UNIVERSIDAD DE LA REPÚBLICA
Facultad de Ciencias Económicas y de Administración
Instituto de Estadística

Modelos de Origen Destino: una primera aproximación aplicada a la Encuesta de Movilidad de Montevideo

María Eugenia Riaño
Gerardo Martínez
Guillermo Zoppolo

Diciembre, 2017

Serie Documentos de Trabajo

DT (17 / 04) - ISSN : 1688-6453

Forma de citación sugerida para este documento:

Riaño, María Eugenia; Martínez, Gerardo; Zoppolo, Guillermo (2017).
Modelos de Origen- Destino: una primera aproximación aplicada a la Encuesta de Movilidad de Montevideo. [en línea]
Serie Documentos de Trabajo, DT (17/04). Instituto de Estadística, Facultad de Ciencias Económicas y de Administración, Universidad de la República, Uruguay.

Modelos de Origen - Destino: una primera aproximación aplicada a la Encuesta de Movilidad de Montevideo

María Eugenia Riaño [1]; Gerardo Martínez[2]; Guillermo Zoppolo[3]
Departamento de Métodos Cuantitativos, Instituto de Estadística, Facultad de Ciencias Económicas y de Administración, Universidad de la República

RESUMEN

Las encuestas de movilidad proveen información acerca del movimiento de personas en un área geográfica. Son un insumo relevante para las entidades encargadas de la planificación, principalmente de la infraestructura y servicios de transporte. El objetivo de los modelos de Origen Destino es entender y explicar los flujos de entidades tangibles o intangibles en un conjunto de regiones delimitadas. Los modelos intentarán explicar los flujos utilizando características de la región de origen, de la región de destino y variables asociadas a la distancia entre regiones. Históricamente, el análisis de las matrices de Origen - Destino se realizó con modelos gravitacionales que no consideran autocorrelación espacial. En los últimos años se ha incorporado este componente al modelo gravitacional, mostrando en algunos casos mejoras en el ajuste del modelo. Existen diversos enfoques para incorporar la autocorrelación espacial a los modelos de Origen Destino. Uno de estos enfoques le da una especificación endógena al modelo. Se intenta explicar cómo los cambios en las características de las regiones vecinas al origen y vecinas al destino impactan los flujos entre regiones mediante un efecto derrame. En el presente trabajo se estima el modelo gravitacional clásico y el modelo endógeno para los casos de Montevideo y Canelones de la Encuesta de Movilidad del Área Metropolitana, específicamente para los viajes con motivo laboral.

Palabras clave: Econometría Espacial, Modelos gravitacionales, Modelos de regresión.
CÓDIGOS JEL: C31, C83

Clasificación MSC2010: 62P25, 62H11,62D05

[1] *email:* eugenia@iesta.edu.uy, ORCID:0000-0003-3451-8249
[2] *email:* gmartinez@iesta.edu.uy
[3] *email:* gzoppolo@iesta.edu.uy

Origin - Destination models: a first approach applied to the Montevideo's Household Travel Survey

María Eugenia Riaño [1]; Gerardo Martínez[2]; Guillermo Zoppolo[3]
Departamento de Métodos Cuantitativos, Instituto de Estadística, Facultad de Ciencias Económicas y de Administración, Universidad de la República

ABSTRACT

Household travel surveys provide information about people's daily mobility in a Geographical area. The information gathered is an important input for transportation planners and policy makers who need comprehensive data on travel and transportation patterns in the city. The goal of the origin - destination models is to understand and explain the flows of tangible or intangible entities in a set of delimited regions. These models will explain the flows between regions using characteristics of the origin, of destination and variables related with the distance between regions. Historically, the analysis of an origin - destination matrix was done with gravitational models that do not consider correlation in space. Recently, correlation has been incorporated to the gravitational model, showing in some cases improvements in the adjustment of the model. There are several approaches to incorporate spatial autocorrelation to the origin - destination models. One of these approaches gives an endogenous specification to the model. This approach try to explain how changes in characteristics of the regions neighboring the origin and regions neighboring to the destination impact the flows between regions through a spillover effect. In the present work the gravitational and endogenous models are estimated for the household travel survey of Montevideo and Metropolitan Area, specifically for work related travels.

Key words: Gravitational models, Regression, Spatial Econometrics.

JEL CODES: C31, C83
Mathematics Subject Classification MSC2010: 62P25, 62H11,62D05

[1] *email:* `eugenia@iesta.edu.uy`, ORCID:0000-0003-3451-8249
[2] *email:* `gmartinez@iesta.edu.uy`
[3] *email:* `gzoppolo@iesta.edu.uy`

1. Introducción

El presente trabajo es una primera aproximación de los modelos de interacción espacial (o de origen - destino) usando los datos de la Encuesta de Movilidad del Área Metropolitana de Montevideo, realizada en el año 2016. Existen diversos enfoques para el ajuste de modelos de interacción espacial. En el marco de los modelos de regresión con perturbaciones normales, el primero en aparecer es el denominado modelo *gravitacional*, que asume independencia entre los flujos de origen - destino. Modelos más avanzados incorporan un término espacial autorregresivo, que modeliza la dependencia de los flujos entre regiones vecinas. Dentro de estos modelos se distinguen los denominados *endógenos* y *exógenos* (Patuelli, 2016), dependiendo de si el rezago espacial se aplica a las variables endógenas o exógenas. Los modelos de origen - destino también pueden ser tratados como modelos de conteo y ser estimados con un modelo Poisson. En este último caso la incorporación de la autocorrelación espacial no se realiza con rezagos autorregresivos, si no que se incorpora con técnicas asociadas al filtrado espacial (Fischer y Griffith, 2008).

Los resultados que se presentan son solamente para los modelos gravitacional y endógeno, quedando pendiente como aplicación futura los modelos de conteo. Se consideran sólo los viajes por motivo laboral. De las localidades de la zona metropolitana no se incluyen a las del departamento de San José: la cantidad de viajes hacia el departamento de San José desde Montevideo y Canelones por motivo laboral es muy baja y se generan celdas vacías en la matriz de flujos. Los modelos que se estiman no pretenden ser definitivos, si no que tienen como objetivo ser un ejercicio para una mejor comprensión de los modelos de origen - destino, la interpretación de sus parámetros y ser la base para futuros refinamientos.

2. Modelos de interacción espacial

El objetivo de estos modelos es explicar los flujos de entidades tangibles (migración de poblaciones, tráfico diario de personas en una ciudad, exportaciones) o intangibles (movimientos en la bolsa de valores, tráfico de información) en un conjunto de regiones delimitadas.

Los modelos buscan explicar estos flujos utilizando:

1. Características de la región de **origen** del flujo que permitan explicar por qué los flujos comienzan allí. Por ejemplo: la cantidad de personas jubiladas y la tasa de empleo de la región.

2. Características de la región de **destino** del flujo que permitan explicar cuán atractiva es esa región. Por ejemplo: para explicar los viajes por esparcimiento en una ciudad,

podemos considerar el número de parques y zonas verdes.

3. Variables que caractericen la **distancia** entre el origen y el destino.

2.1. Modelo Gravitacional sin interacciones

Dadas n regiones, tendremos n^2 combinaciones de posibles flujos entre estas regiones. Llamaremos $Y(i,j)$ al flujo entre la región de origen i y la de destino j, con $i,j = 1,...,n$.

Se asume que el flujo en el par (i,j) es **independiente** del flujo en el par (k,l). El modelo planteado, de forma general, es:

$$Y(i,j) = c\, f(X(i))g(X(j))h(G(i,j)) \tag{1}$$

donde c es una constante, $X(i)$ y $X(j)$ representan características del origen y del destino respectivamente, y $h(G(i,j))$ es una medida de resistencia o *disuasión* que refleja la separación del origen y del destino.

Las funciones f y g se definen de la siguiente manera:

- $f(X(i)) = X(i)^{\beta_0}$ y $g(X(j)) = X(j)^{\beta_d}$, donde $X(i)$ es una variable que representa cuán *propulsor* es el origen y $X(j)$ representa cuán *atractivo* es el destino. Los coeficientes β_0 y β_d deberán ser estimados.
- La especificación de $h(G(i,j))$ será

  $$h(G(i,j)) = [G(i,j)]^{\gamma},$$

  donde $G(i,j)$ es una medida de distancia entre el origen i y el destino j y γ es un parámetro negativo a estimar.

El modelo entonces puede ser escrito como

$$Y(i,j) = c\, X(i)^{\beta_o} X(j)^{\beta_d}[G(i,j)]^{\gamma} \tag{2}$$

Aplicando logaritmos

$$\log(Y(i,j)) = \log(c) + \beta_o \log(X(i)) + \beta_d \log(X(j)) + \gamma \log(G(i,j)) \tag{3}$$

LeSage y Pace (LeSage y Pace 2008) plantean el modelo de forma matricial, similar a la forma de un modelo de regresión clásico.

Sea $\mathbf{Y} = ((Y(i,j))_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n}$ la matriz que en la entrada (i,j) tiene el flujo entre el origen i y el destino j. Podemos escribir a esta matriz como un vector de tamaño $N \times 1$ (con $N = n \times n$) de dos formas, llamadas *origin-centric* y *destination-centric*.

2.1.1. Vector de flujos

El vector *origin-centric* tendrá en los n primeros elementos los flujos entre el origen 1 y los n destinos, en los n siguientes elementos tendrá los flujos entre el origen 2 y los n destinos, y así sucesivamente.

$$\begin{array}{ccc} l^{(o)} & o^{(o)} & d^{(o)} \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ n & 1 & n \\ \vdots & \vdots & \vdots \\ N-n+1 & n & 1 \\ \vdots & \vdots & \vdots \\ N & n & n, \end{array}$$

Denominamos $\mathbf{y} \in \mathbb{R}^N$ al vector *origin-centric*.

2.1.2. Producto de Kronecker

Dadas dos matrices $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{p\times q}$ se define el *producto de Kronecker* $A \otimes B$ a la matriz bloque de tamaño $mp \times nq$ tal que:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}$$

2.1.3. Variables regresoras

Sean $X_1, X_2, ..., X_k$ variables sobre n regiones que pretenden explicar los flujos entre las regiones. Sea $\mathbf{X} = (X_1, X_2, ..., X_k) \in \mathbb{R}^{n\times k}$ la matriz que recoge estas características.

Dada la forma del vector y, donde las posiciones 1 a n reflejan los flujos del origen 1 a los n destinos, repetiremos la matriz $\mathbf{X}$ n veces para producir una matriz de tamaño $N \times k$, que llamaremos $\mathbf{X}_d \in \mathbb{R}^{N\times k}$ y refleja las características de los *destinos*:

$$\mathbf{X}_d = \mathbf{I}_n \otimes \mathbf{X} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \otimes \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nk} \end{pmatrix}$$

De la misma forma, construiremos la matriz $\mathbf{X}_o \in \mathbb{R}^{N\times k}$ que representa las características de los *orígenes*:

$$\mathbf{X}_o = \mathbf{X} \otimes \mathbf{I}_n = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nk} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

2.1.4. Función de distancia

El modelo gravitacional añade la distancia entre las regiones como variable explicativa de los flujos y_{ij}. Sea $\mathbf{G} \in \mathbb{R}^{n\times n}$ tal que

$$G_{ij} = d(i, j)$$

con $d(i, j)$ una distancia entre la región i y la región j. Así la matriz $\mathbf{G}$ es simétrica y tiene elementos nulos en su diagonal principal.

Consideraremos $\mathbf{g} = \text{vec}(\mathbf{G})$ el vector de $\mathbb{R}^{\mathbf{N}}$ que resulta de apilar las filas de la matriz $\mathbf{G}$.

El modelo resultante es, entonces:

$$\mathbf{y} = \alpha 1_N + log(\mathbf{X}_d)\boldsymbol{\beta}_d + \log(\mathbf{X}_o)\boldsymbol{\beta}_o + \gamma \log(\mathbf{g}) + \varepsilon, \tag{4}$$

con $\varepsilon \sim N(\vec{0}_N, \sigma^2 \mathbf{I}_N)$ y donde todas las variables están expresadas en desvíos respecto a su media.

Si bien el modelo anteriormente planteado puede ser adecuado en algunas situaciones, puede ser inadecuado en situaciones donde los flujos en una región están afectados por los flujos de sus regiones vecinas. Por ejemplo, la creación de nuevas industrias o centros comerciales en una región puede generar un efecto de *derrame* sobre las regiones vecinas. Es así que surgen los modelos que incorporan un término espacial autorregresivo, con el fin de modelizar ese efecto *derrame* que puede darse entre regiones cercanas en el espacio.

3. Modelos con interacción espacial

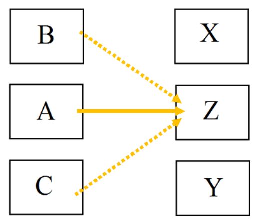
Intuitivamente, los cambios en las características de una región i impactarán los flujos de entrada y de salida de las regiones vecinas a la región i considerada como origen o como destino. En este contexto, se proponen dos enfoques para la modelización de la autocorrelación espacial, los denominados modelos endógenos y exógenos (Patuelli y Arbia 2016). A continuación se describen ambas especificaciones.

3.1. Modelo endógeno

(LeSage y Pace, 2008) y (LeSage y Pace, 2009) sugieren que los flujos de origen - destino exhiben dependencia espacial de tres maneras:

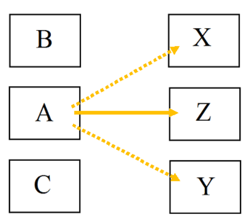
- Dependencia en el Origen:

  Flujos de la región de origen A a la región de destino Z van acompañados por flujos de las regiones vecinas de A, a la región de destino Z.



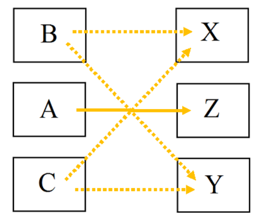
- Dependencia en el Destino:

  Flujos de la región de origen A a la región de destino Z van acompañados por flujos de la región A a regiones vecinas de Z.



- Dependencia en Origen - Destino:

  Flujos de la región de origen A a la región de destino Z van acompañados de flujos de las regiones vecinas de A a las regiones vecinas de Z.



El modelo endógeno pretende modelizar estos tres tipos de interacciones considerando a la propia variable de respuesta rezagada como variable explicativa del modelo. Para ello deben considerarse los tres tipos de interacción planteados, y el rezago espacial, que se reflejará por medio de la matriz de pesos espaciales.

3.2. La matriz de pesos W

Para modelar estas interacciones entre regiones vecinas, se considerará una matriz de pesos, $\mathbf{W} = ((w_{ij}))_{i,j=1,\ldots,n}$ donde w_{ij} indica el grado de influencia entre distintas regiones.

Si bien pueden considerarse casos más sofisticados, en lo que sigue se utiliza una matriz de pesos binaria tal que

$$w_{ij} = \begin{cases} 1 & \text{si } i \text{ y } j \text{ son contiguos} \\ 0 & \text{si } i \text{ y } j \text{ no son contiguos} \end{cases}$$

Los elementos de la diagonal principal de $\mathbf{W}$ son nulos y las filas se normalizan para que sumen uno. Podemos repetir n veces esta cantidad utilizando el producto de Kronecker para obtener las matrices de pesos en origen y en destino:

$$\mathbf{W}_d = \mathbf{I}_n \otimes \mathbf{W}$$

y

$$\mathbf{W}_o = \mathbf{W} \otimes \mathbf{I}_n.$$

De esta manera $\mathbf{W}_d$ captura la dependencia entre regiones en el destino, mientras que $\mathbf{W}_o$ captura la dependencia entre las regiones de origen.

3.3. Especificación del Modelo

El modelo se completa incorporando tres parámetros, ρ_d, ρ_o, ρ_w que indican la intensidad de la dependencia espacial entre orígenes y entre destinos. Así $\rho_d\mathbf{W}_d$ refleja la dependencia entre los orígenes, $\rho_o\mathbf{W}_o$ refleja la dependencia entre los destinos, y $\rho_w\mathbf{W}_w$ la dependencia origen - destino ($\mathbf{W}_w = \mathbf{W}_o \otimes \mathbf{W}_d$).

De esta manera, el modelo queda:

$$(\mathbf{I}_N - \rho_d\mathbf{W}_d)(\mathbf{I}_N - \rho_o\mathbf{W}_o)y = \mathbf{X}\boldsymbol{\beta} + \varepsilon$$

y despejando

$$y = (\mathbf{I}_N - \rho_d\mathbf{W}_d - \rho_o\mathbf{W}_o + \rho_w\mathbf{W}_w)^{-1}(\mathbf{X}\boldsymbol{\beta} + \varepsilon),$$

donde $\varepsilon \sim N(\vec{0}_N, \sigma^2\mathbf{I}_N)$ y $\rho_w = \rho_o\rho_d$ y $\mathbf{X} = (\mathbf{X}_d, \mathbf{X}_o, \mathbf{g})$. Este modelo puede verse como la aplicación de dos filtros espaciales, $(\mathbf{I}_n - \rho_d\mathbf{W}_d)$ y $(\mathbf{I}_n - \rho_o\mathbf{W}_o)$ al modelo sin interacción espacial.

3.4. Modelo exógeno

El modelo exógeno se aplica en el caso en el que el objeto de análisis sean los efectos de *derrame* que provocan los cambios en las características de una región. Por ejemplo, si aumenta la oferta de empleo en una región i, los viajes por trabajo de regiones vecinas j hacia otras regiones disminuirán, aumentando los flujos hacia la región i. La especificación de la interacción espacial en el modelo exógeno se encuentra caracterizada por los rezagos espaciales sobre las variables exógenas X_o y X_d, de la siguiente forma:

$$y = \alpha\iota_{n^2} + X_o\beta_o + X_d\beta_d + g\gamma + W_oX_o\theta_o + W_dX_d\theta_d + \varepsilon$$

donde $\varepsilon \sim N(\vec{0}_N, \sigma^2\mathbf{I}_N)$. Los parámetros θ_o y θ_d dan una medida de la intensidad de la dependencia espacial entre las regiones vecinas en origen y destino respectivamente.

La diferencia con el modelo endógeno es que no hay efectos de retroalimentación sobre la variable de respuesta, provocados por los cambios en las características de las regiones.

En el presente trabajo el modelo exógeno no será estimado, dado que por su especificación presenta problemas de multicolinealidad en muchas de las variables que serán consideradas como explicativas.

4. Interpretación de los parámetros para el Modelo Gravitacional y Endógeno

Los cambios en las características de una región i producen potencialmente impactos en todos los elementos de la matriz de flujos. Se tiene un total de n matrices de dimensión

$n \times n$ de derivadas parciales asociadas con los cambios en cada variable explicativa del modelo.

Las estimaciones de los coeficientes $\hat{\beta}_o$, $\hat{\beta}_d$ no deben ser interpretadas como si fueran de una regresión lineal convencional.

En el Modelo Gravitacional, si por ejemplo $\hat{\beta}_o = 1$ y $\hat{\beta}_d = -0{,}5$ (LeSage y Thomas-Agnan, 2015) el incremento en una unidad del logaritmo en la región 3 de la variable X^r impacta en los flujos de la matriz de origen - destino de la siguiente manera:

$$\Delta Y/\Delta X_3^r = \begin{pmatrix} 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0{,}5 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0{,}5 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

En el Modelo Endógeno, en cambio el impacto sobre la variable de respuesta viene dado por

$$(\mathbf{I}_n - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o + \rho_w \mathbf{W}_w)^{-1} (\Delta Y/\Delta X_3^r)$$

Existen medidas de resumen que pueden interpretarse en forma similar a los coeficientes de una regresión lineal convencional. Para ello es necesario que las variables a utilizar se encuentren como explicativas tanto en la matriz X_o como en la X_d. Puede suceder que algunas variables sean explicativas sólo en origen o sólo en destino y en este caso las medidas de resumen propuestas no son válidas.

Una propuesta alternativa para la intepretación de los parámetros del modelo es estimar los coeficientes del modelo aumentando (o disminuyendo) en una unidad las variables explicativas y comparar el resultado de la predicción con el del modelo original. Esta forma puede resultar engorrosa si son muchos los parámetros estimados, pero da más flexibilidad para la elección de variables a incorporar como explicativas en la matriz de origen o de destino.

5. Datos sobre los que se trabaja

Los datos provienen de la Encuesta de Movilidad en el Área Metropolitana de Montevideo. Se trabajará sólo con los viajes por trabajo. Como se mencionó anteriormente, no se incluyen las localidades del departamento de San José. Las unidades geográficas consideradas son los Municipios en el caso de Montevideo, y el área metropolitana del departamento de Canelones en su totalidad.

La matriz de origen - destino es:

	A	B	C	CH	D	E	F	G	AMC
A	26823	23410	12440	17422	4180	3116	2809	5552	1325
B	499	30987	4467	7948	968	2253	1271	902	1795
C	901	12253	10933	8791	466	922	921	1272	0
CH	429	21159	1481	15542	1288	2684	0	1206	997
D	1269	15327	7415	18080	23229	9382	7607	3658	2248
E	170	22330	4765	11946	1297	7882	1785	913	1878
F	4438	15801	7906	9393	6686	7510	19080	1227	3067
G	4295	9071	4424	2544	1725	2263	1407	13740	440
AMC	335	13488	4438	6925	3743	6597	3317	1811	69129

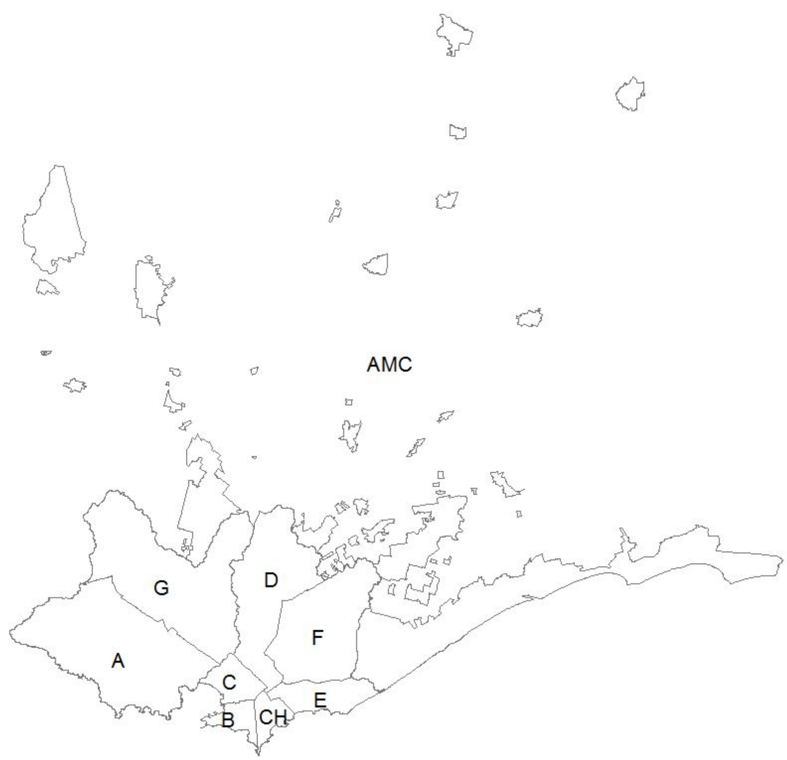


Figura 1: Unidades Geográficas de la matriz de Origen - Destino

En las siguientes figuras se muestran dos gráficos realizados a partir de la matriz de flujos. En la Figura 1 se muestra el que se denomina "líneas de deseo". El ancho de la

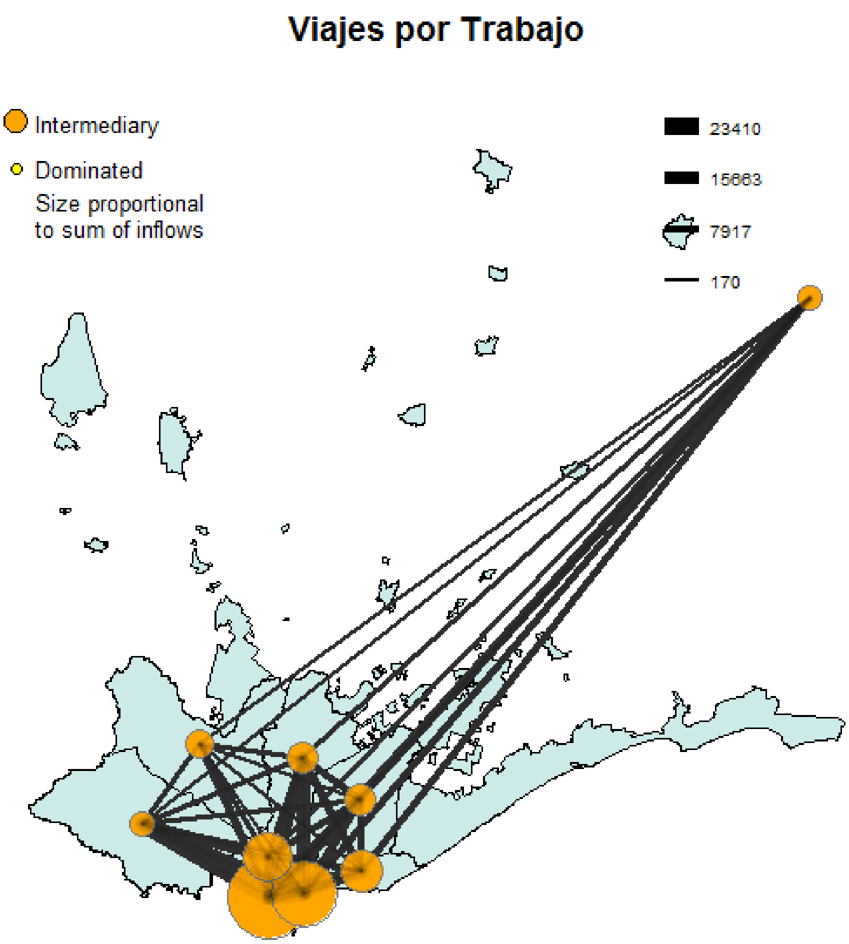


Figura 2: Líneas de deseo

línea es proporcional a la cantidad de flujos, y el tamaño de los puntos es proporcional a la cantidad de viajes de entrada al municipio.

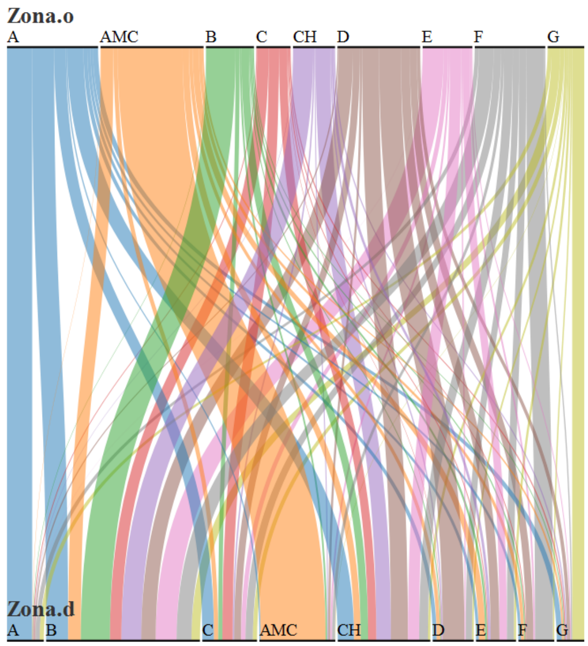


Figura 3: Matriz de flujos de Origen - Destino

En la Figura 3 el tamaño de los segmentos es proporcional a la cantidad de viajes de salida en origen, y de entrada en destino. Los municipios A y D y el Área Metropolitana de Canelones son los que muestran más salidas, mientras que el B es el que recibe más cantidad de viajes. También se aprecia que en la mayoría de los casos, los viajes dentro del mismo municipio representan la mayor proporción como destino dentro de cada municipio.

Las variables explicativas se obtienen del sistema de información geográfica de la Intendencia de Montevideo y de la Encuesta Continua de Hogares del año 2016. La selección de variables no pretende ser definitiva. Las incorporadas en esta instancia son la cantidad de empresas por Municipio, de tres clases de actividad, el ingreso promedio de los hogares del municipio, y la cantidad de ocupados. Las variables a incorporar al modelo deben estar georreferenciadas, lo que al momento es una limitante, ya que en muchos casos no es posible obtener información a nivel de las unidades geográficas.

La matrix $\mathbf{X} \in \mathbb{R}^{8\times5}$ es:

	industria	finanzas	comercio	Ingreso Promedio	Ocupados
A	52	5	6	54287	95943
B	114	110	19	75152	75456
C	97	10	15	76950	88250
CH	40	50	21	116531	94941
D	40	7	9	51167	94171
E	34	17	18	91611	83702
F	63	9	6	51462	83066
G	63	3	5	61201	70272
AMC	445	149	6	58352	275191

- *Industria*: cantidad de empresas del sector industrial.
- *Comercio*: cantidad de comercios dentro de la categoría de Centros Comerciales (Shoppings) y Grandes Supermercados.
- *Finanzas*: cantidad de empresas del sector financiero.
- *Ingreso Promedio*: ingreso promedio del hogar.
- *Ocupados*: cantidad de personas ocupadas.

La matriz $\mathbf{W}$ se construye en base a contigüidad.

6. Resultados

En una primera instancia se ajustan los modelos Gravitacional y Endógeno incluyendo todas las variables como explicativas en las matrices X_o y X_d. Luego se realiza una selección *stepwise* obteniendo los modelos que se presentan a continuación:

Tabla 1: Resultados del ajuste del Modelo Gravitacional

	Estimación	Error Estándar	Valor t	p- valor	Significación
(Intercept)	41.53187	9.92336	4.185	7.94e-05	***
Industria Origen	-0.31499	0.12261	-2.569	0.012273	*
Comercio Origen	-1.26404	0.30031	-4.209	7.30e-05	***
Ingreso Origen	2.51861	0.61419	4.101	0.000107	***
Industria Destino	0.49964	0.16609	3.008	0.003618	**
Comercio Destino	2.24438	0.30412	7.380	2.19e-10	***
Ingreso Destino	-4.04746	0.62045	-6.523	8.26e-09	***
Ocupados Destino	-1.54777	0.49506	-3.126	0.002552	**
Distancia	-0.20235	0.02901	-6.974	1.23e-09	***

Tabla 2: Resultados del ajuste del Modelo Endógeno

	Estimación	Error Estándar	Valor t	p- valor	Significación
(Intercept)	53.3333	11.1524	4.782	9.11e-06	***
ρ_d	0.3738	0.1750	2.135	0.03618	*
Industria Origen	-0.1937	0.1325	-1.462	0.14805	
Comercio Origen	-0.8083	0.3626	-2.229	0.02897	*
Ingreso Origen	1.5559	0.7501	2.074	0.04170	*
Industria Destino	0.5158	0.1623	3.178	0.00220	**
Comercio Destino	2.3616	0.3019	7.822	3.57e-11	***
Ingreso Destino	-4.4901	0.6401	-7.014	1.11e-09	***
Ocupados Destino	-1.6412	0.4852	-3.382	0.00117	**
Distancia	-0.2108	0.0286	-7.372	2.43e-10	***

Algunas de las variables es incluida sólo en origen o sólo en destino. Aunque no sea significativa la variable "pareja" omitida en la selección *stepwise*, puede ser necesario incluirla a la hora de realizar predicciones. Si no se incluyen las dos variables en origen y en destino, no hay efectos que contrarresten los impactos provocados por cambios en las variables explicativas del modelo. Se probó incluir las variables "pareja" en los modelos encontrando que los signos de los coeficientes estimados de las variables agregadas al modelo eran iguales al de su pareja en origen o en destino, no encontrando así una mejora para la predicción en el sentido de que al mantenerse el signo en origen y en destino, no hay un efecto que contrarreste los impactos sobre la matriz de flujos. En el caso del modelo endógeno esto se podría dar si el coeficiente ρ_o fuera significativo.

El R^2 del modelo gravitacional es 0.6346 y en el modelo endógeno 0.6519. Al incorporar los rezagos espaciales, el parámetro asociado al destino ρ_d es significativo. Si existe correlación espacial en alguno de los sentidos, el modelo endógeno es el más adecuado. La correlación espacial en destino significa que cuando una región es atractiva por motivo de trabajo, también lo son sus vecinas. En cuanto a las variables explicativas, la única no incluida en la selección *stepwise* ni en origen ni en destino es la de las empresas del sector financiero. Llama la atención el signo negativo de la variable *Ocupados* en destino, que debería de interpretarse como que a mayor cantidad de ocupados en la región, menos viajes por trabajo hacia ella. En este caso puede suceder que la estimación se vea afectada por la cantidad de viajes intrarregióm que muestra la matriz de origen - destino. Queda como línea futura de trabajo analizar métodos de estimación para el modelo que considere el efecto que puede tener una diagonal con alta frecuencia en la matriz de origen - destino.

Como se anticipó en los apartados anteriores, los coeficientes no deben interpretarse como los de una regresión lineal convencional. Sí se le puede dar una lectura inicial a los signos de los coeficientes. En el modelo gravitacional, un incremento en el logaritmo de la cantidad de industrias en la región hace que caigan los flujos de salida hacia otras regiones por trabajo. Esto se refleja en el parámetro negativo. Desde el punto de vista del destino, un incremento en la cantidad de empresas del sector industrial en una determinada región hace que los viajes por trabajo aumenten hacia esa región.

6.1. Interpretación de los parámetros

Dado que en los modelos ajustados las variables incluidas no aparecen todas como explicativas en origen y en destino, el impacto sobre los flujos de un cambio en las variables explicativas se medirá ajustando los modelos con un incremento (o baja) de una unidad en la variable en una determinada región. Se estimará la matriz de origen - destino como predicción del modelo ajustado, y se presentarán los efectos en forma matricial.

El primer escenario a analizar es el de un aumento de una unidad de la cantidad de empresas industriales en el municipio D, que es uno de los que tiene menor cantidad de industrias en Montevideo. En el modelo Gravitacional el impacto en los flujos sería el siguiente:

	A	B	C	CH	D	E	F	G	AMC
A	0	0	0	0	216	0	0	0	0
B	0	0	0	0	99	0	0	0	0
C	0	0	0	0	109	0	0	0	0
CH	0	0	0	0	86	0	0	0	0
D	-3	-164	-11	-8	71	-10	-21	-2	-9
E	0	0	0	0	224	0	0	0	0
F	0	0	0	0	148	0	0	0	0
G	0	0	0	0	117	0	0	0	0
AMC	0	0	0	0	32	0	0	0	0

En total los viajes diarios por trabajo hacia otros municipios bajaría en aproximadamente 157 viajes, y se atraerían unos 1102 viajes por trabajo hacia el municipio D. Estas cantidades representan un 0.2 % y un 2.5 % de los viajes totales desde y hacia el municipio D respectivamente. Los aumentos y las bajas no se compensan. Esto se debe a que el modelo no tiene ningún tipo de restricción. En un futuro, de acuerdo al modelo de cuatro etapas del transporte, se debería de incorporar una restricción de forma que las predicciones obtenidas sumen el total de viajes de las marginales de la matriz de origen - destino.

En el caso del modelo endógeno, con el mismo escenario, el impacto en los flujos es:

	A	B	C	CH	D	E	F	G	AMC
A	3	3	19	19	234	21	18	19	21
B	1	1	9	8	107	9	8	8	9
C	2	2	10	9	118	10	9	9	10
CH	1	1	8	7	93	8	7	7	8
D	-3	-104	-21	-18	121	-1	-5	4	1
E	3	3	20	19	243	21	19	19	21
F	2	2	13	13	161	14	13	13	14
G	2	2	10	10	126	11	10	10	11
AMC	0	0	3	3	34	3	3	3	3

A diferencia del modelo gravitacional, en donde el aumento de la cantidad de empresas repercute sólo en las filas y columnas correspondientes al municipio D, en el caso del modelo endógeno toda la matriz se ve afectada. Los impactos son mayores cuanto más cercanas sean las regiones al municipio D.

El segundo escenario propuesto es el de la baja de un 10 % del ingreso en el municipio CH. Utilizando el modelo Gravitacional, el impacto en la matriz de flujos es el siguiente:

	A	B	C	CH	D	E	F	G	AMC
A	0	0	0	705	0	0	0	0	0
B	0	0	0	955	0	0	0	0	0
C	0	0	0	115	0	0	0	0	0
CH	-78	-3144	-1034	12083	-1614	-872	-1538	-773	-422
D	0	0	0	530	0	0	0	0	0
E	0	0	0	1196	0	0	0	0	0
F	0	0	0	999	0	0	0	0	0
G	0	0	0	1631	0	0	0	0	0
AMC	0	0	0	234	0	0	0	0	0

El coeficiente de la variable *Ingreso* en origen es positivo y en destino negativo. A mayores ingresos menos atracción por trabajo. Una caída en el ingreso implica un aumento de los viajes por trabajo hacia el municipio, y una baja de los viajes por trabajo hacia otras regiones desde el municipio CH. Con el modelo endógeno:

	A	B	C	CH	D	E	F	G	AMC
A	9	86	104	840	100	88	16	17	18
B	13	117	141	1138	135	120	22	24	25
C	2	14	17	137	16	14	3	3	3
CH	66	488	1889	25541	1520	1811	-745	-265	-128
D	7	65	78	632	75	67	12	13	14
E	16	146	177	1425	169	150	27	30	31
F	13	122	148	1191	142	125	23	25	26
G	22	200	241	1944	231	205	37	40	42
AMC	3	29	35	279	33	29	5	6	6

Al igual que en el primer escenario, los cambios en el municipio CH repercuten en toda la matriz de origen - destino. Respecto al modelo gravitacional se observa además que los viajes de salida bajan sólo para tres municipios: F, G y Área Metropolitana de Canelones, generando un efecto "derrame" positivo en toda la matriz. Esto se debe a la magnitud del coeficiente de la variable ingreso en destino, que casi triplica al coeficiente de origen. Es decir, la atracción del municipio CH por motivos de trabajo es mucho más fuerte que los viajes que genera cuando se trata de la variable ingreso, los cambios en esta variable tienen un mayor impacto en destino.

7. Discusión y lineamientos a futuro

Los modelos presentados permiten analizar escenarios hipotéticos que pueden ser útiles a la hora de la planificación urbana. Dada la significación del parámetro ρ_d, el modelo endógeno es el más adecuado, incluyendo la autocorrelación espacial de los municipios en destino.

Como se mencionó anteriormente, este trabajo es un ejercicio de aplicación, con el fin de contribuir a un mejor entendimiento de los modelos propuestos. A futuro se trabajará en los siguientes lineamientos:

- Selección de variables a incorporar: el hecho de que las variables deban estar georreferenciadas es una limitante ya que en muchos casos esta información no es de dominio público, y los modelos ajustados quedan sesgados hacia aquellas variables a las que se accede con facilidad.

- Imputación de las celdas vacías de la matriz de origen - destino: a nivel de municipios fueron imputados dos valores de acuerdo al ponderador correspondiente al estrato al que pertenecía en la muestra. En la literatura la imputación de los ceros es un problema, ya que afecta la estimabilidad del modelo. En nuestro caso deben explorarse las alternativas propuestas como solución, y es el paso previo a realizarse si se quiere cambiar al municipio como unidad de análisis.

- CCZ como unidad de análisis: los municipios son unidades geográficas muy grandes y algunas correlaciones pueden no detectarse en ese nivel. Los CCZ son unidades más chicas, y es de interés comparar los resultados de los modelos con distintas unidades de información.

- Probar con otras funciones de distancia y otros centroides. Los centroides con los que se trabajaron fueron los geométricos, pero se entiende que el centro nuclear de los municipios puede encontrarse en otro lugar. Se probará mover el centroide hacia estos puntos y se compararán los resultados con los del modelo inicial.

- Incorporar restricciones al modelo de forma que los viajes en origen y destino se compensen en la predicción, y que reproduzcan las marginales de la tabla.

Referencias Bibliográficas

Fischer, M. M. y Griffith, D. A. (2008). Modeling spatial autocorrelation in spatial interaction data: an application to patent citation data in the European Union. *Journal of Regional Science*, 48(5):969–989.

LeSage, J. P. y Pace, R. K. (2008). Spatial econometric modeling of origin-destination flows. *Journal of Regional Science*, 48(5):941–967.

LeSage, J. P. y Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Chapman & Hall.

LeSage, J. P. y Thomas-Agnan, C. (2015). Interpreting spatial econometric origin-destination flow models. *Journal of Regional Science*, 55(2):188–208.

Patuelli, R. (2016). *Spatial Econometric Interaction Modelling*. Springer, Cham, Switzerland.

Instituto de Estadística

Documentos de Trabajo

Eduardo Acevedo 1139. CP 11200 Montevideo, Uruguay
Teléfonos y fax: (598) 2410 2564 - 2418 7381
Correo: ddt@iesta.edu.uy
www.iesta.edu.uy
Área Publicaciones

Diciembre, 2017
DT (17/04)