

INSTITUTO DE ESTUDIOS LABORALES Y DEL DESARROLLO ECONÓMICO (ielde)
Facultad de Ciencias Económicas, Jurídicas y Sociales
Universidad Nacional de Salta (UNSa)
Salta
Argentina

Documentos de Trabajo

Serie de documentos técnicos N° 1

Econometría especial usando Stata. Breve guía aplicada para datos de corte transversal

Marcos Herrera

Otoño de 2015
N° 13

ielde – Facultad de Ciencias Económicas, Jurídicas y Sociales - UNSa

<http://www.economicas.unsa.edu.ar/ielde>

UNSa: Av. Bolivia 5150, A4408FVY, Salta, Argentina
ISSN 1852-1118 (impreso), ISSN 1852-1223 (en línea)

Econometría espacial usando Stata.

Breve guía aplicada para datos de corte transversal*

Marcos Herrera

CONICET-IELDE

Universidad Nacional de Salta (Argentina)

(E-mail: mherreragomez@gmail.com)

Abstract

This document aims to present the common tools within Stata to estimate and test spatial econometric models. This paper contains an exploratory analysis of spatial data and a range of simple and complex spatial regression models. The guide can be used as a manual to apply spatial econometrics in the context of Stata software. The database and codes used in the different sections are available to replicate step by step procedure.

JEL codes: C21, R12, C52.

Keywords: Spatial econometrics, Spatial effects, Stata.

Resumen

Este documento tiene por objetivo presentar las herramientas habituales dentro de Stata para estimar y contrastar modelos econométricos espaciales. Este documento contiene un análisis exploratorio de datos espaciales y una gama de modelos de regresión espacial simples y complejos. La guía puede ser usada como un manual para realizar econometría espacial bajo el contexto del programa Stata. La base de datos y los códigos utilizados en las diferentes secciones se encuentran disponibles para replicar el paso a paso del procedimiento.

Códigos JEL: C21, R12, C52.

Palabras clave: Econometría espacial, Efectos espaciales, Stata.

*Este documento ha sido elaborado con algunas secciones del trabajo preparado para el Taller sobre “Técnicas de Investigación en Economía Aplicada”, XL Reunión de Estudios Regionales, 19 de Noviembre de 2014, Zaragoza (España). El autor agradece las diferentes opiniones recogidas del 7th seminar Jean Paelinck (Zaragoza, 2014) que han permitido mejorar la presente versión, así como de los comentarios realizados por Jorge Paz y Maribel Jiménez. Todos los errores y omisiones son responsabilidad del autor.

1. Introducción

Este documento pretende ser una guía para los economistas aplicados que desean trabajar con datos georeferenciados. Este tipo de datos, por propia naturaleza, ha generado nuevas áreas de investigación y de especialización siendo una de ellas la econometría espacial.

El área de la econometría espacial se ocupa del uso de técnicas estadísticas y econométricas que permiten el tratamiento explícito del espacio en modelos económicos. Este nombre fue dado por Jean Paelinck, quien utilizó por primera vez el término para destacar la necesidad de desarrollar una nueva rama de la econometría que sirviese de fundamentación metodológica a los modelos regionales y urbanos (*Annual Meeting of the Dutch Statistical Association*, Tilburg, 1974). Anselin (2010) ubica el inicio del área en el año 1979 debido a un conjunto de publicaciones que sentaron las primeras bases metodológicas distinguibles: *Spatial econometrics* (Paelinck & Klaassen, 1979), *Exploratory and explanatory statistical analysis of spatial data* (Bartels & Ketellapper, 1979), *Spatial time series* (Bennett, 1979) y *Problems in estimating econometric relations in space* (Hordijk, 1979). Una década más tarde, con la publicación del libro de texto: *Spatial econometrics: Methods and models* (Anselin, 1988), se consolidan las líneas metodológicas fundamentales que han marcado la agenda de investigación en los posteriores años.

Un punto crucial sobre el desarrollo y difusión de la econometría espacial ha sido el relacionado a los programas econométricos. En los inicios, década del 70' y 80', la falta de un programa apropiado para el análisis espacial fue uno de los principales impedimentos en la realización de investigaciones empíricas. Los investigadores utilizaban códigos de Fortran o de algún otro tipo de lenguaje de programación ya que los programas disponibles, tales como SAS y SPSS, solo proveían herramientas estadísticas estándares no considerando la presencia de dependencia espacial. La autocorrelación espacial era vista más bien como un problema a ser removido o evitado (Goodchild et al., 1992). En la década del 90', estas dificultades comienzan a sortearse debido al avance tecnológico de los sistemas de información geográfica (SIG) y al desarrollo de programas estadísticos específicos para manipular y modelizar estos datos como SpaceStat, primera vez lanzado en 1991 (Anselin, 1992).

En la actualidad, existen varias alternativas para realizar un análisis econométrico espacial. Sin ser exhaustivos, pueden mencionarse: Econometrics Toolbox desarrollado en MATLAB (LeSage, 1999); el paquete spdep de R desarrollado en el año 2001 (Bivand, 2002); el programa GeoDa, lanzado en el año 2002 (Anselin et al., 2006), con gran impacto a nivel usuario; y GeoDaSpace de reciente desarrollo basado en el código PySAL (Rey & Anselin, 2010).

El programa Stata no ha sido un precursor del análisis espacial, aunque diferentes usuarios han aportado herramientas espaciales muy útiles bajo su entorno. El primer aporte importante data del año 2001, *Tools for spatial data analysis* (Pisati, 2001). Desde entonces, se han desarrollado gradualmente diferentes comandos que permiten alcanzar un nivel de análisis adecuado para realizar investigaciones en econometría espacial. En el presente documento, se presentarán los códigos disponibles más relevantes hasta la fecha utilizando Stata, versión 12.

En orden de facilitar la reproducibilidad de los resultados aquí presentados, se deja a disposición un conjunto de archivos comprimidos bajo el nombre Guia_Stata_2015.rar en el siguiente link: https://www.researchgate.net/publication/274347889_Dataset._Gua_de_Econometra_espacial_usando_Stata. Estos archivos son: comandos_2015.do (archivo con los comandos de Stata), migr_unemp_10_12.xls (base de datos con variables de Eurostat), el shape nuts2_164 (archivo de formas espaciales). El archivo shape es un formato multiarchivo, es decir, son

archivos asociados con igual nombre pero diferentes extensiones: .shp, .dbf y .shx. En el próximo capítulo se verá en detalle cómo trabajar con esta información.

Datos a utilizar

A lo largo de la guía se utilizará un único conjunto de datos. Estos datos corresponden a 164 regiones NUTS 2¹ de la primera formación de la Unión Europea, para el periodo 2010-2012. Si bien solo se utilizará un único año, los demás años se dejan disponibles para que el lector pueda utilizar alternativamente alguno de ellos. Con la finalidad de mantener la discusión lo más simple posible, se han seleccionado dos variables regionales: la tasa de desempleo y la tasa de migración neta. Ambas variables han sido construidas con información proveniente de la Oficina Europea de Estadísticas, más conocida como Eurostat.

La tasa de desempleo es definida por Eurostat, siguiendo los principales lineamientos de la Organización Internacional del Trabajo, como el porcentaje de personas desempleadas respecto a la fuerza laboral regional, es decir, población en edad de trabajar entre 15-74 años.

La tasa de migración neta se define como la razón de la migración neta durante el año (incluyendo el ajuste estadístico) respecto a la población promedio en ese año, expresada por cada mil habitantes. La migración neta es estimada como la diferencia entre el número de personas que ingresan y el número de personas que egresan de una región por lo que la variable puede asumir valores positivos o negativos. Un valor positivo implica un exceso de personas ingresantes y es referida como inmigración neta (por ejemplo, 18,3 migrantes por cada 1000 ha., valor máximo en el año 2010). Un valor negativo implica un exceso de personas que egresan de la región y es referido como emigración neta (por ejemplo, -4,1 migrantes por cada 1000 ha., valor mínimo en el año 2010).

Cuadro 1: Estadísticos Descriptivos. Desempleo y Migración Neta

Año	Tasa de Desempleo (%)				Tasa de Migración Neta			
	Media	Des. est.	Mín.	Max.	Media	Des. est.	Mín.	Max.
2010	8,87	4,26	2,7	28,0	2,67	3,55	-4,1	18,3
2011	8,83	4,78	2,8	30,4	2,32	3,74	-11,4	21,6
2012	9,55	5,88	2,7	34,6	2,69	4,26	-8,0	18,9

Un número escaso de datos faltantes han sido predichos mediante extrapolación espacio-temporal con la finalidad de poder utilizar todas las regiones y años. Estos datos, en su práctica totalidad, no han sido revisados minuciosamente por lo que las conclusiones que puedan derivarse son solo a fines ilustrativos. Sin embargo, la elección de ambas variables no ha sido al azar y se intenta, mediante la aplicación empírica, incentivar el estudio de la relación entre migración neta y desempleo contribuyendo al debate entre analistas económicos y, también, entre la opinión pública.

Desde la teoría económica, existen algunas posiciones que argumentan que la inmigración neta es la causa del elevado desempleo en las regiones receptoras de migrantes. La escuela neoclásica es la que apoya tal conclusión, suponiendo que la mano de obra es homogénea y que existe competencia perfecta en el mercado de bienes. Bajo estos supuestos, los trabajadores se mueven hacia regiones

¹Las siglas NUTS significan Nomenclatura de las Unidades Territoriales Estadísticas y son utilizadas por la Unión Europea para establecer demarcaciones territoriales. Las NUTS se dividen en tres niveles: El nivel 1 contiene regiones con una población entre 7 y 3 millones de habitantes. El nivel 2 son regiones con una población entre 800 mil y 3 millones. El nivel 3 son regiones de población mínima de 150 mil y un máximo de 800 mil habitantes. Para mayor precisión de los criterios utilizados, puede consultarse la página web de Eurostat: <http://ec.europa.eu/eurostat>.

prósperas y esto genera un aumento en la oferta de trabajo en la región receptora (efecto directo). Simultáneamente, en dicha región receptora, la inmigración genera un aumento en el consumo de bienes locales llevando a las empresas a incrementar la demanda de trabajo (efecto indirecto). La posición neoclásica afirma que el efecto directo es superior al indirecto con el consiguiente incremento del desempleo: existe una relación positiva entre migración neta y desempleo.

Una corriente teórica alternativa apoya la misma dirección de causalidad, desde la migración neta hacia el desempleo. Esta posición es la adoptada por la Nueva Geografía Económica bajo la dinámica centro-periferia. Considerando los supuestos de competencia imperfecta en el mercado de bienes y rigideces en el mercado laboral, puede deducirse que las fuerzas generadoras de aglomeración determinan las disparidades espaciales que producen persistencia interregional en la tasa de desempleo. Si se considera que la integración regional es producto de la caída de los costos de transporte, entonces se espera un mayor flujo migratorio desde las regiones retrasadas hacia las regiones prósperas. El efecto migratorio estimulará las economías de aglomeración (efecto home-market) llevando a un incremento de los beneficios empresariales y por consiguiente una mayor demanda laboral. En este caso, los efectos indirectos de la inmigración sobre la demanda laboral prevalecerán sobre el efecto directo. En conclusión, existe una relación negativa entre ambas variables, donde mayor migración neta (inmigración) produce una caída en la tasa de desempleo de la región receptora.

Ambas teorías dan pie a postular una relación inicial en donde la variable dependiente es la tasa de desempleo y la variable explicativa es la tasa de migración neta. Sin embargo, el signo y la magnitud del impacto de la migración neta sobre el desempleo es una cuestión empírica. Se explorará dicha relación mediante las técnicas espaciales para corte transversal.

La guía se encuentra estructurada de la siguiente manera. En la sección 2 se realiza un análisis exploratorio espacial de los datos, habitual etapa inicial de los estudios de datos espaciales. Debido a la particularidad de los mismos, se utilizan herramientas puntuales como son los diferentes tipos de mapas y diagramas de dispersión de Moran. En la sección 3 se presentan los modelos más utilizados de econometría espacial para datos de corte transversal. La sección 4 se dedica a la interpretación de los resultados. Esta interpretación considera cuidadosamente los efectos espaciales y las implicancias para cada modelo. En todas las secciones se combina la presentación metodológica de la herramienta y los comandos disponibles en Stata para su estimación.

2. Análisis Exploratorio de Datos Espaciales

Esta sección muestra las herramientas disponibles para realizar un análisis exploratorio de datos espaciales (AEDE). El AEDE es un conjunto de técnicas que permite detectar asociaciones espaciales, concentraciones locales y detección de datos anómalos (outliers). La idea detrás del AEDE es que los datos hablen por si mismos, sin imponer una estructura conceptual previa, similar al concepto que surge del Análisis Exploratorio de Datos (AED) (Anselin, 1999). Sin embargo, las herramientas usuales del AED no son válidas para datos espaciales ya que por naturaleza los mismos muestran autocorrelación y es por ello que se necesitan desarrollos específicos como los que se presentan a continuación.

La sección mostrará el potencial de los mapas de coropletas y algunos indicadores simples que permiten detectar patrones espaciales y sugerir hipótesis de trabajo.

2.1. Convirtiendo archivos shapes a Stata

Para iniciar el análisis espacial es necesario tener información georeferenciada. Un problema habitual es la obtención de dicha información. Sin embargo, esta restricción se ha vuelto menos importante a medida que diferentes sitios webs han ido ofreciendo archivos georeferenciados de manera gratuita. Entre los sitios gratuitos puede mencionarse Global Administrative Areas (GADM), que provee archivos de mapas base para casi todos los países²; y el sitio DIVA-GIS³.

El tipo de archivo estándar en el que se almacena la información geográfica es denominado shape (shapefile), originalmente desarrollado por la empresa ESRI para su programa ArcView GIS. Si bien existen otros formatos alternativos, el formato shape se ha convertido en el tipo de archivo más utilizado para el intercambio de información geográfica.

Un archivo shape es un formato vectorial que almacena la localización de elementos geográficos y sus atributos. La diferencia con otro tipo de archivo es que es un formato multiarchivo: conjunto relacionado de ficheros informáticos. El número mínimo y habitual es de tres ficheros con las siguientes extensiones:

.shp es el archivo que almacena las entidades geométricas de los objetos.

.shx es el archivo que almacena el índice de las entidades geométricas.

.dbf es la base de datos, en formato dBASE, donde se almacena la información de los atributos de los objetos.

Existen herramientas desarrolladas por usuarios para convertir los shapes en archivos .dta de Stata. El comando **shp2dta** (Crow, 2013) permite crear dos conjuntos de datos: uno contendrá la información de los atributos del .dbf, y el otro contendrá la información geográfica.

El primer paso es cargar la información georeferenciada mediante el siguiente comando:

```
. shp2dta using nuts2_164, database(datos_shp) coordinates(coord) genid(id) genc(c)
```

Este comando lee el archivo shape “nuts2_164”, generando dos archivos: “datos_shp.dta” y “coord.dta”. En el archivo de datos, “datos_shp”, se genera adicionalmente una variable identificadora de cada unidad geográfica, **id**, y los centroides, puntos que definen el centro geométrico de los polígonos, identificados por la latitud (**y_c**) y la longitud (**x_c**).

```
. use datos_shp, clear
. describe
```

```
Contains data from datos_shp.dta
```

```
obs:          164
vars:           9                15 Oct 2014 01:44
size:          5,904
```

```
-----
              storage  display  value
variable name  type    format   label    variable label
-----
id              int     %12.0g          Area ID
```

² <http://www.gadm.org/about>

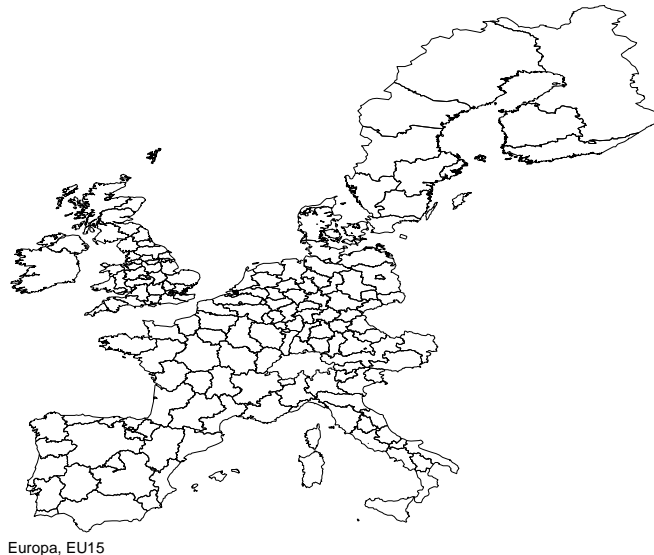
³ <http://www.diva-gis.org/gdata>

x_c	float %9.0g	x-coordinate of area centroid
y_c	float %9.0g	y-coordinate of area centroid
POLY_ID	int %10.0g	POLY_ID
OBJECTID	int %10.0g	OBJECTID
NUM	int %10.0g	NUM
NUTS_ID	str4 %9s	NUTS_ID
X_COORD	double%10.0g	X_COORD
Y_COORD	double%10.0g	Y_COORD

El archivo `datos_shp.dta` contiene la información sobre los atributos del archivo `dbf`. El archivo de atributos tiene únicamente las variables que identifican a cada NUTS 2, el nombre del objeto, y las coordenadas geográficas.

Pisati (2008) provee el comando `spmap` que permite obtener un gráfico del mapa con los límites de cada área, es decir, el mapa base.

```
. spmap using datos_shp, id(id)
```



Puede suceder que el shape contenga polígonos que no son de interés o de los que no se tenga información de las variables. En este caso puede utilizarse el comando `drop` para eliminar las regiones sin información. Por ejemplo, si no hay información sobre Finlandia y Suecia:

```
. drop if id==3|id==5|id==6|id==164|id==7|id==8|id==12|id==4|id==2|id==1|id==11|id==12
. spmap using coord, id(id)
```



Europa sin Finlandia y Suecia, EU15

Estos mapas base solo contienen información geográfica, no hay variables que puedan ser analizadas. Para incorporar la información descargada de Eurostat, el primer paso ha sido guardar las variables en formato excel e incluir una variable que identifica a cada NUTS 2 con su correspondiente forma geográfica. Esta variable servirá de link entre el archivo excel y el shape.

En Stata, se importa el archivo excel y se guarda en formato .dta:

```
. import excel "C:\.\.\migr_unemp_10_12.xls", firstrow
. save "C:\.\.\migr_unemp.dta"
. use migr_unemp
. describe
```

Contains data from migr_unemp.dta

```
obs:          164
vars:          26                15 Oct 2014 01:47
size:         32,472
```

```
-----
      storage display      value
variable name  type  format  label      variable label
-----
POLY_ID        int   %10.0g          POLY_ID
NUTS_ID        str4  %9s            NUTS_ID
NM2010         double%10.0g      NM2010
NM2011         double%10.0g      NM2011
NM2012         double%10.0g      NM2012
U2010          double%10.0g      U2010
U2011          double%10.0g      U2011
U2012          double%10.0g      U2012
-----
```

La información ha sido guardada en formato wide: el número de filas se corresponde al número de NUTS 2 y, para cada año, existe una variable de migración neta y desempleo. Con la información guardada en archivos Stata, solo queda unir ambas bases de datos. Obsérvese que hay una variable repetida en ambas bases que permitirá la unión: POLY_ID.

```
. use datos_shp, clear
. merge 1:1 POLY_ID using migr_unemp, gen(union) force
```



```

Result                                # of obs.
-----
not matched                            0
matched                                164 (union==3)
-----

. assert union==3
. save migr_unemp_shp.dta
file migr_unemp_shp.dta saved

```

Mediante el comando `merge` la información de ambas bases es unida y la función `assert union==3` verifica que la condición sea verdadera. Si esta condición se cumple, entonces Stata no devolverá ninguna línea adicional.

2.2. Manipulación y visualización de datos espaciales

Una vez unidas las bases de datos, es posible visualizar las variables mediante diferentes mapas de coropletas. Este tipo de mapa es un mapa temático en el que las áreas se colorean con una intensidad proporcional al valor que toma una variable cuantitativa (Slocum et al., 2009). Para ello utilizaremos el comando antes presentado, `spmap`.

En `spmap`, un mapa de coropletas es un mapa base cuyos polígonos estarán coloreados de acuerdo al valor tomado de un atributo (variable) que debe ser numérico. La versión actual de `spmap` ofrece seis tipos de coropletas:

- Cuantiles: las clases corresponden a los cuantiles de la distribución de la variable, cada clase tendrá aproximadamente el mismo número de polígonos.
- Intervalos iguales: las clases se dividen según el valor que divide a la distribución de la variable en k intervalos iguales. El valor predefinido de k es 4.
- Diagrama de caja: las clases son seis y el número de polígonos en cada clase depende de la distribución de la variable elegida. Las clases son definidas de la siguiente manera: $[min, p25 - 1,5 \cdot ri]$, $(p25 - 1,5 \cdot ri, p25]$, $(p25, p50]$, $(p50, p75]$, $(p75, p75 + 1,5 \cdot ri]$, $(p75 + 1,5 \cdot ri, max]$, donde p es el percentil y ri es el rango intercuartílico.
- Desviación estándar: la distribución de la variable es dividida en k clases ($2 \leq k \leq 9$) cuya amplitud es definida como una fracción p de su desvío estándar, $s.d.$. El número de clases predefinido es $k = 4$: $[min, media - s.d.]$, $(media - s.d, media]$, $(media, media + s.d.)$, $(media + s.d., max.]$.
- k-medias: la distribución de la variable es dividida en k clases usando el análisis de partición de cluster de k-medias. El procedimiento es iterativo, hasta alcanzar la solución que maximiza la bondad de ajuste de la varianza.
- Custom: los intervalos de clase son especificados por el usuario.

Las anteriores opciones solo permiten trabajar con variables cuantitativas. En el caso de las variables categóricas puede usarse la opción `unique`, que permite graficar variables categóricas asignando un color diferente a cada una de las categorías.

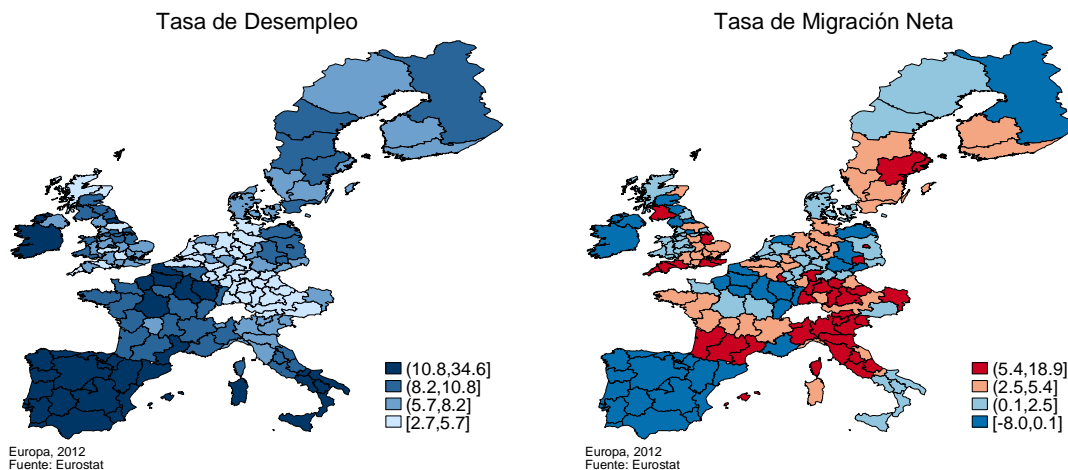
A continuación se presentan los mapas de coropletas aplicados a los datos de EU15 utilizando la información del último año disponible.

```

. format U2012%12.1f
. spmap U2012 using coord, id(id) clmethod(q) title("Tasa de Desempleo") ///
  legend(size(medium) position(5)) fcolor(Blues2) note("Europa, 2012" "Fuente: Eurostat")
. format NM2012%12.1f
. spmap NM2012 using coord, id(id) clmethod(q) title("Tasa de Migración Neta") ///
  legend(size(medium) position(5)) fcolor(BuRd) note("Europa, 2012" "Fuente: Eurostat")

```

Figura 1: Mapa de Cuantiles



El mapa de cuantiles se genera utilizando el subcomando `clmethod(q)`. Además puede modificarse el número de clases usando `cln(#)`, así como los colores, la posición de la leyenda, títulos, notas, etc. La primera línea previa al comando del gráfico, `format`, permite controlar el formato numérico de la leyenda.

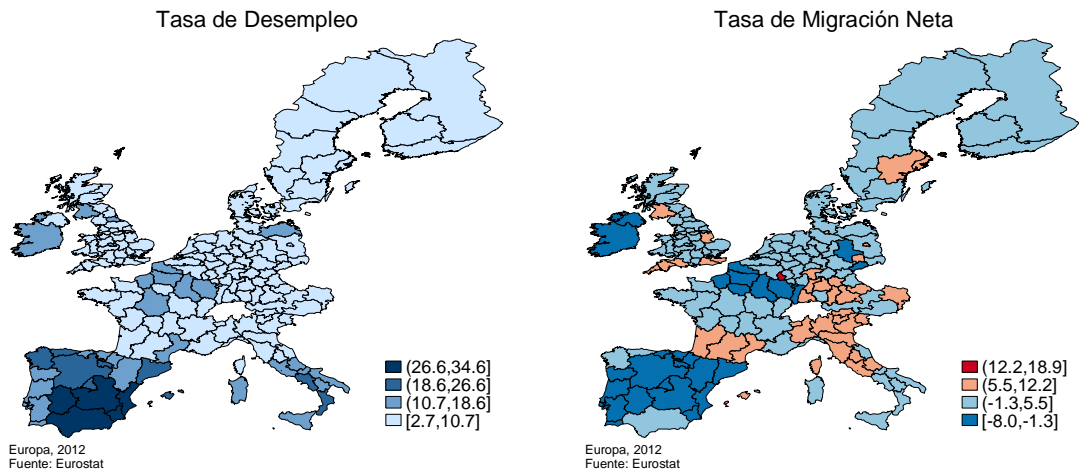
Cada tipo de mapa destaca características diferentes de los datos. A continuación se presentan los mapas de intervalos iguales. Es notorio el cambio en comparación a los mapas de cuantiles. Cuando las variables tienen distribuciones de sus valores igualmente espaciados, o equidistantes, ambos mapas tenderán a asemejarse, pero no es este el caso.

```

. spmap U2012 using coord, id(id) clmethod(e) title("Tasa de Desempleo") ///
  legend(size(medium) position(5)) fcolor(Blues2) note("Europa, 2012" "Fuente: Eurostat")
. spmap NM2012 using coord, id(id) clmethod(e) title("Tasa de Migración Neta") ///
  legend(size(medium) position(5)) fcolor(BuRd) note("Europa, 2012" "Fuente: Eurostat")

```

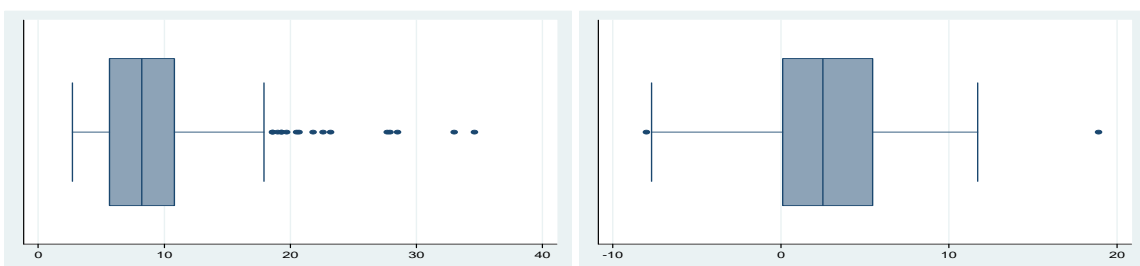
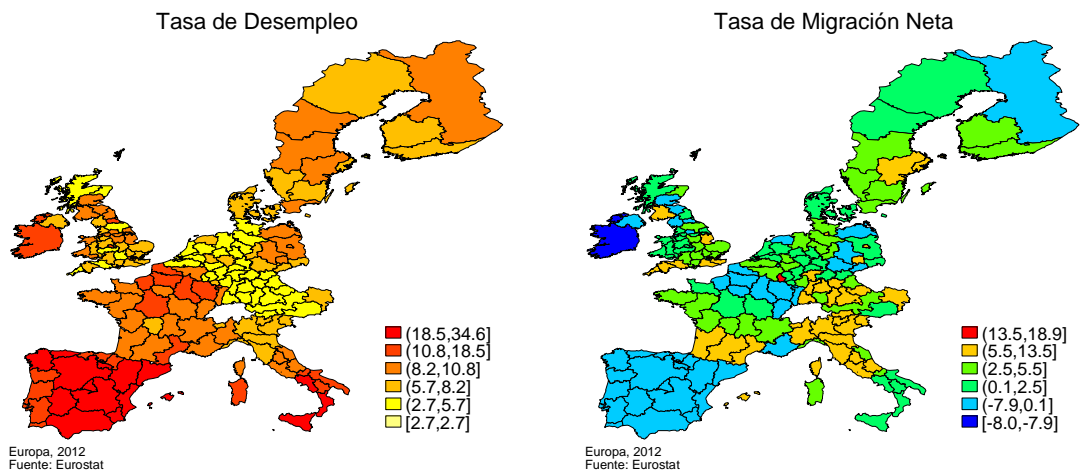
Figura 2: Mapa de Intervalos Iguales



El mapa de diagrama de caja permite visualizar valores extremos o atípicos. Este mapa puede presentarse conjuntamente con el diagrama de caja simple para poder vincular las observaciones:

```
. spmap U2012 using coord, id(id) clmethod(boxplot) title("Tasa de Desempleo") ///
  legend(size(medium) position(5)) fcolor(Heat) note("Europa, 2012" "Fuente: Eurostat")
. spmap NM2012 using coord, id(id) clmethod(boxplot) title("Tasa de Migración Neta") ///
  legend(size(medium) position(5)) fcolor(Rainbow) note("Europa, 2012" "Fuente: Eurostat")
```

Figura 3: Mapa de Diagrama de Caja

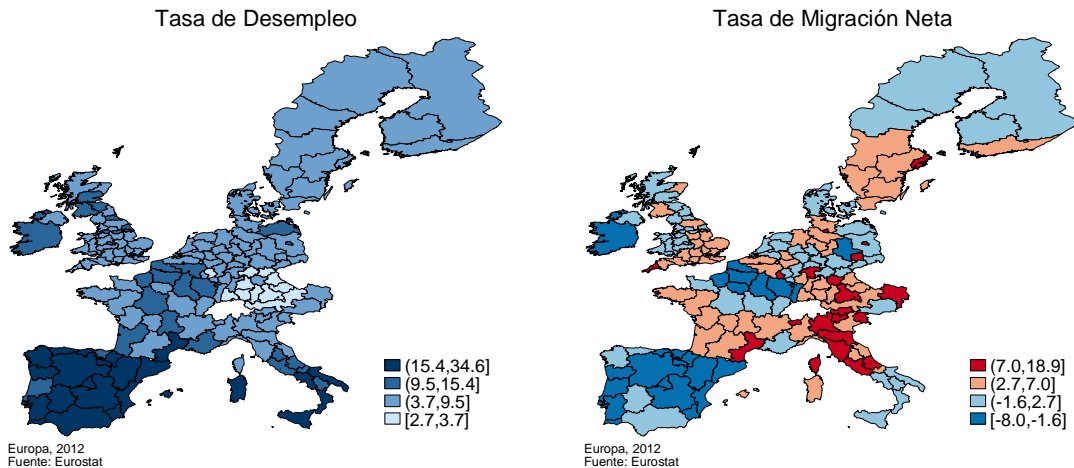


La tasa de desempleo muestra valores extremos altos por encima del 75% de la distribución muestral. Estos valores se encuentran clasificados por la clase de rojo intenso y contiene a la mayoría de las regiones españolas y al sur de Italia. En el caso de la migración neta hay dos observaciones atípicas, una en cada extremo. Se puede distinguir una región con emigración neta extrema (mayor valor negativo) identificada en Irlanda y otra región que experimenta una alta inmigración neta que representa a Luxemburgo.

Otro tipo de mapa es el mapa de desvíos estándares respecto a la media muestral:

```
. spmap U2012 using coord, id(id) clmethod(s) title("Tasa de Desempleo") ///
  legend(size(medium) position(5)) fcolor(Blues2) note("Europa, 2012" "Fuente: Eurostat")
. spmap NM2012 using coord, id(id) clmethod(s) title("Tasa de Migración Neta") ///
  legend(size(medium) position(5)) fcolor(BuRd) note("Europa, 2012" "Fuente: Eurostat")
```

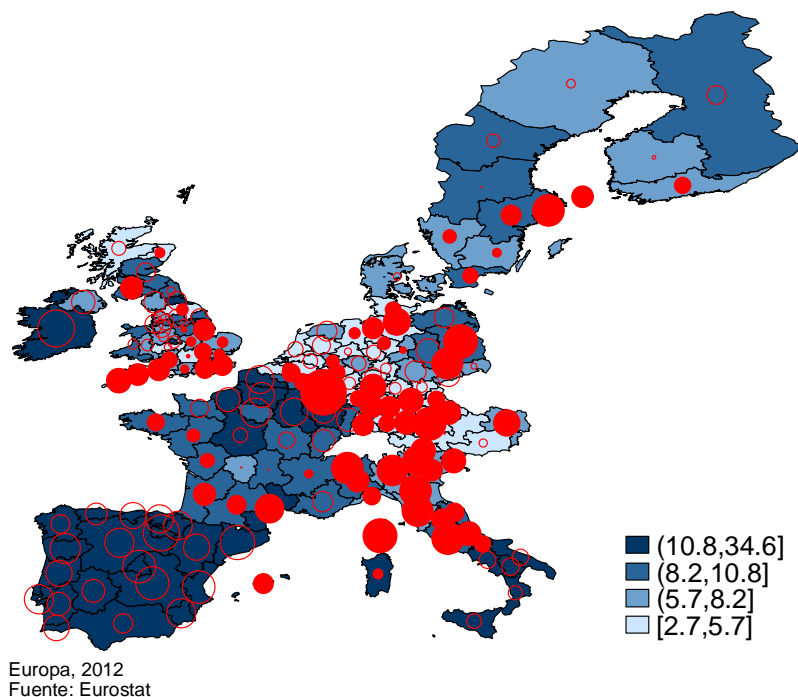
Figura 4: Mapa de Desviaciones



Existen opciones adicionales y la posibilidad de combinar información de áreas con puntos. Por ejemplo, si se desea visualizar en el mapa la relación entre la distribución de cuantiles del desempleo y las desviaciones respecto a la media de la tasa de migración neta, puede aplicarse el siguiente comando:

```
. spmap U2012 using coord, id(id) fcolor(Blues2) point(data(migr_unemp_shp) xcoord(x_c)
  ycoord(y_c) deviation(NM2012) fcolor(red) size(*0.6)) legend(size(medium) position(5))
  note("Europa, 2012" "Fuente: Eurostat")
```

Figura 5: Mapa de Cuantiles del Desempleo y de Desviaciones de Migración Neta



Los círculos huecos representan las desviaciones por debajo de la media de la migración neta y los círculos rellenos representan desviaciones por encima de la media. A mayor tamaño del círculo, más alejado se encuentra el valor de la región respecto a la media europea. Es fácil observar que los círculos huecos y de gran tamaño, que identifican a la emigración neta, se ubican en regiones de alto desempleo. Por lo contrario, los círculos que identifican altos valores de inmigración neta se ubican en áreas de bajo desempleo.

En general puede observarse que ambas variables muestran un patrón de agrupamiento espacial. Para el caso de la tasa de desempleo, los altos valores se encuentran al sur de Europa, en especial en las regiones de España, Portugal y parte de Italia. La distribución de los valores de migración neta refleja que hay zonas de alta inmigración como el norte de Italia, sur de Alemania, sur de Francia y sur del Reino Unido, principalmente. Este análisis visual permite anticipar que el comportamiento de las regiones NUTS 2 no es independiente del espacio.

2.3. Creación de matrices de pesos espaciales

La subsección previa ha brindado los primeros indicios de dependencia o autocorrelación⁴ espacial, sin embargo, no hay una cuantificación de esta autocorrelación y no se conoce si la misma es relevante. Es entonces necesario el uso de herramientas inferenciales para poder obtener una cuantificación de la misma. En este punto surge el problema de la matriz de pesos espaciales.

Siguiendo la información obtenida por el AEDE, se tiene cierta evidencia de que la distribución de la tasa de desempleo y la migración neta depende en cierta medida del valor que asume la variable en las regiones vecinas. Para simplificar la idea, supóngase que la variable de estudio es y ,

⁴En un sentido estricto, dependencia espacial y autocorrelación espacial no son sinónimos. Sin embargo, la literatura ha tendido a utilizar ambos términos de manera intercambiable. Véase Anselin & Bera (1998) para mayor detalle.

distribuida en solo tres regiones y que se ha detectado dependencia espacial. La formalización de esta dependencia puede realizarse de la siguiente manera:

$$\begin{aligned}
y_i &= \alpha_{ij}y_j + \alpha_{ik}y_k + u_i, \\
y_j &= \alpha_{ji}y_i + \alpha_{jk}y_k + u_j, \\
y_k &= \alpha_{ki}y_i + \alpha_{kj}y_j + u_k, \\
u_i; u_j; u_k &\sim i.i.d. (0; \sigma^2),
\end{aligned} \tag{1}$$

es decir, el valor de y que se observa en la región i depende del valor de y en la región j y k , más un término aleatorio que se distribuye de manera idéntica e independiente entre las localizaciones. Lo mismo puede decirse de las otras regiones, j y k . En términos matriciales, la expresión (1) se reduce a:

$$\begin{aligned}
\begin{bmatrix} y_i \\ y_j \\ y_k \end{bmatrix} &= \begin{bmatrix} 0 & \alpha_{ij} & \alpha_{ik} \\ \alpha_{ji} & 0 & \alpha_{jk} \\ \alpha_{ki} & \alpha_{kj} & 0 \end{bmatrix} \begin{bmatrix} y_i \\ y_j \\ y_k \end{bmatrix} + \begin{bmatrix} u_i \\ u_j \\ u_k \end{bmatrix}, \\
y &= Ay + u,
\end{aligned} \tag{2}$$

donde

$$A = \begin{bmatrix} 0 & \alpha_{ij} & \alpha_{ik} \\ \alpha_{ji} & 0 & \alpha_{jk} \\ \alpha_{ki} & \alpha_{kj} & 0 \end{bmatrix}.$$

El problema con el sistema (2) es que, bajo un corte transversal, hay más parámetros que observaciones (3 observaciones y 6 parámetros). Puede imponerse simetría, reduciendo el número de parámetros pero, de todas formas, no podrá estimarse (LeSage & Pace, 2009, p. 8). Por lo tanto, es necesario buscar una solución alternativa para capturar la dependencia subyacente en los datos.

La solución más ampliamente utilizada en econometría espacial es imponer un conjunto de restricciones sobre las relaciones de dependencia. De esta forma, la estructura de la matriz A es reparametrizada de la siguiente manera:

$$A = \rho W,$$

donde ρ es un parámetro a estimar que captura el efecto espacial promedio de los vecinos y W es una matriz de pesos espaciales (también denominada matriz de contigüidades, ponderaciones, distancias o interacciones espaciales). Los elementos de W son:

$$W = \begin{bmatrix} 0 & w_{ij} & w_{ik} \\ w_{ji} & 0 & w_{jk} \\ w_{ki} & w_{kj} & 0 \end{bmatrix},$$

tal que se ha intercambiado a los parámetros α 's (parámetros originales del modelo) por w_{ij} (coeficientes exógenos al modelo).

Generalizando, la matriz W será de orden $n \times n$, siendo n el tamaño muestral. Cada elemento de W es denominado peso espacial, w_{ij} . Los pesos espaciales capturan la vecindad, siendo diferentes de cero cuando las regiones i y j son consideradas vecinas. Por convención, ninguna región puede

ser vecina de si misma, dando como resultado que la diagonal principal de W tiene todos sus elementos iguales a cero, $w_{ii} = 0$.

La matriz de pesos espaciales ocupa una posición central en econometría espacial ya que define el conjunto de vecinos para cada localización y su elección condicionará al resto del análisis. El problema central radica en cómo se construyen los pesos de la matriz W .

En la práctica econométrica, la matriz de contactos es construida mediante diferentes criterios. Estos criterios van desde el uso de la posición geográfica hasta el uso de flujos que capturan interacciones sociales y utilizan otras fuentes de información socio-económica.

En este punto se procederá de la forma tradicional utilizando criterios geográficos, siguiendo la idea de la primera ley de la geografía (Tobler, 1970): “Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes.” Entre los criterios geográficos, puede definirse vecindad por contigüidad, por alguna función de distancia, por k vecinos más cercanos o alguna combinación de estas opciones.

El criterio contigüidad considera como vecinos a los polígonos que comparten límites. Hay varias opciones dentro de este criterio: tipo torre (vecinos al norte-sur y este-oeste), tipo alfil (vecinos al noreste-suroeste y noroeste-sureste) y tipo reina (todos los polígonos que comparten límite). Estos nombres hacen alusión a los movimientos del ajedrez, bajo un mapa con cuadrícula regular.

El criterio distancia utiliza como punto de referencia al centroide de cada polígono, identificado por la latitud y la longitud. Es posible construir diversas medidas, por ejemplo: (1) todas las unidades son vecinas mediante una función de distancia, como la función la inversa: $w_{ij} = f(d_{ij}) = 1/d_{ij}$, tal que a mayor distancia menor relación entre i y j ; o (2) vecinos son aquellas unidades que se encuentran a una distancia inferior a un determinado umbral, definido por un radio d_{ij}^* .

El problema con los criterios de contigüidad y de distancia por umbral es que pueden dar lugar a regiones aisladas que no tienen vecinos. Esto puede suceder debido a que la densidad de los centroides no es regular en el mapa o, bajo contigüidad, cuando se trabaja con mapas que contienen islas.

En el caso de contigüidad, el problema de regiones sin vecinos puede salvarse asignando manualmente vecinos a estas unidades. Bajo el criterio de distancia por umbral, puede determinarse un umbral d cuya amplitud asegure que cada observación contenga al menos un vecino. Este criterio es conocido como max-min, el umbral será la máxima distancia del vecino más cercano de todas las unidades. Un umbral elevado puede ocasionar que existan unidades con una excesiva cantidad de vecinos.

Otro criterio de uso habitual es el de k vecinos más cercanos. Utilizando la información de los centroides, se va eligiendo como vecino al centroide más cercano hasta obtener el número de vecinos establecidos, k . Bajo este criterio todas las regiones poseerán la misma cantidad de vecinos evitando el problema de unidades aisladas o unidades con excesiva cantidad de vecinos.

Una vez definido el criterio de vecindad, se puede representar al conjunto de vecinos mediante una elección binaria, con $w_{i,j} = 1$ cuando i y j son vecinos, y $w_{i,j} = 0$ cuando no lo son. Si el criterio utilizado es el de distancia los pesos binarios pueden reemplazarse por alguna función de distancia entre unidades espaciales (pueden combinar la distancia, el perímetro u otras características geográficas).

Una vez elegidos los pesos espaciales, lo habitual es trabajar con alguna transformación de la matriz debido a que mejora las propiedades estadísticas de los estimadores y los contrastes. La

transformación más utilizada es la normalización por fila, en donde los nuevos pesos se obtienen como $w_{ij}^* = w_{ij} / \sum_j w_{ij}$, tal que la suma de cada fila de W será sea igual a la unidad: $\sum_j w_{ij}^* = 1$.

Además, es posible considerar el orden de vecindad. Por ejemplo, puede seleccionarse para cada región a los vecinos y a los vecinos de los vecinos. En el caso del criterio por umbral, estos serán las regiones que se encuentran a una distancia mayor al umbral pero inferior al doble del umbral. No es de uso común utilizar ordenes superiores a 1, pero si fuese el caso, se suele distinguir mediante un supra-índice en la matriz: $W^{(j)}$, $\forall j \geq 2$, siendo j el orden de vecindad.

Por lo general la elección del criterio de vecindad es un a priori del investigador aunque, en algunos casos, puede provenir de modelizaciones teóricas. Sin importar cuál sea el caso, el desarrollo econométrico asume que W contiene elementos exógenos a la especificación econométrica y la elección de los vecinos no debería realizarse mediante variables consideradas en el modelo analizado, ya sean endógenas o explicativas. Al respecto, Griffith (1996) establece un conjunto de lineamientos útiles para la elección de la matriz: (i) Es mejor utilizar una especificación razonable de la matriz de pesos geográficos que considerar todas las conexiones nulas; (ii) modelos con bajos ordenes deben ser preferidos por sobre los modelos de altos ordenes; (iii) en general, es mejor emplear una matriz de pesos subidentificada que una sobreidentificada.

En Stata, las matrices espaciales pueden generarse mediante diferentes comandos. Uno de ellos es `spmat` (Drukker et al., 2013b) que permite crear, importar, manipular y guardar matrices W . Las matrices son almacenadas como objetos `spmat`, estructura requerida para el uso de otros comandos como `spreg` (Drukker et al., 2013d) y `spivreg` (Drukker et al., 2013c).

El comando `spatwmat` (Pisati, 2001) está integrado a las herramientas de análisis exploratorio que incluye los contrastes de dependencia espacial global y, además, al análisis de dependencia local. Si la matriz ha sido generada por `spmat` no podrá ser usada directamente para los comandos del análisis exploratorio de Pisati.

Una tercera opción para matrices espaciales es `spwmatrix` de Jeanty (2014). Este comando permite generar matrices similares a `spmat` pero además genera matrices de k vecinos más cercanos y matrices con criterios socio-económicos. Tanto `spwmatrix` como `spmat` permiten importar matrices creadas por GeoDa, programa que genera de forma muy simple las matrices de criterios geográficos. Además, las matrices generadas por `spwmatrix` pueden ser directamente utilizadas en los comandos de Pisati.

El comando `spmat` ofrece varias ventajas. Entre ellas, permite la visualización de las matrices así como alternativas avanzadas de especificación. Los detalles de este comando no se verán aquí, pero se encuentran debidamente explicados por Drukker et al. (2013b).

Veamos cómo trabajan estos comandos bajo Stata. Por ejemplo, puede intentarse generar una matriz de contigüidad tipo reina:

```
. spmat contiguity Wcont using "coord.dta", id(id)
warning: spatial-weighting matrix contains 5 islands
```

Tal como pudo visualizarse, el mapa de Europa contiene 5 islas y las mismas no tendrán vecinos mediante el criterio de contigüidad. El problema que plantea la falta de vecinos es que se asume que las unidades se comportan de forma independiente a las demás. Además, las observaciones que no tienen vecinos no serán incluidas en el cómputo de los contrastes de autocorrelación espacial (más detalle en sección 2.4). El otro problema, algo más técnico, es que no podrán calcularse, para la matriz W , los autovalores necesarios para estimar modelos espaciales (esto se verá en la próxima sección).

El criterio alternativo que se seguirá es el de k vecinos más cercanos. Este criterio es aplicado mediante el comando `spwmatrix`:

```
. spwmatrix gecon y_c x_c, wn(W5st) knn(5) row
```

```
Nearest neighbor (knn = 5) spatial weights matrix (164 x 164) calculated successfully
and the following action(s) taken:
```

- Spatial weights matrix created as Stata object(s): W5st.
- Spatial weights matrix has been row-standardized.

Se ha generado una matriz de 5 vecinos más cercanos, estandarizada por fila. Dado que esta matriz es creada como un objeto de Stata, no podrá ser utilizada directamente bajo el comando `spmat`. Para solucionar este problema, puede generarse una matriz no estandarizada y exportarla a un archivo `.txt` con la idea de manipular su estructura. La secuencia de comandos es la siguiente:

```
. spwmatrix gecon y_c x_c, wn(W5bin) knn(5) xport(W5bin,txt)
```

```
Nearest neighbor (knn = 5) spatial weights matrix (164 x 164) calculated successfully
and the following action(s) taken:
```

- Spatial weights matrix created as Stata object(s): W5bin.
- Spatial weights matrix saved to `.txt` file, `C:\..\W5bin.txt`, for use with other Stata packages.

Ahora, ya generado el archivo `.txt`, puede leerse como una base de datos de Stata y modificar el archivo para que sea legible en `spmat`:

```
. insheet using "W5bin.txt", delim(" ") clear
(165 vars, 165 obs)
. drop in 1
(1 observation deleted)
. rename v1 id
. save "W5bin.dta"
file W5bin.dta saved
. spmat dta W5_st v*, id(id) norm(row)
```

La matriz binaria es estandarizada por fila y guardada como objeto `spmat`. Esto permite aprovechar algunos subcomandos como `summarize` y `graph`:

```
. spmat summarize W5_st, links
```

```
Summary of spatial-weighting object W5_st
```

```
-----
```

Matrix	Description
Dimensions	164 x 164
Stored as	164 x 164

```
-----
```

```
Links
```

```

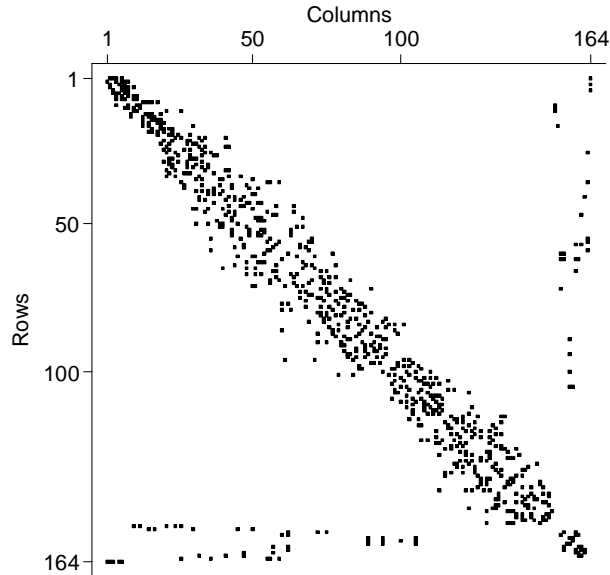
total |           820
  min |           5
  mean |           5
  max |           5

```

```

-----
. spmat graph W5_st

```



La información brindada por estos comandos permite chequear que todas las regiones contienen 5 vecinos, el número total de links es de 820, cerca del 3.07% de todos los posibles links. El gráfico muestra la vinculación de las regiones considerando la identificación, `id`.

2.4. Detección de autocorrelación espacial

Una vez definida la matriz de contactos, pueden utilizarse diferentes estadísticos univariantes que permiten detectar autocorrelación espacial. Los mismos pueden clasificarse como medidas de dependencia globales o locales.

Las medidas globales utilizan la información completa del conjunto de datos con el fin de obtener un valor promedio para todo el espacio geográfico. Al resumir en un único valor toda la información, no es posible detectar la variabilidad de la dependencia ni la localización de estos patrones. Por su parte, las medidas locales examinan la autocorrelación espacial en un subconjunto de datos. En este escrito nos focalizaremos en medidas globales, dejando para próximos avances las medidas locales.

2.4.1. Estadísticos globales

Posiblemente el contraste más ampliamente utilizado sea el estadístico I de Moran (1950):

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j (y_i - \bar{y}) w_{ji} (y_j - \bar{y})}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (3)$$

donde $S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij} = 1'W1$, siendo 1 un vector $(n \times 1)$ de unos, con \bar{y} siendo la media muestral. Si la matriz W contiene islas, las filas de estas observaciones serán nulas y el elemento S_0 será incorrectamente calculado.

Cuando el contraste I de Moran toma un valor positivo existe autocorrelación positiva implicando que los valores de cada observación y sus vecinos se asemejan. Si el I asume un valor negativo entonces esto implica autocorrelación negativa tal que el valor de los vecinos son altos cuando la observación tiene un valor bajo y si es alto entonces sus vecinos asumen valores bajos.

Los momentos del estadístico de Moran, asumiendo que y se distribuye normal, son:

$$E[I] = -\frac{1}{n-1}, \quad (4)$$

$$V[I] = \frac{(3S_0^2 + S_1n^2 - nS_2)}{S_0(n+1)(n-1)} - \frac{1}{(n-1)^2}, \quad (5)$$

con $S_1 = (1/2) \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2$ y $S_2 = \sum_{i=1}^n \left(\sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji} \right)^2$. Es decir, bajo normalidad, los momentos del estadístico están determinados únicamente por elementos de la matriz espacial.

Su distribución asintótica es normal:

$$\sqrt{n}[I - E(I)] \underset{as}{\sim} \mathcal{N}[0; V(I)]. \quad (6)$$

Otro contraste de uso habitual es el estadístico c de Geary (1954):

$$c = \frac{n-1}{2S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - y_j)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (7)$$

donde S_0 y \bar{y} ya han sido definidos en la I de Moran.

Los momentos del estadístico de Geary, bajo la hipótesis nula de aleatoriedad, son:

$$E[c] = 1, \quad (8)$$

$$V[c] = \frac{(2S_1 + S_2)(n-1) - 4S_0^2}{2(n+1)S_0^2}, \quad (9)$$

donde cada elemento ha sido definido para el contraste I de Moran.

Su distribución asintótica es normal:

$$\sqrt{n}[c - 1] \underset{as}{\sim} \mathcal{N}[0; V(c)]. \quad (10)$$

Para el estadístico c , un valor negativo (es decir, por debajo del valor esperado) implica que existe autocorrelación positiva y un valor de c positivo, por encima del valor esperado, implica autocorrelación negativa.

Por último, otro estadístico de utilidad es el G de Getis & Ord (1992):

$$G = \frac{\sum_i^n \sum_{j \neq i}^n w_{ij} y_i y_j}{\sum_i^n \sum_{j \neq i}^n y_i y_j}, \quad (11)$$

donde cada elemento ya ha sido definido previamente.

Bajo permutación aleatoria, el valor esperado de G es:

$$E[G] = \frac{\sum_i^n \sum_{j \neq i}^n w_{ij}}{n(n-1)}.$$

La particularidad de este contraste es que necesita que los pesos w_{ij} sean de tipo binario. Getis & Ord (1992) muestran que la distribución de G tiende a la normal a medida que el tamaño muestral crece. De esta forma, puede utilizarse la aproximación asintótica de la distribución:

$$z[G] = \frac{G - E(G)}{\sqrt{V(G)}} \underset{as.}{\sim} \mathcal{N}(0, 1). \quad (12)$$

En Stata, estos estadísticos pueden ser calculados por el comando `spatgsa` de Pisati (2001).

```
. spatgsa U2012, w(W5st) moran geary two

Measures of global spatial autocorrelation
Weights matrix
-----
Name: W5st
Type: Distance-based (binary)
Distance band: c1.c2 < d <= c3.c4
Row-standardized: Yes
-----

Moran's I
-----

```

Variables	I	E(I)	sd(I)	z	p-value*
U2012	0.767	-0.006	0.045	17.084	0.000

```
-----
Geary's c
-----

```

Variables	c	E(c)	sd(c)	z	p-value*
U2012	0.228	1.000	0.054	-14.282	0.000

```
-----
*2-tail test
```

Ambos estadísticos, el I de Moran como el c de Geary, detectan autocorrelación espacial bajo la matriz de contactos de 5 vecinos más cercanos para la tasa de desempleo de 2012. Lo mismo puede decirse sobre la tasa de migración neta.

```
. spatgsa NM2012, w(W5st) moran geary two
```

Measures of global spatial autocorrelation

Weights matrix

Name: W5st

Type: Distance-based (binary)

Distance band: c1.c2 < d <= c3.c4

Row-standardized: Yes

Moran's I

Variables	I	E(I)	sd(I)	z	p-value*
NM2012	0.431	-0.006	0.046	9.563	0.000

Geary's c

Variables	c	E(c)	sd(c)	z	p-value*
NM2012	0.559	1.000	0.050	-8.840	0.000

*2-tail test

Para obtener los resultados del estadístico de Getis y Ord se necesita la matriz binaria, por lo que se genera mediante `spwmatrix` y posteriormente se aplica el comando para ambas variables.

```
. spwmatrix gecon y_c x_c, wn(W5bin) knn(5)
```

Nearest neighbor (knn = 5) spatial weights matrix (164 x 164) calculated successfully and the following action(s) taken:

- Spatial weights matrix created as Stata object(s): W5bin.

```
. spatgsa U2012, w(W5bin) go two
```

Measures of global spatial autocorrelation

Weights matrix

Name: W5bin

Type: Distance-based (binary)

Distance band: c1.c2 < d <= c3.c4

Row-standardized: No

Getis & Ord's G

Variables	G	E(G)	sd(G)	z	p-value*
U2012	0.039	0.031	0.001	11.864	0.000

*2-tail test

```

. spatgsa NM2012, w(W5bin) go two
Measures of global spatial autocorrelation
Weights matrix
-----
Name: W5bin
Type: Distance-based (binary)
Distance band: c1.c2 < d <= c3.c4
Row-standardized: No
-----
Getis & Ord's G
-----
Variables |      G      E(G)   sd(G)    z    p-value*
-----
  NM2012 |  0.066   0.031   0.004   9.441   0.000
-----
*2-tail test

```

Alternativamente, el estadístico de Moran puede calcularse por el comando `splagvar` (Jeanty, 2012). Este comando ofrece la versión aleatorizada como alternativa a la aproximación asintótica, versión similar a la ofrecida por GeoDa. Otra ventaja de este comando es la posibilidad de generar el diagrama de dispersión de Moran, habitual en investigaciones aplicadas.

El diagrama de dispersión de Moran muestra la relación entre la variable y su rezago espacial, Wy . En vez de utilizar la variable original, es preferible trabajar con desviaciones respecto a la media ya que simplifica las fórmulas del contraste de Moran:

$$z_i = (y_i - \bar{y}),$$

donde y_i , en este caso es: `U2012` o `NM2012`.

La relación entre la variable estandarizada y el rezago espacial de dicha variable puede representarse mediante una regresión lineal simple, sin constante:

$$Wz_i = \beta z_i + u_i,$$

donde u_i es un término de error que cumple con los supuestos habituales. El coeficiente estimado $\hat{\beta}$ es igual al valor I de Moran:

$$\hat{\beta} = (X'X)^{-1} X'y = \frac{z'Wz}{z'z} = \hat{I}.$$

Los resultados para ambas variables se obtienen aplicando `splagvar`, y utilizando el subcomando `plot`:

```

. splagvar U2012, wname(W5st) wfrom(Stata) ind(U2012) order(1) plot(U2012) moran(U2012)
(permute U2012 : splagvar_randper)
Moran's I Statistics Under Normal Approximation and Randomization Assumptions
-----
Statistics      | Normal Approximation   Randomization
-----

```

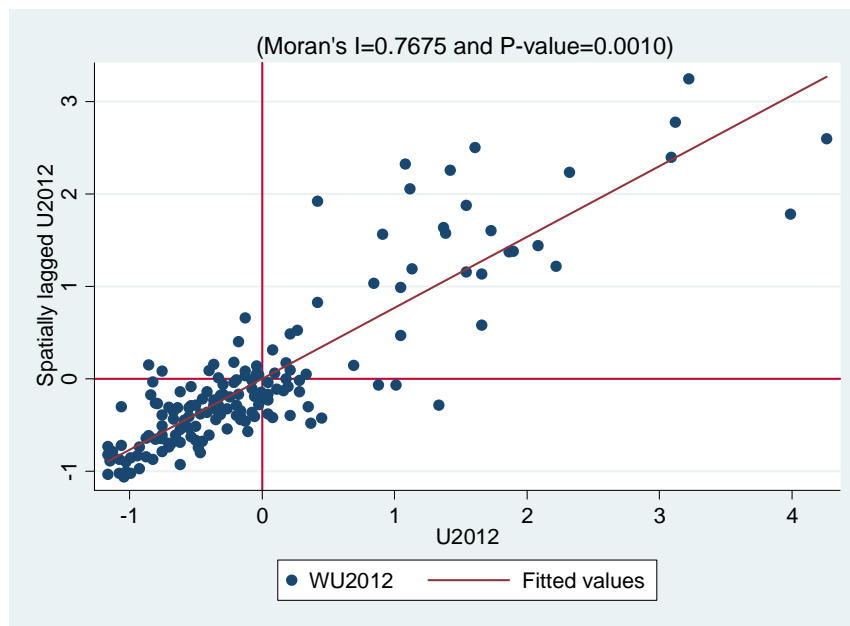
Moran's I		0.7675		0.7675
Mean		-0.0061		-0.0061
Std dev		0.0458		0.0453
Z-score		16.8866		17.0840
P-value*		0.0000		0.0000

*: Two-tailed test

Note: Under the random permutation procedure:

Mean = -0.0046 and Standard deviation = 0.0480

Spatially lagged variable(s) calculated successfully and/or all requests processed.



El diagrama de Moran divide al gráfico en cuatro cuadrantes respecto al valor promedio: valores altos de desempleo con valores altos de sus vecinos (Alto-Alto), valores bajos de desempleo con valores bajos de sus vecinos (Bajo-Bajo), valores altos de desempleo con valores bajos de sus vecinos (Alto-Bajo), y valores bajos del desempleo con valores altos de los vecinos (Bajo-Alto).

Los resultados de la tasa migración neta se muestran a continuación:

```
. splayvar NM2012, wname(W5st) wfrom(Stata) ind(NM2012) order(1) plot(NM2012) moran(NM2012)
(permute NM2012 : splayvar_randper)
```

Moran's I Statistics Under Normal Approximation and Randomization Assumptions

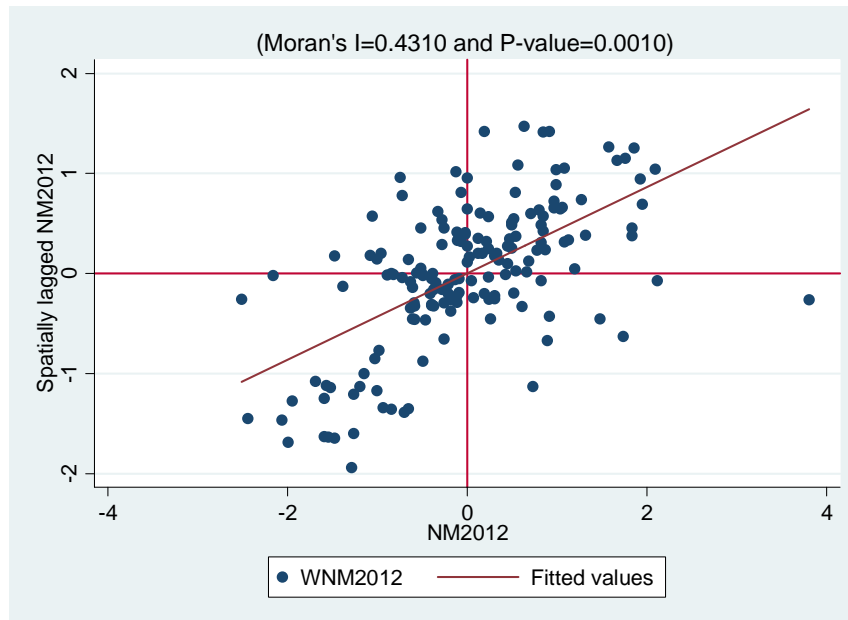
Statistics		Normal Approximation		Randomization
Moran's I		0.4310		0.4310
Mean		-0.0061		-0.0061
Std dev		0.0458		0.0457
Z-score		9.5417		9.5634
P-value*		0.0000		0.0000

*: Two-tailed test

Note: Under the random permutation procedure:

Mean = -0.0067 and Standard deviation = 0.0466

Spatially lagged variable(s) calculated successfully and/or all requests processed.



El rechazo de la hipótesis nula de aleatoriedad de los estadísticos de autocorrelación espacial permite fundamentar un análisis más detallado sobre los determinantes de las variables bajo estudio. En la próxima sección el interés se centrará en incorporar la dependencia espacial global a los modelos de regresión. Es decir, detectada la importancia del espacio sobre las variables, se introducirán efectos espaciales explícitos en la especificación del modelo econométrico.

3. Modelos de Econometría Espacial

En esta sección se desarrollan los principales modelos espaciales bajo un corte transversal. Se comenzará con un modelo simple de regresión general múltiple y se incorporará estructura espacial, por medio de la dirección de los diferentes contrastes, hasta alcanzar los modelos más complejos.

Una clara diferencia respecto a lo expuesto en la sección 2 radica en que el establecimiento de un modelo de regresión lineal supone un tratamiento asimétrico de las variables bajo estudio. El modelo de regresión asume que existe una variable dependiente que es de principal interés y que existe un conjunto de variables explicativas que condicionan su comportamiento. Siguiendo con el ejemplo desempleo-migración, se utilizará como variable dependiente a la tasa de desempleo y como variable explicativa a la tasa de migración neta.

Hasta el momento se ha podido detectar que ambas variables poseen estructura espacial y que la misma debería considerarse de manera explícita. La detección de autocorrelación espacial puede ser real, debido a la propia estructura de difusión de la variable, o puede ser aparente, debido a la existencia de otras variables que pueden explicar la dependencia espacial detectada. Por lo tanto, incorporar variables explicativas puede generar que se reduzca o aún se anule la dependencia espacial de la variable dependiente. A continuación se desarrollan los pasos para la detección y especificación del modelo econométrico.

3.1. Detección de estructura espacial bajo MCO

El modelo estático más simple considera que existe una variable dependiente y un conjunto de variables explicativas. La ecuación de trabajo es la siguiente:

$$\begin{aligned} y &= X\beta + u, \\ u &\sim (0, \sigma^2 I_n), \end{aligned} \tag{13}$$

siendo la variable dependiente y un vector de dimensión $(n \times 1)$, X es una matriz de variables explicativas, incluyendo una constante, de orden $(n \times k)$, β es un vector de parámetros desconocidos de orden $(k \times 1)$ y u es el término de error de dimensión $(n \times 1)$. Se asume que el modelo cumple con los supuestos habituales de la parte sistemática: forma funcional correcta, variables explicativas exógenas, no multicolinealidad.

La ecuación (13) es la hipótesis nula de los estadísticos de dependencia espacial. Típicamente, el término de error es asumido independiente e idénticamente distribuido, *i.i.d.*, pero claramente este no será el caso bajo la hipótesis alternativa. El objetivo es encontrar evidencia tal que permita mantener la hipótesis nula de no autocorrelación espacial. En caso contrario, se incorporarán elementos espaciales al modelo.

La presencia de estructura espacial en el modelo (13) puede contrastarse en base a estadísticos simples que utilizan los resultados de la estimación por mínimo cuadrados ordinarios (*MCO*). El primer contraste que puede utilizarse es el test I de Moran aplicado sobre los residuos *MCO* del modelo (13), sugerido por Cliff & Ord (1972):

$$I = \frac{n}{S_0} \frac{\hat{u}' W \hat{u}}{\hat{u}' \hat{u}}, \tag{14}$$

donde \hat{u} es el vector de residuos *MCO*, n es el número de observaciones y S_0 es la suma de todos los elementos de W , tal como se ha definido en la sección previa. Bajo una matriz estandarizada por fila, el término $n/S_0 = 1$ y puede ser ignorado.

La distribución probabilística del estadístico I de Moran es desconocida para muestras finitas por lo que comúnmente se utiliza una aproximación empírica por permutación. Otra opción es plantear la inferencia estadística bajo el supuesto de normalidad asintótica.

Recuérdese que la hipótesis nula del contraste es no autocorrelación espacial. El problema con el test I de Moran es que el rechazo de la hipótesis nula no brinda información sobre el posible modelo a especificar. La hipótesis alternativa es general y no da una guía clara sobre el tipo de estructura espacial que se encuentra en el proceso generador de datos.

Como alternativa al contraste I de Moran, existe un conjunto de contrastes de Multiplicadores de Lagrange, LM , que resultan de la aplicación del principio de máxima verosimilitud. Estos contrastes tienen la ventaja de que la hipótesis alternativa se encuentra bien definida o restringida.

Una primera hipótesis alternativa proviene de plantear la presencia de autocorrelación espacial en el término de error. Esta estructura de dependencia puede deberse a un proceso autoregresivo o a uno de medias móviles. Sin pérdida de generalidad, se concentrará el análisis en la estructura autoregresiva dado que ante ambas alternativas el estadístico es el mismo (Burrige, 1980). Es así que un modelo de error espacial autoregresivo asume que el término de error aleatorio de (13) obedece a un proceso como:

$$u = \rho W u + \varepsilon, \quad (15)$$

donde ρ es el parámetro espacial autoregresivo, W es una matriz de pesos espaciales no estocástica de orden $(n \times n)$ y ε es un vector de innovaciones con media nula y varianza constante $(0, \sigma^2 I)$.

El contraste LM para detectar autocorrelación espacial en el error establece las siguientes hipótesis nula y alternativa:

$$\begin{aligned} H_0 : \quad & \rho = 0, \\ H_1 : \quad & \rho \neq 0, \end{aligned}$$

siendo su fórmula:

$$LM_{ERROR} = \frac{1}{T_1} \left(\frac{\hat{u}' W \hat{u}}{\hat{\sigma}^2} \right)^2 \underset{as}{\sim} \chi_{(1)}^2, \quad (16)$$

donde T_1 es igual a $\text{tr} \left[(W' + W) W \right]$ y \hat{u} son los residuos MCO y $\hat{\sigma}^2 = \hat{u}' \hat{u} / n$.

Una segunda hipótesis alternativa proviene de plantear un modelo con estructura espacial sustantiva, modelo de rezago espacial (SLM , spatial lag model). Este modelo incorpora un rezago espacial de la variable dependiente, Wy , como explicativa:

$$y = \lambda W y + X \beta + u, \quad (17)$$

donde λ es un parámetro espacial autoregresivo, u es un vector de errores de dimensión $(n \times 1)$ y $u \sim (0, \sigma^2 I_n)$, con los demás elementos ya definidos.

El contraste para la detección de dependencia sustantiva plantea como hipótesis nula y alternativa:

$$\begin{aligned} H_0 : \quad & \lambda = 0, \\ H_1 : \quad & \lambda \neq 0, \end{aligned}$$

tal que el contraste tiene la siguiente estructura:

$$LM_{LAG} = \frac{\left(\frac{\hat{u}' W y}{\hat{\sigma}^2} \right)^2}{n \hat{J}_{\lambda\beta}} \underset{as}{\sim} \chi_{(1)}^2, \quad (18)$$

con $\hat{J}_{\lambda\beta} = 1/n \hat{\sigma}^2 \left[(W X \hat{\beta})' M (W X \hat{\beta}) + T_1 \hat{\sigma}^2 \right]$, con $M = I - X (X' X)^{-1} X'$.

Estos contrastes LM presentan como inconveniente que son sensibles a diferentes tipos de errores de especificación. Por ejemplo, el LM_{ERROR} detecta autocorrelación espacial debido a la presencia de un rezago espacial de la variable endógena (Wy), y lo mismo puede decirse del LM_{LAG} , que brinda falsos positivos cuando el término de error contiene un rezago espacial (Wu).

Ante estos inconvenientes, Anselin et al. (1996) proponen dos nuevos multiplicadores de Lagrange diseñados para que su comportamiento sea robusto a dichos errores de especificación. El LM_{ERROR}^* analiza la falta de correlación en los residuos, siendo robusto a la omisión del término Wy :

$$LM_{ERROR}^* = \frac{\left[\left(\frac{\hat{u}' W \hat{u}}{\hat{\sigma}^2} \right) - T_1 \left(n \hat{J}_{\lambda\beta} \right)^{-1} \left(\frac{\hat{u}' W y}{\hat{\sigma}^2} \right) \right]^2}{\left[T_1 - T_1^2 \left(n \hat{J}_{\lambda\beta} \right)^{-1} \right]} \underset{as}{\sim} \chi_{(1)}^2. \quad (19)$$

Alternativamente, la versión robusta del contraste LM_{LAG} , LM_{LAG}^* , permite detectar la autocorrelación espacial sustantiva en presencia de estructura espacial en el término de error:

$$LM_{LAG}^* = \frac{\left[\left(\frac{\widehat{u}W\widehat{y}}{\widehat{\sigma}^2} \right) - \left(\frac{\widehat{u}W\widehat{u}}{\widehat{\sigma}^2} \right) \right]^2}{n\hat{J}_{\lambda\beta} - T_1} \underset{as}{\sim} \chi_{(1)}^2. \quad (20)$$

Para mayor detalle de la derivación de estos contrastes puede consultarse a Anselin & Florax (1995, p. 23-26). El conjunto de contrastes habituales para la detección de autocorrelación espacial se resume en el Cuadro 2.

Estos contrastes permiten incorporar elementos espaciales de acuerdo al rechazo o no de cada una de las hipótesis nulas. Tal es así que puede establecerse una estrategia de especificación:

- Si I de Moran rechaza $H_0 \rightarrow$ Evidencia a favor de inclusión de elementos espaciales.
- Si LM_{ERROR} y LM_{ERROR}^* rechazan $H_0 \rightarrow$ Evidencia a favor de un modelo de error espacial (SEM).
- Si LM_{LAG} y LM_{LAG}^* rechazan $H_0 \rightarrow$ Evidencia a favor de un modelo de rezago espacial (SLM).
- Si no se rechaza H_0 bajo ninguno de los contrastes \rightarrow Evidencia a favor del modelo lineal general no espacial.
- Si ambos contrastes robustos, LM_{ERROR}^* y LM_{LAG}^* , rechazan H_0 entonces se deberán incorporar elementos espaciales en la parte sistemática (Wy) y aleatoria (Wu).

Utilizando estos contrastes conjuntamente puede especificarse el modelo más adecuado.

Cuadro 2: Contrastes espaciales

Hipótesis Nula, H_0	Parámetros en H_1		Contraste
	Rezago espacial	Error espacial	
$\rho = 0$	-	si	LM_{ERROR}
	si	si	LM_{ERROR}^*
$\lambda = 0$	si	-	LM_{LAG}
	si	si	LM_{LAG}^*
No autocorrelación espacial			I de Moran

Ahora, se verá cómo obtener estos resultados en Stata. El modelo (13) se puede estimar por el comando `reg` que implementa la estimación por mínimos cuadrados ordinarios.

```
. use migr_unemp_shp, clear
. reg U2012 NM2012
```

Source	SS	df	MS	Number of obs =	164
-----+-----					
Model	1453.94714	1	1453.94714	F(1, 162) =	56.32
Residual	4182.20231	162	25.8160636	Prob > F =	0.0000
-----+-----					
Total	5636.14945	163	34.577604	R-squared =	0.2580
				Adj R-squared =	0.2534
				Root MSE =	5.081

U2012	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
NM2012	-.7011928	.0934347	-7.50	0.000	-.8856998	-.5166859
_cons	11.43504	.4697136	24.34	0.000	10.50749	12.36259

Una vez estimado el modelo, puede realizarse el diagnóstico de los residuos mediante el comando `spatdiag` de Pisati (2001). La implementación del comando requiere una matriz de contactos. Siguiendo con lo elaborado en la sección previa, se utilizará la matriz de pesos de 5 vecinos más cercanos estandarizada por filas:

```
. spwmatrix gecon y_c x_c, wn(W5st) knn(5) row
. spatdiag, weights(W5st)
```

Diagnostic tests for spatial dependence in OLS regression

Fitted model

U2012 = NM2012

Weights matrix

Name: W5st

Type: Distance-based (binary)

Distance band: c1.c2 < d <= c3.c4

Row-standardized: Yes

Diagnostics

Test	Statistic	df	p-value
Spatial error:			
Moran's I	12.703	1	0.000
Lagrange multiplier	148.081	1	0.000
Robust Lagrange multiplier	0.750	1	0.386
Spatial lag:			
Lagrange multiplier	182.220	1	0.000
Robust Lagrange multiplier	34.889	1	0.000

Los resultados reportados muestran que existe autocorrelación espacial en los residuos según el contraste I de Moran, LM_{ERROR} y el LM_{LAG} . Ambas versiones del contraste de dependencia sustantiva detectan estructura espacial. No sucede lo mismo con los LM del error, ya que el LM_{ERROR}^* no rechaza la hipótesis de $\rho = 0$. La lectura conjunta de estos resultados implica que debe estimarse un modelo con autorrelación espacial sustantiva, modelo SLM .

Obsérvese que el comando de Stata ubica al contraste I de Moran como un estadístico de error espacial. Esto no implica que un valor significativo del estadístico brinde evidencia de estructura

espacial en el error. Tal como ya se ha mencionado, la hipótesis alternativa del I de Moran no se encuentra restringida y un valor significativo del contraste puede sugerir una o ambas estructuras espaciales.

Una vez detectada la posible estructural espacial, surge el problema de cómo estimar este tipo de modelos. Lamentablemente, la estimación del modelo SLM o SEM no puede realizarse por MCO debido a que la estimación será inconsistente y/o ineficiente, dependiendo del caso (Anselin, 1988, p. 57-59).

Las alternativas de estimación han sido tradicionalmente de dos tipos. Una alternativa es por máxima verosimilitud, MV , asumiendo distribución normal de la perturbación aleatoria. La otra alternativa es por variables instrumentales o, su versión generalizada, método de momentos generalizado, GMM , que recurre a la teoría asintótica sin necesidad del supuesto de normalidad.

Además del modelo SLM y el SEM , existen otros modelos espaciales que son de uso habitual. A continuación se desarrollarán con mayor detalle para que el lector pueda conocer y estimar cada uno de ellos en Stata.

3.2. Estimación por máxima verosimilitud

La estimación por máxima verosimilitud ha sido extensamente tratada. Aquí solo se presentarán los elementos más importantes de dicho procedimiento para un conjunto de modelos espaciales. Para una mayor revisión de este método puede consultarse a LeSage & Pace (2009, capítulo 3).

3.2.1. Modelo de rezago espacial

El modelo de rezago espacial o SLM posee la siguiente estructura:

$$\begin{aligned} y &= \lambda W y + X\beta + u, \\ u &\sim \mathcal{N}(0, \sigma^2 I_n), \end{aligned} \tag{21}$$

donde el supuesto de normalidad del término de error permite plantear la habitual función de log-verosimilitud no-concentrada, esto es:

$$l(y|\varphi) = -\frac{n}{2} \ln(2\pi\sigma^2) + \ln|A| - \frac{1}{2\sigma^2} \left[(Ay - X\beta)' (Ay - X\beta) \right], \tag{22}$$

siendo $A = I - \lambda W$.

Existen importantes detalles a destacar en la maximización de (22). Uno de ellos es la necesidad de imponer alguna restricción sobre el rango de valores que puede asumir λ . Sin restricciones sobre este parámetro, el espacio paramétrico de la función será discontinuo e inestable. Para evitar estos problemas y también facilitar la estimación de λ , Ord (1981) propuso restringir los posibles valores del parámetro en un rango $1/\tau_{min} < \lambda < 1/\tau_{max}$, donde τ_{min} y τ_{max} son los autovalores, mínimo y máximo, de la matriz W . Bajo una matriz espacial estandarizada por filas automáticamente se impone una restricción sobre el máximo autovalor tal que λ se encontrará restringido entre los valores $1/\tau_{min} < \lambda < 1$. Algunos autores sugieren limitar el espacio paramétrico de λ al intervalo $(-1, 1)$. Para mayor revisión puede consultarse a Elhorst (2014, p. 13-17) y Kelejian & Prucha (2010).

Otro detalle de importancia es que la maximización de (22) involucra el determinante de una matriz de dimensión $n \times n$, $\ln|I - \lambda W|$. Ord (1975) sugiere estimar dicho término mediante: $\ln|I - \lambda W| = \sum_{i=1}^N \ln(1 - \lambda\tau_i)$ siendo τ_i los autovalores de la matriz W .

Para maximizar (22) se necesitan obtener las condiciones necesarias de primer orden. Estas condiciones conforman un sistema no lineal, imposibilitando la obtención de una solución analítica. La estimación de este sistema puede realizarse por medio de algoritmos numéricos o por el procedimiento basado en la log-verosimilitud concentrada. La aplicación de este último procedimiento es sencilla, condicionando a los estimadores máximo-verosímiles al parámetro de dependencia λ :

$$\begin{aligned} \beta_\lambda &= [X'X]^{-1} X' Ay, \\ &= [X'X]^{-1} X' y - \lambda [X'X]^{-1} X' y = \beta_{mco} - \lambda \beta_{Wy, mco}, \end{aligned} \quad (23)$$

$$\sigma_\lambda^2 = \frac{\tilde{u}'\tilde{u}}{n} = \frac{1}{n} [u_{mco} - \lambda u_{Wy, mco}]' [u_{mco} - \lambda u_{Wy, mco}], \quad (24)$$

siendo $\beta_{mco} = [X'X]^{-1} X' y$, $\beta_{Wy, mco} = [X'X]^{-1} X' Wy$, $u_{mco} = y - X\beta_{mco}$ y $u_{Wy, mco} = Wy - X\beta_{Wy, mco}$.

Sustituyendo estos resultados en (22), se obtiene la función de log-verosimilitud concentrada que depende únicamente del parámetro λ . Esta función es optimizada mediante iteraciones hasta alcanzar convergencia. Posteriormente, el valor $\hat{\lambda}$ es utilizado para obtener estimaciones de las ecuaciones (23)-(24). Los errores estándares de los parámetros se pueden calcular a partir del Hessiano aproximado por una función de optimización numérica o por el cálculo analítico asintótico.

En Stata, el modelo *SLM* puede estimarse por el comando `spreg ml`, sub-opción `dlmat(W5_st)` (Drukker et al., 2013c):

```
. spreg ml U2012 NM2012, id(id) dlmat(W5_st)
```

```
Performing a grid search... finished
```

```
Iteration 0: log likelihood = -419.42774
```

```
Iteration 1: log likelihood = -419.34967
```

```
Iteration 2: log likelihood = -419.34966
```

```
Optimizing unconcentrated log likelihood
```

```
Iteration 0: log likelihood = -419.34966
```

```
Spatial autoregressive model
```

```
Number of obs = 164
```

```
(Maximum likelihood estimates)
```

```
Wald chi2(1) = 10.5058
```

```
Prob > chi2 = 0.0012
```

```
-----+-----
```

	U2012	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
U2012						
NM2012		-.1898462	.0585715	-3.24	0.001	-.3046444 -.0750481
_cons		2.407902	.5260658	4.58	0.000	1.376832 3.438972

```
-----+-----
```

lambda |

	_cons		.8174715	.0411827	19.85	0.000	.7367548	.8981882

sigma2								
	_cons		8.182474	.9309406	8.79	0.000	6.357864	10.00708

A la hora de interpretar los resultados es importante tener presente la forma reducida del modelo *SLM*: $E(y) = (I - \lambda W)^{-1} X\beta$. Este valor esperado contiene como elemento diferencial respecto al modelo de regresión lineal: $(I - \lambda W)^{-1}$.

Considerando que $|\lambda| < 1$ y aplicando la expansión de Leontief, puede obtenerse el multiplicador espacial:

$$(I - \lambda W)^{-1} = I + \lambda W + \lambda^2 W^2 + \dots,$$

que introducido en el valor esperado del *SLM* genera:

$$E(y) = (I - \lambda W)^{-1} X\beta = X\beta + \lambda W X\beta + \lambda^2 W^2 X\beta + \dots.$$

Este resultado permite destacar algunos puntos de interés:

- El resultado en una localización i no solo será afectado por las variables exógenas de i , sino que será retroalimentado por las otras localizaciones a través de la inversa de la transformación espacial $(I - \lambda W)^{-1}$. Este efecto irá decayendo debido a $|\lambda| < 1$.
- Un shock aleatorio en una localización i no solo afectará el resultado de i , además tendrá un impacto en todas las localizaciones.

La presencia de la matriz W permite detectar efectos de interacción espacial y efectos de impacto directo. Los detalles sobre estos efectos, y su obtención, serán desarrollados en la sección 4.

3.2.2. Modelo de error espacial

Para el caso de estructura espacial en el error, se tendrá el modelo de error espacial (*SEM*, spatial error model):

$$\begin{aligned} y &= X\beta + u, \\ u &= \rho W u + \varepsilon \Rightarrow u = B^{-1} \varepsilon, \\ \varepsilon &\sim \mathcal{N}(0, \sigma^2 I_n), \end{aligned} \tag{25}$$

siendo $B = [I - \rho W]$ y ε un vector de innovaciones.

Este modelo tiene una matriz de varianzas y covarianzas igual a:

$$\begin{aligned} E[uu'] &= E\left\{ [B^{-1} \varepsilon] [B^{-1} \varepsilon]' \right\}, \\ &= E\left[B^{-1} \varepsilon \varepsilon' B^{-1'} \right], \\ &= \sigma^2 B^{-1} B^{-1'}, \\ E[uu'] &= \sigma^2 \Omega(\rho)^{-1}, \end{aligned}$$

siendo $\Omega(\rho) = (I - \rho W)'(I - \rho W) = B'B$. Esta estructura genera que los errores se desborden afectando a todo el sistema.

Asumida la distribución normal de las innovaciones, la función de log-verosimilitud del *SEM* será:

$$L = -\frac{n}{2} \ln(2\pi\sigma^2) + \ln|B| - \frac{1}{2\sigma^2} \left[(y - X\beta)' \Omega(\rho) (y - X\beta) \right]. \quad (26)$$

Tal como sucedía en el modelo *SLM*, las condiciones de primer orden son expresiones altamente no lineales y por lo tanto la log-verosimilitud no puede maximizarse directamente.

Para un ρ dado, la optimización de la función de log-verosimilitud (consúltese LeSage & Pace, 2009) muestra que:

$$\beta_\rho = \left[X' \Omega(\rho) X \right]^{-1} X' \Omega(\rho) y,$$

y

$$\sigma_\rho^2 = \frac{e' \Omega(\rho) e}{n},$$

donde $e = y - X\beta_\rho$ y $\Omega(\rho) = (I - \rho W)'(I - \rho W)$, con ambos parámetros siendo funciones de ρ posibilitando la obtención de una log-verosimilitud concentrada:

$$L_c = C + \ln|I - \rho W| - \frac{n}{2} \ln \left[e' \Omega(\rho) e \right].$$

Igual que el procedimiento implementado para el *SLM*, se maximiza la log-verosimilitud concentrada hasta lograr convergencia y posteriormente con el valor $\hat{\rho}$ se obtienen los demás parámetros del modelo. Una vez logrado la convergencia de L_c , puede sustituirse en las expresiones previas tal que:

$$\begin{aligned} \hat{\beta}_\rho &= \left[X' \Omega(\hat{\rho}) X \right]^{-1} X' \Omega(\hat{\rho}) y, \\ \hat{\sigma}_\rho^2 &= \frac{1}{n} \left(e' \Omega(\hat{\rho}) e \right). \end{aligned}$$

El modelo *SEM* por máxima verosimilitud puede estimarse por medio del comando `spreg ml`, sub-opción `elmat(W5_st)`:

```
. spreg ml U2012 NM2012, id(id) elmat(W5_st)
Performing a grid search.... finished
Iteration 0:  log likelihood = -449.50614
Iteration 1:  log likelihood = -422.50561
Iteration 2:  log likelihood = -422.28688
Iteration 3:  log likelihood = -422.28639
Iteration 4:  log likelihood = -422.28639
Iteration 5:  log likelihood = -422.28639
Iteration 6:  log likelihood = -422.28639
Iteration 7:  log likelihood = -422.28639
```

```
Optimizing unconcentrated log likelihood
Iteration 0:  log likelihood = -422.28639
```

```
Spatial autoregressive model                Number of obs   =       164
```



```

(Maximum likelihood estimates)                                Wald chi2(1)    =  4.82871
-----+-----
      U2012 |          Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
U2012      |
  NM2012 |  -.1465469   .0666901   -2.20   0.028   - .277257   - .0158368
  _cons |   10.96933   1.611852    6.81   0.000    7.810155    14.1285
-----+-----
rho        |
  _cons |    .8584009   .0367768   23.34   0.000    .7863196    .9304822
-----+-----
sigma2     |
  _cons |    8.230577   .9412256    8.74   0.000    6.385808   10.07534
-----+-----
. estimates store modelSEM

```

La interpretación de estos resultados no difiere del modelo de regresión lineal. En este modelo el efecto espacial solo afecta a los errores estándares.

3.2.3. Modelo SARAR

El modelo *SARAR* puede ser indicado ante el rechazo conjunto de los contrastes *LM* robustos de error y lag. Este modelo incorpora, simultáneamente, estructura sustantiva y residual:

$$\begin{aligned}
 y &= \lambda W y + X \beta + u, \\
 u &= \rho W u + \varepsilon,
 \end{aligned}
 \tag{27}$$

asumiendo $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$.

Una forma útil de presentar este modelo es la siguiente:

$$\begin{aligned}
 Ay &= X \beta + u, \\
 Bu &= \varepsilon,
 \end{aligned}$$

donde $A = (I - \lambda W)$ y $B = (I - \rho W)$.

Bajo esta notación, asumiendo $|\lambda| < 1$ y $|\rho| < 1$, la función de log-verosimilitud es:

$$L = -\frac{n}{2} \ln(\sigma^2) + \ln |A| + \ln |B| - \frac{1}{2\sigma^2} v' v,
 \tag{28}$$

con $v' v = (Ay - X\beta)' \Omega(\rho) (Ay - X\beta)$ como la suma de cuadrados de los errores transformados, siendo $\Omega(\rho) = B' B$.

Puede concentrarse la log-verosimilitud primero maximizando (28) con respecto a β y σ^2 obteniendo:

$$\beta_{\lambda\rho} = \left[X' \Omega(\rho) X \right]^{-1} X' \Omega(\rho) Ay,$$

y

$$\sigma_\rho^2 = \frac{1}{n} [Ay - X\beta_{\lambda\rho}]' \Omega(\rho) [Ay - X\beta_{\lambda\rho}].$$

Reemplazando estas expresiones en (28) se obtiene la log-verosimilitud concentrada que depende de λ y ρ . Una vez obtenidos los valores $\hat{\lambda}$ y $\hat{\rho}$ que maximizan la función, puede calcularse las estimaciones de β y σ^2 como $\hat{\beta}_{\hat{\lambda}\hat{\rho}}$ y $\hat{\sigma}_{\hat{\lambda}\hat{\rho}}^2$.

En Stata puede estimarse de la siguiente manera:

```
. spreg ml U2012 NM2012, id(id) dlmata(W5_st) elmata(W5_st)
Performing a grid search.... finished
Iteration 0:  log likelihood = -418.21798
Iteration 1:  log likelihood = -418.07666
Iteration 2:  log likelihood = -418.07652
Iteration 3:  log likelihood = -418.07652

Optimizing unconcentrated log likelihood
Iteration 0:  log likelihood = -418.07652
Spatial autoregressive model                Number of obs   =       164
(Maximum likelihood estimates)              Wald chi2(1)    =     9.31253
                                              Prob > chi2     =     0.0023

-----+-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012          |
  NM2012       |   - .1626519   .0532998   -3.05  0.002   - .2671176   - .0581863
  _cons        |   1.738683    .5200708    3.34  0.001    .7193631    2.758003
-----+-----
lambda         |
  _cons        |   .8798585    .0422453   20.83  0.000    .7970593    .9626577
-----+-----
rho            |
  _cons        |  - .3348529   .1988064   -1.68  0.092   - .7245062    .0548004
-----+-----
sigma2         |
  _cons        |   7.552715    .9240574    8.17  0.000    5.741596    9.363834
-----+-----
```

El coeficiente espacial del error, $\hat{\rho}$, no es significativo al 5%, favoreciendo la estimación de un modelo sin estructura espacial en el error.

Si bien este modelo es utilizado en diferentes investigaciones de contenido teórico (principalmente en las investigaciones de Kelejian y Prucha), su utilización empírica no es habitual prefiriéndose otros modelos como el de Durbin.

3.2.4. Modelo espacial de Durbin

El modelo espacial de Durbin, *SDM*, posee la siguiente especificación:

$$y = \lambda W y + X\beta + W X\gamma + u, \tag{29}$$

donde WX captura el efecto espacial local de las variables exógenas.

Ante una situación de rechazo de ambos contrastes de Lagrange robustos, la elección de este modelo se sustenta en que incorpora dependencia sustantiva y además incorpora nuevas variables explicativas que pueden haberse omitido. Es decir, la detección de estructura espacial en el error puede ser causada por una omisión de variables relevantes y no por efectos espaciales en el error.

El procedimiento para estimar este modelo no difiere del presentado por el *SLM*, ya que solo se incorporan nuevas variables exógenas, considerando que W es, por construcción, no estocástica.

En Stata, para estimar el modelo *SDM*, se necesita generar un rezago espacial de la variable migración neta. Esta nueva variable se genera fácilmente por medio de las opciones de `spmat`: el subcomando es `lag`, seguido por el nombre de la nueva variable, la matriz de pesos y la variable original de la que se desea obtener el rezago espacial:

```
. spmat lag wx_NM2012 W5_st NM2012
```

Una vez generada la variable, se procede a estimar el modelo *SLM*, pero incorporando `wx_NM2012` como variable explicativa:

```
. spreg ml U2012 NM2012 wx_NM2012, id(id) dmat(W5_st)
Performing a grid search... finished
Iteration 0:  log likelihood = -418.88804
Iteration 1:  log likelihood = -418.88134
Iteration 2:  log likelihood = -418.88133

Optimizing unconcentrated log likelihood
Iteration 0:  log likelihood = -418.88133
Spatial autoregressive model                Number of obs   =       164
(Maximum likelihood estimates)              Wald chi2(2)    =    11.1669

-----+-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012          |
      NM2012 |   -.1599277   .0664974   -2.41   0.016   - .2902602   -.0295951
      wx_NM2012 |  -.1014938   .1057227   -0.96   0.337   - .3087066   .1057189
      _cons |    2.820703   .6965794    4.05   0.000    1.455432    4.185973
-----+-----
lambda         |
      _cons |    .7953717   .0487782   16.31   0.000    .6997682    .8909752
-----+-----
sigma2         |
      _cons |    8.250679   .9431208    8.75   0.000    6.402197   10.09916
-----+-----

. estimates store modelSDM
```

Es interesante mencionar que el modelo *SDM*, a pesar de contener una estructura más compleja que el *SLM*, puede reducirse a un modelo *SEM*. Esta relación se hace evidente mediante el contraste denominado *COMFAC* (de factores comunes).

Suponiendo que se ha estimado el modelo *SDM*:

$$y = \lambda W y + X\beta + WX\gamma + u. \quad (30)$$

El estadístico de factores comunes plantea las siguientes hipótesis, nula y alternativa:

$$H_0 : \gamma + \lambda\beta = 0,$$

$$H_1 : \gamma + \lambda\beta \neq 0.$$

Bajo H_0 , $\gamma = -\lambda\beta$, entonces reemplazando en el modelo (30):

$$\begin{aligned} y &= \lambda W y + X\beta + WX(-\lambda\beta) + u, \\ &= \lambda W y + X\beta - \lambda WX\beta + u, \\ (I - \lambda W)y &= (I - \lambda W)X\beta + u. \end{aligned}$$

La última expresión puede resumirse en el *SEM*:

$$y = X\beta + (I - \rho W)^{-1} \varepsilon,$$

donde se ha reemplazado a λ por ρ .

El modelo bajo hipótesis alternativa es el *SDM* y el modelo restringido, bajo hipótesis nula, es el *SEM*. Bajo estimación *MV*, puede plantearse el estadístico *LR*:

$$LR_{COMFAC} = 2 [l_{|H_1} - l_{|H_0}] \underset{as}{\sim} \chi_q^2 \quad (31)$$

siendo $l_{|H_1}$ la log-verosimilitud obtenida en la estimación del modelo amplio y $l_{|H_0}$ la correspondiente al modelo de la hipótesis nula, q es el número de restricciones e igual al número de parámetros incluidos en β , sin considerar la constante.

Para aplicar este contraste en Stata, previamente se han guardado las estimaciones del modelo *SDM* y *SEM*. Luego, se aplica el siguiente comando:

```
. lrtest modelSDM modelSEM
Likelihood-ratio test                LR chi2(1) =      6.81
(Assumption: modelSEM nested in modelSDM)  Prob > chi2 =    0.0091
```

donde `modelSDM` y `modelSEM` son las estimaciones almacenadas en Stata del modelo *SDM* y *SEM*, respectivamente.

En conclusión, se rechaza la hipótesis nula, arrojando evidencia a favor de un modelo *SDM*.

Este modelo es muy utilizado en estimaciones empíricas debido a varias ventajas. En primer lugar, permite incorporar efectos espaciales globales por medio de la estimación del parámetro λ y efectos espaciales locales capturados por γ . En segundo lugar, es un modelo que anida al *SLM*, al *SEM* y al *SLX*. El modelo *SLX* es simplemente un modelo con efectos locales, WX :

$$y = X\beta + WX\gamma + u, \quad (32)$$

que puede ser estimado por *MCO*.

3.2.5. Modelo de Cliff-Ord

Existe un modelo más completo que los anteriores y es denominado modelo de Cliff-Ord. Este modelo incorpora dependencia sustantiva, dependencia residual y dependencia local debido a las variables exógenas, siendo su expresión:

$$\begin{aligned} y &= \lambda W y + X\beta + W X \gamma + u, \\ u &= \rho W u + \varepsilon. \end{aligned} \tag{33}$$

Imponiendo restricciones en la ecuación (33) sobre γ , ρ y λ pueden obtenerse los anteriores modelos:

- $\gamma = 0, \rho = 0, \lambda \neq 0 \rightarrow SLM$.
- $\gamma = 0, \rho \neq 0, \lambda = 0 \rightarrow SEM$.
- $\gamma = 0, \rho \neq 0, \lambda \neq 0 \rightarrow SARAR$.
- $\gamma \neq 0, \rho = 0, \lambda \neq 0 \rightarrow SDM$.

El método de estimación *MV* para el modelo de Cliff-Ord no difiere del presentado por el modelo *SARAR*, ya que solo incorpora un elemento adicional en las variables explicativas, *WX*. Tal como sucedía con el modelo *SARAR*, es poco frecuente la estimación de este tipo de modelo en aplicaciones empíricas debido a problemas de identificación de los parámetros.

La estimación en Stata se realiza con la siguiente secuencia:

```
. spreg ml U2012 NM2012 wx_NM2012, id(id) dmat(W5_st) elmat(W5_st)
Performing a grid search... finished
Iteration 0:  log likelihood = -418.2163
Iteration 1:  log likelihood = -418.03811
Iteration 2:  log likelihood = -418.03731
Iteration 3:  log likelihood = -418.03731

Optimizing unconcentrated log likelihood
Iteration 0:  log likelihood = -418.03731
Spatial autoregressive model                Number of obs   =       164
(Maximum likelihood estimates)              Wald chi2(2)    =    8.77638
                                              Prob > chi2     =    0.0124

-----+-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012          |
      NM2012 |   -.1515091   .0670909   -2.26  0.024   - .2830048   -.0200134
      wx_NM2012 |  -.0286924   .1035134   -0.28  0.782   - .231575   .1741901
      _cons |    1.86788   .7290904    2.56  0.010   .4388886   3.296871
-----+-----
lambda         |
      _cons |    .871513   .0543806   16.03  0.000   .764929   .9780969
-----+-----
rho            |
```

	_cons		-.3104278	.220725	-1.41	0.160	-.7430409	.1221853

sigma2								
	_cons		7.622098	.9685032	7.87	0.000	5.723867	9.52033

Para nuestro ejemplo, obsérvese que el parámetro estimado ρ es no significativo, reduciendo el modelo de Cliff-Ord a un modelo *SDM*. Por otra parte, en la estimación del modelo *SDM*, el efecto local de las variables explicativas no es significativo, reduciendo el modelo a un *SLM*, tal como se detectó mediante los contrastes *LM*.

Metodológicamente, nuestra búsqueda de especificación espacial comenzó con una estimación *MCO* y, luego, se fueron incorporando elementos espaciales hasta alcanzar el modelo más complejo. Alternativamente, puede estimarse el modelo más complejo e ir eliminando los elementos no significativos hasta alcanzar una especificación adecuada. La primera estrategia es denominada “de lo particular a lo general” y la segunda “de lo general a lo particular”. Para una discusión sobre las ventajas de una u otra estrategia en econometría espacial puede consultarse el trabajo de Mur & Angulo (2009) y las referencias allí citadas.

3.3. Estimación por método de momentos generalizados

Los principales precursores de este método han sido los trabajos de Kelejian & Prucha (1998, 1999), con generalizaciones recientes realizadas por Kelejian & Prucha (2010); Arraiz et al. (2010); Drukker et al. (2013a).

3.3.1. Modelo de rezago espacial

Para implementar la estrategia de estimación se reexpresará al modelo *SLM* como:

$$y = Z\delta + u, \quad (34)$$

donde $Z_{n \times (k+1)} = [X, Wy]$ y $\delta' = [\beta', \lambda]$, con $u \sim (0, \sigma_u^2 I_n)$.

La estimación de este modelo se basa en el enfoque de variables instrumentales (*IV*). La idea es simple partiendo de que el modelo (34) no puede estimarse por los métodos habituales debido a que existe una variable endógena, Wy , que se correlaciona con el término de error. La alternativa es encontrar un instrumento, H , que se encuentre correlacionado con la variable que genera endogeneidad Wy pero no correlacionado con el error, asintóticamente hablando. Es decir, el instrumento H debe cumplir con:

$$plim \left(\frac{1}{n} \right) H' Wy = M_{HWy},$$

y

$$plim \left(\frac{1}{n} \right) H' u = 0,$$

siendo M_{HWy} una matriz finita no-singular.

Recordando que la forma reducida del modelo *SLM* es: $E(y) = (I - \lambda W)^{-1}(X\beta + u)$, y considerando que $(I - \lambda W)^{-1}$ puede expresarse como una progresión: $(I - \lambda W)^{-1} = I + \lambda W + \lambda^2 W^2 + \lambda^3 W^3 + \dots$. Entonces, el valor esperado de y puede expresarse como:

$$E(y) = X\beta + \lambda WX\beta + \lambda^2 W^2 X\beta + \lambda^3 W^3 X\beta + \dots,$$

y, además,

$$E(Wy) = WX\beta + W^2 X(\lambda\beta) + W^3 X(\lambda^2\beta) + \dots.$$

Luego, $E(Wy)$ se relaciona linealmente con $WX, W^2X, W^3X, \dots, W^qX, \dots$, siendo instrumentos directos de la variable endógena Wy . Diferentes experimentos Monte Carlo muestran un buen comportamiento para $q = 2$, por lo que habitualmente se utilizan WX y W^2X como instrumentos. El uso de ordenes mayores tiende a generar problemas de multicolinealidad entre los regresores.

Aplicando esta idea al modelo (34) se obtiene el estimador de variables instrumentales (IV):

$$\hat{\delta} = (\hat{Z}'Z)^{-1} \hat{Z}'y,$$

donde $\hat{Z} = P_H Z = [X, \widehat{W}y]$, $\widehat{W}y = P_H Wy$ y $P_H = H(H'H)^{-1}H'$, siendo los instrumentos elegidos $H = [X, WX, W^2X]$.

Kelejian & Prucha (1998) demuestran que

$$n^{1/2}(\hat{\delta} - \delta) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 \text{plim}_{n \rightarrow \infty} [\hat{Z}'\hat{Z}]^{-1}\right).$$

La estimación IV de estos parámetros puede ser obtenida como un estimador de mínimos cuadrados en dos etapas ($2SLS$):

1. Estimar la relación entre los instrumentos $H = [X, WX]$ y $Z = [X, Wy]$:

$$Z = H\theta + \epsilon,$$

$$\hat{\theta} = (H'H)^{-1} H'Z,$$

tal que las predicciones del modelo son: $\hat{Z} = H\hat{\theta} = H(H'H)^{-1}H'Z = P_H Z$.

2. Estimar $y = Z\delta + \epsilon$, usando los resultados del paso anterior:

$$\hat{\delta} = (\hat{Z}'\hat{Z})^{-1} \hat{Z}'y,$$

donde $\hat{Z} = P_H Z = [X, \widehat{W}y]$, donde $\widehat{W}y = P_H Wy$.

En Stata, el comando que permite realizar esta estimación es **spivreg** desarrollado por Drukker et al. (2013c):

```
. spivreg U2012 NM2012, dl(W5_st) id(id)
Spatial autoregressive model          Number of obs   =       164
(GS2SLS estimates)
```

	U2012	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
U2012							
	NM2012	-.1421611	.071088	-2.00	0.046	-.281491	-.0028311
	_cons	1.566084	.9010135	1.74	0.082	-.1998704	3.332038
lambda							
	_cons	.8937041	.0781357	11.44	0.000	.740561	1.046847

Si se asume que el término de error es independiente pero heteroscedástico, entonces se modifica la matriz de varianzas-covarianzas asintótica de $\hat{\delta}$, adquiriendo la forma *sandwich*:

$$\left(\hat{Z}'\hat{Z}\right)^{-1}\hat{Z}'\hat{\Sigma}\hat{Z}\left(\hat{Z}'\hat{Z}\right)^{-1},$$

donde $\hat{\Sigma}$ es una matriz diagonal cuyos elementos i –ésimos son \tilde{u}_i^2 con $\tilde{u}_i = y_i - Z_i\hat{\delta}$.

Para realizar esta corrección en Stata basta agregar al comando anterior la sub-opción **het**:

```
. spivreg U2012 NM2012, dl(W5_st) id(id) het
Spatial autoregressive model          Number of obs   =       164
(GS2SLS estimates)
```

	U2012	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
U2012							
	NM2012	-.1421611	.0693011	-2.05	0.040	-.2779888	-.0063333
	_cons	1.566084	1.099695	1.42	0.154	-.5892783	3.721446
lambda							
	_cons	.8937041	.1109302	8.06	0.000	.6762849	1.111123

Comparando las estimaciones bajo homoscedasticidad y heteroscedasticidad, puede observarse que la única corrección se da en los errores estándares de los coeficientes. La principal corrección se produce en el error del parámetro espacial, sin embargo, los resultados principales no se modifican.

3.3.2. Modelo de error espacial

El modelo *SEM* puede expresarse como:

$$\begin{aligned} y &= Z\delta + u, \\ u &= \rho Wu + \varepsilon, \end{aligned} \tag{35}$$

donde $Z = X$, $\delta = \beta$, ρ es un parámetro autoregresivo espacial, siendo las innovaciones $\varepsilon \sim (0, \sigma^2 I)$, con I siendo una matriz identidad de orden $n \times n$.

El objetivo es obtener un estimador consistente de ρ . Este paso genera un estimador inicial *GMM* para ρ que surge de las condiciones de momentos cuadráticos de las innovaciones:

$$E\left(\varepsilon' A_s \varepsilon\right) = 0, \quad s = 1, \dots, S,$$

donde las matrices A_s satisfacen $\text{tr}(A_s) = 0$.

El estimador inicial *GMM* viene dado por

$$\tilde{\rho} = \arg \min \left\{ \left[\tilde{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right]' \left[\tilde{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right] \right\},$$

siendo

$$\tilde{\Gamma} = \frac{1}{n} \begin{bmatrix} \hat{u}' (A_1 + A_1') \hat{u} & -\hat{u}' A_1 \hat{u} \\ \vdots & \vdots \\ \hat{u}' (A_S + A_S') \hat{u} & -\hat{u}' A_S \hat{u} \end{bmatrix}, \quad \tilde{\omega} = \frac{1}{n} \begin{bmatrix} \hat{u}' A_1 \hat{u} \\ \vdots \\ \hat{u}' A_S \hat{u} \end{bmatrix}, \quad \text{con } \hat{u} = y - Z\hat{\beta} \text{ y } \hat{u} = W\hat{u}.$$

Así planteado, el estimador *GMM* proviene de un estimador de mínimos cuadrados no-lineales. Generalmente, se utilizan dos ecuaciones ($S = 2$).

Bajo homoscedasticidad de las innovaciones:

$$A_1 = \left[1 + \left\{ \frac{1}{n} \text{tr}(W'W) \right\}^2 \right]^{-1} \left\{ W'W - \frac{1}{n} \text{tr}(W'W) I_n \right\},$$

y

$$A_2 = W.$$

Obtenida la estimación inicial de ρ , se procede a obtener los coeficientes del modelo (35). Para ello se utiliza la transformación Cochrane-Orcutt:

$$\begin{aligned} (I - \rho W) y &= (I - \rho W) Z\beta + \varepsilon, \\ y_\rho &= Z_\rho \beta + \varepsilon. \end{aligned} \tag{36}$$

Bajo esta especificación, se utiliza la estimación *GMM* $\tilde{\rho}$:

$$\hat{\delta} = \left(\hat{Z}'_\rho \hat{Z}_\rho \right)^{-1} \hat{Z}'_\rho y_\rho,$$

donde $y_\rho = (I - \tilde{\rho}W) y$, $Z_\rho = (I - \tilde{\rho}W)Z$, y $\hat{Z}_\rho = P_H Z_\rho = X (X'X)^{-1} X' Z_\rho$.

El procedimiento completo es conocido como mínimos cuadrados espaciales en dos etapas generalizados (*GS2SLS*).

Posterior a la estimación de los parámetros del modelo, puede darse un paso adicional para obtener un estimador eficiente de ρ . Para ello se utilizan los residuos del procedimiento *GS2SLS*:

$$\hat{\rho} = \arg \min \left\{ \left[\hat{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right]' \left\{ \hat{\Psi}_\rho^{\rho\rho} \right\}^{-1} \left[\hat{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right] \right\},$$

donde $\widehat{\Psi}_{\rho}^{\rho\rho}$ es un estimador de la matriz de varianzas-covarianzas del vector de momentos muestrales (normalizados) basados en los residuos *GS2SLS*. Bajo homoscedasticidad los elementos r, s (con $r, s = 1, 2$) de $\widehat{\Psi}_{\rho}^{\rho\rho}$ están dados por:

$$\begin{aligned}\widehat{\Psi}_{\rho}^{\rho\rho} &= \left\{ \widehat{\sigma}_{\rho}^2 \right\}^2 \frac{1}{2n} \text{tr} \left\{ \left(A_r + A_r' \right) \left(A_s + A_s' \right) \right\} + \frac{1}{n} \widehat{\sigma}_{\rho}^2 \widetilde{a}_{\rho r}' \widetilde{a}_{\rho s} \\ &+ \frac{1}{n} \left[\widetilde{\mu}_{\rho}^{(4)} - 3 \left\{ \widehat{\sigma}_{\rho}^2 \right\}^2 \right] \text{vec}_D \left(A_r \right)' \text{vec}_D \left(A_s \right) \\ &+ \frac{1}{n} \widetilde{\mu}_{\rho}^{(3)} \left\{ \widetilde{a}_{\rho r}' \text{vec}_D \left(A_s \right) + \widetilde{a}_{\rho s}' \text{vec}_D \left(A_r \right) \right\}\end{aligned}$$

siendo

$$\begin{aligned}\widetilde{a}_{\rho r} &= \widehat{T}_{\rho} \widetilde{\alpha}_{\rho r}, \\ \widehat{T}_{\rho} &= H \widehat{P}_{\rho}, \\ \widehat{P}_{\rho} &= \widehat{Q}_{HH}^{-1} \widehat{Q}_{\rho HZ} \left\{ \widehat{Q}_{\rho HZ}' \widehat{Q}_{HH}^{-1} \widehat{Q}_{\rho HZ}' \right\}^{-1}, \\ \widehat{Q}_{HH} &= \left(\frac{1}{n} H' H \right), \\ \widehat{Q}_{\rho HZ} &= \left(\frac{1}{n} H' Z_{\rho} \right), \\ Z_{\rho} &= (I - \widetilde{\rho} W) Z, \\ \widetilde{a}_{\rho r} &= -\frac{1}{n} \left\{ Z_{\rho}' \left(A_r + A_r' \right) \widehat{\varepsilon}_{\rho} \right\}, \\ \widehat{\varepsilon}_{\rho} &= (I - \widetilde{\rho} W) \widehat{u}, \\ \widehat{\sigma}_{\rho}^2 &= \frac{1}{n} \widehat{\varepsilon}_{\rho}' \widehat{\varepsilon}_{\rho}, \\ \widetilde{\mu}_{\rho}^{(3)} &= \frac{1}{n} \sum_{i=1}^n \widehat{\varepsilon}_{i\rho}^3, \\ \widetilde{\mu}_{\rho}^{(4)} &= \frac{1}{n} \sum_{i=1}^n \widehat{\varepsilon}_{i\rho}^4.\end{aligned}$$

En Stata, la estimación de este modelo se obtiene mediante:

```
. spivreg U2012 NM2012, el(w5_st) id(id)
Estimating rho using 2SLS residuals
Iteration 0:   GMM criterion = .07399364
Iteration 1:   GMM criterion = .03891457
Iteration 2:   GMM criterion = .03890692
Iteration 3:   GMM criterion = .03890692

Estimating rho using GS2SLS residuals
Iteration 0:   GMM criterion = .06371936
Iteration 1:   GMM criterion = .00222485
Iteration 2:   GMM criterion = .00209885
Iteration 3:   GMM criterion = .00209884
Spatial autoregressive model                Number of obs   =       164
(GS2SLS estimates)

-----+-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012          |
NM2012 |   - .2031552   .0669429   -3.03   0.002   - .3343609   - .0719496
   _cons |    10.4998    1.328934    7.90   0.000    7.895143    13.10447
-----+-----
```

```

rho          |
      _cons |   .8279244   .0371811   22.27   0.000   .7550508   .9007979
-----

```

La estimación obtenida asume innovaciones homoscedásticas. En el caso de heterocedasticidad, se debe modificar en el estimador inicial el elemento A_1 : $A_1 = W'W - \text{diag}(W'W)$.

En la segunda etapa, que realiza una estimación eficiente de ρ usando los residuos *GS2SLS*, se modifica el elemento r, s de $\widehat{\Psi}_\rho^{\rho\rho}$:

$$\widehat{\Psi}_\rho^{\rho\rho} = \frac{1}{2n} \text{tr} \left\{ (A_r + A'_r) \widehat{\Sigma}_\rho (A_s + A'_s) \right\} + \frac{1}{n} \widehat{\sigma}_\rho^2 \widetilde{a}_{\rho r} \widehat{\Sigma}_\rho \widetilde{a}_{\rho s},$$

donde $\widehat{\Sigma}_\rho$ es una matriz diagonal cuyo elemento i -ésimo es $\widehat{\varepsilon}_{i\rho}^2$, y $\widetilde{a}_{\rho r}$ y $\widetilde{a}_{\rho s}$ son definidos como en el caso homoscedástico.

En Stata las correcciones heteroscedásticas son implementadas por la sub-opción **het**:

```

. spivreg U2012 NM2012, el(W5_st) id(id) het
Estimating rho using 2SLS residuals
Iteration 0:   GMM criterion = .02255292
Iteration 1:   GMM criterion = .00726521
Iteration 2:   GMM criterion = .00726363
Iteration 3:   GMM criterion = .00726363
Estimating rho using GS2SLS residuals
Iteration 0:   GMM criterion = .02134596
Iteration 1:   GMM criterion = .00014747
Iteration 2:   GMM criterion = .00010296
Iteration 3:   GMM criterion = .00010296
Spatial autoregressive model                Number of obs   =       164
(GS2SLS estimates)

-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012          |
  NM2012 |   -.2057361   .0662701   -3.10   0.002   - .3356232   -.0758491
  _cons |   10.49601   1.279068    8.21   0.000    7.989081   13.00293
-----+-----
rho          |
  _cons |   .8216508   .0681744   12.05   0.000    .6880314   .9552702
-----

```

3.3.3. Modelo SARAR

Este modelo es planteado como:

$$\begin{aligned} y &= Z\delta + u, \\ u &= \rho Wu + \varepsilon, \end{aligned} \tag{37}$$

donde $Z = [X, Wy]$ y $\delta' = [\beta', \lambda]$, con $u \sim (0, \sigma_u^2 I_n)$.

El procedimiento de estimación del modelo *SARAR* es una combinación de los pasos dados en los anteriores modelos y puede ser resumido en cuatro pasos:

Paso 1: Estimar por *2SLS* la ecuación $y = Z\beta + u$, obteniendo

$$\hat{\delta} = (\hat{Z}'Z)^{-1} \hat{Z}'y,$$

donde $\hat{Z} = P_H Z = [X, \widehat{W}y]$, $\widehat{W}y = P_H W y$ y $P_H = H(H'H)^{-1}H'$, siendo los instrumentos elegidos $H = [X, WX]$.

Paso 2: Obtener $\hat{u} = y - Z\hat{\delta}$, $\hat{u} = W\hat{u}$. Usando estos vectores residuales obtener el estimador inicial *GMM* minimizando:

$$\tilde{\rho} = \arg \min \left\{ \left[\tilde{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right]' \left[\tilde{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right] \right\},$$

Paso 3: Estimador *GS2SLS*. Se utiliza la estimación inicial *GMM*, $\tilde{\rho}$, en la transformación Cochrane-Orcutt para obtener

$$\hat{\delta} = (\hat{Z}'_{\tilde{\rho}} Z_{\tilde{\rho}})^{-1} \hat{Z}'_{\tilde{\rho}} y_{\tilde{\rho}},$$

donde $y_{\tilde{\rho}} = (I - \tilde{\rho}W)y$, $Z_{\tilde{\rho}} = (I - \tilde{\rho}W)Z$, y $\hat{Z}_{\tilde{\rho}} = P_H Z_{\tilde{\rho}} = H(H'H)^{-1}H'Z_{\tilde{\rho}}$.

Paso 4: Estimador eficiente y consistente *GMM*. Utilizando los residuos del procedimiento *GS2SLS*, se estima:

$$\hat{\rho} = \arg \min \left\{ \left[\hat{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right]' \left\{ \hat{\Psi}_{\rho}^{\rho\rho} \right\}^{-1} \left[\hat{\Gamma} \begin{pmatrix} \rho \\ \rho^2 \end{pmatrix} - \tilde{\omega} \right] \right\},$$

Habiendo computado los estimadores $\hat{\delta} = (\hat{\beta}, \hat{\lambda})$ y $\hat{\rho}$, el próximo paso es computar un estimador consistente de la matriz de varianzas-covarianzas, Υ . El estimador viene dado por $n\hat{\Upsilon}$ donde:

$$\hat{\Upsilon} = \begin{pmatrix} \hat{\Upsilon}^{\delta\delta} & \hat{\Upsilon}^{\delta\rho} \\ \hat{\Upsilon}^{\delta\rho'} & \hat{\Upsilon}^{\rho\rho} \end{pmatrix},$$

con

$$\hat{\Upsilon}^{\delta\delta} = \hat{P}'_{\tilde{\rho}} \hat{\Psi}_{\tilde{\rho}}^{\delta\delta} \hat{P}_{\tilde{\rho}},$$

$$\hat{\Upsilon}^{\delta\rho} = \hat{P}'_{\tilde{\rho}} \hat{\Psi}_{\tilde{\rho}}^{\delta\rho} \hat{P}_{\tilde{\rho}} \left\{ \hat{\Psi}_{\tilde{\rho}}^{\rho\rho} \right\}^{-1} \hat{J} \left[\hat{J}' \left\{ \hat{\Psi}_{\tilde{\rho}}^{\rho\rho} \right\}^{-1} \hat{J} \right]^{-1},$$

$$\hat{\Upsilon}^{\rho\rho} = \hat{J}' \left[\hat{J}' \left\{ \hat{\Psi}_{\tilde{\rho}}^{\rho\rho} \right\}^{-1} \hat{J} \right]^{-1},$$

$$\hat{J} = \hat{\Gamma} \begin{pmatrix} 1 \\ 2\hat{\rho} \end{pmatrix}.$$

Bajo el supuesto de homoscedasticidad de las innovaciones, se utilizan:

$$\hat{\Psi}_{\tilde{\rho}}^{\delta\delta} = \hat{\sigma}_{\tilde{\rho}}^2 \hat{Q}_{HH},$$

$$\widehat{\Psi}_{\rho}^{\delta\rho} = \frac{1}{n} \widehat{\sigma}_{\rho}^2 H' \left\{ a_{1\rho}, a_{2\rho} \right\} + \frac{1}{n} \mu_{\rho}^3 H' \left\{ \text{vec}_D(A_1), \text{vec}_D(A_2) \right\}.$$

En Stata, el comando para estimar un *SARAR* homoscedástico es:

```
. spivreg U2012 NM2012, dl(W5_st) el(W5_st) id(id)
Estimating rho using 2SLS residuals
Iteration 0:   GMM criterion = .02184886
Iteration 1:   GMM criterion = .02116691
Iteration 2:   GMM criterion = .0211669

Estimating rho using GS2SLS residuals
Iteration 0:   GMM criterion = .00570952
Iteration 1:   GMM criterion = .00383306
Iteration 2:   GMM criterion = .00382727
Iteration 3:   GMM criterion = .00382726
Spatial autoregressive model           Number of obs   =       164
(GS2SLS estimates)

-----+-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012          |
  NM2012       |   - .151962    .0704003   -2.16   0.031   - .2899441   - .01398
  _cons        |   1.564904    .8197337    1.91   0.056   - .0417443   3.171553
-----+-----
lambda         |
  _cons        |   .8951628    .0691381   12.95   0.000    .7596545    1.030671
-----+-----
rho            |
  _cons        |  - .2720242    .2404617   -1.13   0.258   - .7433204    .199272
-----+-----
```

Cuando se especifica heteroscedasticidad, la estructura de los siguientes elementos se modifica (además de las modificaciones presentadas en los pasos previos):

$$\widehat{\Psi}_{\rho}^{\delta\delta} = \frac{1}{n} H' \widehat{\Sigma}_{\rho} H,$$

$$\widehat{\Psi}_{\rho}^{\delta\rho} = \frac{1}{n} H' \widehat{\Sigma}_{\rho} \left\{ a_{1\rho}, a_{2\rho} \right\}.$$

Nuevamente, en Stata basta con incorporar al comando `spivreg`, la sub-opción `het`:

```
. spivreg U2012 NM2012, el(W5_st) dl(W5_st) id(id) het
Estimating rho using 2SLS residuals
Iteration 0:   GMM criterion = .04395599
Iteration 1:   GMM criterion = .04069686
Iteration 2:   GMM criterion = .04069677

Estimating rho using GS2SLS residuals
Iteration 0:   GMM criterion = .00615584
Iteration 1:   GMM criterion = .00349947
Iteration 2:   GMM criterion = .0034926
Iteration 3:   GMM criterion = .00349258
Spatial autoregressive model           Number of obs   =       164
(GS2SLS estimates)

-----+-----
```

U2012	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----						
U2012						
NM2012	-.1523475	.0698285	-2.18	0.029	-.2892088	-.0154861
_cons	1.566555	1.031593	1.52	0.129	-.4553304	3.58844
-----+-----						
lambda						
_cons	.8950652	.1007308	8.89	0.000	.6976365	1.092494
-----+-----						
rho						
_cons	-.1763782	.3552156	-0.50	0.620	-.8725881	.5198316
-----+-----						

3.3.4. Modelo espacial de Durbin

El modelo *SDM*, incorpora al *SLM* las variables WX . La introducción de estas variables exógenas rezagas tiende a generar problemas de multicolinealidad en la estimación *IV*. El motivo se debe a los instrumentos utilizados para la variable endógena Wy :

$$H = [X, WX, W^2X].$$

Al introducir como variables explicativas WX , el conjunto de instrumentos pasa a ser:

$$H = [X, WX, W^2X, WX, W(WX), W^2(WX)],$$

en donde puede observarse que WX y W^2X se encuentran repetidos en el conjunto de instrumentos. Luego, los instrumentos disponibles para la estimación del *SDM* serán:

$$H = [X, WX, W^2X, W^3X]. \quad (38)$$

Siguiendo con el ejemplo aplicado en Stata:

```
. spivreg U2012 NM2012 wx_NM2012, dl(W5_st) id(id)
note: 2 instruments omitted from the H matrix because of collinearity
      see e(H_omitted) for names of the omitted instruments
Spatial autoregressive model          Number of obs   =       164
(GS2SLS estimates)

-----+-----
      U2012 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
U2012      |
  NM2012   |  -.1417927   .0733571   -1.93   0.053    - .28557   .0019845
  wx_NM2012 |  .0174259   .247244    0.07   0.944    - .4671634 .5020151
  _cons    |  1.398022   2.772848   0.50   0.614    -4.03666   6.832704
-----+-----
lambda     |
  _cons    |  .9062994   .2149423   4.22   0.000    .4850202   1.327579
-----+-----
```

El mensaje posterior al comando informa que se han omitido dos instrumentos. Mediante el comando `ereturn list` puede visualizarse a cada instrumento omitido:

```
e(H_omitted) : "NM2012*W NM2012*W^2"
```

Para comprobar que el conjunto de instrumentos utilizados son los planteados en (38) puede aplicarse un comando de estimación por variables instrumentales ya predefinido en Stata. Para ello se necesitan crear los rezagos espaciales de migración neta:

```
. spmat lag wx2_NM2012 W5_st wx_NM2012
. spmat lag wx3_NM2012 W5_st wx2_NM2012
```

Con estas nuevas variables como instrumentos, se aplica el siguiente comando. Obsérvese que entre paréntesis se ubica la variable considerada endógena y, a continuación del signo igual, los instrumentos para la misma:

```
. ivregress 2sls U2012 NM2012 wx_NM2012 (wx_U2012 = wx2_NM2012 wx3_NM2012)
Instrumental variables (2SLS) regression                Number of obs =    164
                                                         Wald chi2(3) =   316.35
                                                         Prob > chi2 =   0.0000
                                                         R-squared =     0.7716
                                                         Root MSE =     2.8017

-----+-----
          U2012 |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      wx_U2012 |    .9062994    .2149423     4.22  0.000    .4850202    1.327579
      NM2012 |   -.1417927    .0733571    -1.93  0.053   -.28557    .0019845
      wx_NM2012 |   .0174259    .247244     0.07  0.944   -.4671634   .5020151
        _cons |   1.398022    2.772848     0.50  0.614   -4.03666    6.832704
-----+-----

Instrumented:  wx_U2012 Instruments:  NM2012 wx_NM2012 wx2_NM2012 wx3_NM2012
```

Los resultados obtenidos por variables instrumentales son iguales a los reportados por el comando `spivreg`. El coeficiente que acompaña a `wx_U2012` es el `lambda`.

Tal como sucedía en el modelo *SLM*, puede corregirse por heteroscedasticidad de las innovaciones (resultados que no se presentan).

3.3.5. Modelo de Cliff-Ord

Sin mayor detalle que el expuesto en el modelo *SARAR*, la estimación del modelo de Cliff-Ord supone utilizar el conjunto de instrumentos del modelo *SDM*.

```
. spivreg U2012 NM2012 wx_NM2012, e1(W5_st) d1(W5_st) id(id)
note: 2 instruments omitted from the H matrix because of collinearity
see e(H_omitted) for names of the omitted instruments
Estimating rho using 2SLS residuals
Iteration 0:  GMM criterion = .0275166
Iteration 1:  GMM criterion = .02714823
Iteration 2:  GMM criterion = .02714822

Estimating rho using GS2SLS residuals
```

```

Iteration 0:  GMM criterion =  .0071405
Iteration 1:  GMM criterion =  .00026778
Iteration 2:  GMM criterion =  .00026684
Iteration 3:  GMM criterion =  .00026684

```

```

Spatial autoregressive model                Number of obs =    164
(GS2SLS estimates)
-----+-----
          U2012 |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
U2012          |
      NM2012 |  -.1558511   .073233   -2.13   0.033   - .2993852   - .012317
      wx_NM2012 | -.0566468   .2485094   -0.23   0.820   - .5437163   .4304228
      _cons |    2.16067   2.740343    0.79   0.430   -3.210303   7.531643
-----+-----
lambda         |
      _cons |    .8494096   .2121095    4.00   0.000    .4336827   1.265136
-----+-----
rho            |
      _cons |   -.0868985   .3627544   -0.24   0.811   - .7978842   .6240871
-----+-----

```

Un último punto a considerar es la posibilidad de incorporar variables endógenas no espaciales, además del tratamiento por heteroscedasticidad. El comando `spivreg` permite trabajar simultáneamente con ambos tipos de endogeneidad.

Por ejemplo, puede pensarse que la tasa de migración neta no es una variable estrictamente exógena al modelo planteado. En este caso, la migración neta en el 2012 puede instrumentalizarse por medio de un rezago temporal del 2010. Las condiciones para que sea un instrumento válido ya se han planteado al inicio de la discusión de *GMM*. En Stata, asumiendo el modelo *SLM* (por elegir alguno), esto se implementa fácilmente:

```

. spivreg U2012 NM2012 (NM2012 = NM2010), d1(W5_st) id(id) het
  (omitted product)

```

4. Interpretación de los Resultados

De especial interés es la interpretación de los resultados obtenidos. El modelo *SLM*, y también el de Durbin, han sido de los modelos más difundidos en el campo aplicado posiblemente debido a la cuantificación de los efectos espaciales o de desbordamiento deducidos de su estimación. En las investigaciones pioneras la interpretación del modelo se realizaba casi exclusivamente considerando los valores de los coeficientes estimados para cada variable explicativa y en particular Wy . Un valor significativo de λ era clara evidencia de la importancia de los efectos de desbordamiento o contagio concluyendo el análisis en este punto. LeSage & Pace (2009) llaman la atención sobre la necesidad de profundizar la interpretación de estos modelos espaciales considerando las derivadas parciales del impacto de un cambio en una unidad de las variables explicativas.

Para ver más claramente el punto destacado por LeSage & Pace (2009), debe tenerse presente que los efectos desbordamiento en un modelo *SDM* (por considerar el más complejo) se deben a la presencia de Wy (efecto global) y WX (efecto local):

$$y = \lambda W y + X\beta + W X \gamma + u. \quad (39)$$

Supongamos que se desea analizar el impacto del incremento unitario de la k –ésima variable explicativa, este incremento es asumido igual para todas las localizaciones. Bajo un modelo de regresión no espacial, el efecto total sobre la variable dependiente será igual a la estimación del coeficiente β_k , ceteris paribus, sin importar la localización. Pero ante un modelo como el presentado en (39), el efecto total depende de las unidades vecinas de cada localización y de la magnitud de los coeficientes que acompañan a las variables espaciales.

El efecto total puede ser descompuesto en un efecto directo y otro indirecto. El efecto directo es el impacto del cambio de la variable explicativa sobre la variable dependiente en cada localidad, este efecto tenderá a ser similar al obtenido por un modelo de regresión no espacial si λ y γ_k son cercanos a 0. El efecto indirecto se debe a la dinámica espacial generada por la presencia de λ y γ que afectará a todas las unidades el modelo.

Bajo un modelo de Durbin pueden distinguirse dos tipos de efectos indirectos producto de la interdependencia entre las unidades. Uno es de tipo global, afectando a todas las unidades: y_i aumenta inicialmente β_{ik} unidades, y esto genera un nuevo incremento producto del cambio en las y_j 's de los J vecinos, capturado por el término endógeno $\lambda W y$. Es decir, para cada i –ésima localidad existirá un nuevo impacto igual a $\lambda \sum_{j \neq i}^J w_{ij} \beta_{jk}$ unidades, retroalimentando el cambio en la variable dependiente y que solo cesará al converger (asumiendo que $|\lambda| < 1$). Este efecto es identificado como un *efecto espacial global* para distinguirlo del *efecto espacial local* que surge del incremento unitario en la k –ésima variable explicativa de los J vecinos de la localidad i : $\sum_{j \neq i}^J w_{ij} \gamma_{jk}$. Este efecto espacial refuerza al efecto inicial β_k y es de tipo local en el sentido que no posee un efecto dinamizador como el generado por la presencia de λ . El efecto espacial local posee similitudes al estimado en otros contextos econométricos conocido como *efecto de los pares*.

En términos matriciales, y de forma más compacta, el *efecto marginal total* en el modelo *SDM* puede ser representado por la matriz de derivadas parciales de y respecto al cambio en una unidad de la k –ésima variable explicativa en X :

$$\begin{aligned} \left[\frac{\partial E(y)}{\partial x_{1k}} \quad \dots \quad \frac{\partial E(y)}{\partial x_{nk}} \right] &= \begin{bmatrix} \frac{\partial E(y_1)}{\partial x_{1k}} & \dots & \frac{\partial E(y_1)}{\partial x_{nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial E(y_n)}{\partial x_{1k}} & \dots & \frac{\partial E(y_n)}{\partial x_{nk}} \end{bmatrix}, \\ &= (I_n - \lambda W)^{-1} \begin{bmatrix} \beta_k & w_{12}\gamma_k & \dots & w_{1n}\gamma_k \\ w_{21}\gamma_k & \beta_k & \dots & w_{2n}\gamma_k \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}\gamma_k & w_{n2}\gamma_k & \dots & \beta_k \end{bmatrix}, \\ &= S [\beta_k I_n + \gamma_k W], \end{aligned} \quad (40)$$

donde w_{ij} es el elemento (i, j) –ésimo de W , β_k es el k –ésimo elemento del vector β , γ_k es el elemento k –ésimo del vector γ y S es igual a $(I_n - \lambda W)^{-1}$. Obsérvese que para la estimación de los efectos marginales se utiliza el valor esperado de la forma reducida del modelo, $E(y) = (I - \lambda W)^{-1} [X\beta + W X \gamma]$.

En la expresión (40) puede observarse la presencia del efecto directo y otro indirecto. El *efecto directo* es representado por los elementos de la diagonal principal de la ecuación (40). Este cambio

será diferente para cada región por lo que no existe un único efecto directo. Por su parte, el *efecto indirecto* (spatial spillover), proviene de los elementos fuera de la diagonal principal de (40) y, nuevamente, será diferente para cada región. En este caso se tiene una sumatoria de efectos para cada localidad.

En el caso de un *SLM*, $y = \lambda W y + X \beta + \varepsilon$, los *efectos marginales totales* se simplifican tal que:

$$\begin{aligned}
 \left[\frac{\partial y}{\partial x_{1k}} \quad \dots \quad \frac{\partial y}{\partial x_{nk}} \right] &= \begin{bmatrix} \frac{\partial y_1}{\partial x_{1k}} & \dots & \frac{\partial y_1}{\partial x_{nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_{1k}} & \dots & \frac{\partial y_n}{\partial x_{nk}} \end{bmatrix}, \\
 &= (I_n - \lambda W)^{-1} \begin{bmatrix} \beta_k & 0 & \dots & 0 \\ 0 & \beta_k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \beta_k \end{bmatrix}, \\
 &= \beta_k S,
 \end{aligned} \tag{41}$$

donde β_k es el k -ésimo elemento del vector β y S definido como en (40). El *efecto directo* será igual al que surge en el modelo *SDM*, pero el *efecto indirecto* se torna más simple siendo igual a la suma de los elementos fuera de la diagonal principal de $\beta_k S$, conteniendo solo efectos globales.

Ya sea que se estime un modelo *SLM* o de Durbin, siempre se tendrá efecto marginal que no es único para todas las localidades (en contraposición a lo que sucede bajo un modelo no-espacial). Debido a esta característica, LeSage & Pace (2009) proponen una medida resumen para cada uno de los efectos considerando valores promedio:

$$\begin{aligned}
 \overline{M}_{Total} &= \frac{\beta_k}{n} \mathbf{1}'_n S \mathbf{1}_n = \frac{\beta_k}{n} \sum_i \sum_j S_{i,j}. \\
 \overline{M}_{Directo} &= \frac{\beta_k}{n} \text{tr}(S) = \frac{\beta_k}{n} \sum_i S_{i,i}. \\
 \overline{M}_{Indirecto} &= \frac{\beta_k}{n} \mathbf{1}'_n S \mathbf{1}_n - \frac{\beta_k}{n} \text{tr}(S) = \frac{\beta_k}{n} \sum_i \sum_{j \neq i} S_{i,j}.
 \end{aligned}$$

La significancia de estos efectos puede ser obtenida mediante una simulación Monte Carlo por medio de shocks aleatorios en el término de error. Este tipo de procedimiento se ha implementado en programas como Matlab y R, pero no en el caso de Stata.

Es importante destacar que la interpretación de estos efectos debe realizarse con precaución ya que tienen un fuerte componente causal. En el modelo de Durbin, el efecto indirecto es generado por una combinación de efectos globales y locales. El efecto global generado por la presencia de λ ha sido interpretado en algunos casos como un parámetro causal (véase Gibbons & Overman, 2012, p. 175). Sin embargo, la interpretación de causalidad habitual supone una relación entre variables (por ejemplo, x causa a y) y el autoregresivo espacial captura solo la dinámica espacial de la propia variable endógena. La posición adoptada por Gibbons & Overman (2012) no parece ser adecuada porque sería como, por establecer un ejemplo paralelo, que bajo un proceso AR(1) en el tiempo, $y_t = \lambda y_{t-1} + \beta x_{t-1} + \varepsilon_t$, se interprete a λ como parámetro causal. Por su parte, el efecto local

capturado por el parámetro γ puede ser interpretado como un efecto causal, previo contraste de causalidad espacial. Para una discusión sobre este punto véase Herrera et al. (2014).

En Stata, Drukker et al. (2013d) proponen una secuencia de comandos para obtener la interpretación adecuada de los parámetros. A continuación aplicamos dicha secuencia.

En primer lugar, se estima el modelo final elegido, *SLM*:

```
. spreg ml U2012 NM2012, dl(W5_st) id(id)
(omitted product)
```

Drukker et al. (2013d) proponen un camino alternativo para obtener el efecto total de un cambio en la variable explicativa utilizando la predicción de la forma reducida mediante el comando `predict`:

```
. predict y0
(option rform assumed)
```

Asumiendo un cambio unitario en la migración neta (1 inmigrante neto por cada mil ha.):

```
. quietly replace NM2012 = NM2012 + 1
. predict y1
(option rform assumed)
```

Ahora el impacto total en la tasa de desempleo es igual a la diferencia entre la predicción previa y posterior al cambio unitario:

```
. gen delta_y = y1-y0
. sum delta_y
```

Variable	Obs	Mean	Std. Dev.	Min	Max
delta_y	164	-1.040091	4.15e-07	-1.040092	-1.040091

```
. quietly replace NM2012 = NM2012 - 1
```

Este resultado implica que el cambio total esperado es una disminución cercana a 1 punto porcentual en la tasa de desempleo. Obsérvese que el cambio unitario en la migración neta es similar al cambio promedio experimentado entre los años 2010 y 2012⁵. Este impacto sobre el desempleo no debe interpretarse como un impacto inmediato, ya que requiere que el efecto difusión sobre el espacio se complete implicando un periodo temporal de mediano a largo plazo.

Si se utilizan las fórmulas desarrolladas por LeSage & Pace (2009), entonces:

```
. spmat getmatrix W5_st W
. mata
----- mata (type end to exit) -----
: b = st_matrix("e(b)")
: b
```

	1	2	3	4
1	-.1898462498	2.40790211	.8174714987	8.182474325

```
: lambda = b[1,3]
: lambda
```

⁵Para observar esto, puede realizarse en Stata la siguiente secuencia: `gen dif_NM1210=NM2012-NM2010` y luego: `sum dif_NM1210`.

```
.8174714987
: S = luinv(I(rows(W))-lambda*W)
: end
```

Estos comandos permiten computar $S = (I - \hat{\lambda}W)^{-1}$ para utilizar en posteriores cálculos.

Ahora, puede computarse el efecto total propuesto por LeSage & Pace (2009) mediante el siguiente comando:

```
. mata: b = st_matrix("e(b)")
      -1.040090991
```

Este resultado es el mismo que se obtiene usando las predicciones. El cómputo del efecto directo se obtiene de la siguiente manera:

```
. mata: (b[1,1]/rows(W))*trace(S)
      -.2414164099
```

Este efecto es superior al coeficiente beta estimado, representando casi un 25 % del efecto total sobre el desempleo. El restante efecto es debido a los desbordamientos espaciales capturado por el efecto indirecto:

```
. mata: (b[1,1]/rows(W))*sum(S) - (b[1,1]/rows(W))*trace(S)
      -.7986745812
```

5. Resumen

En los últimos años el avance de los programas informáticos ha posibilitado que las técnicas de econometría espacial se encuentren disponibles para los economistas aplicados.

Esta guía ha buscado utilizar el programa Stata íntegramente para el análisis de datos de corte transversal. Las secciones desarrolladas han incluido el análisis espacial exploratorio como un primer paso para aproximarse a la modelización espacial. Este tipo de análisis brinda información importante para desarrollar posibles hipótesis y detectar datos anómalos. En la sección 3, se presentó una batería de contrastes espaciales que permitieron definir la adecuada especificación espacial. Además, se presentaron las alternativas de estimación dependiendo del tipo de especificación elegida. La revisión ha mostrado las principales especificaciones así como los métodos usuales de estimación. Finalmente, la sección 4 mostró una discusión sobre la interpretación de los resultados finales.

Por otra parte, el documento no compara los resultados obtenidos por Stata respecto a otros programas estadísticos. Otras investigaciones han realizado dicha comparación respecto a Matlab, R y Pysal (Bivand & Piras, 2013). Algunos investigadores solo se han enfocado en un modelo en particular, *SEM*, explorando las alternativas de estimación de Stata respecto a GeoDaSpace, Pysal y R (Anselin et al., 2012). Consúltense estas referencias para conocer las diferencias entre los programas disponibles.

Referencias

- Anselin, L. (1988). *Spatial econometrics: Methods and models*, volume 4. Springer.
- Anselin, L. (1992). Spacestat tutorial: A workbook for using SpaceStat in the analysis of spatial data. *Urbana-Champaign: University of Illinois*, 263.
- Anselin, L. (1999). Interactive techniques and exploratory spatial data analysis. In P. Longley, M. Goodchild, D. Maguire, & D. Rhind (Eds.), *Geographical Information Systems: Principles, Techniques, Management and Applications*. Cambridge: Geoinformation Int.
- Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in regional science*, 89(1), 3–25.
- Anselin, L., Amaral, P. V., & Arribas-Bel, D. (2012). Technical aspects of implementing GMM estimation of the spatial error model in PySAL and GeoDaSpace. *GeoDa Center for Geospatial Analysis and Computation*.
- Anselin, L. & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs*, 155, 237–290.
- Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional science and urban economics*, 26(1), 77–104.
- Anselin, L. & Florax, R. J., Eds. (1995). *New directions in spatial econometrics*. Springer, New York.
- Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical analysis*, 38(1), 5–22.
- Arraiz, I., Drukker, D. M., Kelejian, H. H., & Prucha, I. R. (2010). A spatial cliff-ord-type model with heteroskedastic innovations: Small and large sample results. *Journal of Regional Science*, 50(2), 592–614.
- Bartels, C. & Ketellapper, R. (1979). *Exploratory and explanatory statistical analysis of spatial data*. Martinus Nijhoff, Boston, MA.
- Bennett, R. (1979). *Spatial time series*. Pion, London.
- Bivand, R. (2002). Spatial econometrics functions in R: Classes and methods. *Journal of geographical systems*, 4(4), 405–421.
- Bivand, R. & Piras, G. (2013). *Comparing Implementations of Estimation Methods for Spatial Econometrics*. Technical report, Regional Research Institute, West Virginia University.
- Burridge, P. (1980). On the cliff-ord test for spatial correlation. *Journal of the Royal Statistical Society. Series B (Methodological)*, (pp. 107–108).
- Cliff, A. & Ord, K. (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, 4(3), 267–284.
- Crow, K. (2013). Shp2dta: Stata module to converts shape boundary files to stata datasets. *Statistical Software Components*.

- Drukker, D. M., Egger, P., & Prucha, I. R. (2013a). On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors. *Econometric Reviews*, 32(5-6), 686–733.
- Drukker, D. M., Peng, H., Prucha, I. R., & Raciborski, R. (2013b). Creating and managing spatial-weighting matrices with the `spmat` command. *Stata Journal*, 13(2), 242–286.
- Drukker, D. M., Prucha, I. R., & Raciborski, R. (2013c). A command for estimating spatial-autoregressive models with spatial-autoregressive disturbances and additional endogenous variables. *The Stata Journal*, 13(2), 287–301.
- Drukker, D. M., Prucha, I. R., & Raciborski, R. (2013d). Maximum-likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. *Stata Journal*, 13(2), 221–241.
- Elhorst, J. P. (2014). *Spatial Econometrics. From Cross-sectional data to Spatial Panels*. SpringerBriefs in Regional Science, Springer.
- Geary, R. C. (1954). The contiguity ratio and statistical mapping. *The incorporated statistician*, (pp. 115–146).
- Getis, A. & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical analysis*, 24(3), 189–206.
- Gibbons, S. & Overman, H. G. (2012). Mostly pointless spatial econometrics? *Journal of Regional Science*, 52(2), 172–191.
- Goodchild, M., Haining, R., & Wise, S. (1992). Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems*, 6(5), 407–423.
- Griffith, D. A. (1996). Some guidelines for specifying the geographic weights matrix contained in spatial statistical models. *Practical handbook of spatial statistics*, (pp. 65–82).
- Herrera, M., Mur, J., & Ruiz, M. (2014). Detecting causal relationships between spatial processes. *Papers in Regional Science*.
- Hordijk, L. (1979). Problems in estimating econometric relations in space. *Papers in Regional Science*, 42(1), 99–115.
- Jeanty, P. W. (2012). `Splagvar`: Stata module to generate spatially lagged variables, construct the moran scatter plot, and calculate moran's *i* statistics. *Statistical Software Components*.
- Jeanty, P. W. (2014). `Spwmatrix`: Stata module to generate, import, and export spatial weights. *Statistical Software Components*.
- Kelejian, H. H. & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1), 99–121.
- Kelejian, H. H. & Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, 40(2), 509–533.

- Kelejian, H. H. & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1), 53–67.
- LeSage, J. & Pace, R. (2009). *Introduction to spatial econometrics*. CRC press.
- LeSage, J. P. (1999). *Spatial econometrics*. Regional Research Institute, West Virginia University.
- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, (pp. 17–23).
- Mur, J. & Angulo, A. (2009). Model selection strategies in a spatial setting: Some additional results. *Regional Science and Urban Economics*, 39(2), 200–213.
- Ord, J. (1981). Towards a theory of spatial statistics: a comment. *Geographical Analysis*, 13(1), 86–91.
- Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70(349), 120–126.
- Paelinck, J. & Klaassen, L. (1979). *Spatial econometrics*. Saxon House, Farnborough.
- Pisati, M. (2001). Sg162: Tools for spatial data analysis. *Stata Technical Bulletin*, 60, 21–37.
- Pisati, M. (2008). Spmap: Stata module to visualize spatial data. *Statistical Software Components*.
- Rey, S. J. & Anselin, L. (2010). Pysal: A Python library of spatial analytical methods. In F. M. & A. Getis (Eds.), *Handbook of applied spatial analysis* (pp. 175–193). Springer.
- Slocum, T., McMaster, R., Kessler, F., & Howard, H. (2009). *Thematic cartography and geovisualisation*. Prentice Hall, New Jersey, NJ.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography*, (pp. 234–240).