 (5.3)
ACC TM0	λ_{ACCTM0}	-0.902 (-76.5)	-1.246 (-66.7)	0.186 (23.3)	0.368 (12.1)
ACC TM1	ϕ_{ACCTM1}	-0.654 (-58.0)	-0.535 (-51.7)	-2.333 (-47.7)	0.350 (8.7)
ACC TM1	λ_{ACCTM1}	0.142 (28.1)	1.002 (75.8)	0.438 (8.9)	0.069 (7.8)
ACC TM2	ϕ_{ACCTM2}			0.019 (1.0)	-2.122 (-42.2)
ACC TM2	λ_{ACCTM2}			0.305 (38.6)	9.720 (33.4)
LL		-36.889	-35.962	-34.028	-17.148
Observaciones		1062	1005	1007	438
R^2 ajustado (%)		39.3	21.6	26.0	61.9

Como conclusión se puede apreciar que los parámetros resultan ser estadísticamente significativos salvo uno de la accesibilidad (con tasa de motorización mayor que uno para la categoría 3). En cuanto a los signos de los efectos de la matriz de pesos espaciales, distancia y perímetro, estos resultan ser correctos ya que el exponente de la distancia es negativo y tanto el parámetro lineal como exponencial de la proporción del perímetro son positivos.

Por otro lado, en cuanto a la proporción de la varianza explicada por el modelo se puede apreciar que en general no es muy alto, observándose los mejores grados de ajuste en la categoría de ingreso 4 y 5. Sin embrago más importante que analizar el grado de ajuste del efecto espacial de la disposición a pagar, es el tema de testear la capacidad para reproducir la estructura de localización observada, vale decir efecto directo más espacial. Para esto se calculó la estructura de disposiciones a pagar corregidas dada por la ecuación (13) y a partir de ésta se calculó la estructura de localización. Posteriormente se calculó el número predicho $\hat{N}H_{nz}$ y observado NH_{nz} de hogares de cada categoría de ingreso *n* localizada en cada una de las 264 zonas a través de la siguiente formulación:

$$\hat{N}H_{nz} = \sum_{v} f_v \cdot \sum_{h \in n} P_{h/v} \tag{19}$$

$$NH_{nz} = \sum_{v} f_v \cdot \sum_{h \in n} 1_{h/v} \tag{20}$$

donde $1_{h/v}$ = 1 si el hogar **h** se localizó en la vivienda - zona **v** en 1991; 0 si no. f_v corresponde al factor de expansión de la vivienda tipo **v**, o sea el número de viviendas que representa ese tipo en la realidad.. A continuación se estimó por mínimos cuadrados el siguiente modelo:

$$NH_{nz} = \alpha_0 + \alpha_1 \cdot \hat{N}H_{nz} \tag{21}$$

tanto a nivel global, como por categoría de ingreso, lo que permitirá verificar si la corrección del efecto espacial mejora el modelo de localización. Los resultados obtenidos fueron los siguientes:

Cuadro 3
Resumen bondad de ajuste (R^2) de los modelos, c/r a la localización observada [%]

MODELO	GLOBAL	Categoría 1	Categoría 2	Categoría 3	Categoría 4	Categoría 5
MUSSA	83.9	89.6	89.6	81.4	45.6	62.0
Corrección Efecto Espacial	85.8	92.1	88.6	86.1	66.5	61.1

De estos resultados se puede concluir que la corrección realizada es positiva, ya que el ajuste global muestra que éste modelo explica, en total, 1.9% más de la varianza en la localización residencial que el modelo MUSSA. Las diferencias en los ajustes entre categorías muestra que los efectos espaciales no contribuyen homogéneamente a explicar los errores de MUSSA, lo que puede mejorarse mediante formas funcionales específicas por categoría.

Cabe ahora preguntarse si el sesgo producto de la dependencia espacial ha desaparecido a raíz de este nuevo modelo. Para responder esto se aplica nuevamente el test de autocorrelación de Kelejian y Robinson al error producto de la diferencia de la nueva especificación corregida para la disposición a pagar (considerando el efecto espacial) y la disposición a pagar solución del problema matemático de punto fijo. Para esto se calculó el estadístico S obteniendo un valor S = 2.6, que es menor que el valor crítico de la distribución chi - cuadrado con 6 grados de libertad (12.6), por lo que se puede concluir que se ha superado el problema producido por la autocorrelación en el sesgo sobre los parámetros del modelo MUSSA.

5. CONCLUSIONES

En el presente trabajo se ha desarrollado una metodología mejorar la capacidad del modelo de localización residencial desarrollado para la ciudad de Santiago (MUSSA) mediante el tratamiento de los errores analizando autocorrelación espacial. Esta mejora se logra modelando los errores de la función de disposición a pagar a través de un término explícito que de cuenta del

efecto del espacio. Como los errores no son observables (sólo lo es la estructura de localización), estos se estimaron a partir de la solución de un problema matemático de punto fijo, lo que no sólo permitió realizar la estimación de la corrección del modelo, sino que además permitió realizar un test de hipótesis para verificar la ausencia del problema generado por los errores autocorrelacionados, originalmente presente en el modelo MUSSA.
Una vez realizada la modelación del efecto espacial se agregó este término a la disposición a pagar del modelo MUSSA, realizándose posteriormente el cálculo de las probabilidades de localización y estimándose el número de hogares localizados por zona y desagregado por categoría de ingreso del hogar. Estos datos se compararon con los hogares observados determinándose de esta forma el ajuste del modelo y comparándose éste con el obtenido anteriormente por MUSSA. Como conclusión se puede señalar que el ajuste del nuevo modelo resulta ser mejor que el original.

Si bien en el modelo finalmente propuesto no existen problemas de autocorrelación espacial, se debe considerar que los parámetros que se mantuvieron de la estimación de MUSSA resultan ser sesgados por lo que el modelo final propuesto mantiene este problema. Para superar el sesgo en los parámetros originales de MUSSA se requiere estimar el modelo en forma íntegra nuevamente, vale decir tanto los parámetros del efecto directo (calibrado en MUSSA) como el efecto espacial. Esto se intentó, sin embargo el ajuste de la localización arrojó peores resultados, lo que encuentra su explicación en el alto número de parámetros del modelo conjunto y la no linealidad del modelo, lo que condujo a pobres convergencias. Esta extensión es materia de próxima investigación.

El tratamiento de los errores por autocorrelación espacial tiene un impacto en la modelación que trasciende la calibración misma. En efecto, el modelo finalmente obtenido es de mucho mayor complejidad que el original, lo que cobra relevancia en su aplicación para predicción. Su mayor complejidad radica en la mutua dependencia de las disposiciones apagar, ya no solamente entre consumidores, sino que ahora también espacialmente.

AGRADECIMIENTOS

Los autores agradecen el financiamiento parcial del estudio a través del proyecto FONDECYT N° 1950280.

REFERENCIAS

Anselin, L.(1988) **Spatial Econometrics: Methods and Models**, Kluwer Academic, Dordrecht.

Anselin, L.(1992) Space and applied econometrics: Introduction, **Regional Science and Urban Economics 22**, 307-316.

Can, A. (1992) Specification and estimation of hedonic housing price models, **Regional Science and Urban Economics 22**, 453-474.

Case, A. (1992) Neighborhood influence and technological change, **Regional Science and Urban Economics 22**, 491-508.

Dubin, R.A. (1992) Spatial autocorrelation and neighborhood quality, **Regional Science and Urban Economics 22**, 433-452.

Johnston, J. (1984) **Econometric Methods**. McGraw-Hill, New York.

Kelejian, H. y D. Robinson (1992) Spatial Autocorrelation: A new Computationally Simple Test with an Application to per capita County Police Expenditures, **Regional Science and Urban Economics 22**, 317-331.

Kmenta, J.(1971) Elements of Econometrics, en Anselin, L. (1988) **Spatial Econometrics: Methods and Models**, Kluwer Academic, Dordrecht.

Martínez, F. (1992) The Bid - Choice Land Use Model: an integrated economic framework, **Environment & Planning 24A,** 871-885.

Martínez, F. y P. Donoso (1995a) El Modelo de Uso de Suelo de Santiago I: La teoría , **Actas del VII Congreso Chileno de Ingeniería de Transporte,** 202-216.

Martínez, F. y P. Donoso (1995b) El Modelo de Uso de Suelo de Santiago II: El modelo empírico, **Actas del VII Congreso Chileno de Ingeniería de Transporte,** 217-235.

Maddala, G. (1977) **Econometrics**. McGraw-Hill, New York.

Pindyck, R y D. Rubinfeld (1981) **Econometric Models and Economic Forecast.** McGraw-Hill, New York.

Rosen, S. (1974) Hedonic prices and implicit markets: product differentiation in pure competition, **Journal of Political Economy 82** (1), 34-55.

Stetzer, F. (1982) Specifying weights in spatial forecasting models: the results of some experiments. **Environment and Planning 14A**, 571 - 584.

U. de Chile - CIS (1996) Informe Final Estudio Análisis de Usos de Suelo: Modelo MUSSA. Ministerio de Planificación Nacional (MIDEPLAN), Chile.