

UNIVERSIDAD DE SANTIAGO DE COMPOSTELA



**Geografía electoral: una radiografía
de las elecciones autonómicas de
Andalucía de 2018.**

*Xeografía electoral: unha radiografía das eleccións
autonómicas de Andalucía de 2018*

*Electoral geography: an x-ray of the Andalusian regional
elections of 2018.*

Autor: Castiñeira Vázquez, Pablo

Director(a) del trabajo: Pereira López, María

Grado en Ciencias Políticas y de la Administración

Julio 2020

Trabajo de Fin de Grado

Grado de Ciencias Políticas y de la Administración cursado en la Universidad de
Santiago de Compostela

AGRADECIMIENTOS

Después de un intenso período hoy llega a su fin una etapa más. Escribo este apartado de agradecimientos para finalizar este ciclo. Ha sido un proceso de aprendizaje intenso tanto a nivel académico como personal. Escribir este trabajo me ha llevado a descubrir nuevas disciplinas de la ciencia política que hasta ahora no concebía y es por eso que me gustaría agradecer a todas aquellas personas que me han ayudado y apoyado durante este proceso, desde compañeros de todos los cursos a profesorado, en especial a mi tutora en este proceso, Dra. María Pereira López, por darme las piezas necesarias para adentrarme en este tema de investigación. Tampoco olvidar a mi familia quienes han sido apoyos fundamentales en el proceso.

Castiñeira Vázquez, Pablo

Santiago de Compostela, julio de 2020

RESUMEN | La geografía electoral, como rama de la geografía política, nos brinda la oportunidad de explicar las orientaciones y preferencias de voto de los electores en tanto que organización territorial ocurren los comicios electorales. Este tipo de estudios parten de la premisa de que es el comportamiento de voto individual está condicionado de alguna manera por las condiciones que se dan en un determinado espacio/territorio, es decir, son dependientes del contexto.

Este trabajo trata de explicar la existencia de una geografía del voto, mediante el software diseñado por el Dr. Luc Anselin, *GeoDa* y mediante el Análisis Exploratorio de Datos Espaciales (AEDE) y posterior Análisis Confirmatorio de Datos Espaciales (ACDE), en las elecciones al Parlamento andaluz celebradas el domingo 2 de diciembre del año 2018 para elegir a los 109 diputados de la XI legislatura del Parlamento autonómico. Los hallazgos aquí encontrados permiten ver como existen ciertas tendencias de voto a determinados partidos en función de las características del contexto en el que se desenvuelvan los individuos.

Palabras clave: Geografía electoral, GeoDa, AEDE, elecciones, Andalucía, voto, ACDE.

RESUMO | A xeografía electoral, como rama da xeografía política, dámos a oportunidade de explicar as orientacións e preferencias dos votos dos electores, no tocante á organización territorial das eleccións electorais. Este tipo de estudos supón que o comportamento do voto individual está condicionado dalgún xeito polas condicións que se dan nun determinado territorio, é dicir, non son independentes do contexto.

Este traballo trata de explicar a existencia dunha xeografía do voto, a través do software deseñado por Dr. Luc Anselin, *GeoDa* e a través da Análise Exploratoria de Datos Espaciais (AEDE) e posterior Análise Confirmatorio de Datos Espaciais (ACDE), nas eleccións ao parlamento andaluz celebradas o domingo 2 de decembro, no ano 2018 para elixir aos 109 deputados da XI lexislatura do Parlamento autonómico. Os achados atopados aquí permiten ver como hai certas tendencias de voto para certos partidos segundo as características do contexto no que os individuos operan.

Palabras clave: Xeografía electoral, *GeoDa*, AEDE, eleccións, Andalucía, voto, ACDE.

ABSTRACT | Electoral geography, as a branch of political geography, offers us the opportunity to explain the orientations and voting preferences of voters as territorial organization of electoral elections. These types of studies start from the premise that the behavior of individual voting is conditioned in some way by the conditions that exist in a given space / territory, that is, they are dependent on the context.

This work tries to explain the existence of a voting geography, through the software designed by Dr. Luc Anselin, GeoDa and through the Exploratory Analysis of Spatial Data (EASD) and subsequent Confirmatory Analysis of Spatial Data (CASD), in the elections to the Andalusian Parliament held on Sunday, December 2, 2018 to elect the 109 deputies of the XI legislature of the autonomous Parliament. The findings found here allow us to see how certain voting trends exist for certain parties depending on the characteristics of the context in which the individuals operate.

Keywords: Electoral geography, *GeoDa*, EASD, elections, Andalusia, vote, CASD.

ÍNDICE DE CONTENIDO

CAPÍTULO 1. Bases teóricas y objetivos.....	1
1. INTRODUCCIÓN.....	1
1.1 Objetivo y preguntas clave	1
1.2 Estructura del trabajo.....	2
1.3 ¿Por qué estas elecciones?.....	2
2. REVISIÓN BIBLIOGRÁFICA	5
2.1 Geografía Política.....	5
2.2 ¿Qué es la Geografía electoral? Consideraciones teóricas	5
2.3 Limitaciones de la GE. Mapas electorales.....	9
2.4 Evolución de la disciplina y dónde está ahora.....	10
2.5 La relevancia del contexto y econometría espacial como base	13
2.6 La clave es el votante (“efecto vecindario”).....	16
CAPÍTULO 2. Bases metodológicas	18
1. METODOLOGÍA A UTILIZAR.....	18
1.1 Diseño de la estrategia de trabajo	18
1.2 Instrumentos de investigación	18
1.3 Métodos de investigación: el AEDE y ACDE.....	20
1.4 Variables a utilizar.....	21
CAPÍTULO 3. Desarrollo del análisis de caso.....	26
1 ANÁLISIS DE CASO	26
1.1 Análisis del ciclo electoral.....	26
2. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES.....	27
2.1 Análisis LISA	30
2.1.1 Índice I de Moran local	30
2.1.2 Factores económicos como explicación. Renta y paro	32
2.1.3 Edad como factor explicativo.....	44
2.1.4 ¿Influye la nacionalidad?	48

3. ANÁLISIS CONFIRMATORIO DE DATOS ESPACIALES.....	52
3.1 Partido Socialista Obrero Español (PSOE)	54
3.2 Partido Popular	58
3.3 Ciudadanos (Cs)	61
3.4 Vox	63
3.5 Adelante Andalucía	65
4. CONCLUSIONES.....	71
BIBLIOGRAFÍA	74
1. Referencias en papel:.....	74
2. Páginas web:	76
ANEXOS.....	77
1. Legislación.....	77
2. Relativo a conceptos	78
3. Modelos de regresión espacial redefinidos en términos de error y de retardo	79
3.1 Partido Socialista Obrero Español.....	79
3.2 Partido Popular	81
3.3 Ciudadanos	83
3.4 Vox	85
3.5 Adelante Andalucía	87

INDICE DE TABLAS

Tabla 1. Resultados de las Elecciones al Parlamento de Andalucía en el período autonómico 1982-2018.....	4
Tabla 2. Descripción de variables	23
Tabla 3. Elecciones 2015 & 2018	27
Tabla 4. I de Moran Local Bivariante (retardo espacial: rentamedia)	38
Tabla 5. Coeficientes de I Morán (retardos: “porc65y” & “menos20”).....	44
Tabla 6. Coeficientes de I Morán -con criterios de contigüidad 2- (retardos: “porc65y” & “menos20”).....	44
Tabla 7. Modelo de regresión lineal básico por MCO; (VD:Tasa_PSOE)	55
Tabla 8. Comparativas modelos	58
Tabla 9. Modelo de regresión lineal básico por MCO; (VD: Tasa_PP)	59
Tabla 10. Comparativa de modelos.....	60
Tabla 11. Modelo de regresión lineal básico por MCO; (VD: Tasa_Cs).....	62
Tabla 12. Comparativa de modelos.....	63
Tabla 13. Modelo básico de regresión lineal por MCO; (VD: Tasa_Vox)	64
Tabla 14. Comparativa de modelos.....	65
Tabla 15. Modelo básico de regresión lineal por MCO; (VD: Tasa_AA)	66
Tabla 16. Comparativa de modelos.....	67
Tabla 17. Comparativas modelos espaciales (I).....	68
Tabla 18. Comparativa de modelos espaciales (II)	69
Tabla 19. Recordatorio Coeficientes I de Morán	69
Tabla 20. Modelo redefinido en términos de error; (VD: Tasa_PSOE).....	79
Tabla 21. Modelo redefinido en términos de retardo; (VD: Tasa_PSOE).....	80
Tabla 22. Modelo redefinido en términos de retardo espacial; (VD: Tasa_PP)	81
Tabla 23. Modelo redefinido en términos de error; (VD: Tasa_PP).....	82
Tabla 24. Modelo redefinido en términos de error; (VD: Tasa_Cs)	83
Tabla 25. Modelo redefinido en términos de retardo espacial; (VD: Tasa_Cs).....	84
Tabla 26. Modelo redefinido en términos de error (VD: Tasa_Vox).....	85
Tabla 27. Modelo redefinido en términos de retardo espacial (VD: Tasa_Vox).....	86
Tabla 28. Modelo redefinido en términos de error; (VD: Tasa_AA).....	87
Tabla 29. Modelo redefinido en términos de retardo espacial; (VD: Tasa_AA)	88

INDICE DE ILUSTRACIONES

Ilustración 1. Geografía electoral: puente entre lo cualitativo y cuantitativo	8
Ilustración 2. Autocorrelación espacial positiva	15
Ilustración 3. Autocorrelación espacial negativa	15
Ilustración 4. Shapefile de municipios andaluces	19
Ilustración 5. Distribución del total de población	28
Ilustración 6. Distribución del porcentaje de extranjeros.....	28
Ilustración 7. Distribución de edad mayores de 20	28
Ilustración 8. Distribución edad mayores de 65	28
Ilustración 9. Distribución renta neta declarada por habitante.....	29
Ilustración 10. Distribución Tasa de paro municipal	29
Ilustración 11. Criterios de vecindad.....	30
Ilustración 12. Mapa de clusters I de Morán Tasa_PSOE w/ rentamedia.....	33
Ilustración 13. Gráfico de dispersión I de Morán Tasa_PSOE w/ rentamedia	33
Ilustración 14. Mapa de clusters I de Morán Tasa_PP w/ rentamedia	34
Ilustración 15. Diagrama de dispersión I de Morán Tasa_PP w/ rentamedia	34
Ilustración 16. Diagrama de dispersión I de Morán Tasa_Cs w/ rentamedia	35
Ilustración 17. Mapa de clusters I de Morán Tasa_Cs w/ rentamedia	35
Ilustración 18. Diagrama de dispersión I de Morán Tasa_Vox w/ rentamedia.....	36
Ilustración 19. Mapa de clusters I de Morán Tasa_Vox w/ rentamedia.....	36
Ilustración 20. Gráfico de dispersión I de Morán Tasa_AA w/ rentamedia	37
Ilustración 21. Mapa de clusters Tasa_AA w/ rentamedia.....	37
Ilustración 22. Mapa cluster PSOE w/ rentamedia	38
Ilustración 23. Mapa cluster PP w/ rentamedia.....	38
Ilustración 24. Mapa cluster Cs w/ rentamedia	39
Ilustración 25. Mapa cluster Vox w/ rentamedia	39
Ilustración 26. Mapa cluster AA w/ rentamedia	39
Ilustración 27. Mapa de desviación estándar de la variable rentamedia	39
Ilustración 28. Diagrama de dispersión I de Morán Tasa_PSOE w/ Paro_Mun.....	40
Ilustración 29. Diagrama de dispersión I de Morán Tasa_PP w/ Paro_Mun	40
Ilustración 30. Diagrama de dispersión I de Morán Tasa_Cs w/ Paro_Mun	41
Ilustración 31. Diagrama de dispersión I de Morán Tasa_Vox w/ Paro_Mun.....	41
Ilustración 32. Diagrama de dispersión I de Morán Tasa_AA w/ Paro_Mun.....	41
Ilustración 33. Mapa de clusters I de Morán Tasa_PSOE w/ Paro_Mun.....	42

Ilustración 34. Mapa de clusters I de Morán Tasa_PP w/ Paro_Mun	42
Ilustración 35. Mapa de clusters I de Morán Tasa_Cs w/ Paro_Mun	43
Ilustración 36. Mapa de clusters I de Moran Tasa_Vox w/ Paro_Mun.....	43
Ilustración 37. Mapa de clusters I de Moran Tasa_AA w/ Paro_Mun.....	43
Ilustración 38. Mapa de clusters Tasa_PSOE w/ Porc_65y	46
Ilustración 39. Mapa de clusters Tasa_PP w/ Porc_65y	46
Ilustración 40. Mapa de clusters Tasa_Cs w/ Porc_65y.....	46
Ilustración 41. Mapa de clusters Tasa_Vox w/ Porc_65y	46
Ilustración 42. Mapa de clusters Tasa_AA w/ Porc_65y	46
Ilustración 43. Mapa de clusters Tasa_PSOE w/ menos20.....	47
Ilustración 44. Mapa de clusters Tasa_PP w/ menos20	47
Ilustración 45. Mapa de clusters Tasa_Cs w/ menos20	47
Ilustración 46. Mapa de clusters Tasa_Vox w/ menos20.....	47
Ilustración 47. Mapa de clusters Tasa_AA w/ menos20	47
Ilustración 48. Diagrama de dispersión de Morán Tasa_PSOE w/ porc_extra.....	48
Ilustración 49. Mapa de clusters de Morán Tasa_PSOE w/ porc_extra.....	48
Ilustración 50. Diagrama de dispersión I Morán Tasa_PP w/ porc_extra.....	49
Ilustración 51. Mapa de clusters I de Morán Tasa_PP w/ porc_extra.....	49
Ilustración 52. Diagrama de dispersión I de Morán Tasa_Cs w/ porc_extra	49
Ilustración 53. Mapa de clusters de Morán Tasa_Cs w/ porc_extra	49
Ilustración 54. Diagrama de dispersión de Morán Tasa_Vox w/ porc_extra.....	50
Ilustración 55. Mapa de clusters de Morán Tasa_Vox w/ porc_extra.....	50
Ilustración 56. Diagrama de dispersión de Morán Tasa_AA w/ porc_extra	51
Ilustración 57. Mapa de clusters de Morán Tasa_AA w/ porc_extra.....	51
Ilustración 58. Proceso de selección de modelo.....	53

SUMARIO DE ABREVIATURAS

1. G.E (Geografía electoral)
2. AEDE (Análisis Exploratorio de Datos Espaciales)
3. *Shape* (en referencia a los archivos *shapefile*)
4. PSOE (Partido Socialista Obrero Español)
5. PP (Partido Popular)
6. Cs (Ciudadanos)
7. AA (Adelante Andalucía)
8. ACDE (Análisis Confirmatorio de Datos Espaciales)

CAPÍTULO 1. Bases teóricas y objetivos.

1. INTRODUCCIÓN

El presente trabajo se enfoca en determinar si existen o no variables sociodemográficas que influyan en el voto a los cuatro principales partidos que compusieron el Parlamento andaluz tras los comicios celebrados el 02/12/18, considerando así el efecto espacial.

El votante es uno de los actores más importantes (sino el que más) en una democracia. No se entiende una democracia plena sin ellos, por ello, entender al elector en sus niveles individual y colectivo nos puede llevar a entender su decisión de voto y a determinar qué variables pueden haber incidido en hacer efectivo el derecho al sufragio y por qué partido votar (Charles-Leija, García, & Valadez, 2015)

1.1 Objetivo y preguntas clave

Por todo ello en este trabajo el objetivo principal va a ser determinar las características sociodemográficas del voto en las elecciones andaluzas del 2018 asumiendo factores contextuales. Para ello debemos tratar de dar respuesta a una serie de preguntas de investigación: ¿se pueden correlacionar factores económicos de un municipio con el voto? ¿es la edad una variable fundamental? ¿varía el voto a determinados partidos, en función de la cantidad de personas extranjeras residentes en un municipio?

Con las nuevas metodologías y herramientas que emanan cada año de la disciplina podemos permitirnos no solo una descripción de los hechos como ocurría en los primeros años del nacimiento de la geografía electoral. La principal empresa, por ende, del presente trabajo es determinar el comportamiento electoral, de una serie de individuos, en un área geográfica concreta y en un momento dado (las elecciones al Parlamento andaluz del año 2018).

Para la consecución de las bases teóricas se ha procedido a realizar una revisión bibliográfica de trabajos realizados mediante técnicas similares a la empleada en éste, además de los “artículos” con la información correspondiente y necesaria para la construcción de base informativa acerca de lo qué es y para qué sirve la geografía electoral y sus herramientas metodológicas.

El hallazgo principal de esta investigación se puede considerar el hecho de que existen ciertas tendencias de voto influidas por las variables con las que en este trabajo se aplican para el análisis haciendo ver que existen tendencias en la orientación de voto hacia ciertos partidos

según el contexto. Aunque esto es evidente que es más difícil asegurarlo cuanto más alejados estén unos municipios de otros, pero sí que se encuentra dependencia espacial en algunas de las variables.

1.2 Estructura del trabajo

Para esclarecer los datos presentados, el trabajo cuenta con una división por capítulos dentro de los cuales se recoge la información correspondiente a cada uno de ellos con sus subapartados pertinentes presentados en el Índice de contenido.

Así pues, el primer capítulo de los tres totales existentes corresponde a la presentación del trabajo, su justificación, contextualización dentro de la disciplina y presentación de objetivos del mismo. En el segundo (“-*Bases metodológicas*-”) se verán cuáles de las herramientas correspondientes a la geografía electoral se utilizarán para la correcta realización del trabajo. Por su parte en el tercero (“-*Desarrollo de análisis de caso*-”) se presentarán los resultados obtenidos gracias a la metodología elegida, finalizando en el último apartado, con las conclusiones definitivas del trabajo.

1.3 ¿Por qué estas elecciones?

Estos comicios tienen un interés politológico alto, pues desde la I hasta la X legislatura, el resultado de las elecciones autonómicas en Andalucía se había decantado por el bando socialista¹. En esta última el resultado fue diferente (obsérvese tabla 1). El partido con más votos obtenidos volvió a ser el *Partido Socialista Obrero Español*, pero esta vez los pactos cayeron del lado de la derecha, a diferencia del 2012, y es que el *Partido Popular*, *Ciudadanos* y *Vox*, llegaron a un acuerdo con el que acabaron con la hegemonía del PSOE tras años y años de mandato y luego de los ya famosos escándalos de corrupción.

¹ En las elecciones del 2012 el partido que se alzó con más cantidad de votos fue el Partido Popular, sin embargo, se quedaron a 5 escaños de la mayoría absoluta y un pacto del PSOE con IU hizo que siguiese la hegemonía socialista en el poder otra legislatura más.

Dentro de este pacto procedente del *lado derecho* destaca la irrupción del partido político *Vox*, lo cual no asombra si nos ponemos unas lentes europeas, pero en el lado español era la primera vez que la derecha radical ² tenía representación en algún tipo de Parlamento. A día de hoy, suele ser de interés para los científicos sociales y políticos el estudio del auge de *Vox*, y suelen plantearse algunas preguntas de gran interés como: ¿por qué en ese momento? y ¿por qué surge inicialmente en Andalucía? ¿hay algún factor que determine este hecho? Pero esto se les deja a otros estudios, como ya se dijo, el presente trabajo se va a focalizar en otro tipo de análisis, -tratar de explicar cuál es la relación entre el voto y las características de una determinada área geográfica-.

² Se sigue esta terminología y no se utiliza “ultaderecha” o “externa derecha” en función de los últimos trabajos de C. Mudde (2019) en donde se dirige a *Vox* en este término.

Otros autores también caen en diferenciar dos corrientes: “la extrema derecha” o “vieja extrema derecha” y la nueva extrema derecha o “derecha radical populista” (Ignazi, 1992). Pippa Norris por su parte también alude al uso del término “derecha radical”

Tabla 1. Resultados de las Elecciones al Parlamento de Andalucía en el período autonómico 1982-2018

Elecciones al Parlamento de Andalucía							
Legislatura	Participación (%)	Año	Partido más votado	Votos	%	Escaños / May. Absoluta	Partido(s) resultante(s)
I	66,31	1982	PSOE	1.496.522	52,10	66 / 55	PSOE
II	70,55	1986	PSOE	1.581.354	46,53	60 / 55	PSOE
III	54,72	1990	PSOE	1.368.529	49,61	62 / 55	PSOE
IV	67,28	1994	PSOE	1.395.132	38,71	45 / 55	PSOE
V	77,94	1996	PSOE	1.903.160	44,40	52 / 55	PSOE
VI	68,71	2000	PSOE	1.790.653	44,18	52 / 55	PSOE
VII	75,85	2004	PSOE	2.260.545	50,36	61 / 55	PSOE
VIII	72,67	2008	PSOE	2.148.328	48,41	56 / 55	PSOE
IX	60,78	2012	PP	1.570.833	40,66	50 / 55	PSOE + IU
X	63,94	2015	PSOE	1.408.566	35,43	47 / 55	PSOE
XI	58,65	2018	PSOE	1.009.243	33	33 / 55	PP + Cs + VOX

Fuente: elaboración propia a partir de datos oficiales del Ministerio del Interior.

2. REVISIÓN BIBLIOGRÁFICA

A la hora de buscar información para responder a las demandas de la investigación, se ha procedido a revisar y posteriormente comentar las indagaciones previas de los estudiosos destacados de la disciplina. Con toda esa cantidad de información se trata de responder claramente en los siguientes apartados a preguntas cuyas soluciones clarifiquen las bases sobre las que se sustenta el presente trabajo.

2.1 Geografía Política

En el resumen previo se han mencionado un par de conceptos que por el bien de la continuidad del trabajo deben aclararse. “*Geografía electoral como rama de la geografía política*”; este segundo concepto hace referencia a la parte de la geografía humana³ que trata de la distribución del territorio con respecto al espacio que ocupa el ser humano. Su principal área de estudio son las instituciones políticas y todos los aspectos afines a ellas, aunque, el espacio geográfico es también un tema de especial interés en esta disciplina, tanto que se puede llegar a observar una rama dentro de ella, construyendo así la geografía electoral.

2.2 ¿Qué es la Geografía electoral? Consideraciones teóricas

“Hoy más que nunca, los ciudadanos hacemos elección de nuestros gobernantes, pero no escogemos las condiciones ni los lugares en los que dicha elección se hace”
(Sonnleinter, 2013)

La mayoría de los enfoques teórico-metodológicos del comportamiento político se centra hoy en día en el análisis de los individuos (Sonnleinter, 2013). Los principales modelos explicativos del voto tienen en cuenta las preferencias de los individuos agregadas sin tener en cuenta lo que pueden suponer los aspectos geográficos, socioculturales y/o contextuales. Así pues, asumen que los procesos sociopolíticos son solamente la suma de las decisiones individuales, condicionadas por las categorías sociológicas (edad, género, ingresos, etc) sin tener en cuenta dinámicas grupales y espaciales.

³ La geografía humana como disciplina se encarga del estudio las sociedades humanas desde una perspectiva espacial, la relación entre estas sociedades y el medio físico en el que habitan, así como los paisajes culturales y las regiones humanas que éstas construyen.

No obstante, otros muchos estudios asumen que el voto también es una conducta social y colectiva en donde elementos espaciales pueden intervenir con un grado de importancia importante. El trabajo de Darmofal (2006), uno de los más importantes de este siglo, da cuenta de la existencia de ciertos factores políticos que han jugado un papel crítico en la producción de diferencias a nivel macro de participación entre regímenes de alta participación como la posguerra civil y bajos regímenes de participación, como el sur posterior a la reconstrucción. En este estudio se desafía a algunos de los supuestos con los que se miran los sistemas de participación política americana tratando de analizar la participación política a nivel macro utilizando técnicas espaciales, rompiendo con las tendencias de estudio de voto nivel micro, determinando la existencia de factores sociopolíticos dependiendo del área.

Así a todo, esto no significa que nuestros comportamientos políticos estén predeterminados, ni mucho menos, sobre todo en las épocas que corren en donde la afiliación e identificación partidista está en continuo debilitamiento, estamos ante un aumento de la fragmentación política, de votos de “sanción”, voto “útil” y votantes volátiles, pero asumir que en el voto solo influyen factores individuales, es quedarse a medio camino. En suma, no escogemos las condiciones en las que operamos, pero hoy más que nunca los ciudadanos hacemos la elección de nuestros gobernantes.

La geografía electoral emerge como subdisciplina de las ciencias sociales, situándose entre la geografía, la historia, la antropología, la sociología y ciencia política y tiene como objeto de estudio la dimensión espacial de los procesos político electorales, en especial el análisis del voto como un acto social territorializado (Sonnleinter, 2013), aun así, en palabras de Michel Bussi (1998) esta se trata de una subdisciplina segmentada y despreciada a lo largo de la historia. En sus inicios fue desarrollada por eruditos alejados del ámbito de la geografía, renovándose posteriormente desde dicha disciplina, para ser marginada por la ciencia política en su etapa institucionalista y por la sociología electoral (cada vez más individualista). En la primera mitad del siglo XX fue desplazada por la innovación de las encuestas de opinión, enfoques psicosociales y racionales del voto, desde la ciencia política y no es hasta 1975 cuando se redescubre y se enriquece de los nuevos enfoques de la geografía política y humana, abriendo camino a novedosas técnicas y campos en la geografía electoral como el AEDE y la cartografía.

Tenemos, por tanto, dos vertientes a la hora de analizar el voto en ciencias sociales. La primera se centra en el comportamiento electoral a partir de datos agregados en unidades territoriales, en diferentes niveles y escalas de la organización territorial. A medida que nos

vamos interesando en el contexto y en el entorno donde se desarrolla la acción de votar, se le denomina como “análisis ecológico”, metodología base de la geografía electoral. Este es un enfoque territorial y colectivo que contrasta con la segunda vertiente más individualista sustentada en el análisis de las encuestas de opinión omitiendo la información que nos brinda el contexto o espacio.

Joaquín Bosque Sendra, uno de los adalides en investigaciones geográficas en España nos define la Geografía Electoral como “(...) una subdisciplina geográfica -muy minoritaria en cuanto a números de practicantes- como una de las partes de la Geografía Humana que más se aproxima a una verdadera ciencia “dura” semejante a cualquier ciencia experimental” (Bosque Sendra, 1998). A su vez otros autores nos la definen como “un método cartográfico para definir el reparto regional de las fuerzas partidarias (...) nos permite identificar las zonas donde tienen influencia los Partidos y, de esta manera, observar las relaciones entre las tendencias políticas y las características de la población que vive en esas zonas geográficas” (Gómez Tagle y Valdez, 2000)

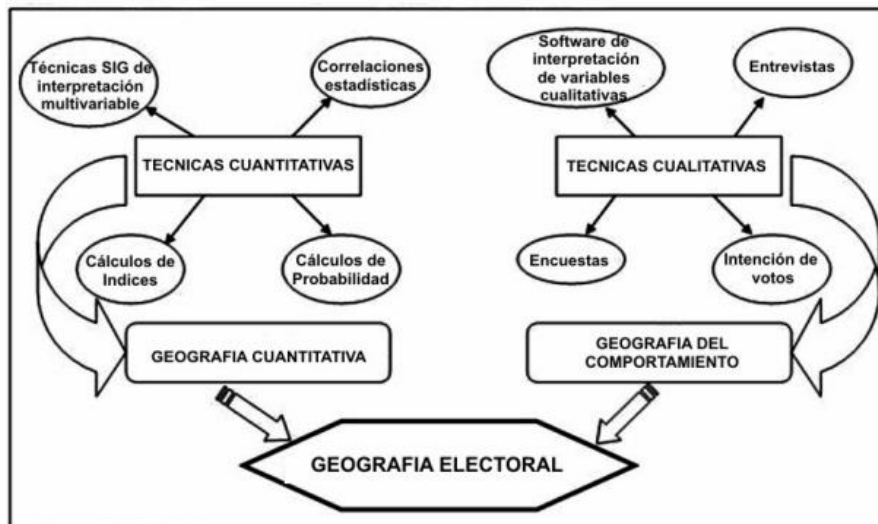
Como vemos esta “estudia el proceso electoral que valida la continuidad del aparato estatal poniendo en función a sus autoridades” (Monzón, 2013). Los estudios de la geografía electoral se diferencian de otras ramas de la geografía política en que estos ponen énfasis en el comportamiento electoral de los ciudadanos, en un momento electoral concreto y en relación a un espacio geográfico específico, es decir, desde esta disciplina se pretende entender que relación existe entre la manera de votar de los ciudadanos y las características del contexto en el que votan.

Monzón nos cuenta que la geografía electoral nace como un producto de la revolución cuantitativa y que, a lo largo de su vida, esta disciplina se ha analizado desde dos puntos de vista. El enfoque corológico y el ecológico. El primero de ellos da cuenta de que la geografía electoral en sus inicios, quedaba enfocada únicamente a un aspecto puramente estadístico, casi como si fuese una ciencia dura, y con el tiempo se vio la necesidad de explicar el porqué de la distribución espacial del voto. Se necesitaba responder a preguntas tan lógicas como ¿Por qué este electorado voto de esta manera? ¿Por qué tal candidato perdió por tanta diferencia?

Fue así como se empezó a hacer cruces con la situación social, cultural y económica de los ciudadanos dando paso al segundo enfoque (Bosque, 1988). El enfoque *ecológico* combina el uso de técnicas cuantitativas y cualitativas pretendiendo hallar una explicación del

comportamiento electoral de un grupo específico votantes, en un momento determinado y para un espacio geográfico delimitado.

Ilustración 1. Geografía electoral: puente entre lo cualitativo y cuantitativo.



Fuente: (Monzón, 2013)

Es por eso que no debemos olvidar que la construcción del voto de cada elector se realiza de manera individual y personal, en donde interviene ya no solo la ideología, sino también la carga subjetiva de sus rasgos intelectuales, culturales, socioeconómicos y del compromiso con su espacio vital (esto es barrio, escuela, iglesia, etc). El elector recoge y analiza de alguna manera el conjunto de todos estos rasgos y le da su visión electoral. Fue Bosque (1988) quien clasificó dicha subjetividad que influye en el comportamiento electoral en dos grupos de factores: los primeros los llamados individuales y los segundos factores contextuales y es la suma de estos dos de donde surge la decisión de voto. Los primeros hacen referencia a circunstancias individuales de cada elector. Por su parte los factores referentes al segundo tipo se refieren a las condiciones medioambientales geográficas del entorno del dicho elector. La diferencia entre ellos según el propio autor radica en que los llamados factores individuales se refieren al individuo o elector o a un grupo de ellos que no tienen por qué estar necesariamente distribuidos homogéneamente en el espacio. Por su parte los factores contextuales implican la idea de continuidad espacial, esto es hacen referencia a unos mismos elementos mantenidos en un mismo contexto o espacio.

A día de hoy, están disponibles diversos trabajos que pueden afirmar que los resultados de unos comicios electorales no son independientes del espacio o área geográfica en el que se disputan los votos. De los trabajos más conocidos en este campo nos encontramos algunos como

los de (Ward et al., 1996; O'Loughlin, 2002, 2003; Darmofal, 2006; Klos, 2008; Tam Cho & Nicley, 2008; Soares & Terron, 2008; Chen & Rodden, 2009; Seabrook, 2009; Lefebvre & Robin, 2009; Cutts & Webber, 2010; Sue Wing & Walker, 2010; Rodden, 2010; Crespín, Darmofal & Eaves, 2011) que lo que pretenden a grandes rasgos es contrastar la hipótesis de no aleatoriedad de los resultados electorales.

Por todo ello, el espacio constituye una dimensión fundamental del voto y para estudiarla, disciplinas como la cartografía y la geografía nos proporcionan herramientas que nos permiten explorar el voto desde una perspectiva de la organización territorial.

2.3 Limitaciones de la GE. Mapas electorales

Autores como François Goguel (1990) advierten que la cartografía automática puede llegar a ser una catástrofe, ya que esta nos permite generar cientos de mapas sin ninguna reflexión o metodología científica. Es pues evidente que la cartografía electoral también conoce límites que pueden llegar a convertirse en trampas e inducir a errores, llegando a convertirse en armas de doble filo.

Entre los primeros problemas a manejar, siguiendo a Sonnleinter (2013) se encuentra la posible distorsión que se puede producir por efectos gráficos y visuales que se produce en, muchos mapas en donde se atribuyen mayor visibilidad a las entidades extensas, pero escasamente pobladas en detrimento de las zonas urbanas y metropolitanas más pequeñas, pero más habitadas, aunque afortunadamente existen técnicas para contrarrestar este efecto visual.

Encontramos otra limitación en el carácter estático de un mapa. Aunque es verdad que existe la posibilidad de construir representaciones sintéticas de los procesos políticos representando tendencias agregadas de continuidades o rupturas ya sean temporales o espaciales, no debemos olvidar que seguimos trabajando con modelos parciales de comportamiento y procesos complejos cuyos contenidos cambian inevitablemente en el espacio y tiempo.

Ahora bien, no podemos no mencionar los problemas relacionados a la falacia *ecológica*. Así, la falacia ecológica se refiere a pensar que las relaciones observadas en los grupos se mantienen para los individuos, dicho de otro modo, se da cuando se trata de inferir observaciones observadas a nivel colectivo a nivel individual. Si la agrupación de unidades

espaciales es homogénea en el interior, no presentará un problema de falacia ecológica; no obstante, eso no ocurre en la mayoría de las ocasiones (Openshaw, 1984). Una de las principales críticas al determinismo geográfico fue dicha falacia, la cual mantuvo a la disciplina en un proceso de *stand by* durante unos años. W. S. Robinson (1950) trató de analizar la correlación entre las tasas de analfabetismo y la proporción de ciudadanos afroamericanos en distintas escalas de la geografía norteamericana. Dicho autor advirtió de una correlación muy fuerte (0,946) al nivel de las 9 divisiones geográficas ⁴, pero que tendía a reducirse al nivel de los 48 estados (0,773) alcanzando solamente 0,203 al nivel de los 97.3 millones de individuos mayores de 10 años censados en 1930. Robinson concluyó que había que privilegiar estudios basados en las correlaciones individuales en vez de las ecológicas que no podían ser sustitutas de las primeras en ningún caso.

Desde esos años el argumento fue (y sigue siendo) el mismo a la hora de atacar los fundamentos metodológicos de la geografía electoral. Se han desarrollado dos aproximaciones para hacer frente al problema de la falacia ecológica. Por un lado se han formulado modelos estadísticos para reducir el sesgo de agregación, como el trabajo de Gotway y Young (2002). La otra, ha sido controlar la forma en que se agregan áreas, tal como señala Openshaw (1977), ya que usualmente se utilizan áreas definidas administrativamente para realizar inferencias estadísticas. Sin embargo, Sonnleinter (2013) nos propone una tercera vía. Es verdad que al cambiar el nivel de observación las fuerzas de las correlaciones pueden variar, sin embargo para poder superar esta falacia sin caer en la reducción de toda conducta social a un conjunto de comportamientos individuales e independientes (falacia *atomística*) pasa por una reflexión y un estudio sobre las razones y los significados empíricos de las correlaciones cambiantes que se producen en diferentes niveles de análisis, lo que nos exige un enfoque multidireccional que nos permita articular las diferentes escalas de observación.

2.4 Evolución de la disciplina y dónde está ahora

Los primeros adalides de la Geografía electoral conocida como tal no se dieron hasta entrado el siglo XX. Antes del comienzo del siglo anterior a este no se registran demasiados elementos que nos conduzcan a la cuestión que aquí residen lugar.

⁴ La Oficina del Censo de los Estados Unidos divide el país en cuatro regiones y un total de 9 divisiones.

Es pues a finales del siglo XIX lo que se destaca como la consolidación del lugar de la Geografía en el mundo académico y la distinción entre lo humano y lo físico, siguiendo las palabras de Norma Medus (2005). Con la entrada del siglo XX se distinguieron dos conceptos en la disciplina: la región y el determinismo ambiental. Aun así, no sería hasta 1913 con los trabajos e André Siegfried, representante de la escuela francesa de Geografía Regional, cuando la disciplina se empezase a interesar por conocer cómo funcionan los mecanismos que conllevan a la representación del pueblo. Siegfried, considerado el padre de la geografía electoral, a partir de mapas electorales que el mismo elaboró en base a resultados electorales en donde buscó relaciones causales entre los mismos.

A raíz del mencionado autor se empezaron a congregarse diversos estudios en la materia, pero salvo estos casos en la primera mitad del siglo XX los profesionales en geografía electoral se dedicaron a estudiar las relaciones externas de los estados olvidando el plano interno (dejando de lado procesos como los electorales, por ejemplo). En la década de los 70 podemos hablar de un primer cambio de paradigma. Ackerman (1958) ya había comentado que toda investigación que tenga fines utilitarios debe apoyarse en la cuantificación. Hasta entonces los estudios se estaban realizando bajo las consignas de los trabajos de Sauer ⁵ en un marco cualitativo que cambia cuando penetran las tendencias cuantitativas. El estilo cambia y se combinan los estudios demográficos con los electorales. La incidencia de lo cuantitativo llevó a la Geografía Humana a declinar los estudios regionales cualitativos. Es una época en la que todos los estudios basados en la cuantificación están sobre-prestigiados. Este es el comienzo pues del denominado Paradigma Cuantitativo.

Hablar de los años 70 y Geografía Cuantitativa es prácticamente hablar de Geografía Electoral. Taylor (1994) destaca tres formas de abordar los estudios electorales: 1- el análisis estadístico estándar; 2- las influencias geográficas sobre el comportamiento del voto; 3- modelos de probabilidades en la distribución espacial de los representados. Es tal el impacto de los análisis estadísticos que los trabajos realizados sustituyen la línea de Siegfried. Aun así, esta es una línea muy criticada en la geografía electoral puesto que los análisis estadísticos por sí solos poco aportan a la explicación de los comicios electorales. En el momento en el que surge

⁵ En un momento similar al de Siegfried, Sauer aportó en Estados Unidos elementos para definir los distritos para las elecciones a representantes en el Congreso y es comúnmente asociado a la fundación de la Escuela de Geografía Regional-Cultural.

este Paradigma de lo Cuántico, la disciplina se ve agitada por elementos ajenos a ella que llegan a influir en la gestación de nuevas ideas y surge lo que se conoce como Geografía Radical.

Ya instalada, a lo largo de la existencia de la geografía electoral han existido diversos enfoques que es pertinente comentar. El enfoque corológico implica cartografiar los resultados electorales para cada unidad electoral. Este enfoque da cuenta de que la geografía electoral en sus inicios, quedaba enfocada únicamente a un aspecto puramente estadístico, casi como si fuese una ciencia dura, y con el tiempo se vio la necesidad de explicar el porqué de la distribución espacial del voto (Monzón, 2013). Analizando la ubicación geográfica pretende diferenciar áreas territoriales homogéneas. La explicación es netamente cualitativa y el padre de este enfoque podemos decir que es Siegfried (1913). Joaquín Bosque Sendra (1988) es muy crítico con este enfoque y nos cuenta que lo importante también es tener en cuenta que existe una aparente oposición entre esta continuidad espacial del voto y la notable variabilidad del comportamiento electoral de los individuos.

Por su parte el enfoque ecológico trata de comprender los datos cuantitativos, utilizando conceptos biológicos para ello, intentando obtener una relación entre los resultados electorales y otros hechos socioeconómicos o sociopolíticos. Lo que más destaca Bosque Sendra (1988) de este enfoque es la debilidad teórica. Lo que trata de hacer es efectuar una selección de algunas variables y correlacionarlas con las que mesuran en el comportamiento político.

El enfoque sociológico es tratado por sociólogos que trabajan únicamente con datos individuales procedentes de las encuestas de intención de voto. No logran explicar más que la muestra obtenida. Según el ya citado Bosque Sendra (1988) existe una falacia individualista al asumir que los datos agregados son la suma de los datos individuales.

Llegamos ahora a los años 2000 y las investigaciones realizadas nos permiten dar cuenta de que más allá de la creación de un nuevo paradigma, podemos observar que, de hecho, se consolidan lo ya existentes, -sin haber uno que despunte-. Podemos hablar de la coexistencia de los diferentes enfoques teóricos.

A modo de síntesis podemos concluir que para explicar el recorrido de la geografía electoral nos debemos ir al siglo XX, momento el cual se destacan los estudios de André Siegfried en el año 1913. En los 70 se produce el cambio de lo cualitativo a lo cuantitativo que arrasa con los estudios de la línea del mencionado autor y se instaura el Paradigma Cuantitativo que incide de manera clave en las ciencias sociales, siendo la geografía electoral objeto de enorme interés. Con la instauración de la geografía electoral con el transcurso de las décadas

posteriores a 1970 podemos ver los diferentes enfoques que van surgiendo y que van a coexistir a día de hoy, pero no se estaría hablando de esto en este trabajo si no hicieran su papel las nuevas tecnologías que trajeron con ellas nuevas metodologías de análisis renovando la disciplina.

2.5 La relevancia del contexto y econometría espacial como base

“Los estudios espaciales parten de la idea de que la distancia entre los individuos es importante, es decir, todo tiene relación con todo, pero las cosas cercanas están más relacionadas entre sí que las lejanas” (Tobler, 1970).

En relación con la cita anterior, con el paso de los años han ido apareciendo técnicas que incorporan el estudio del contexto o del espacio a sus análisis. Aquí es donde debemos hablar de la econometría espacial. Esta es un conjunto de técnicas estadísticas que incorporan los efectos de la ubicación de las unidades de estudio en las estimaciones de sus relaciones (Charles-Leija, García, & Valadez, 2015).

Esta disciplina -la econometría espacial- es de gran utilidad en investigaciones en las que existe una dependencia espacial entre las observaciones, así como cuando existe heterogeneidad espacial en las relaciones que se modelan (LeSage, citado en Sonnleinter, 2013). Esta disciplina de estudio tiene como ventaja sobre la estadística tradicional la riqueza de información que puede otorgar el estudio del espacio, ya que en palabras de Ávila-Eggleton y Gutiérrez Sánchez (2017), al ignorar el espacio se pierde información valiosa y se incumple el supuesto de independencia entre las observaciones. Y es que, así las cosas, el uso de este tipo de datos lleva consigo dos características denominadas comúnmente por la literatura como *efectos espaciales*: la heterogeneidad y la dependencia espacial mencionadas al principio de este párrafo.

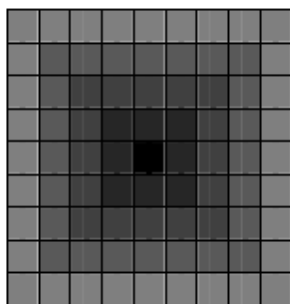
Estos efectos explicados en palabras de Serrano y Valcarce (2004) el primero de ellos aparece cuando se utilizan datos de unidades espaciales diferentes entre sí que pretenden explicar un mismo fenómeno, lo que lleva a problemas de heterocedasticidad ⁶ o inestabilidad

⁶ En estadística se dice que un modelo de regresión lineal presenta heterocedasticidad cuando la varianza de los errores no es constante en todas las observaciones realizadas. Esto implica el incumplimiento de una de las hipótesis básicas sobre las que se asienta el modelo de regresión lineal.

estructural, pudiendo ser resueltas por técnicas econométricas para series temporales, aunque dicho problema en este trabajo no viene al caso pues la serie de datos corresponden a una misma serie temporal. Por su parte, la dependencia o autocorrelación espacial, se da siempre que el valor de una variable en un lugar concreto del espacio está relacionado con su valor en otro u otros lugares del espacio, la dependencia espacial nos plantea el hecho de que una observación de la localidad A puede estar correlacionada con otra de la localidad B cuando $A \neq B$. Sin embargo, esta autocorrelación espacial no puede ser tratada mediante la econometría estándar, debido a la multidireccionalidad de las relaciones de interdependencia entre las unidades espaciales, esto es, a la hora de hablar de autocorrelación espacial se nos plantea una disyuntiva, y es que a diferencia de una autocorrelación temporal, en donde el tiempo transcurre en línea recta, en donde el pasado explica el presente, el espacio se ve afectado desde múltiples direcciones, siendo que una región puede verse afectada por más de una región contigua a ella es por ello que la econometría espacial utiliza retardos (o rezagos) espaciales, esto es, efecto de vecinos, para poder estimar con precisión la influencia que las unidades de observación ejercen entre sí (Charles-Leija, García, & Valadez, 2015).

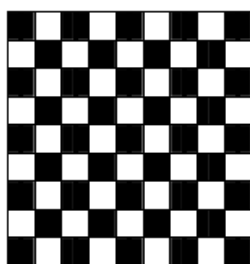
Más en detalle dentro de los efectos especiales para el caso que ocupa este trabajo el de mayor preocupación recae en el segundo. La dependencia o autocorrelación espacial aparece como consecuencia de la existencia de una relación funcional entre lo que ocurre en un punto determinado del espacio y lo que ocurre en otro (Cliff y Ord, citados en Sonnleinter, 2013). Esta puede ser positiva o negativa en función de si la presencia de un fenómeno determinado en una región lleva a que se extienda ese mismo fenómeno hacia otras regiones que la rodean, haciendo que se concentre dicho fenómeno. En este caso dicha autocorrelación se denominará positiva. Por la contra, si la presencia de un fenómeno en una región o espacio impide o dificulte su extensión a otras regiones o espacios hablaremos de una autocorrelación negativa. Por otra parte, cuando la variable a estudiar se distribuya de manera aleatoria, no podemos hablar de autocorrelación espacial.

Ilustración 2. Autocorrelación espacial positiva



Fuente: (Baronio, Vianco, & Rabanal, 2012)

Ilustración 3. Autocorrelación espacial negativa



Fuente: (Baronio, Vianco, & Rabanal, 2012)

La literatura en su conjunto afirma que la solución a la multidireccionalidad de los sucesos recae en una matriz de pesos espaciales, W , una matriz cuadrada no estocástica cuyos elementos W_{ij} reflejan la intensidad de la interdependencia entre las regiones i y j (Baronio, Vianco, & Rabanal, 2012)

Así es pues, que la econometría espacial surge como subdisciplina de la econometría general que proporciona las técnicas de contrastación y de estimación necesarias para trabajar con datos espaciales y poder resolver sus problemáticas características (Serrano & Valcarce, 2004), que además en el presente trabajo vamos a acompañar con las técnicas aportadas por el AEDE.

2.6 La clave es el votante (“efecto vecindario”)

Fue el trabajo de Tingsten *Political Behavior: Studies in Election Statistics*, (1937), inmerso en el auge de estudios de carácter *behaviourista* el adalid de una abundante literatura dirigida a analizar la influencia de los contextos locales sobre la conducta de los electores, llegando a ser tantos los estudios posteriores ⁷ que consolidaron una literatura relevante tanto en Estados Unidos como en Reino Unido, sobre todo. Y es que es a través del ejercicio electoral, como el ciudadano manifiesta su influencia en el entorno político (Mata, 2013), basando su orientación de voto en decisiones tanto individuales como colectivas. Una de las principales líneas de investigación contextual se ha dirigido a analizar el llamado *efecto del vecindario* o *amigos y vecinos*, esto es, la influencia de la composición de los contextos o espacios locales, a través de la interacción social informal, sobre la conducta en las urnas. La perspectiva de análisis contextual no se presenta como alternativa a los análisis individuales, sino que pretende complementarlos (Books & Prysby, citados en Villodres & Cerezo, 2013).

El efecto *amigos y vecinos* o *del vecindario* se basa en el supuesto de que la interacción social, particularmente en las zonas residenciales más próximas al domicilio, influye en la conducta de los electores en su decisión de voto en las urnas en período de comicios electorales. Así pues, desde esta perspectiva analítica del contexto, la interacción social entre individuos son el producto conjunto de las preferencias individuales y de las constricciones ambientales (Baybeck & Huckfeldt, 2002). Por tanto, se puede decir que el contexto proporciona los potenciales compañeros de socialización y de ahí su significancia política. Por ejemplo, los residentes en áreas conservadoras tendrán más papeletas para relacionarse con conservadores que los de residan en localidades progresistas. Por otro lado, esta óptica de análisis también debe asumir que los individuos no aceptan cualquier fuente de información a su alcance, sino que seleccionan a sus colegas de conversación en función de sus preferencias políticas.

Cuando pretendemos entender el comportamiento de un votante no nos debemos limitar a comprender los motivos por los que vota o deja de votar y por quién votará, también debemos atender a lo que antecede y precede a esa decisión (Moreno, 2003). Y es que, precisamente la

⁷ La investigación contextual arranca con los estudios de Miller (1956), Putnam (1966), Katz & Eldersfeld (1961), Foldare 1968) y Cox (1969), siendo las investigaciones electorales más importantes recientes (Books & Prysby, 1991; Carmines & Huckfeldt, 1996; Franklin & Wlezien, 2002; Zuckerman, 2005; Johnston & Pattie 2006).

estadística apodada tradicional omite el concepto de dependencia espacial entre observaciones de la que se ha hablado antes.

Como ya se ha dicho, en este trabajo se procede a investigar la relación entre votos a los cuatro partidos que conforman el Parlamento andaluz y una serie de variables contextuales o sociodemográficas, pero es importante recalcar la importancia del estudio del votante desde una óptica contextual, porque un elector únicamente puede ejercer su voto dentro del territorio al que pertenece, dándose aquí el efecto “*vecindario*”, puesto que las personas más interesadas en la política hablan con las de su mismo territorio pudiendo condicionarlos a la hora de decidir la orientación del voto, asimismo, los individuos de manera natural prefieren vivir en un entorno donde los vecinos comparten ideología política. En suma, y en palabras de Sonnleitner (2013): “*el voto es una conducta social e interactiva, colectiva y territorializada*”.

CAPÍTULO 2. Bases metodológicas

1. METODOLOGÍA A UTILIZAR

1.1 Diseño de la estrategia de trabajo

En este apartado se tratará de esclarecer los pasos a seguir para la consecución del objetivo de la manera más exitosa posible. Como ya se ha dicho, lo primero es conseguir un archivo **.shp**, el cual se descargará desde la base de datos oficial del Instituto de Estadística y Cartografía de Andalucía. La siguiente tarea consistirá en la revisión del propio archivo **.shp** por si falta algún municipio y actuar en consecuencia.

Una vez comprobado y teniendo listo el *shapefile* proseguiremos a rellenar la base de datos con las variables que se crean pertinentes. Como se ha dicho en el punto anterior los datos deberán ser sacados de fuentes oficiales y fiables para evitar errores o sesgos.

Ya establecidas las variables con las que vamos a trabajar solo nos queda proceder al análisis pertinente mediante el cual se va a basar las ideas plasmadas en el trabajo y las conclusiones. La información ofrecida por el programa, esto es, tablas o imágenes, serán presentadas en el trabajo acompañadas de una justificación y una explicación, así como de una conclusión final.

1.2 Instrumentos de investigación

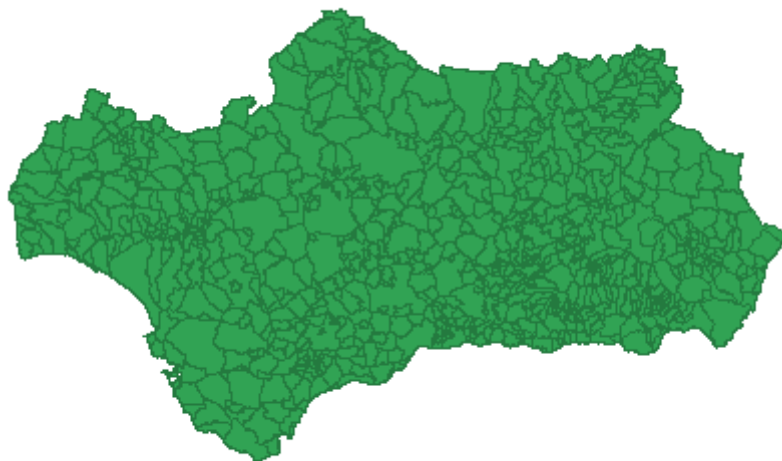
El software utilizado para el análisis de los comicios andaluces celebrados a finales del año pasado el programa informático *GeoDa*, desarrollado por el Dr. Luc. Anselin que es uno de los principales desarrolladores del campo de la econometría espacial.

Esta es una herramienta de software gratuita y de código abierto que sirve como introducción al análisis de datos espaciales. Está diseñada para facilitar nuevos conocimientos del análisis de datos mediante la exploración y el modelado de patrones espaciales.

El programa proporciona una interfaz gráfica y sencilla de utilizar para los métodos de análisis exploratorio de datos espaciales (de ahora en adelante AEDE), así como estadísticas de autocorrelación espacial para datos agregados.

Para trabajar con *GeoDa* lo primero que se ha de realizar es conseguir (crearlo u obtenerlo en una página web) un archivo *shape* (**.shp**). Un *shape* o *shapefile* es un formato sencillo y no topológico que se utiliza para almacenar la ubicación geométrica y la información de atributos de las entidades geográficas. Es decir, nos presenta el mapa de la ubicación geográfica que nosotros hayamos elegido, con el objetivo de trabajar en el mediante una serie de variables.

Ilustración 4. Shapefile de municipios andaluces



Fuente: Imagen obtenida del programa GeoDa

El siguiente paso a seguir sería el de completar una base de datos con las variables correspondientes que para el presente trabajo se consideren las más importantes y oportunas, para después proceder al análisis correspondiente.

Para completar los datos que se precisan para construir cada una de las variables se ha indagado en las diferentes páginas de datos oficiales como por ejemplo el Instituto Nacional de Estadística. El *shapefile*, dividido por municipios, por su parte se ha descargado del Instituto de Estadística y Cartografía de Andalucía.

1.3 Métodos de investigación: el AEDE y ACDE

La econometría y estadística espacial como norma general no incorpora herramientas o técnicas que den validez a un modelo econométrico, ofreciéndonos una descripción y visualización de distribuciones espaciales, pues estas omiten el factor espacial con todo lo que ello conlleva. Es precisamente de este problema donde surge el *sino* del Análisis Exploratorio de Datos Espaciales.

El Análisis Exploratorio de Datos Espaciales en palabras de Anselin (1999) puede definirse como el conjunto de técnicas que describen y visualizan las distribuciones espaciales, identifican localizaciones atípicas o “atípicos espaciales”, descubren esquemas de asociación espacial, agrupamientos (“*clusters*”) o puntos calientes (“*hot spots*”) y sugieren estructuras espaciales u otras formas de heterogeneidad espacial.

Para entender un poco mejor el AEDE se debe concebir esta metodología como inmersa dentro del Análisis Exploratorio de Datos, cuyos métodos gráficos y visuales se utilizan para identificar las propiedades de los datos con la esperanza de detectar ciertos patrones de comportamiento en los datos o de formular hipótesis a partir de dichos datos.

Para el presente trabajo se va a dar uso de las principales técnicas del AEDE en combinación del análisis estadístico gráfico, lo que hace posible entre otras cosas el estudio de las distribuciones espaciales y sus valores atípicos, esquemas de asociación espacial y agrupamientos espaciales.

Por ende, con lo visto hasta ahora del AEDE podemos ver que es más una técnica descriptiva (estadística) que confirmativa (econométrica) (Chaco Yrigoyen, 2003). Por ello, el análisis exploratorio de datos es la fase previa al análisis confirmatorio de datos espaciales, en los que se formularán modelos de regresión y estimaciones y se realizará la estimación de los parámetros muestrales.

En muchas ocasiones, el análisis de datos comienza con el objetivo de contrastar o rechazar una hipótesis o modelo inicial. En este caso esta circunstancia es inexistente. En este caso, un servidor no parte de una hipótesis inicial a contrastar, sino que la naturaleza del AEDE es exploratoria como ya se mencionó por lo que las técnicas que se utilicen deberán ser con el objetivo de encontrar alguna dependencia en los datos, se utiliza para poder encontrar relaciones sistemáticas entre variables cuando no existe una certeza del germen de dichas relaciones.

Cuando un estudio trabaja series geográficas o datos espaciales se deben asumir explícitamente las características propias de esta información, esto es, georreferenciación, multidireccionalidad y multidimensionalidad (Coro)⁸. Y es que un espacio geográfico está georreferenciado, esto es, te exige conocer la posición relativa o absoluta de donde se están produciendo los hechos que se pretenden analizar, siendo las diferentes relaciones que se producen de tipo multidireccional, a diferencia de las relaciones temporales, por ejemplo, las cuales son unidireccionales (pasado>presente>futuro). Estas particularidades, producen los mismos hechos, particularidades de dependencia o autocorrelación espacial y heterogeneidad espacial, conocidas como **efectos espaciales**.⁹

En suma, el AEDE, es una técnica descriptiva que permite la combinación de herramientas estadísticas, con métodos de análisis gráfico, ayudando, por ende, a identificar y analizar la estructura de la distribución espacial en un espacio univariante o multivariante, al poder esclarecer cómo es la regularidad de los en los patrones de los esquemas espaciales, que nos sugieren una idea del comportamiento de las variables.

1.4 Variables a utilizar

A la hora de establecer las variables a utilizar, debemos tener en cuenta, primero: el tipo de análisis que vamos a elaborar y segundo: que nos interesa medir. En cuanto a lo primero, al estar utilizando técnicas de geografía electoral, cualquier variable establecida tendrá que medirse obligatoriamente en función de la delimitación territorial que nosotros deseemos. España se divide primero por Comunidades Autónomas, que a su vez están delimitadas por Provincias divididas en municipios. En este caso vamos a establecer los análisis en función de esta última delimitación territorial.

Andalucía es una comunidad autónoma reconocida como nacionalidad histórica por su Estatuto de Autonomía (1981), compuesta por las provincias de Almería, Cádiz, Córdoba, Granada, Huelva, Jaén, Málaga y Sevilla. En total cuenta con 786 municipios que se dividen en: 103 para la provincia de Almería, 45 para Cádiz, 77 para Córdoba, 80 Huelva, 97 Jaén, 174 Granada, 103 Málaga y 106 Sevilla (la capital).

⁸ No se ha encontrado el año de publicación del artículo, pero aun así se incluye en la bibliografía.

⁹ Estas son las mismas características de la econometría espacial mencionadas en el capítulo anterior.

Es preciso aclarar que en los mapas presentados en el capítulo 3 se indican como indefinidos 8 casos, esto es que, son 8 municipios que a pesar de existir cuentan con sus casilleros vacíos en la base de datos puesto que el 02/10/2018 el Consejo de Gobierno aprobó la creación de los nuevos municipios de San Martín del Tesorillo, en la provincia de Cádiz; Fuente Carreteros y La Guijarrosa, en Córdoba; Fornes y Torrenueva Costa (Granada); Tharsis y La Zarza-Perrunal (Huelva), y El Palmar de Troya, en Sevilla. Tras una prolongada experiencia de autogobierno como Entidades Locales Autónomas (ELA)¹⁰, su constitución se llevará a cabo por segregación de los términos municipales matrices de Jimena de la Frontera, Fuente Palmera, Santaella, Arenas del Rey, Motril, Alonso, Calañas y Utrera, respectivamente. Con esta decisión, el número de municipios que integran la comunidad autónoma se elevó a 786¹¹. Debido a la proximidad de las elecciones (03/12/2018) y la inexistencia de gobierno¹² en cada uno de estas nuevas localidades no existen datos de voto para los municipios creados, los cuales seguirán integrados bajo los municipios a los que respondían antes en términos electorales.¹³

En el presente trabajo es de nuestro interés las variables sociodemográficas¹⁴, con lo que podremos comprobar en función de estas características la relación entre voto y área geográfica.

¹⁰ En Andalucía (España) una entidad local autónoma es una administración pública de carácter local creada sobre la base de la *Ley 7/1993, de 27 de julio, Reguladora de la Demarcación Municipal de Andalucía*, para la administración descentralizada de los intereses propios de un núcleo separado de población dentro de un término municipal.

Las entidades locales autónomas que se creen al amparo de la mencionada ley, tendrán la condición y tratamiento de entidades locales, igualmente gozarán de tal condición aquellas entidades creadas con anterioridad a la aprobación de esta Ley por el Parlamento de Andalucía, como las EATIM reguladas por el art. 45 de la *Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local*, salvo que los vecinos en forma mayoritaria y directa muestren su voluntad contraria.

¹¹ A pesar de que el número de municipios aumente, en ninguna de las bases de datos consultadas se encuentra rastro alguno de datos electorales de estos nuevos municipios y tampoco se cuenta en este trabajo con respuesta oficial puesto que las entidades ofrecen una respuesta.

¹² No formarán gobierno hasta las elecciones municipales del 26 de mayo de 2019

¹³ Las sentencias serán citadas en el anexo 1

¹⁴ Las variables sociodemográficas las componen indicadores sociales, económicos y demográficos que permiten segmentar la población en grupos homogéneos.

Con motivo de esclarecer ahora conceptos que más adelante se tratarán en lo referente a las variables, se presentarán las variables a utilizar en la siguiente tabla, divididas en función del aspecto a tratar, y acto seguido se procede a una definición más detallada tanto de las variables utilizadas en el análisis como las incluidas en la base de datos (estas últimas por mera acción informativa, pues, aunque no tienen relevancia en el resultado final, han sido utilizadas para la formación de las variables que finalmente sí se utilizarán ¹⁵):

Tabla 2. Descripción de variables

Aspecto	Variable	Descripción
Inmigración	<i>Num_Extra</i>	Número de extranjeros
Edad	<i>65_years</i>	Población mayor de 65 años
	<i>less20years</i>	Población menor de 20 años
Económico	<i>rentamedia</i>	Renta neta media declarada por habitante
	<i>Paro_Mun</i>	Tasa de paro municipal
Voto	<i>Tasa_PSOE</i>	Porcentaje de votos a Partido Socialista Obrero Español
	<i>Tasa_PP</i>	Porcentaje de votos a Partido Popular
	<i>Tasa_VOX</i>	Porcentaje de votos a VOX
	<i>Tasa-Cs</i>	Porcentaje de votos a Ciudadanos
	<i>Tasa_AA</i>	Porcentaje de votos a Adelante Andalucía

Fuente: Elaboración propia

- **Población total** (*Pobl_TOT*) hace referencia a la cantidad de habitantes totales del municipio en cuestión.
- **Hombres** (*Hombres*); para designar la cantidad de hombres existente en dicho municipio.
- **Mujeres** (*Mujeres*); para designar la cantidad de mujeres existente en dicho municipio.
- **Número de extranjeros** (*Num_Extra*); designa la cantidad de personas de otras nacionalidades que viven en el municipio (medido en porcentaje en función del total de la población)
- **Población mayor de 65 años** (*65_years*); se refiere a la cantidad de personas que sobrepasen los 65 años (o los igualeen) que vivan en el municipio (medido en porcentaje en función de la población total)

¹⁵ Por ejemplo, para la formación de la variable *Num_extra* han sido necesarios el número de extranjeros de cada municipio y su población total, ambas incluidas en la base de datos.

- **Población menor de 20 años** (*less20year*); designa a la cantidad de ciudadanos del municipio menores de 20 años (medido en porcentaje en función de la población total)
- **Renta media declarada** (*rentamedia*); esta es un indicador de riqueza que se define como el cociente entre la renta neta total declarada y el número de declaraciones. Se debe tener en cuenta que existe un umbral mínimo de renta por debajo del cual no es obligatorio presentar declaración por I.R.P.F.
- **Tasa de desempleo municipal** (*Paro_Mun*); se calcula como el cociente entre el volumen de demandantes no ocupados en un momento dado en un territorio y la agregación de estos demandantes con las afiliaciones de residentes en este mismo territorio, es decir, el cociente entre no ocupados y habitantes.
- **Censo electoral** (*Censo_Elec*); se refiere a la cantidad de personas en disposición de votar.
- **Número de votos al PSOE** (*Tasa_PSOE*); como porcentaje del número de personas del total del censo que votaron al Partido Socialista.
- **Número de votos al PP** (*Tasa_PP*); como porcentaje del número de personas del total del censo que votaron al Partido Popular.
- **Número de votos a Ciudadanos** (*Tasa-Cs*); como porcentaje del número de personas del total del censo que votaron a Ciudadanos.
- **Número de votos a Vox** (*Tasa_Vox*); como porcentaje del número de personas del total del censo que votaron a Vox.
- **Número de votos a Adelante Andalucía** (*Tasa_AA*); como porcentaje del número de personas del total del censo que votaron a Adelante Andalucía.
- **Abstención** (*Abstencion*); como porcentaje de abstención.

Las variables elegidas para el análisis se decidieron en función de la relevancia que puede llegar a tener una variable sociodemográfica en la captación de votos por parte de un partido en campaña electoral y que representan un incentivo del voto, así pues, se han escogido variables referentes a temas como los cortes de edad, económicos, en relación con la inmigración, por ser los más acudidos en estudios de comportamiento electoral, al ser temas tratados para incentivar el voto. Se debe añadir, además, que en el caso de los grupos de edad se ha elegido simplemente dos de ellos -menores de 20 y mayores de 65- puesto que el interés principal es saber si la presencia o ausencia de un grupo de edad difieren en la tendencia del voto con la del grupo contrario, es decir, interesa más entender desde una perspectiva sociológica el cómo puede llegar a cambiar el voto la presencia o ausencia de gente en un

municipio de un grupo de edad diferente con todo lo que eso conlleva, pues la diferencia generacional entre menores de 20 años y más de 65 años es importante y eso puede llevar a diferentes comportamientos que lleven a diferencias en el voto -de ahí la inexistencia de grupos de edad intermedios-. Como última aclaración, se debe apuntar que en su totalidad los datos integrados en la base de datos son recogidos del Instituto de Estadística y Cartografía de Andalucía.

CAPÍTULO 3. Desarrollo del análisis de caso

1 ANÁLISIS DE CASO

Como ya advertimos en los anteriores puntos del trabajo, vamos a tratar de explicar la tendencia de voto en función de una serie de variables que no son independientes del entorno en el que se desarrolla éste en las elecciones al Parlamento de Andalucía celebradas a finales del año pasado.

1.1 Análisis del ciclo electoral

En un apartado anterior (véase apartado 1. Introducción; *¿por qué estas elecciones?*) se dejó entrever el carácter especial de estos comicios por todo lo que conllevaron -esto es, el fin de la hegemonía socialista la irrupción de Vox por primera vez en un Parlamento autonómico en nuestro país-.

A partir del 2 de diciembre del 2018, con las elecciones al Parlamento andaluz, se disparan una serie de interrogantes para el sistema de partidos y no únicamente a nivel andaluz. La irrupción sorpresiva del partido político Vox en el Parlamento andaluz se interpretó como el anuncio de profundas transformaciones en el mapa electoral español. Por primera vez en España un partido de extrema derecha obtenía unos resultados inéditos, superando prácticamente todas las expectativas, en una arena que parecía serle particularmente poco propicia, y el PSOE, aun habiendo obtenido más votos y escaños que ninguna otra fuerza política, se veía desplazado del Gobierno y relegado a la oposición por primera vez desde los primeros comicios autonómicos.

La irrupción de Vox y la insuficiencia del voto de izquierdas para alcanzar la mayoría absoluta en las elecciones al Parlamento andaluz nos muestran datos significativos, como la continuidad en la pérdida de apoyos del Partido Popular tras su máximo histórico de 2012 (cuando fue el primer partido en número de votos, con un 40 % de los sufragios, pero no pudo gobernar dado que las otras dos únicas fuerzas parlamentarias, PSOE e IULV acordaron una coalición), el crecimiento de Ciudadanos (que dobló sus apoyos, pasando del 9,24 % del voto en 2015 al 18,27 % en 2018, y aumentó considerablemente su representación, de 9 a 21 escaños), y la caída de Podemos, que concurre en 2018 con Izquierda Unida y obtiene en

coalición en torno a 6.000 votos menos de los que había conseguido de manera independiente en 2015 (en dichas elecciones Podemos obtuvo un 14,8% del voto y 15 escaños, mientras Izquierda Unida se hizo con un 6,8 del voto y 5 escaños; en 2018, la coalición de ambas fuerzas, Adelante Andalucía, se queda en un 16,2 % del voto, más de 5 puntos por debajo de lo que sumaban en 2015, y consigue 17 escaños, 12 de Podemos y 5 de Izquierda Unida)

El PSOE perdió más de 400.000 votos respecto a las anteriores elecciones andaluzas, lo que supuso una pérdida de 14 escaños en el Parlamento de Andalucía, aunque mantuvo la mayoría del voto en siete de las ocho provincias andaluzas exceptuando Almería. Por lo tanto, podemos decir a nivel general respecto a las elecciones andaluzas de 2015 que el PSOE, PP y Adelante Andalucía tienen importantes pérdidas de votos, mientras que Ciudadanos con un incremento de 290.000 y Vox de 380.000 votos obtienen unas importantes ganancias, -como vemos en la tabla 3-.

Tabla 3. Elecciones 2015 & 2018

	2015		2018	
	ESCAÑOS	Nº voto / %	ESCAÑOS	Nº voto / %
PSOE	47	1411278 / 35.41 %	33	1009243 / 27.95%
PP	33	1066458 / 32.1%	26	749275 / 20.75%
Cs	9	36994 / 9.24%	21	659631 / 18.27%
PODEMOS ¹⁶	15	590011 / 14.80%	-	-
VOX	0	18017 / 0.45 %	12	395978 / 10.97%
Izquierda Unida Los Verdes-Convocatoria por Andalucía -ADELANTE ANDALUCIA-	5	274518 / 6.86%	17	584040 / 16.18%

Fuente: elaboración propia a partir de datos oficiales del Ministerio del Interior.

2. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES

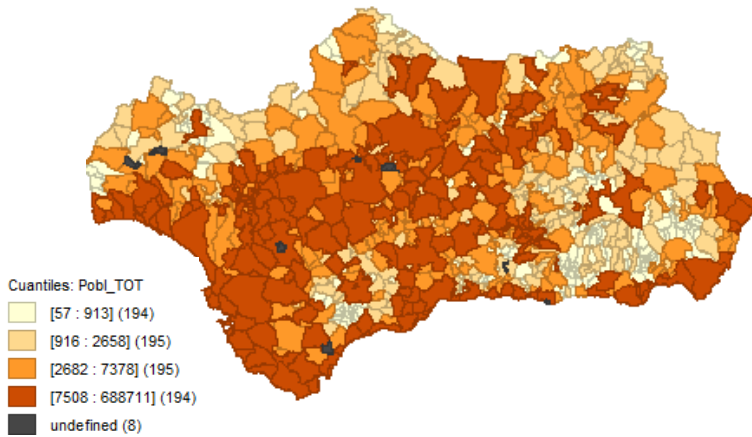
Para comenzar con el análisis primero se procederá al comentario de una serie de mapas descriptivos que contengan la distribución de las variables a lo largo de los municipios del mapa andaluz.

Los mapas temáticos consisten en la representación cartográfica de una variable geográfica. Esta representación puede llevarse a cabo mediante símbolos y colores, que pongan de manifiesto el valor de dicha variable en cada una de las unidades geográficas seleccionadas, para este caso, para cada municipio. Estos son útiles a la hora de determinar o identificar fenómenos geográficos mediante el uso de recursos visuales. Los más utilizados suelen ser los

¹⁶ En el 2018 presenta candidatura en coalición con Adelante Andalucía

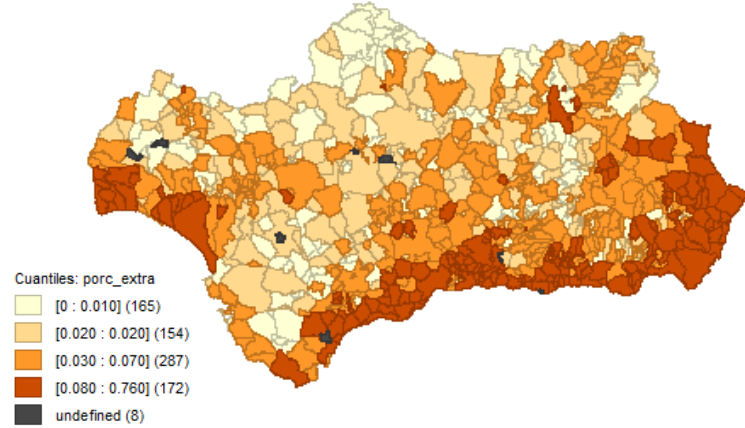
mapas divididos en cuantiles que representan el comportamiento de una variable espacialmente distribuida, para lo que, los datos se dividen y se agrupan en unas categorías denominadas cuantiles que contienen igual número de observaciones en medida de lo posible. En este caso se dividirá el mapa andaluz en cuatro cuantiles.

Ilustración 5. Distribución del total de población



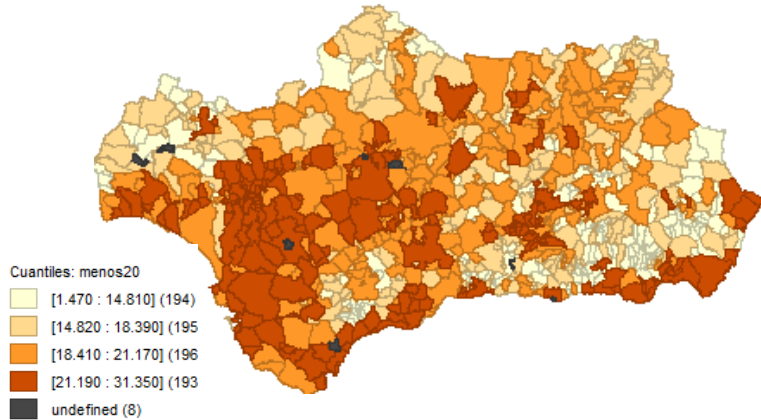
Fuente: datos obtenidos de GeoDa

Ilustración 6. Distribución del porcentaje de extranjeros



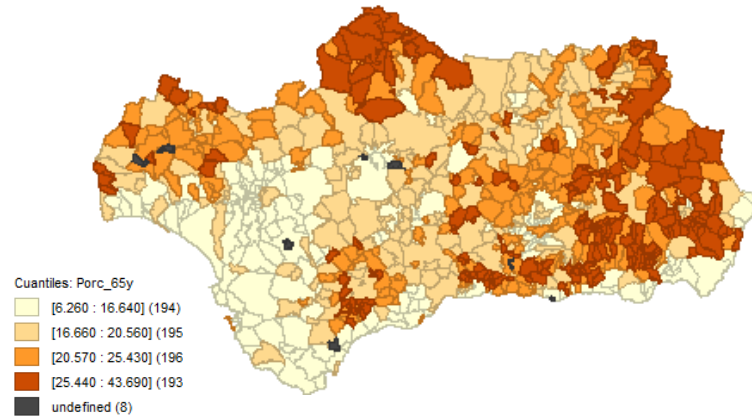
Fuente: datos obtenidos de GeoDa

Ilustración 7. Distribución de edad mayores de 20



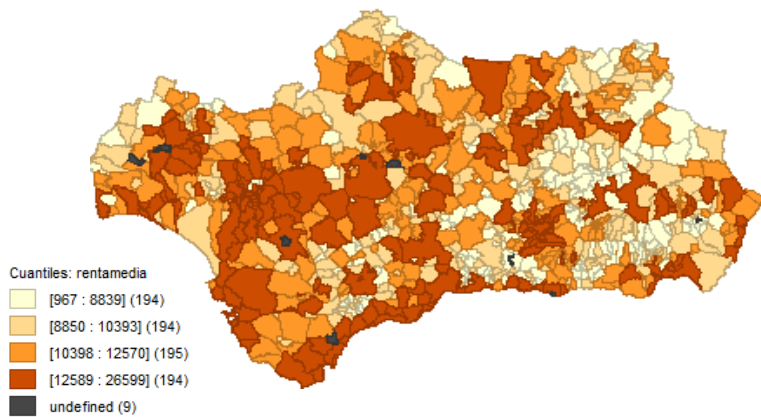
Fuente: datos obtenidos de GeoDa

Ilustración 8. Distribución edad mayores de 65



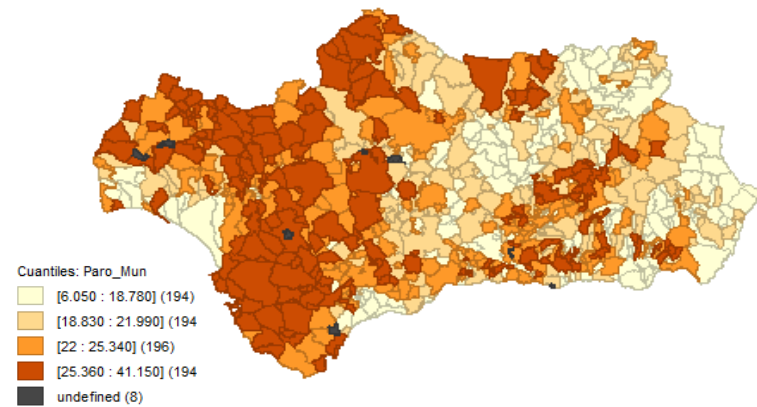
Fuente: datos obtenidos de GeoDa

Ilustración 9. Distribución renta neta declarada por habitante



Fuente: datos obtenidos de GeoDa

Ilustración 10. Distribución Tasa de paro municipal



Fuente: datos obtenidos de GeoDa

Como observamos en la ilustración 5, se muestra la distribución de la población total de cada municipio y se ve en qué municipios se encuentra el mayor número de habitantes. Se puede avistar un fenómeno curioso, y es que, las zonas que más acumulan residentes son los municipios encuadrados en el cuadrante Suroeste del mapa siendo Sevilla el municipio más poblado con 688711 habitantes. En la ilustración 6 se puede encontrar que los municipios con más porcentaje de extranjeros ¹⁷ se encuentran en el cuadrante Sureste del mapa, habiendo una continuidad espacial, como podemos apreciar, a lo largo de toda la costa. Por su parte en la zona norte se avista una proporción de extranjeros baja. En relación con la edad, ilustraciones 7 y 8, vemos como, son los municipios más cercanos al pico sur de la Comunidad autónoma en donde existe más concentración de menores de 20 años, siendo todo lo contrario en la zona Este del territorio andaluz, que es donde más se concentra la población anciana. Por su parte la distribución en la tasa de paro muestra que las tasas más elevadas se encuentran en la parte Este del mapa. A su vez, la distribución de la variable renta, muestra como existe más concentración de casos en la zona oeste.

Visto todos estos mapas descriptivos, se observa que las distribuciones de las variables siguen unos patrones de contigüidad, es decir, no son distribuidas de manera, a priori, aleatoria, sino que si se observan las imágenes cada variable tiene una zona de influencia en donde se

¹⁷ Se debe recordar que cuando se habla de la variable *Porc_Extra* (o *porcentaje de extranjeros*) no se habla de cantidades totales si no de proporciones. Es decir, para el cálculo de dicha variable se ha dividido la cantidad total de extranjeros en un municipio concreto entre la cantidad de habitantes totales de dicho municipio obteniendo un coeficiente.

distribuye en función del municipio vecino. El siguiente paso se trata de investigar si es una distribución aleatoria o es significativa, y para ello hemos de acudir a las preguntas de investigación; ¿se pueden correlacionar factores económicos de un municipio con el voto? ¿pasaría lo mismo con la edad? ¿varía el voto en función de la cantidad de la cantidad de personas extranjeras residentes en un municipio?

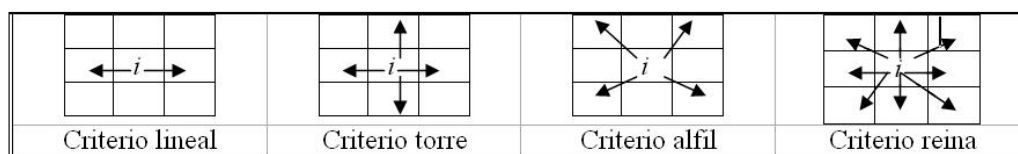
2.1 Análisis LISA

Para completar este enfoque, se observarán índices de correlación de Moran en un contexto local empleando análisis LISA (Local Indicators Spacial Association).

2.1.1 Índice I de Moran local

El análisis exploratorio de datos espaciales (AEDE) busca estimar la dependencia espacial entre las observaciones. La forma más común de incorporar elementos de la configuración espacial consiste en el uso de matrices de pesos espaciales (W). En la matriz de pesos, cada observación (i) se conecta con el grupo de observaciones vecinas (j); por lo tanto, se obtiene una matriz del tipo W_{ij} que refleja la intensidad de la interdependencia entre cada par de observaciones i y j . La manera más sencilla de definir la intensidad de la interdependencia es atribuir el valor $W_{ij}=1$ si las unidades son físicamente adyacentes y $W_{ij}=0$ si no lo son o son elementos de la diagonal principal (Martori, Hoberg, & Madariaga, 2008). Así se habla de una matriz de contactos binaria. Cuando se usa el término adyacente hablando de polígonos se puede entender desde una única perspectiva, lo mejor es pensar en un tablero de ajedrez para definir las unidades adyacentes a la unidad i :

Ilustración 11. Criterios de vecindad



Fuente: (Martori, Hoberg, & Hoberg, La incorporación del espacio en los métodos estadísticos , 2008)

En este trabajo se hablará de una matriz de contactos binaria en donde se utilizará el criterio de vecindad reina con un grado de contigüidad 1, especificándose en los casos en los que se use otra contigüidad. La contigüidad no es más que “los vecinos” que se tratan de

analizar, así, si hablamos de grado de contigüidad 1 debemos pensar en los vecinos de primer orden, los que están alrededor del municipio en cuestión y si pensamos en un grado de contigüidad 2 debemos pensar en “los vecinos de los vecinos”, asumiendo que cuanto más alejados sean menor fuerza de análisis habrá.

A la hora de tratar con el coeficiente de autocorrelación global (I e Morán) se debe tener en cuenta que cuando $H_0 = 0$, el coeficiente es igual cero, por lo que se habla de la inexistencia de un patrón espacial; se esperaría un patrón aleatorio.

Si, por el contrario, $H_1 \text{ IM} \neq 0$ el coeficiente de autocorrelación global (IM) es diferente a cero; por lo tanto, existe autocorrelación espacial y, en consecuencia, es posible afirmar que los resultados obtenidos no se distribuyen aleatoriamente en el espacio geográfico y podemos hablar de una autocorrelación positiva o negativa.

Sin embargo, el I de Morán es un indicador de autocorrelación global, y para visualizar y estimar estadísticamente resultados locales se requiere otro tipo de indicadores. Es decir, para estimar los agrupamientos o polarización del voto se propone utilizar indicadores locales de asociación espacial (*LISA*, por sus siglas en inglés, correspondientes a Local Indicator of Spatial Association). Los indicadores locales de asociación espacial son una herramienta para identificar la autocorrelación o heterogeneidad a nivel local en un área de estudio. Buscan capturar la inestabilidad de los grupos o clústeres locales en el área geográfica. Anselin (1995) plantea que los *clusters* espaciales locales son algunas veces referidos como *puntos calientes*, en los cuales el indicador LISA es significativo, por el hecho de que agrupa valores con características similares en el espacio geográfico bajo análisis.

Así pues, los valores positivos -o negativos- del contraste, estarán indicando la existencia de un *cluster* de valores similares -o disímiles- de la variable analizada alrededor de la región *i*. Es decir, el índice podrá arrojar valores positivos o negativos que estarán representando autocorrelación espacial positiva o negativa¹⁸. El índice tiene a su vez una representación gráfica o diagrama de dispersión de Moran. Este diagrama plantea la relación entre la variable y su retardo espacial, el gráfico se analiza en cuatro cuadrantes, donde el primero de ellos indica las unidades geográficas en las cuales la variable tiene un alto valor y la correlación espacial es elevada, es usualmente conocido como el cuadrante (Alto-Alto). Esta

¹⁸ El I de Moran puede ser clasificado como positivo, negativo y sin autocorrelación espacial. La autocorrelación espacial positiva es cuando valores similares se agrupan en un mapa. La autocorrelación espacial negativa es cuando valores disímiles se agrupan en un mapa.

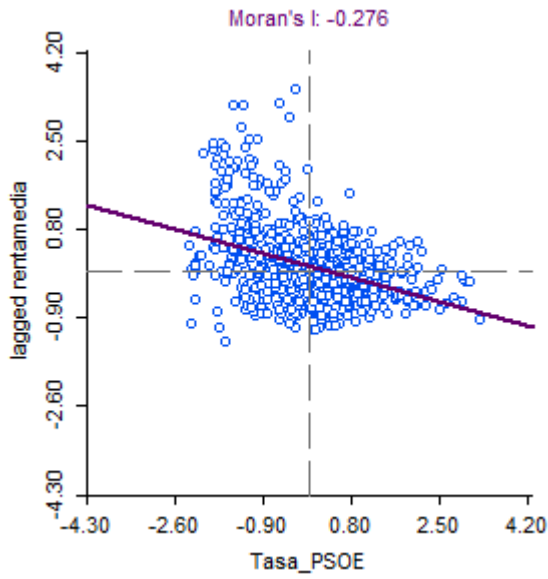
situación se repite para valores bajos de la variable y baja correlación espacial en el tercer cuadrante (Bajo-Bajo). Los cuadrantes 3 y 4, muestran bajos o altos valores de la variable con alta o baja correlación espacial respectivamente.

Los indicadores locales pueden tomar cualquier valor, tanto positivo como negativo. Un valor positivo y estadísticamente significativo indica la presencia de un *clúster* (un municipio similar a sus vecinos). A la inversa, un signo negativo sugiere que un distrito tiene valores diferentes a sus vecinos, es decir, se trata de un *outlier*. Los mapas *LISA* ponen de manifiesto atípicos espaciales cuya mayor o menor intensidad dependerá de la significancia asociada a los estadísticos de prueba (Anselin, 1995).

2.1.2 Factores económicos como explicación. Renta y paro

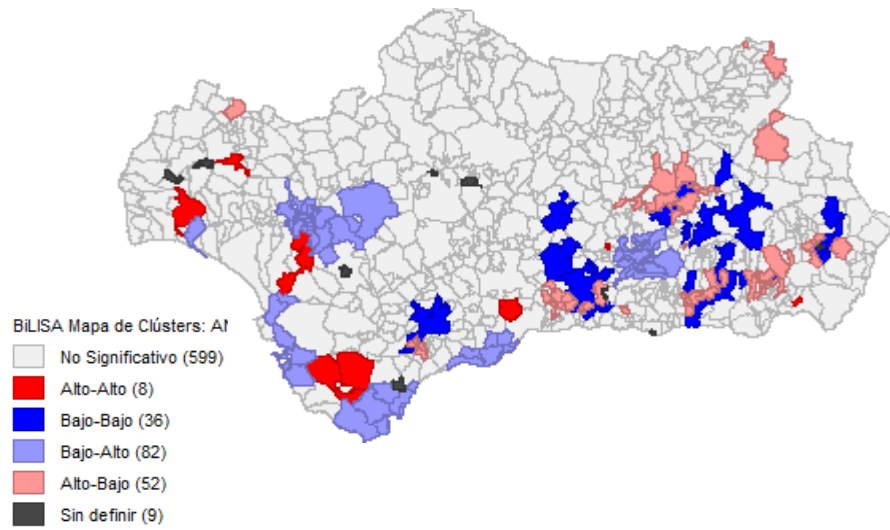
En este caso podemos ver (ilustración 12) un mapa de dispersión. Este nos indica, que, para todos los componentes, después de eliminar algunos *outliers* o valores atípicos, se presencian asociaciones negativas (-0.276) siendo la mayoría de casos que se sitúan en los cuadrantes Bajo-Alto y Alto-Bajo.

Ilustración 12. Gráfico de dispersión I de Morán Tasa_PSOE w/ rentamedia



Fuente: datos obtenidos de GeoDa

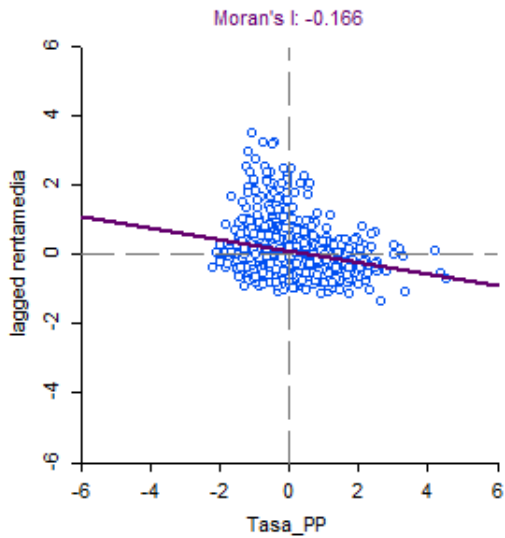
Ilustración 13. Mapa de clusters I de Morán Tasa_PSOE w/ rentamedia



Fuente: datos obtenidos de GeoDa

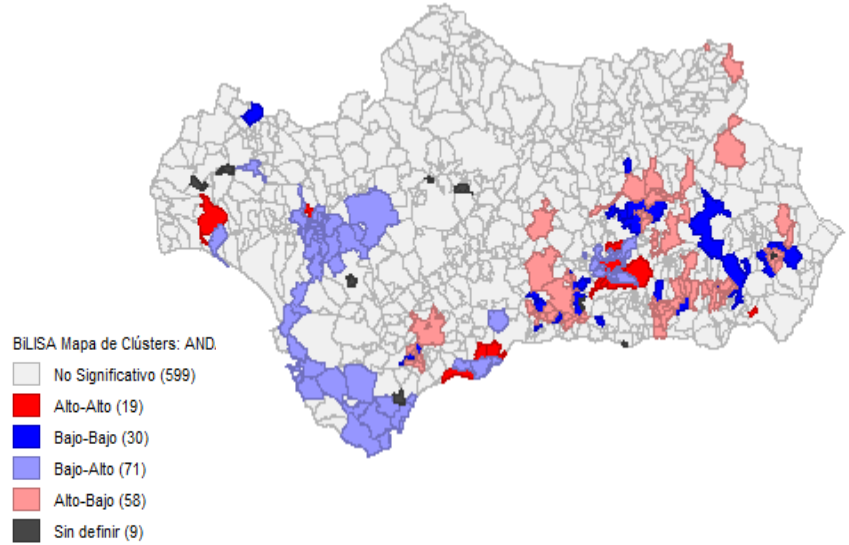
El valor negativo del coeficiente indica que existe más bien una autocorrelación espacial negativa, esto es, si, en pocas palabras, la autocorrelación espacial indica el grado de asociación de una variable a su entorno, una autocorrelación espacial negativa nos indicará que no existe tanta asociación a su entorno, y por ende, en este caso el retardo de la variable a explicar, la renta neta declarada media, no presenta evidencias de continuidad a lo largo de los municipios vecinos de primer orden.

Ilustración 14. Gráfico de dispersión I de Morán Tasa_PP w/ rentamedia



Fuente: datos obtenidos de GeoDa

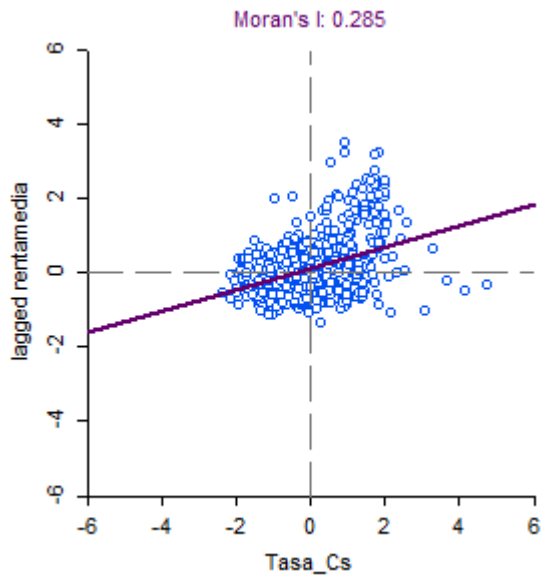
Ilustración 15. Mapa de clusters I de Morán Tasa_PP w/ rentamedia



Fuente: datos obtenidos de GeoDa

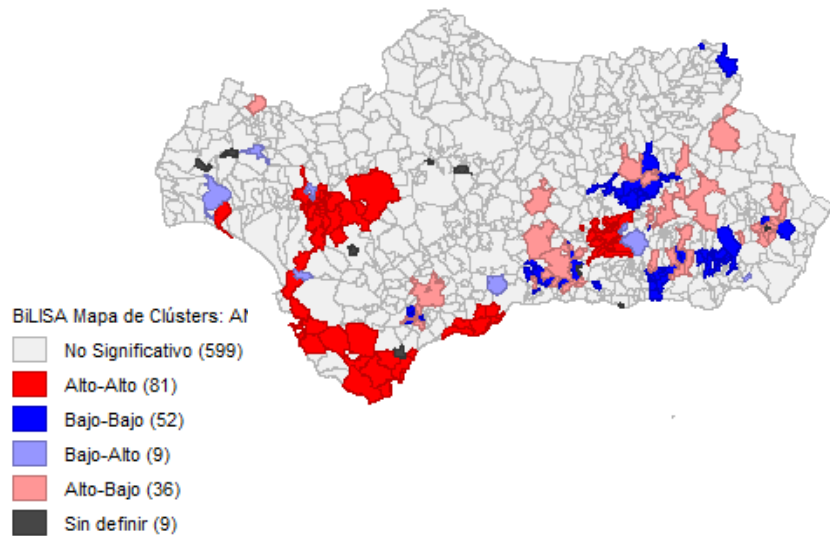
Para el caso del porcentaje de votos del Partido Popular, siendo la renta neta media declarada por habitante en cada municipio su retardo espacial, podemos observar como el índice de I Morán es también negativo (Ilustración 14) con lo que eso conlleva, estamos hablando de una autocorrelación negativa, por lo que no se puede hablar de una extensión de los valores de las variables a lo largo de los municipios. Se puede ver en el diagrama de dispersión del I de Moran y en el mapa de *clusters* la mayoría de casos se concentran alrededor de los cuadrantes Bajo-Alto y Alto-Bajo, por lo que no parece que exista una agrupación de los valores.

Ilustración 16. Diagrama de dispersión I de Morán Tasa_Cs w/ rentamedia



Fuente: datos obtenidos de GeoDa

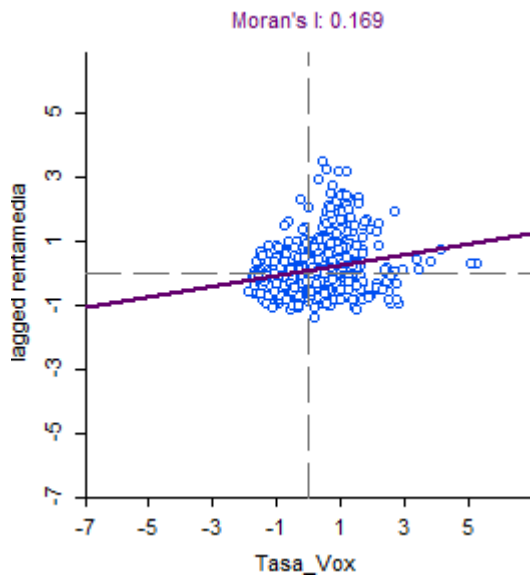
Ilustración 17. Mapa de clusters I de Morán Tasa_Cs w/ rentamedia



Fuente: datos obtenidos de GeoDa

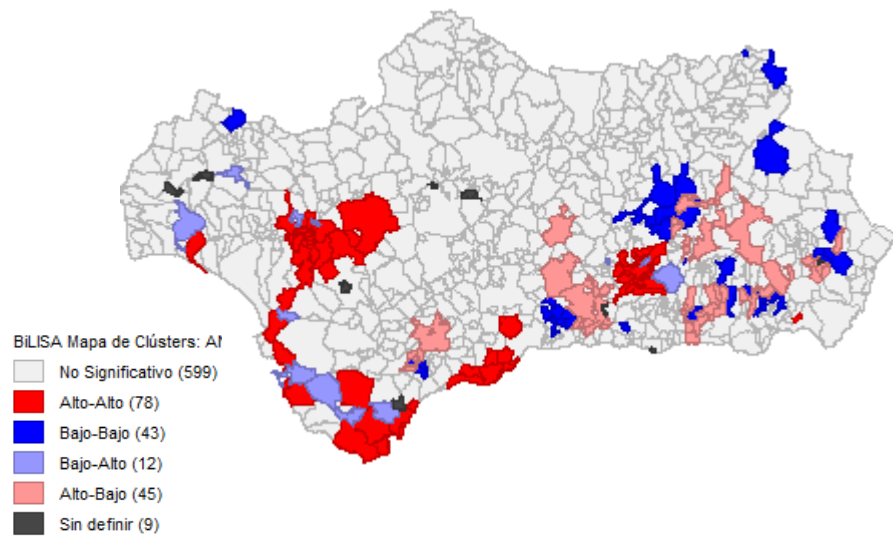
En el caso de Cs ya vemos el primer resultado disonante. Como vemos en la ilustración 16 el I de Morán es positivo con lo que tenemos una autocorrelación positiva. Esto quiere decir que sí que existe cierta asociación entre el retardo -renta neta media declarada por habitante- y la variable -porcentaje de votos del partido Ciudadanos- y una extensión de estos valores a municipios de primer orden. Los casos se sitúan en los cuadrantes Alto-Alto; Bajo-Bajo; seguido de Alto-Bajo, siendo significativo como vemos cuanto más hacia la zona de Cádiz nos desplazamos.

Ilustración 18. Diagrama de dispersión I de Morán Tasa_Vox w/ rentamedia



Fuente: datos obtenidos de GeoDa

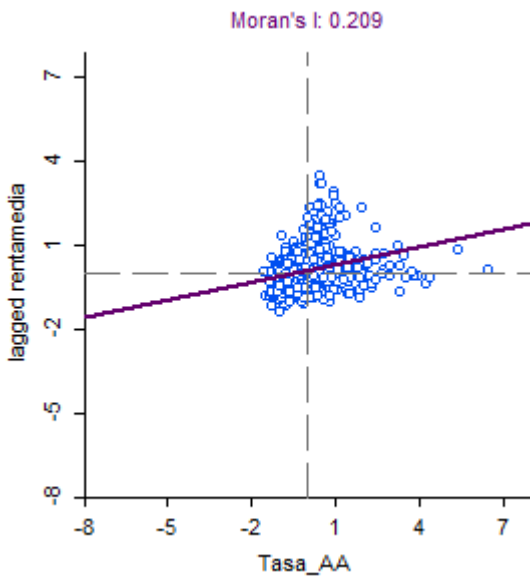
Ilustración 19. Mapa de clusters I de Morán Tasa_Vox w/ rentamedia



Fuente: datos obtenidos de GeoDa

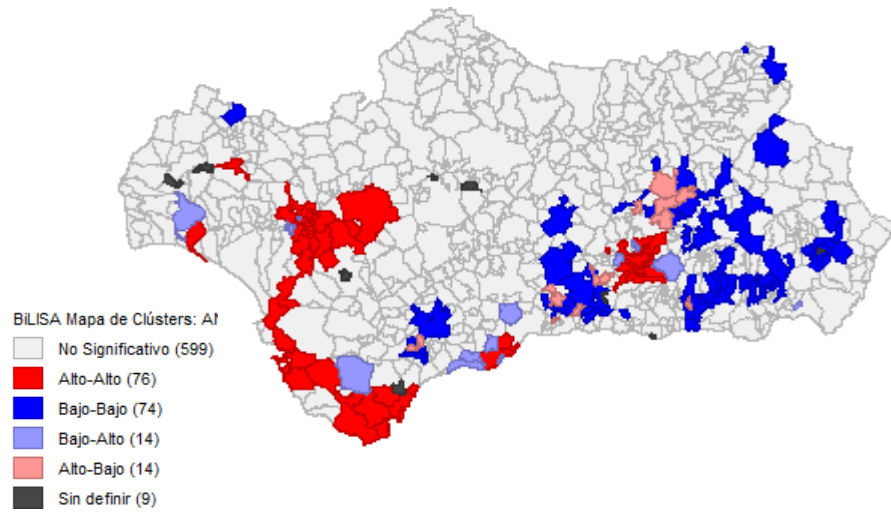
Para el caso de Vox avistamos resultados similares que a los de su vecino ideológico - Ciudadanos-, incluso se ve que el efecto se produce en las mismas zonas, tendencia que se repite con la variable Tasa_AA (porcentaje de votos de Adelante Andalucía) y su retardo espacial (renta neta media declarada por habitante) -como podemos ver en las ilustraciones 20 y 21- Estos resultados al ser positivos indican que el porcentaje de voto de Cs, Vox y AA no se determinan únicamente en el municipio al cual corresponden, sino que por el contrario los aspectos relacionados con estas variables se ven influenciados por la renta neta media declarada de sus vecinos de primer orden.

Ilustración 20. Gráfico de dispersión I de Morán Tasa_AA w/ rentamedia



Fuente: datos obtenidos de GeoDa

Ilustración 21. Mapa de clusters Tasa_AA w/ rentamedia



Fuente: datos obtenidos de GeoDa

Hasta el momento hablamos de un ejercicio de correlación espacial, con una matriz de contactos binaria y criterio de vecindad reina. Pero sería curioso ver como varían los resultados con una contigüidad de segundo orden, -Tabla 4-, esto es los vecinos de los vecinos. Es claro que la dependencia espacial disminuye en la medida que aumenta el orden de vecindad, siendo los vecinos más cercanos aquellos que influyen directamente sobre las condiciones de vida de un municipio, pero como vemos son coeficientes que se mantienen muy cercanos, sobre todo para el PP y no tanto quizás, para Vox, -queriendo remarcar que evidentemente ahora son incluso más cercanos al 0-.

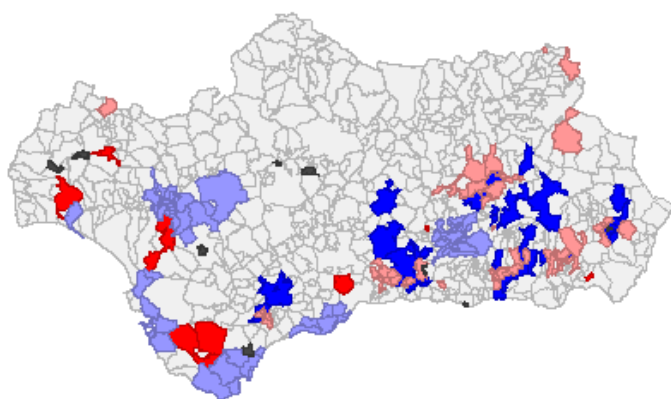
Tabla 4. I de Moran Local Bivariante (retardo espacial: *rentamedia*)

Variable	Primer orden	Segundo orden
Tasa_PSOE	-0,276	-0,176
Tasa_PP	-0,166	-0,123
Tasa_CS	0,285	0,199
Tasa_VOX	0,169	0,084
Tasa_AA	0,209	0,139

Fuente: datos obtenidos de GeoDa

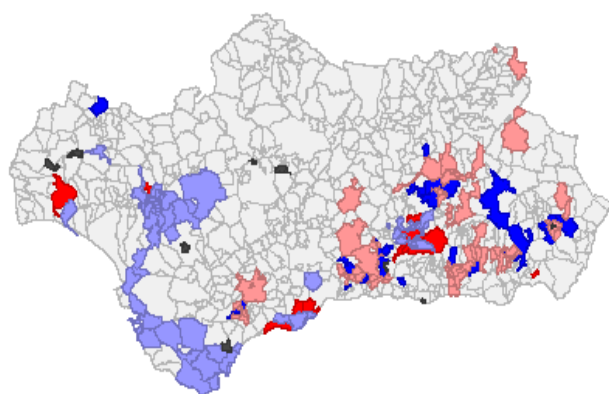
En los mapas de clusters se pueden ver cuales son aquellos municipios con más presencia de de la variable renta media y el voto a los partidos. Es interesante ver como para el caso de los partidos más grandes como PSOE y PP no se encuentran demasiados municipios con un valor alto de ambas variables, pero sí en el resto de partidos, que además presentan esta faceta en los municipios en donde se encuentran las rentas más elevadas, según la ilustración 27 que nos muestra los grupos de renta.

Ilustración 22. Mapa cluster PSOE w/ *rentamedia*



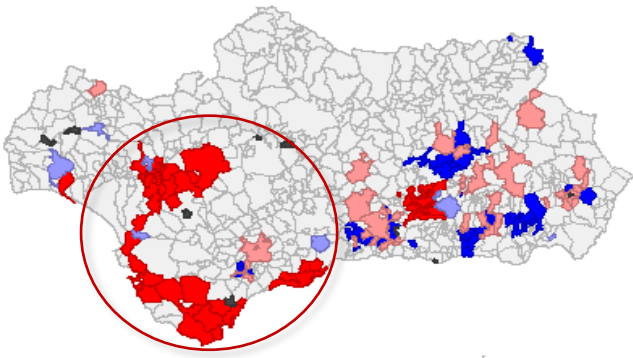
Fuente: datos obtenidos de GeoDa

Ilustración 23. Mapa cluster PP w/ *rentamedia*



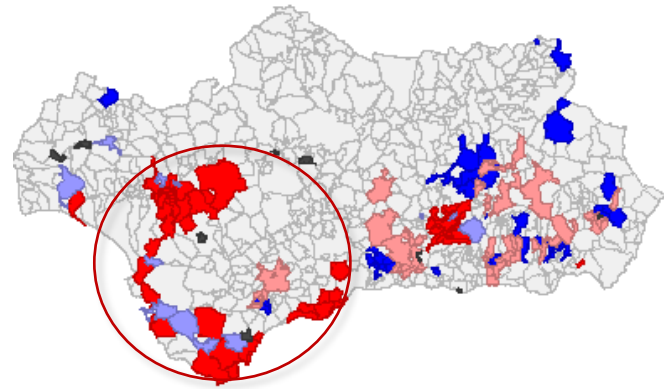
Fuente: datos obtenidos de GeoDa

Ilustración 24. Mapa cluster Cs w/ rentamedia



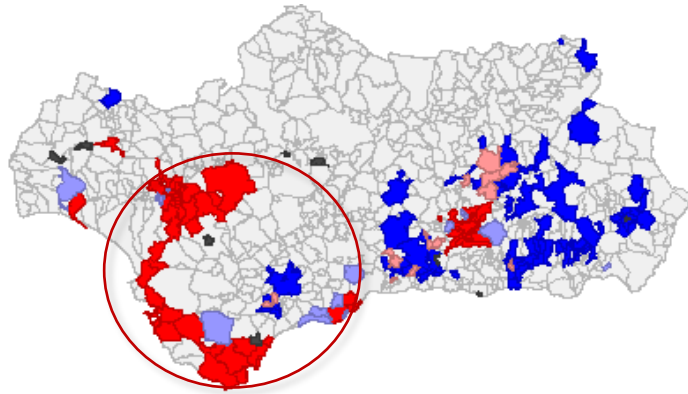
Fuente: datos obtenidos de GeoDa

Ilustración 25. Mapa cluster Vox w/ rentamedia



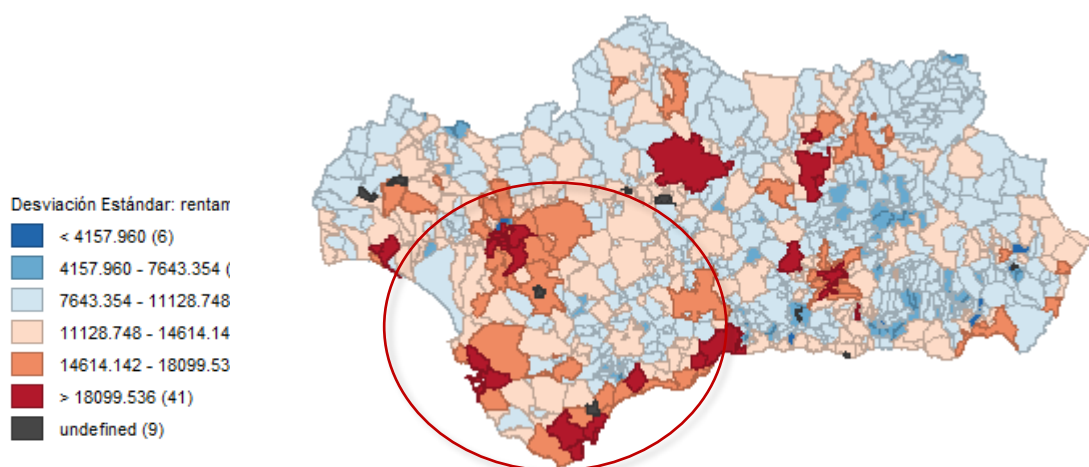
Fuente: datos obtenidos de GeoDa

Ilustración 26. Mapa cluster AA w/ rentamedia



Fuente: datos obtenidos de GeoDa

Ilustración 27. Mapa de desviación estándar de la variable rentamedia

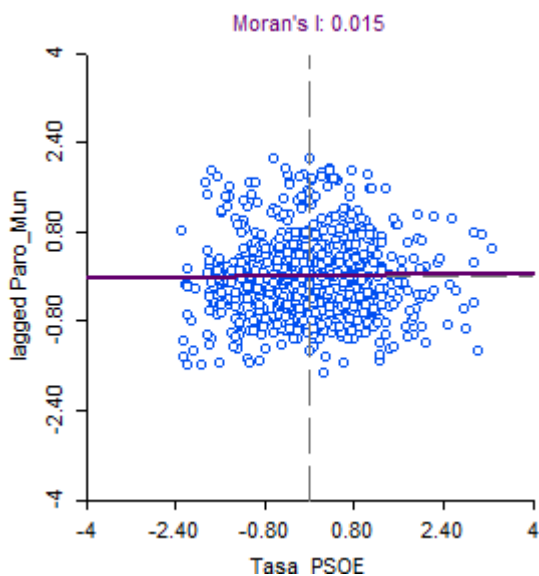


Fuente: datos obtenidos de GeoDa

En este mapa de desviación estándar se pueden ver los diferentes grupos de renta (en la leyenda) y por tanto cuales son los municipios que presentan los mayores niveles de renta. Se puede apreciar que los cluster existentes en el caso de los partidos Cs, Vox y Adelante Andalucía siguen una distribución muy similar, acumulándose en el cuadrante suroeste del mapa autonómico y perteneciendo a los grupos de renta más elevados.

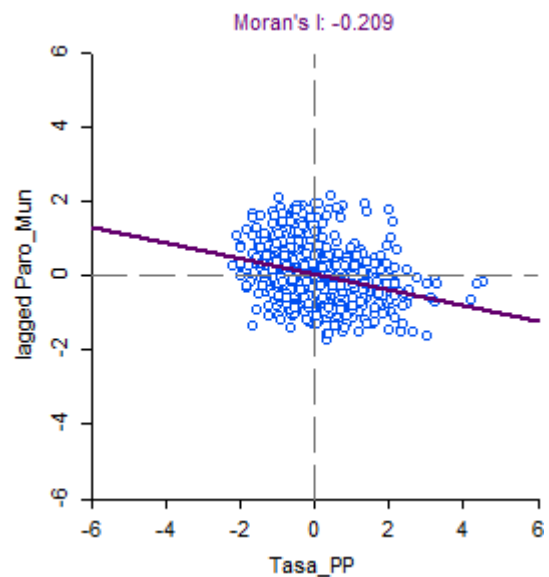
Interesante sería también, para terminar de tratar la pregunta de índole económica, realizar el mismo análisis con el retardo espacial “*Paro_Mun*” -esto es, la tasa de paro municipal-. Se empezará pues comparando los respectivos diagramas de dispersión referidos al I de Morán.

**Ilustración 28. Diagrama de dispersión I de Morán
Tasa_PSOE w/ Paro_Mun**



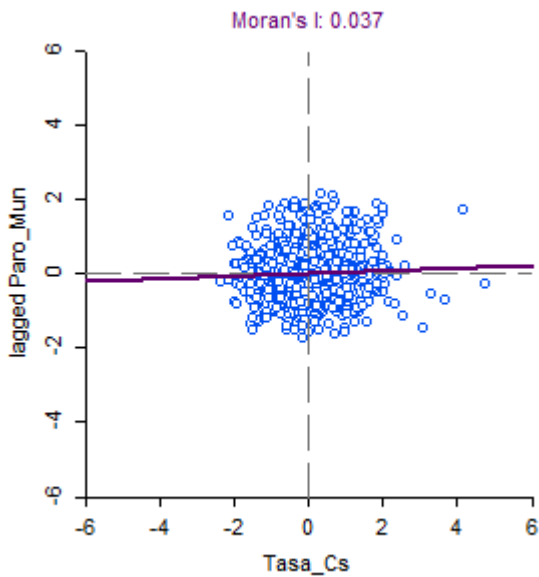
Fuente: datos obtenidos de GeoDa

**Ilustración 29. Diagrama de dispersión I de
Morán Tasa_PP w/ Paro_Mun**



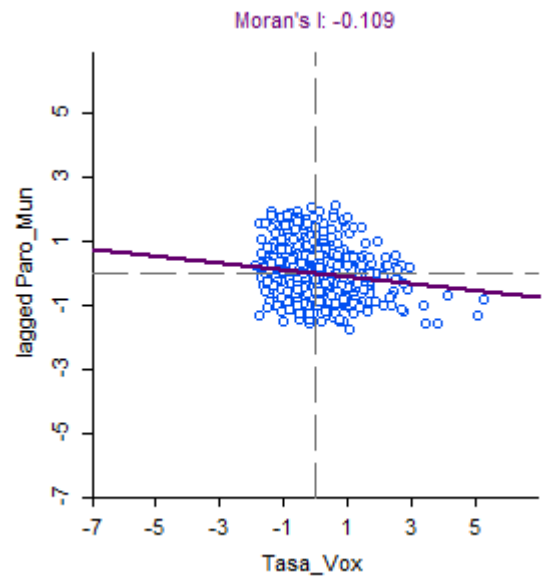
Fuente: datos obtenidos de GeoDa

Ilustración 30. Diagrama de dispersión I de Morán Tasa_Cs w/ Paro_Mun



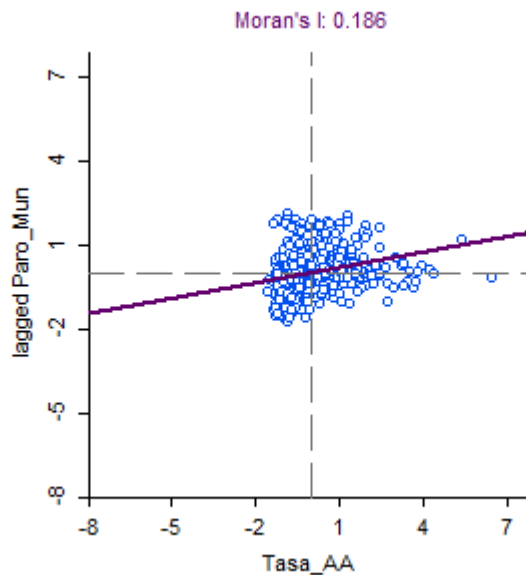
Fuente: datos obtenidos de GeoDa

Ilustración 31. Diagrama de dispersión I de Morán Tasa_Vox w/ Paro_Mun



Fuente: datos obtenidos de GeoDa

Ilustración 32. Diagrama de dispersión I de Morán Tasa_AA w/ Paro_Mun

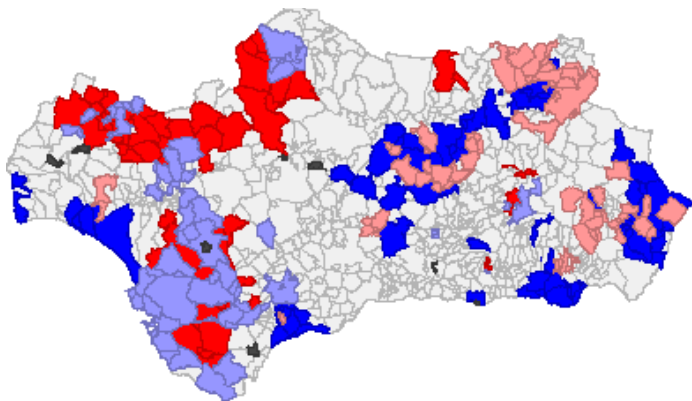


Fuente: datos obtenidos de GeoDa

Se puede apreciar que ninguno de los coeficientes se distancia lo suficiente del 0 como para poder hablar de una autocorrelación positiva o negativa, por lo que en este caso se hablaría más bien de una distribución aleatoria, esto es, el rezago espacial -paro municipal- no influye

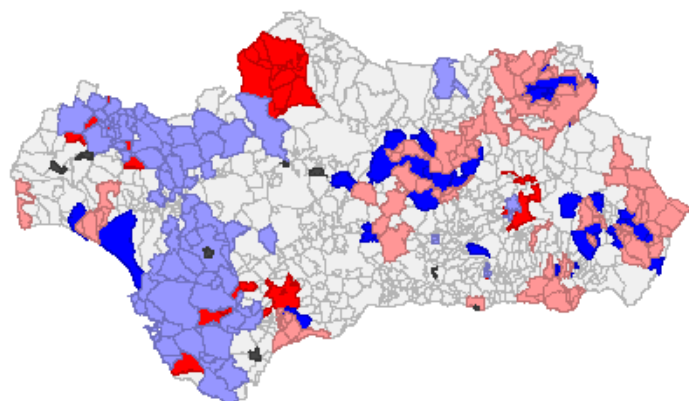
en la distribución espacial de la variable a explicar -porcentaje de votos de los diferentes partidos-. Lo significativo, así como encontrarnos con una distribución aleatoria, es hacerlo en el caso de todos los partidos, solo es el PP el caso en el que parece que tiene más influencia con un coeficiente del (-0,209) siendo de todos modos una autocorrelación negativa, lo que vendría a significar que no presenta evidencias de continuidad a lo largo de los municipios vecinos de primer orden -como se muestra en la ilustración 34-. En cuanto a los otros mapas, estos nos muestran como son los partidos Cs y Adelante Andalucía quienes más agrupaciones en el cuadrante alto-alto presentan, esto es, presentan más casos de municipios contiguos en donde la presencia de la variable y la presencia de una autocorrelación espacial son fuertes y es curioso también que esta presencia se de en el pico de la comunidad autónoma, justo donde no es significativo en el caso de los partidos tradicionales

**Ilustración 33. Mapa de clusters I de Morán
Tasa_PSOE w/ Paro_Mun**



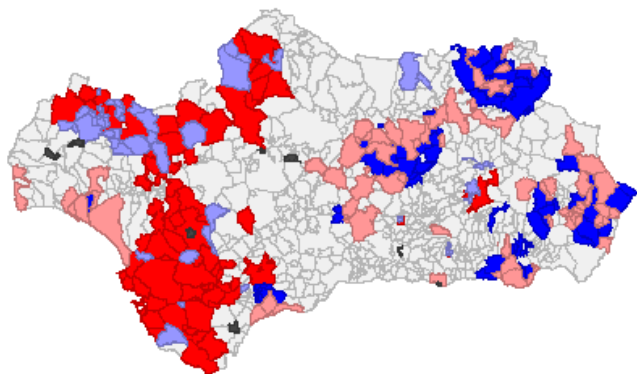
Fuente: datos obtenidos de GeoDa

**Ilustración 34. Mapa de clusters I de Morán
Tasa_PP w/ Paro_Mun**



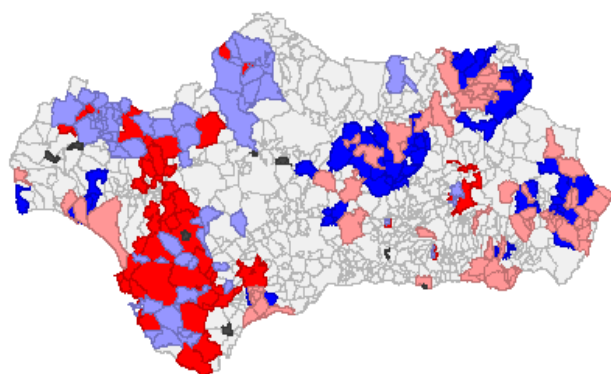
Fuente: datos obtenidos de GeoDa

**Ilustración 35. Mapa de clusters I de Morán
Tasa_Cs w/ Paro_Mun**



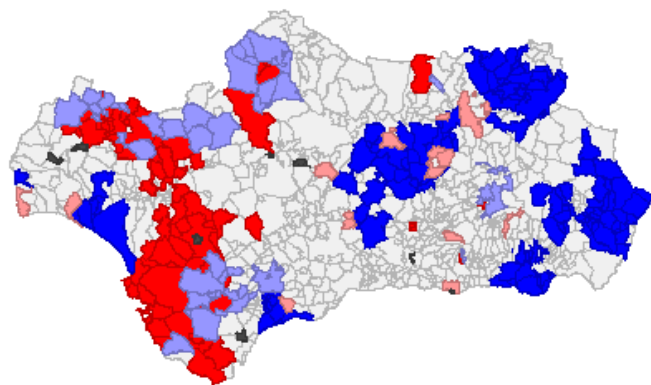
Fuente: datos obtenidos de GeoDa

**Ilustración 36. Mapa de clusters I de Morán
Tasa_Vox w/ Paro_Mun**



Fuente: datos obtenidos de GeoDa

**Ilustración 37. Mapa de clusters I de Morán
Tasa_AA w/ Paro_Mun**



Fuente: datos obtenidos de GeoDa

2.1.3 Edad como factor explicativo

Tras haber finalizado el análisis respecto al factor económico y su relación con el voto, es hora de dar paso al apartado relacionado con la edad de los votantes. Para poder estudiar esta parte contamos con la presencia de dos variables: porcentaje de mayores de 65 años y menores de 20 años.

Tabla 5. Coeficientes de I Morán (retardos: “porc65y” & “menos20”)

	Mayores de 65	Menores de 20
Tasa_PSOE	0.230	-0.218
Tasa_PP	0.234	-0.237
Tasa_Cs	-0.216	0.227
Tasa_Vox	-0.139	0.127
Tasa_AA	-0.265	0.263

Fuente: elaboración propia a partir de datos obtenidos de GeoDa

Se observa que los valores son altos en todos los casos -el que menos Vox- y las líneas de tendencia hacen pensar que los partidos tradicionales obtienen su voto de los municipios con mayor presencia de población anciana (coeficientes positivos, lo que indica autocorrelación espacial positiva) y lo contrario para el resto de sus rivales políticos. Todo esto es teniendo en cuenta el criterio de contigüidad 1, que recordemos, trata a los vecinos más cercanos, y siendo valores tan altos sería interesante saber si con criterios de contigüidad 2 o más se mantendrían.

Tabla 6. Coeficientes de I Morán -con criterios de contigüidad 2- (retardos: “porc65y” & “menos20”)

	Mayores de 65	Menores de 20
Tasa_PSOE	0.129	-0.114
Tasa_PP	0.231	-0.226
Tasa_Cs	-0.156	0.163
Tasa_Vox	0.032	0.009
Tasa_AA	-0.236	0.227

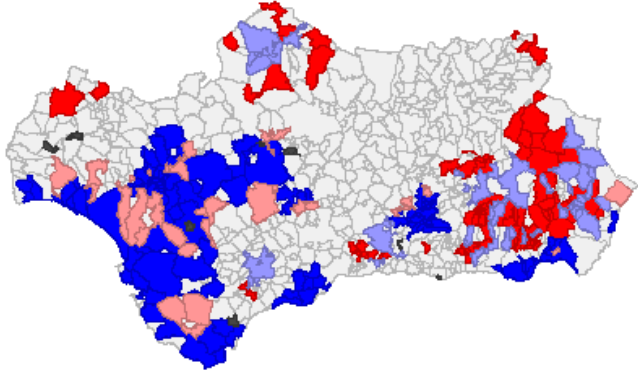
Fuente: elaboración propia a partir de datos obtenidos de GeoDa

En este punto encontramos varios hechos significativos. Los dos grandes partidos tradicionales, como son PSOE y PP, cuentan con valores del coeficiente de I de Morán positivos, cuando hablamos de municipios que ostentan mayor porcentaje de gente de avanzada edad, esto significa que existe una autocorrelación positiva $-(0.129)$ y (0.231) significativamente- y sigue siendo un valor relativamente alto. Al contrario, por ejemplo, que partidos más nuevos como Cs o Adelante Andalucía -este último no es un partido nuevo como tal, pero no debemos olvidar que este año se presentaron junto a Podemos-, en donde nos encontramos con coeficientes de I de Morán negativos cuando hablamos de más presencia de mayores de 65 años -a saber (-0.156) y (-0.236) respectivamente- y positivos cuando hablamos de presencia más juvenil -a saber (0.163) y (0.227) respectivamente-. El cambio más significativo es el caso de Vox quien ya ostenta valores muy cercanos al 0, además de ser visible un cambio de signo en el coeficiente cuando hablamos del retardo “mayores de 65” y es que con el segundo criterio de contigüidad se pasa de una autocorrelación negativa a una positiva, es decir, al hablar de una autocorrelación negativa en los vecinos de primer orden para el caso de Vox es indicador de que en estos municipios el retardo espacial no es explicativo del voto de Vox, sin embargo, en los “vecinos de los vecinos” empieza a serlo. Esto significa que, en los municipios de segundo orden, esto es, los municipios colindantes a aquellos donde existe más presencia de la variable, que ostenten más presencia de gente residente de más de 65 años, para el caso de Vox, empieza a haber una tendencia positiva al voto, una tendencia, que, en los vecinos de primer orden, no existía.

Es de recalcar también cómo se diferencian los “clusters”¹⁹ generados para el caso de Cs, Vox y Adelante Andalucía de los partidos de la vieja política a la hora de hablar de la variable “menos20”, haciendo ver la importancia que ostenta en estos partidos y la ausencia de esta en los partidos más tradicionales.

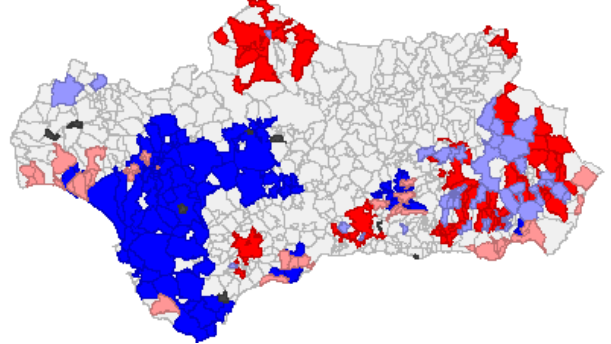
¹⁹ Los siguientes mapas de clusters están generados con criterios de contigüidad 1.

**Ilustración 38. Mapa de clusters Tasa_PSOE w/
Porc_65y**



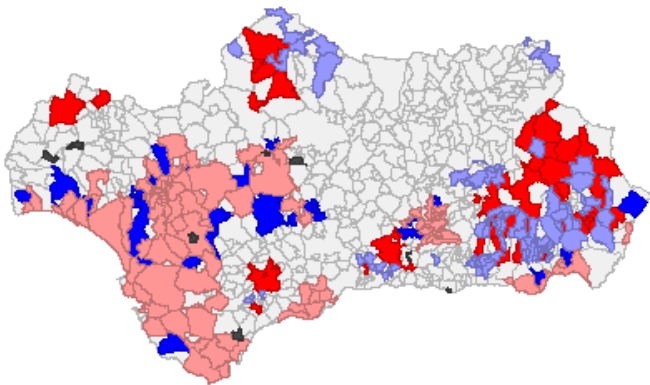
Fuente: datos obtenidos de GeoDa

**Ilustración 39. Mapa de clusters Tasa_PP w/
Porc_65y**



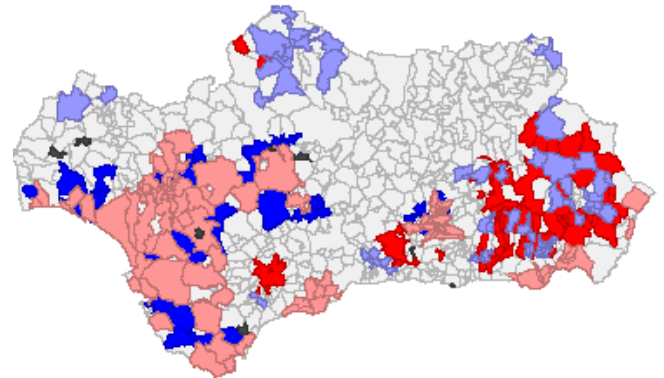
Fuente: datos obtenidos de GeoDa

**Ilustración 40. Mapa de clusters Tasa_Cs w/
Porc_65y**



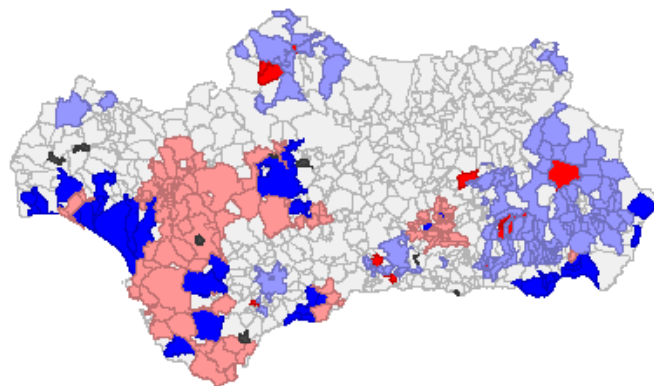
Fuente: datos obtenidos de GeoDa

**Ilustración 41. Mapa de clusters Tasa_Vox w/
Porc_65y**



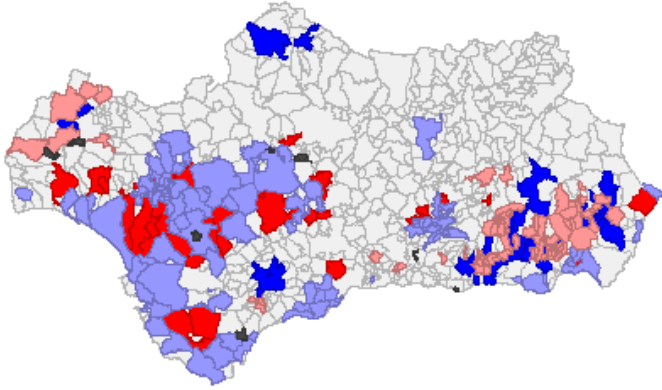
Fuente: datos obtenidos de GeoDa

**Ilustración 42. Mapa de clusters Tasa_AA w/
Porc_65y**



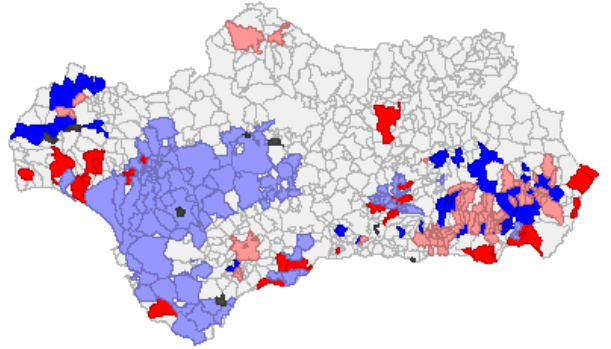
Fuente: datos obtenidos de GeoDa

Ilustración 43. Mapa de clusters Tasa_PSOE w/ menos20



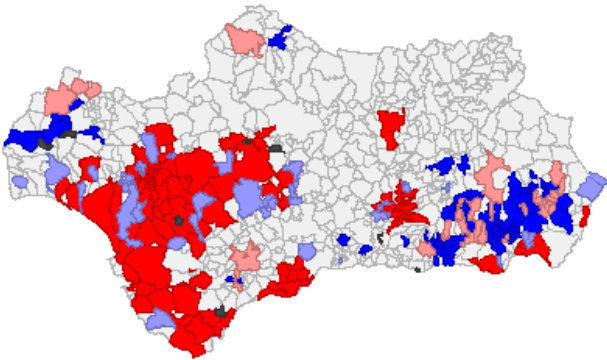
Fuente: datos obtenidos de GeoDa

Ilustración 44. Mapa de clusters Tasa_PP w/ menos20



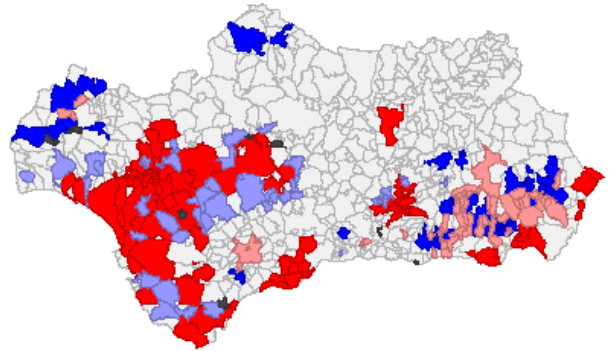
Fuente: datos obtenidos de GeoDa

Ilustración 45. Mapa de clusters Tasa-Cs w/ menos20



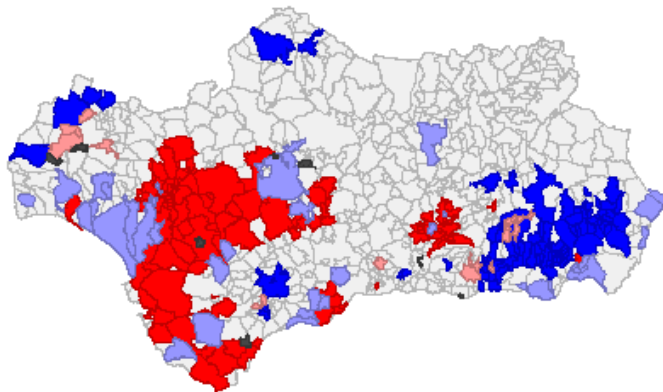
Fuente: datos obtenidos de GeoDa

Ilustración 46. Mapa de clusters Tasa_Vox w/ menos20



Fuente: datos obtenidos de GeoDa

Ilustración 47. Mapa de clusters Tasa_AA w/ menos20

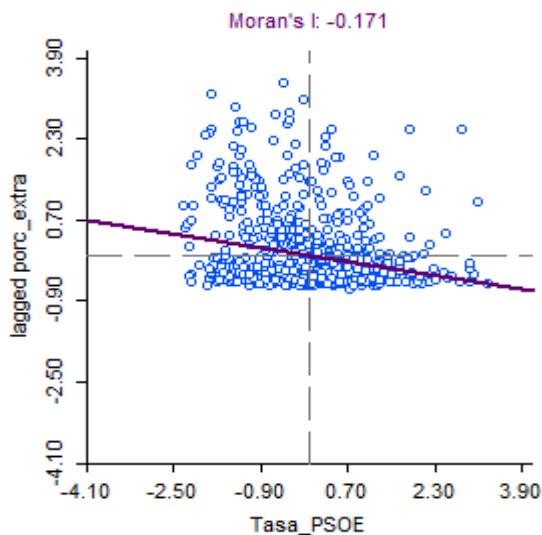


Fuente: datos obtenidos de GeoDa

2.1.4 ¿Influye la nacionalidad?

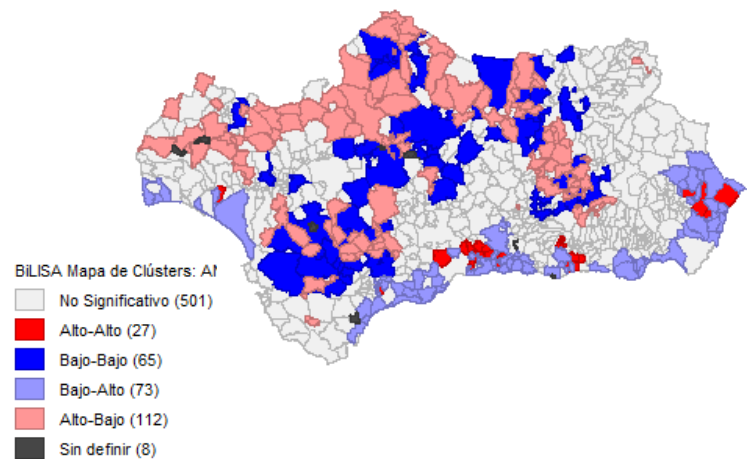
Durante los dos subapartados anteriores se ha analizado la influencia -más alta o más baja- de los factores económicos y de la edad en el proceso de determinación del voto. Ahora, con motivo de responder a la última de las cuestiones planteadas en la introducción de este trabajo, se tratará de analizar la dependencia espacial respecto al porcentaje de extranjeros sobre la población total del municipio, con la finalidad de averiguar si hay algún tipo de relación o patrón de interés. Al igual que se hizo en los apartados anteriores, se utilizará el coeficiente I de Morán, con una matriz de pesos establecida por el criterio de contigüidad reina para vecinos de primer orden estableciendo como variables a explicar los porcentajes de los diferentes partidos y su retardo espacial.

Ilustración 48. Gráfico de dispersión de Morán Tasa PSOE w/ porc extra



Fuente: datos obtenidos de GeoDa

Ilustración 49. Mapa de clusters de Morán Tasa_PSOE w/ porc_extra

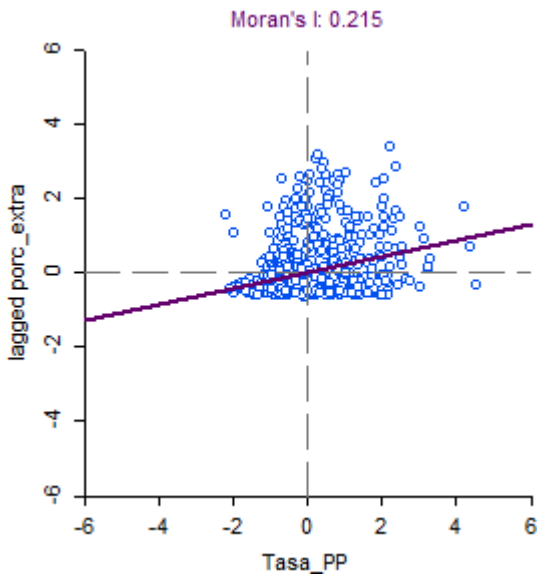


Fuente: datos obtenidos de GeoDa

En el caso del PSOE el coeficiente I de Morán es negativo, siendo indicador de una autocorrelación negativa, en donde sus mayores casos se agrupan en los cuadrantes Alto-Bajo y Bajo-Alto, esto es en la mayoría de casos la variable tiene un alto valor, pero una autocorrelación espacial baja y viceversa. En cambio, en el caso del PP vemos como además de ser más significativo -ilustración 50-, hablamos de una autocorrelación positiva (0.215) en donde la mayoría de los casos se posicionan en el cuadrante Bajo-Bajo, pero seguido del Alto-

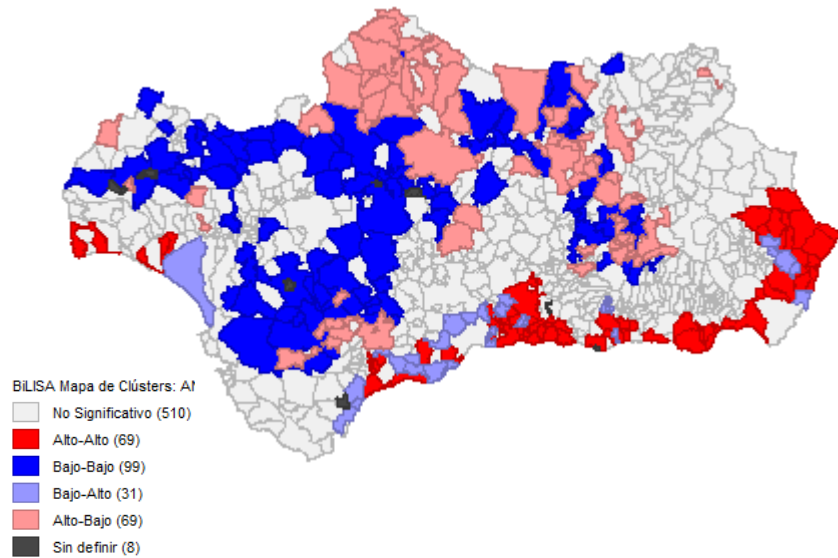
Alto, siendo que, este último cuadrante agrupa a sus municipios contiguos a lo largo de la costa Sureste del mapa.

Ilustración 50. Diagrama de dispersión de Morán Tasa_PP w/ porc_extra



Fuente: datos obtenidos de GeoDa

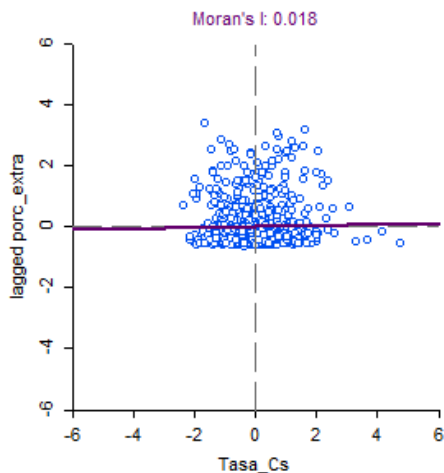
Ilustración 51. Mapa de clusters de Morán Tasa_PP w/ porc_extra



Fuente: datos obtenidos de GeoDa

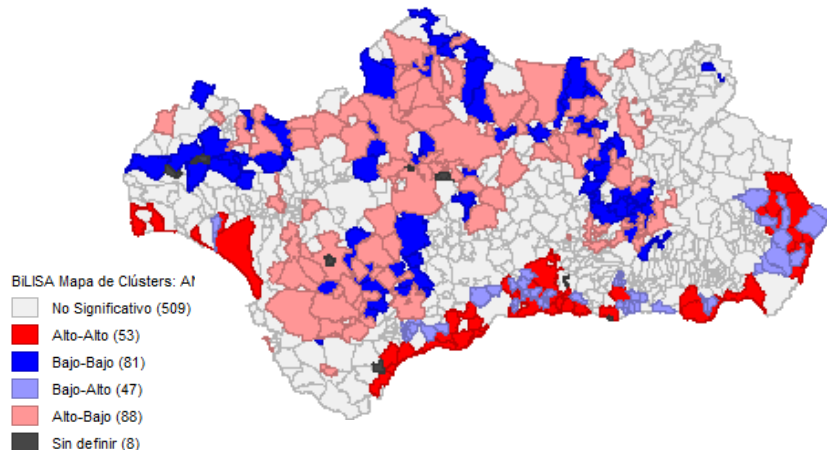
Por su parte el escenario que presenta Cs es el de un acercamiento a una distribución aleatoria (0.018), aun así, vemos que existen 53 municipios que presentan valores altos de la variable, a la vez que una autocorrelación elevada (cuadrante alto-alto), a lo largo de la zona costera, similar a la de la distribución del PP -ilustración 53-.

Ilustración 52. Diagrama de dispersión de Morán Tasa_Cs w/ porc_extra



Fuente: datos obtenidos de GeoDa

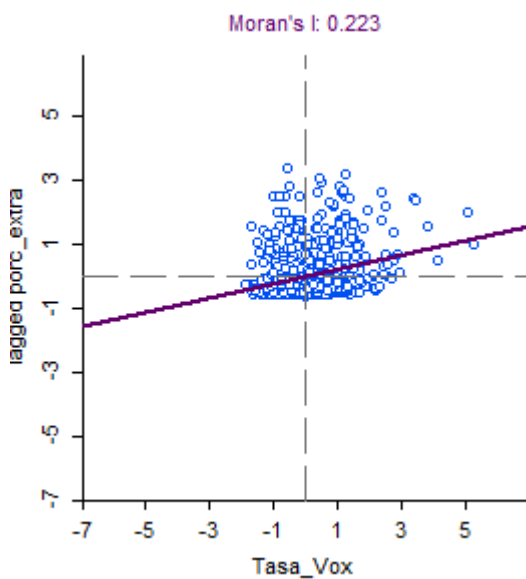
Ilustración 53. Mapa de clusters de Morán Tasa_Cs w/ porc_extra



Fuente: datos obtenidos de GeoDa

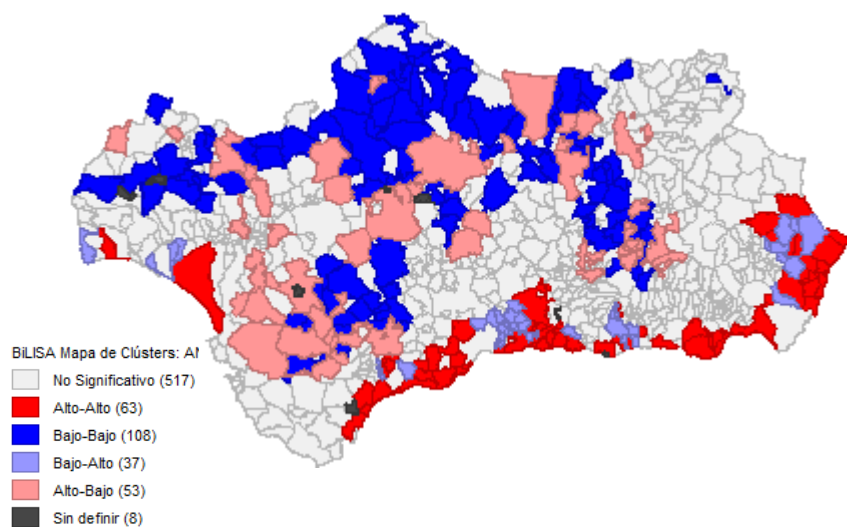
Uno de los casos más llamativos es Vox, el cual presenta el coeficiente más elevado hasta ahora (0.223) frente al (0.215) del PP. Se habla de una autocorrelación positiva, esto es, el partido de derecha radical presenta una tendencia mayor al resto de sus rivales políticos, a conseguir porcentajes de voto más elevados en aquellos municipios en donde existe una mayor presencia de residentes extranjeros, que como se aprecia en la ilustración 55 destacan las zonas costeras similares también a la de los dos partidos anteriores.

Ilustración 54. Diagrama de dispersión de Morán
Tasa_Vox w/ porc_extra



Fuente: datos obtenidos de GeoDa

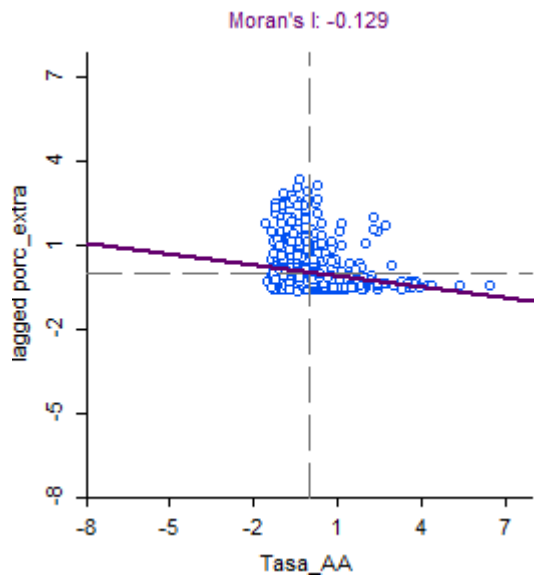
Ilustración 55. Mapa de clusters de Morán
Tasa_Vox w/ porc_extra



Fuente: datos obtenidos de GeoDa

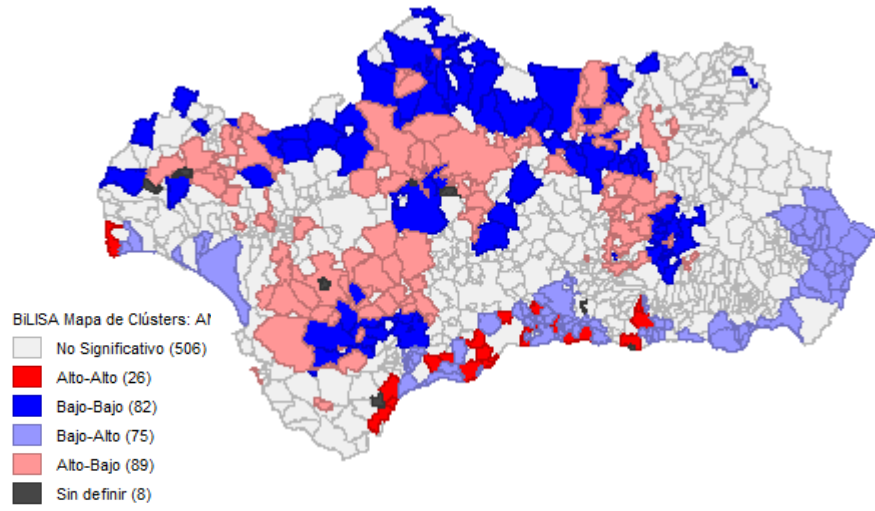
Por último, Adelante Andalucía no presenta signos de autocorrelación positiva, sino todo lo contrario, con un coeficiente negativo de (-0.129) presenta tendencia hacia la autocorrelación negativa, por lo que en los municipios que presenten mayores porcentajes de extranjeros en relación con la población total del municipio, la autocorrelación espacial negativa nos indicará que no existe tanta asociación a su entorno, y por ende, en este caso el retardo de la variable a explicar, porcentaje de extranjeros, no presenta evidencias de continuidad a lo largo de los municipios vecinos de primer orden, para esta formación.

Ilustración 56. Diagrama de dispersión I de Morán Tasa_AA w/ porc_extra



Fuente: datos obtenidos de GeoDa

Ilustración 57. Mapa de clusters de Morán Tasa_AA w/ porc_extra



Fuente: datos obtenidos de GeoDa

3. ANÁLISIS CONFIRMATORIO DE DATOS ESPACIALES

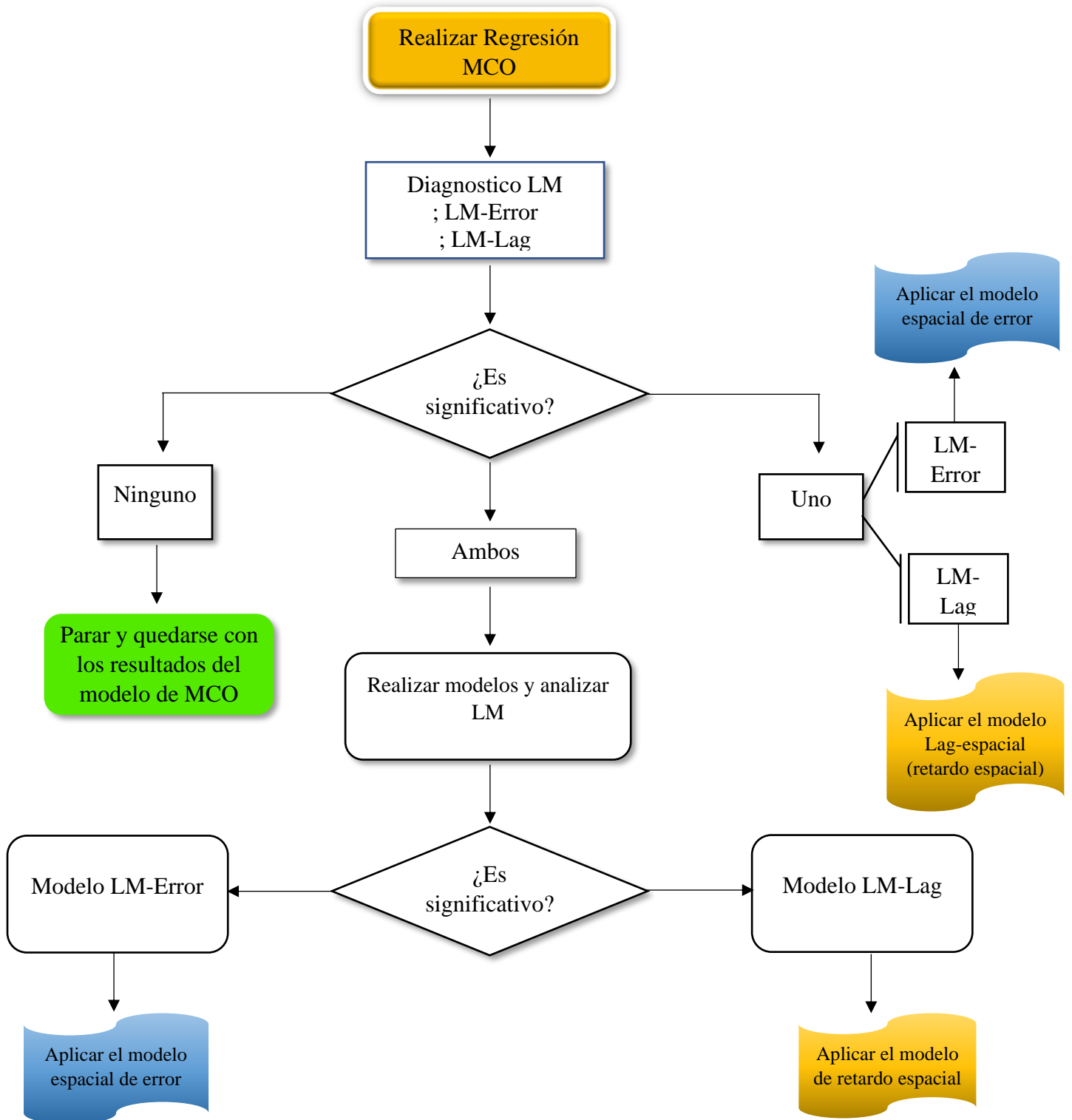
El conjunto de estadísticas de prueba para la autocorrelación espacial puede parecer desconcertante al principio, pero hay una forma bastante intuitiva de proceder con los resultados hacia una especificación de regresión espacial, siguiendo las indicaciones propuestas por Luc Anselin (2003). Este proceso se resume en la Ilustración 58.

El proceso comienza en la parte superior del gráfico. El primer paso es realizar un modelo de regresión lineal estándar y analizar los estadísticos de prueba LM-Error y LM-Lag. Si ninguno rechaza la hipótesis nula de no existencia de autocorrelación espacial (dependencia espacial sustantiva) y de correlación espacial en el error (dependencia residual), respectivamente, se tendrá que quedar con los resultados del modelo realizado por Mínimos Cuadrados Ordinario (MCO). Sin embargo, si una de las estadísticas de la prueba LM rechaza la hipótesis nula, y la otra no, entonces la decisión es fácil también: estimar el modelo alternativo de regresión espacial que coincida con la estadística de prueba que rechaza el valor nulo. Es decir, si, por ejemplo, LM-Error rechaza el valor nulo, pero LM-Lag no, se estima un modelo espacial de error²⁰, y viceversa. En el caso de que ambas estadísticas de prueba LM rechacen la hipótesis nula, se procederá a la parte inferior del gráfico. Realizamos ambos modelos y se procederá a detectar cuál de los dos es más significativo. Si se da el caso de que uno solo uno de ellos sea significativo o tenga órdenes de magnitud más significativos que el otro (por ejemplo: $p < 0.00000$ comparado con $p < 0.03$) se estimará el modelo de regresión espacial que coincida con la estadística más significativa. En el caso de que ambos sean altamente significativos, se procederá con el modelo que posea el valor más grande para la estadística de prueba. También hay casos raros en los que ninguno de las estadísticas de prueba *Robust LM* son significativas. En esos casos, más graves es probable que haya problemas de especificación²¹ errónea y que deben abordarse primero.

²⁰ Los modelos de retardo espacial presentan la correlación -o dependencia- espacial en la variable dependiente (Y). Estos modelos pretenden probar la presencia de dependencia espacial mediante pruebas de diagnóstico. Permiten a las observaciones de la variable dependiente Y en el área *i* depender de las observaciones vecinas. Por su parte, los modelos de error espacial, explican la dependencia espacial en términos de error. La dependencia del error espacial puede surgir de variables latentes no observables que están correlacionadas espacialmente, aunque también puede surgir de los límites del área que no reflejan con precisión la vecindad que dan lugar a las variables del análisis.

²¹ En el análisis de la regresión y campos relacionados, tales como la econometría, la especificación es el proceso de convertir la teoría en un modelo de regresión

Ilustración 58. Proceso de selección de modelo



Fuente: (Anselin, 2005)

3.1 Partido Socialista Obrero Español (PSOE)

Es importante aclarar, que, una vez empezado el análisis de los datos, el esquema de trabajo utilizado va a ser el mismo siempre para cada partido. Los pasos a seguir son los detallados en el apartado anterior (ver Ilustración 58) siguiendo la guía de Luc Anselin. Esto es, realizar un modelo MCO estándar y analizar en él si se rechazan las H_0 de dependencia espacial -entendida como el fenómeno que se presenta cuando los valores observados en una región dependen de los valores de regiones vecinas- de los estadísticos LM-Err y LM-Lag. A partir de ahí, seguir con las indicaciones de la ilustración 58. Las variables a introducir en cada caso son el porcentaje de voto para cada partido como variable dependiente (VD) y como variables independientes (Vis) se utilizarán: la renta neta media declarada por habitante del municipio (*rentamedia*), la tasa de paro municipal (*Paro_Mun*), el porcentaje de población extranjera residente en el municipio (*porc_extra*), el porcentaje de población menor de 20 años (*menos20*) y el porcentaje de población mayor de 65 años (*porc_65y*).

Como último apunte, se debe señalar que las tablas generadas por el programa geostatístico (GeoDa) de los modelos redefinidos en términos de retardo y error para el caso de todos los partidos estarán incluidas en los anexos (los modelos de regresión básicos por MCO sí estarán en el texto) con motivo de no alargar en exceso la extensión pautada para este trabajo, aunque es recomendable tenerlas a mano para seguir la explicación de cada caso - además en los anexos está incluido un cuadro que contiene de forma resumida la descripción de los estadísticos utilizados en el análisis-. Sin embargo, para facilitar la comprensión del texto existen a lo largo de los siguientes subapartados tablas que facilitan los contenidos analizados en dichos modelos y que permiten la comparación, con el objetivo de elegir el modelo más válido en cada supuesto.

Así pues, el primero de los modelos a plantear va a ser un modelo básico de regresión lineal por MCO en donde la VD es Tasa_PSOE y las Vis serían las ya mencionadas: *rentamedia*, *menos20*, *porc_extra*, *Tasa_Paro* y *Porc_65y*. Así pues, obtenemos el siguiente modelo (Tabla 7):

Tabla 7. Modelo de regresión lineal básico por MCO; (VD:Tasa_PSOE)

Data set : **ANDALUCIA**
 Dependent Variable : **Tasa_PSOE** Number of Observations: 777
 Mean dependent var : 38,0808 Number of Variables : 6
 S.D. dependent var : 10,7522 Degrees of Freedom : 771

R-squared : 0,317848 F-statistic : 71,8494
 Adjusted R-squared : 0,313424 Prob(F-statistic) : 0
 Sum squared residual: 61276,9 Log likelihood : -2799,37
 Sigma-square : 79,4772 Akaike info criterion : 5610,75
 S.E. of regression : 8,915 Schwarz criterion : 5638,68
 Sigma-square ML : 78,8634
 S.E of regression ML: 8,88051

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	32,3354	5,27388	6,13125	0,00000
porc_extra	-27,1198	3,60287	-7,52727	0,00000
Porc_65y	0,554599	0,111699	4,96511	0,00000
menos20	0,408082	0,15501	2,63262	0,00864
Paro_Mun	0,0943554	0,0646639	1,45917	0,14493
rentamedia	-0,001231	0,000110664	-11,1237	0,00000

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 44,631047
 TEST ON NORMALITY OF ERRORS
 TEST DF VALUE PROB
 Jarque-Bera 2 83,0308 0,00000

.DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS
 TEST DF VALUE PROB
 Breusch-Pagan test 5 60,0262 0,00000
 Koenker-Bassett test 5 33,6515 0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : ANDALUCIA
 (row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0,1986	9,3103	0,00000
Lagrange Multiplier (lag)	1	89,2486	0,00000
Robust LM (lag)	1	10,1581	0,00144
Lagrange Multiplier (error)	1	81,4723	0,00000
Robust LM (error)	1	2,3818	0,12275
Lagrange Multiplier (SARMA)	2	91,6304	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

La tabla anterior pone de manifiesto la significatividad individual de los parámetros exceptuando el caso de la variable *Paro_Mun* (t-student) y un R de 0,31 (R ajustado 0,31). Pero los resultados quedan invalidados puesto que los contrastes de los multiplicadores de Lagrange

para el retardo (LM-LAG) y para el error (LM-ERR) rechazan la hipótesis nula de no existencia de autocorrelación espacial (dependencia espacial sustantiva ²²) y de correlación espacial en el error (dependencia residual ²³) y siendo además LM-LAG mayor (89,2486) que LM-ERR (81,4723), lo que sugiere una reespecificación del modelo incluyendo un esquema autorregresivo en términos de retardo espacial. En el modelo básico de regresión lineal podemos apreciar como de las variables explicativas son todas significativas a excepción de la tasa de paro, la cual con un $p > 0.05$ se asume que no es significativa de este modelo. Así pues, se procederá a redefinir el modelo incluyendo una estructura autorregresiva espacial en términos de error y de retardo espacial para poder comparar. Como ya se especificó anteriormente, los modelos espaciales de error tratan de explicar la dependencia espacial en términos de error, y por su parte, los segundos modelos -aquellos que tratan de explicar la dependencia en términos de retardo- tratan de explicarla mediante la inclusión de retardos espaciales. Vaya y Montero (2000), plantean el operador de retardos espaciales como solución alternativa a la matriz de datos espaciales W para hacer frente a los problemas derivados del análisis regional. Este consiste en un promedio ponderado de variables aleatorias en localizaciones vecinas -con ponderaciones fijas y exógenas-. El operador de retardos espaciales se forma definiendo para cada localización i a su vecino en la correspondiente columna como un elemento distintivo de 0 w_{ij} en una matriz de pesos espaciales W positiva y no estocástica. Es decir, se añaden retardos espaciales a las variables explicativas con el objetivo de analizar si existe o no dependencia espacial, más allá de la relación existente por parte de cada variable explicativa *per se*.

En el modelo por error ²⁴ se aprecia como un coeficiente en los errores espacialmente correlacionados (LAMBDA) es agregado como un indicador adicional. Este es un estadístico que se basa en la reducción proporcional del error de predicción cuando se utilizan los valores de la variable independiente para predecir los de la dependiente. Tiene un efecto positivo y es relativamente significativo siendo que LAMBDA=0,451165 sobre un valor de 1. Como resultado, la redefinición mejoró el ajuste general del modelo, como se indica en valores más altos de R cuadrado y Log likelihood.

²² La dependencia espacial sustantiva se presenta como un caso en el que el valor de una región depende del valor de sus regiones vecinas (definidas según una matriz de contactos).

²³ La dependencia espacial residual podría explicarse como “la omisión de variables no cruciales que se hallen correlacionadas espacialmente o por la existencia de errores de medida” (Moreno & Vayá, 2000).

²⁴ Se debe recordar que para ver las tablas íntegras de los modelos por retardo y error se debe acudir a los anexos en su correspondiente subapartado o en su defecto acudir a la tabla resumen.

Se puede ver en el modelo que la baja probabilidad ($p < 0.05$) en el test de Breusch-Pagan sugiere que existe heteroscedasticidad en el modelo después de redefinirlo en términos de error. Además, en el Likelihood Ratio Test ²⁵, el resultado es también significativo. Es importante también destacar que ha habido un cambio en la significación de las variables en cada modelo. Mientras que para el modelo realizado por MCO la variable referida al paro no es significativa, en el modelo de error sí lo es ($p = 0,04533$). Por ello se puede concluir que la introducción de una estructura autorregresiva en términos de error mejoró el modelo y además los factores significativos del modelo de MCO lo siguen siendo en este (a excepción del paro, la cual no es significativa en el modelo estándar y sí en el de error)

Para el caso del modelo en términos de retardo, se debe observar que apareció W_Tasa_PSOE como un indicador adicional. Su parámetro (ρ) refleja la dependencia espacial inherente a nuestros datos de muestra, midiendo la influencia promedio en observaciones por sus observaciones vecinas. Tiene un efecto positivo y es relativamente significativo, aunque menos que el estadístico LAMBDA del modelo anterior (0,397162). Como resultado, el ajuste general del modelo mejoró, ya que además indica valores más altos de R cuadrado y Log Likelihood sigue siendo significativo.

Podemos ver que la baja probabilidad en la prueba de Breusch-Pagan sugiere que todavía hay heteroscedasticidad en el modelo después de introducir el término de retardo espacial. Y en el Likelihood Ratio Test of Spatial Lag Dependence, el resultado sigue siendo significativo. Por lo tanto, sí concluimos que, la introducción del término de retardo espacial mejoró el ajuste del modelo sin hacer que los efectos espaciales desaparecieran. Se puede apreciar, por último, como en este caso no se vuelve significativa la variable del paro, a diferencia del modelo de error ($p = 0,11329$).

²⁵ Conocido en español como razón de verosimilitud, el Likelihood Ratio Test se define como la razón entre la posibilidad de observar la ocurrencia de un evento frente a la posibilidad de no ocurrencia del mismo. Por tanto, para este caso, tenemos que si el valor es alto y $p < 0,05$ existirá más probabilidad de encontrar peso de esas variables en los municipios vecinos que de no encontrarla.

Tabla 8. Comparativas modelos

	<i>Modelo espacial en términos de error</i>	<i>Modelo espacial en términos de retardo</i>
R ²	0,404916	0,401325
Log Likelihood	-2761,576086	-2760,21
Breusch-Pagan (p<0.05)	61,9820	67,9638
Likelihood Ratio Test(p<0.05)	75,5950	78,3291
W_TasaPSOE	-	9,41921
LAMBDA (z valor)	9,98223	-
LM-LAG	-	89,2486
LM-ERR	81,4723	-

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Ahora viene la pregunta ¿cuál escoger? Comparando los modelos de retardo espacial y error espacial (tabla 8), podemos ver que ambos modelos alternativos mejoran el rendimiento del modelo MCO original. Por lo tanto, debemos concluir que controlar la dependencia espacial mejorará el rendimiento de nuestro modelo. Ahora la pregunta ¿cuál de los dos modelos es mejor? Hasta cierto punto, esta es una pregunta abierta. En este caso se procederá a comparar el rendimiento del modelo según los parámetros siguiendo las indicaciones de Anselin (2005) en su manual de uso de GeoDa. En este caso, el modelo de error espacial tiene mayores valores de R cuadrado, aunque son subidas muy pequeñas comparado con los valores de los demás estadísticos en los que gana el modelo por retardo, por lo que se quedara con este último.

3.2 Partido Popular

Se vuelve a presentar en primer lugar un modelo de regresión lineal básico por MCO, en donde las variables van a ser: Tasa_PP para el caso de la VD y porc_extra, Porc_65y, menos20, Paro_Mun y rentamedia -y se mantendrán para los modelos redefinidos-.

Tabla 9. Modelo de regresión lineal básico por MCO; (VD: Tasa_PP)

Data set : **ANDALUCIA**
 Dependent Variable : **Tasa_PP** Number of Observations: 777
 Mean dependent var : 22,2762 Number of Variables : 6
 S.D. dependent var : 8,77466 Degrees of Freedom : 771

R-squared : **0,176970** F-statistic : 33,1564
 Adjusted R-squared : **0,171632** Prob(F-statistic) : 1,07139e-030
 Sum squared residual: 49237,7 Log likelihood : -2714,39
 Sigma-square : 63,8621 Akaike info criterion : 5440,78
 S.E. of regression : 7,99138 Schwarz criterion : 5468,72
 Sigma-square ML : 63,3689
 S.E of regression ML: 7,96046

Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANT	34,0269	4,72749	7,19766	0,00000
porc_extra	15,0176	3,2296	4,64997	0,00000
Porc_65y	0,0954389	0,100127	0,95318	0,34080
menos20	-0,41358	0,13895	-2,97646	0,00301
Paro_Mun	-0,337947	0,0579645	-5,83024	0,00000
rentamedia	1,50223e-005	9,91992e-005	0,151436	0,87967

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 44,631047
 TEST ON NORMALITY OF ERRORS
 TEST DF VALUE PROB
 Jarque-Bera 2 77,1596 0,00000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS
 TEST DF VALUE PROB
 Breusch-Pagan test 5 70,2036 0,00000
 Koenker-Bassett test 5 48,2506 0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : ANDALUCIA
 (row-standardized weights)
 TEST MI/DF VALUE PROB
 Moran's I (error) 0,1577 7,4309 0,00000
 Lagrange Multiplier (lag) 1 75,3224 0,00000
 Robust LM (lag) 1 31,6967 0,00000
 Lagrange Multiplier (error) 1 51,3734 0,00000
 Robust LM (error) 1 7,7478 0,00538
 Lagrange Multiplier (SARMA) 2 83,0701 0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

En este caso tanto el coeficiente I de Morán de los residuos como ambos estadísticos LM rechazan las respectivas hipótesis nulas de ausencia de dependencia espacial, siendo en este caso el valor de LM-Lag mayor (75,3224) que el de LM-Error (51,3734), por lo que se asume que se debería redefinir el modelo incluyendo un esquema autorregresivo en términos

de retardo espacial, pero se harán ambos para comprobar igual que en el caso anterior y los posteriores.

Tabla 10. Comparativa de modelos

	<i>Modelo espacial en términos de error</i>	<i>Modelo espacial en términos de retardo</i>
R ²	0,250457	0,265895
Log Likelihood	-2689,143472	-2681,26
Breusch-Pagan (p<0.05)	81,8617	88,2317
Likelihood Ratio Test (p<0.05)	50,4963	66,2722
W_TasaPP	-	8,70781
LAMBDA (z valor)	8,17669	-
LM-LAG	-	89,2486
LM-ERR	81,4723	-

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

En ambos modelos vemos como la heteroscedasticidad, medida con el estadístico Breusch-Pagan sigue siendo significativa (p<0.05), al igual que el coeficiente de Likelihood Radio Test que mide la dependencia espacial. Esto supone que la introducción de estos modelos no eliminó los efectos espaciales y además mejora su rendimiento con respecto al modelo de regresión lineal básico por MCO. En este caso se elegirá el modelo redefinido en términos de retardo espacial, puesto que en el modelo de regresión lineal estándar es el que posee mayor valor y en términos comparados con el modelo redefinido en términos de error casi todos los estadísticos son más significativos.

En el modelo de regresión lineal básico por MCO se nos da cuenta de que los principales factores explicativos de la variable dependiente -a saber, en este caso *Tasa_PP*- son el porcentaje de extranjeros residentes en el municipio, el número de residentes menores a 20 años y la tasa de paro municipal. Aquí se debe matizar, por otro lado, que el signo del coeficiente I de Morán para el caso de *menos20* era negativo, por lo que señalaba una autocorrelación negativa, así pues, se entiende en este sentido, que la ausencia de residentes menores de 20 años en un municipio es un indicador más significativo que la presencia de mayores de 65 años. En

el modelo espacial redefinido en términos de retardo (Tabla 22 en anexo 3) vemos como son los mismos factores los significativos, por lo que en este aspecto no ha variado.

3.3 Ciudadanos (Cs)

En el modelo por MCO de Cs se puede apreciar como también los multiplicadores de Lagrange rechazan las hipótesis nulas de dependencia espacial, siendo el LM-Error (34,4172) mayor que el LM-Lag (28,5132) lo que induce a pensar en una redefinición del modelo en términos de error. Además, el Robusto de LM-Lag no es significativo, lo que es importante tener en cuenta puesto que, un estimador robusto o que tiene la propiedad de robustez, tiene como característica que su validez no se ve alterada como consecuencia de la violación de alguno de los supuestos de partida. Por tanto si uno no es significativo la validez de ese modelo se ve alterada.

Tabla 11. Modelo de regresión lineal básico por MCO; (VD: Tasa_Cs)

Data set : **ANDALUCIA**
 Dependent Variable : **Tasa_Cs** Number of Observations: 777
 Mean dependent var : 13,498 Number of Variables : 6
 S.D. dependent var : 4,95861 Degrees of Freedom : 771

R-squared : **0,306884** F-statistic : 68,2736
 Adjusted R-squared : **0,302389** Prob(F-statistic) : 0
 Sum squared residual: 13241,8 Log likelihood : -2204,18
 Sigma-square : 17,1748 Akaike info criterion : 4420,36
 S.E. of regression : 4,14425 Schwarz criterion : 4448,3
 Sigma-square ML : 17,0422
 S.E of regression ML: 4,12822

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	1,99056	2,45163	0,811932	0,41708
porc_extra	4,66479	1,67484	2,78521	0,00548
Porc_65y	0,0211299	0,0519249	0,406932	0,68415
menos20	0,160377	0,0720583	2,22565	0,02633
Paro_Mun	0,0200814	0,0300598	0,668047	0,50430
rentamedia	0,000668516	5,14438e-005	12,9951	0,00000

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 44,631047

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	515,1407	0,00000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	47,6055	0,00000
Koenker-Bassett test	5	17,0560	0,00439

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : ANDALUCIA

(row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0,1291	6,1155	0,00000
Lagrange Multiplier (lag)	1	28,5132	0,00000
Robust LM (lag)	1	0,1270	0,72152
Lagrange Multiplier (error)	1	34,4172	0,00000
Robust LM (error)	1	6,0311	0,01406
Lagrange Multiplier (SARMA)	2	34,5443	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

En el modelo redefinido en términos de error (tabla 24 del anexo) se observa como la heterocedasticidad sigue siendo significativa, al igual que el coeficiente del test de Likelihood que denota la dependencia espacial, por ello este modelo propone una mejora del estándar por MCO. Por otra parte, el R^2 ha aumentado de valor y el coeficiente LAMBDA es positivo y

relativamente significativo. En el otro lado tenemos el modelo espacial de retardo. Este también mantiene la heterocedasticidad significativa al igual que el test de Likelihood, asumiendo que es un modelo también mejorado y que no ha reducido los efectos de dependencia espacial. Se debe destacar que el modelo realizado por MCO da por significativa la variable de edad *menos20*, al contrario que ambos modelos posteriores (ver tabla 25 del anexo 3)

Tabla 12. Comparativa de modelos

	<i>Modelo espacial en términos de error</i>	<i>Modelo espacial en términos de retardo</i>
R ²	0,340504	0,333406
Log Likelihood	-2190,059375	-2192,41
Breusch-Pagan (p<0.05)	52,0987	54,1952
Likelihood Ratio Test (p<0.05)	28,2452	23,5376
W_TasaCs	-	4,69887
LAMBDA (z valor)	5,26024	-
LM-LAG	-	28,5132
LM-ERR	34,4172	-

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

En este caso ambos modelos son similares, pero se atenderá al modelo redefinido en términos de error puesto que en el modelo estándar por MCO se mostraba como poseía un valor más alto. En dicho modelo encontramos en los factores que más influyen en la VD -en este caso Tasa_Cs- son el porcentaje de extranjeros y la renta neta declarada media por habitante.

3.4 Vox

Tanto el coeficiente I de Morán como los multiplicadores de Lagrange rechazan las respectivas hipótesis nulas de este modelo de regresión lineal básico por MCO, siendo que el estadístico con un valor más alto es el LM-Error (254,0999) sobre el LM-Lag (232,5833) y siendo menos significativo el robusto de LM-Lag se asume que el modelo que más pueda mejorar al estándar por MCO sea el redefinido en términos de error.

Tabla 13. Modelo básico de regresión lineal por MCO; (VD: Tasa_Vox)

```

Data set           : ANDALUCIA
Dependent Variable : Tasa_Vox  Number of Observations: 777
Mean dependent var : 8,23094    Number of Variables   : 6
S.D. dependent var : 4,17472    Degrees of Freedom    : 771

R-squared          : 0,259632    F-statistic           : 54,0747
Adjusted R-squared : 0,254830    Prob(F-statistic)    : 0
Sum squared residual: 10025,9    Log likelihood        : -2096,1
Sigma-square       : 13,0038    Akaike info criterion : 4204,2
S.E. of regression : 3,60608    Schwarz criterion     : 4232,13
Sigma-square ML    : 12,9034
S.E of regression ML: 3,59213
    
```

Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANT	10,3718	2,13326	4,86195	0,00000
porc_extra	15,0868	1,45735	10,3522	0,00000
Porc_65y	-0,157068	0,0451819	-3,47636	0,00054
menos20	-0,0761512	0,0627008	-1,21452	0,22492
Paro_Mun	-0,0784767	0,0261563	-3,0003	0,00278
rentamedia	0,000299535	4,47633e-005	6,69154	0,00000

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 44,631047

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	130,4928	0,00000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	54,9545	0,00000
Koenker-Bassett test	5	33,4713	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : ANDALUCIA

(row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0,3508	16,3014	0,00000
Lagrange Multiplier (lag)	1	232,5833	0,00000
Robust LM (lag)	1	4,1428	0,04181
Lagrange Multiplier (error)	1	254,0999	0,00000
Robust LM (error)	1	25,6594	0,00000
Lagrange Multiplier (SARMA)	2	258,2427	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Ambos modelos -redefinido en términos de error y de retardo- son similares y mantienen la significatividad de los coeficientes Breusch-Pagan test y Likelihood con lo que eso significa además de incrementar el R cuadrado considerablemente. Son modelos que mejoran al primero,

pero se hará énfasis en el redefinido en términos de error, puesto que tiene valores más altos. Este modelo nos muestra como es la variable *menos20* -los residentes menores de 20 años en un municipio- la única que no es significativa, manteniendo esa tendencia del modelo estándar.

Tabla 14. Comparativa de modelos

	<i>Modelo espacial en términos de error</i>	<i>Modelo espacial en términos de retardo</i>
R ²	0,463687	0,333406
Log Likelihood	-2000,088931	-2005,13
Breusch-Pagan (p<0.05)	23,9359	21,7941
Likelihood Ratio Test (p<0.05)	192,0219	181,9468
W_TasaVox	-	14,7864
LAMBDA (z valor)	15,7776	-
LM-LAG	-	232,5833
LM-ERR	258,2427	-

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

3.5 Adelante Andalucía

En este caso lo lógico sería redefinir el modelo en términos de retardo espacial pues LM-Lag > LM-Error, además de que el robusto de este segundo es menos significativo (p=0,01213) aunque sigue siéndolo pues el valor *p* no es >0,05. Esto es interesante de ver puesto que un estimador robusto es aquel cuya validez no se ve alterada como consecuencia de la violación de alguno de los supuestos de partida. Se puede observar como ambos siguen siendo significativos, sin embargo, para el caso de LM-Error se debe entender que pierde un poco de significatividad por lo que la validez de ese modelo se ve alterada.

Tabla 15. Modelo básico de regresión lineal por MCO; (VD: Tasa_AA)

```

Data set           : ANDALUCIA
Dependent Variable : Tasa_AA  Number of Observations: 777
Mean dependent var : 13,5155  Number of Variables : 6
S.D. dependent var : 8,49693  Degrees of Freedom : 771

R-squared         : 0,155553  F-statistic           : 28,4047
Adjusted R-squared : 0,150076  Prob(F-statistic)    : 1,77073e-026
Sum squared residual: 47371,6  Log likelihood        : -2699,38
Sigma-square      : 61,4418  Akaike info criterion : 5410,76
S.E. of regression : 7,83848  Schwarz criterion     : 5438,7
Sigma-square ML   : 60,9673
S.E of regression ML: 7,80816
    
```

Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANT	16,9282	4,63704	3,65064	0,00028
porc_extra	-10,8367	3,16781	-3,42086	0,00066
Porc_65y	-0,398107	0,0982112	-4,05358	0,00006
menos20	-0,0301851	0,136292	-0,221474	0,82481
Paro_Mun	0,22312	0,0568555	3,92434	0,00009
rentamedia	0,000119893	9,73013e-005	1,23218	0,21826

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 44,631047

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	2024,7562	0,00000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	96,8385	0,00000
Koenker-Bassett test	5	22,0977	0,00050

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : ANDALUCIA

(row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0,1804	8,4741	0,00000
Lagrange Multiplier (lag)	1	90,0405	0,00000
Robust LM (lag)	1	29,1045	0,00000
Lagrange Multiplier (error)	1	67,2283	0,00000
Robust LM (error)	1	6,2923	0,01213
Lagrange Multiplier (SARMA)	2	96,3328	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Si se miran las tablas correspondientes a los modelos redefinidos (en anexos o en la tabla comparativa) ambos modelos mejoran al primero sin eliminar los efectos de dependencia espacial, pero en este caso se utilizará el segundo de ellos, el modelo espacial en términos de

retardo (Tabla 29 en el anexo o tabla 16), puesto que todos los valores de los estadísticos son más altos. En la prueba t de este (mirar tabla en anexo), se observa como los factores que más inciden en el voto a Adelante Andalucía y que poseen ese efecto espacial son todos los estudiados menos la variable referida al porcentaje de residentes menores de 20 años en un municipio.

Tabla 16. Comparativa de modelos

	<i>Modelo espacial en términos de error</i>	<i>Modelo espacial en términos de retardo</i>
R ²	0,252902	0,267172
Log Likelihood	-2666,251239	-2658,72
Breusch-Pagan (p<0.05)	99,9316	102,7546
Likelihood Ratio Test (p<0.05)	66,2604	81,3274
W_TasaAA	-	10,0446
LAMBDA (z valor)	9,6448	-
LM-LAG	-	90,0405
LM-ERR	67,2283	-

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

En las siguientes tablas se expondrán una serie de comparativas acerca de los datos obtenidos a lo largo de la investigación y para ello se procederá a realizar en primera instancia una comparativa de los modelos finalmente seleccionados para cada partido.

Tabla 17. Comparativas modelos espaciales (I)

	PSOE	PP	Cs	Vox	AA
R²	0,40	0,26	0,34	0,46	0,26
Log Likelihood	-2760,21	-2681,26	-2190,05	-2000,08	-2658,72
Breusch-Pagan	67,96	88,23	52,09	23,93	102,75
Likelihood Ratio Test	78,32	66,27	28,24	192,02	81,32
W_Tasa *	9,41	8,70	-	-	10,04
LAMBDA (z valor) **	-	-	5,26	15,77	-

*Para aquellos con modelos de retardo

**Para aquellos con modelos de error

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

La tabla 17 nos muestra los estadísticos más significativos de los modelos elegidos para el estudio en cada partido. En primer lugar, vemos como los modelos con un valor de R² más altos son los del Vox y PSOE seguidos de Cs, AA y PP. Esto induciría a pensar que los modelos elegidos para esos dos primeros partidos son más explicativos. También se ve como los estadísticos que se paran a medir los efectos de la dependencia espacial como Log Likelihood o Likelihood Ratio Test, que miden la probabilidad de encontrar las variables explicativas en los vecinos municipales contiguos, tienen efectos más grandes en Vox, Adelante Andalucía seguido del PP respectivamente, siendo Vox y Adelante Andalucía de los que más valor de los coeficientes W o LAMBDA presentan (Vox > AA). Con todos estos datos se puede deducir que el modelo más válido correspondería al caso de Vox, puesto que además de presentar el R² más elevado, su función de verosimilitud (log likelihood), su coeficiente LAMBDA y el test de Likelihood son también de los más elevados. El siguiente modelo más explicativo sería el de Adelante Andalucía. Estos modelos al tener todos estos estadísticos tan significativos se puede deducir que son para los casos de estos partidos en donde podemos encontrar mayores tendencias (o más fuertes) de dependencia espacial de las variables analizadas.

Tabla 18. Comparativa de modelos espaciales (II)

	P valor (prueba t-student)				
	PSOE	PP	Cs	Vox	AA
Porcentaje de extranjeros (<i>porc_extra</i>)	0,00000 *	0,00446 *	0,00469 *	0,00000 *	0,00983 *
Porcentaje de residentes mayores de 65 años (<i>Porc_65y</i>)	0,00002*	0,59097	0,89345	0,00158 *	0,00029 *
Porcentaje de residentes menores de 20 años (<i>menos20</i>)	0,00195*	0,00874 *	0,09313	0,36942	0,38078
Tasa de paro municipal (<i>Paro_Mun</i>)	0,11329	0,00001 *	0,57084	0,03345 *	0,01057 *
Renta neta declarada media por habitante (<i>rentamedia</i>)	0,00000*	0,63251	0,00000 *	0,00000 *	0,55119 *

*Significativos para $p < 0.05$

*Nota I: los valores de las pruebas t-student son los relativos a los modelos redefinidos seleccionados para cada caso, esto es, para el PSOE, PP y AA: modelo por retardo; Cs y Vox: por error

*Nota II: se debe recordar que estas tablas están en el anexo III.

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Tabla 19. Recordatorio Coeficientes I de Morán

	PSOE	PP	Cs	Vox	AA
Renta neta declarada media por habitante (<i>rentamedia</i>)	-0,276	-0,166	0,285	0,169	0,209
Tasa de paro municipal (<i>Paro_Mun</i>)	0,015	-0,209	0,037	-0,109	0,186
Porcentaje de residentes mayores de 65 años (<i>Porc_65y</i>)	0,230	0,234	-0,216	-0,139	-0,265
Porcentaje de residentes menores de 20 años (<i>menos20</i>)	-0,218	-0,237	0,227	0,127	0,263
Porcentaje de extranjeros (<i>porc_extra</i>)	-0,171	0,215	0,018	0,223	-0,129

*Nota: el color rojo indica que no son significativos siguiendo la tabla 18

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

En la tabla 18 se puede apreciar cuáles son las variables más influyentes para cada partido. En el caso del PSOE se puede observar cómo son estadísticamente significativas todas a excepción de la variable referida al paro. Siendo así importantes factores contextuales en la construcción del voto del PSOE la presencia de extranjeros en el municipio, los residentes

mayores de 65 años y menores de 20 y la renta neta declarada media por habitante. En la tabla 19 se pueden ver cuáles eran los signos del coeficiente de Morán, así se podrá saber si existe una tendencia hacia la autocorrelación positiva o hacia la negativa. Para el caso del PSOE, de las variables significativas, tres de ellas presentan una autocorrelación negativa -esto es, no presentan evidencias de continuidad a lo largo de los municipios vecinos de primer orden, pero aun así el ACDE muestra que son significativas- (renta, residentes, menores de 20 años y residentes extranjeros) y solo una positiva (residentes mayores de 65 años). Así pues, se debe interpretar que es significativo en el voto al PSOE aquellos municipios en los que no se encuentran residentes de menos de 20 años, y a su vez, triunfa más en aquellos con mayor presencia de gente de mayor edad. Echando un vistazo al resto de variables que presentan una autocorrelación negativa, se puede decir que no existe una tendencia de voto al PSOE en aquellos municipios en donde existe una presencia de rentas elevadas, ni de un número de extranjeros elevado.

Por otro lado, se puede apreciar como lo más significativo en el caso del PP es la tendencia que existe a su voto en municipios donde existe un porcentaje elevado de extranjeros (0,215), mientras que no se presentan rasgos de continuidad entre municipios de primer orden en las variables relativas al paro y a los menores de 20 años. Esto significa que es significativa la presencia de un porcentaje de voto elevado al PP en municipios que no presentan porcentajes de paro elevados ni tampoco un porcentaje de residentes menores de 20 años elevado.

En el caso de Cs se puede apreciar como las únicas variables significativas presentan una autocorrelación positiva, se puede decir pues, que en aquellos municipios en donde la renta neta media declarada por habitante es elevada, al igual que el porcentaje de número de extranjeros también lo es, existe una fuerte presencia de la variable a explicar -a saber, el porcentaje de voto a Cs-. Es decir, los factores contextuales principales del voto a Cs en las elecciones andaluzas se componían por la presencia de extranjeros y por el nivel de renta.

Para el caso Vox destaca que, los municipios con valores elevados de las variables tasa de paro y número de residentes mayores de 65 años, no son coincidentes con los municipios que presentan mayor porcentaje de voto al partido. Caso contrario en el caso de la renta y el número de extranjeros.

Por último, se puede comprobar que son los municipios con mayores presencias de las variables renta y paro, y aquellos con menor presencia de las variables que aluden a los

residentes mayores de 65 años y el porcentaje de extranjeros, las que presentan mayores porcentajes de voto a Adelante Andalucía.

4. CONCLUSIONES

Una vez relatados los hallazgos de la investigación para cada partido es hora de concluir con la respuesta a cada pregunta principal. En cuanto a la primera *¿se pueden correlacionar factores económicos de un municipio con el voto?* Gracias al AEDE primero y al ACDE después, podemos ver como efectivamente existen tendencias de voto a determinados partidos condicionadas por factores económicos contiguos en diversos municipios (tanto de primer orden como de segundo en algunos casos). En el caso de Cs, Vox o Adelante Andalucía se ha visto como la renta neta media declarada por habitante es un factor presente y continuo a lo largo de los municipios que presentan mayores porcentajes de voto a estos partidos. Incluso para Adelante Andalucía otro factor significativo es, como vimos, la tasa de paro, siendo que cuanto mayor es ésta, mayor es la presencia de voto a esta formación.

Ahora *¿es la edad una variable fundamental?* También. Existen varias tendencias. Primero vemos como los partidos más tradicionales, se alejan de los nichos de votantes jóvenes (menores de 20 años) -coeficiente de Morán para el caso del PSOE: (-0,218); para el caso del PP: (-0.237)-, siendo que, incluso el PSOE guarda estrecha relación con aquellos municipios en donde hay mayor presencia de residentes mayores de 65 años. En el caso de Vox y Adelante Andalucía, resulta ser al contrario: los municipios con mayores porcentajes de voto a estos partidos, difieren de aquellos con mayor presencia de gente mayor -coeficientes de Morán: (-0,139) y (-0,265) respectivamente-.

Por último, *¿varía el voto a determinados partidos, en función de la cantidad de personas extranjeras residentes en un municipio?* Los datos así parecen confirmarlo. En el caso de PP, Cs y Vox, gozan de municipios en donde se encuentran porcentajes de extranjeros más elevados a la vez que también presentan mayor porcentaje de voto a sus partidos. Es decir, presentan una autocorrelación positiva puesto que estas variables presentan valores similares en una serie de municipios -véase tabla 19 para comprobar los coeficientes de Moran-, siendo al contrario en los casos de PSOE y Adelante Andalucía que agrupan valores disímiles. Por lo que se puede decir que en los municipios en donde PSOE y Adelante Andalucía presentan mejores porcentajes de voto, existe menor cantidad de extranjeros en proporción al número total de habitantes del municipio en cuestión.

En líneas de lo publicado por Ryan Weichelt (2017) los detractores de la geografía espacial en su momento, estaban bien fundadas, pero con las tendencias actuales, descartarla ahora sería un error. La revolución en el mundo geoespacial de principios de siglo, junto al *Big Data* vuelve a relanzar en importancia a esta subdisciplina de la geografía humana. El mundo y la tecnología avanzan de la mano y con ello los avances en estudios electorales. Un ejemplo claro son las elecciones estadounidenses del 2008 en las que Barack Obama se alza con el Ejecutivo norteamericano tras una campaña centrada en internet, al igual que el caso del Brexit. Aunque es evidente que el análisis espacial del voto tiene limitaciones.

Recordando a Joaquín Bosque Sendra (1988) la construcción del voto no depende solo de factores contextuales, sino que también de factores individuales tales como la ideología y sería lo propio tenerlo en cuenta para que futuros estudios del voto se tengan en cuenta ambos factores. Como ya se presentó en el marco teórico, igual que Darmofal (2006), en el presente trabajo también se da cuenta de ciertos factores sociopolíticos que explican el voto a diferentes partidos dependiendo del contexto en el que se encuentren. Trabajo relevante también es el de Wing y Walker (2004) en el cual el objetivo principal era presionar los límites del análisis ecológico en un estudio sobre la polarización de los votantes en las elecciones presidenciales de EEUU del año 2004. Operaban bajo la hipótesis de que la existencia de un número de procesos sociales que operan en una fina escala espacial tienden a empujar a las visiones de los votantes individuales hacia un alineamiento cercano con las preferencias ideológicas de una mayoría geográficamente próxima -un fenómeno conocido como atrinchamiento localizado-. Los hechos resultantes son similares a los del presente trabajo, el cual también opera bajo el supuesto de atrinchamiento localizado o efecto vecindario dando cuenta de que si existen tendencias de voto localizado en contextos de similares características. Estos trabajos presentan el hecho de que los resultados de unos comicios electorales no son independientes del espacio o área geográfica en el que se disputan los votos. Los estudios acerca del espacio en líneas de comicios electorales demuestran que en menor o mayor medida se puede contrastar la hipótesis de no aleatoriedad de los resultados electorales.

Este trabajo no pretende más que -a parte de lo expuesto en los primeros párrafos de las conclusiones- poner de relieve que la geografía electoral no está muerta, y menos con las nuevas técnicas de análisis y el crecimiento de las nuevas tecnologías y fuentes de datos. Es una subdisciplina con estudios lo suficientemente rigurosos para tenerla en cuenta como otro amplio espacio de trabajo a la hora de llevar a cabo el análisis del voto. Esto no quiere decir que se deban sustituir los análisis a nivel micro o individual, de hecho, lo óptimo para una mayor

comprensión de la construcción del voto es la complementación de ambos campos con sus respectivas técnicas.

BIBLIOGRAFÍA

1. Referencias en papel:

- Ackerman. (1958). *Geography as a Fundamental Research Discipline*,. Chicago: The University of Chicago Press.
- Anselin, L. (2005). *Exploring Spatial Data with GeoDa: A workbook*. Illinois: Center for Spatially Integrated Social Science.
- Ávila-Eggleton, & Sánchez, G. (2017). El papel del espacio en la explicación del abstencionismo. *Espiral*.
- Baronio, A., Vianco, a., & Rabanal, C. (2012). Una introduccion a la econometría espacial . *Catedra de Econometría* .
- Bussi, M. (1998). Eléments de géographie électorale, à travers l'exemple de la France de l'Ouest. *Collection Nouvelles donnes en géographie, Publications de l'Université de Rouen, Rouen*,, 399.
- Charles-Leija, H., García, A. J., & Valadez, L. M. (2015). Características sociodemográficas del voto para diputados; un análisis de econometría espacial. *El Colegio de San Luis* 17, 107-135.
- Coro, C. Y. (s.f.). Métodos gráficos del análisis exploratorio de datos espaciales. *Universidad Autónoma de Madrid*.
- Corso Sicilia, G. B., Rivera, M. P., & Navarro, J. G. (2017). Métodos gráficos de análisis exploratorio de datos espaciales con variables espacialmente distribuidas. *Cuadernos Latinoamericanos de Administración* , 13, 92-104.
- Darmofal, D. (2006). The political geography of macro-level turnout. *Political Geography*, 25, 123-150.
- Goguel, F. (1990). Géographie électorale et science politique: un itinéraire. *Espaces Temps. Réfléchir les sciences sociales*, 43, 19-24.
- Gómez Tagle, S., & Vega, M. E. (2000). *La geografía del poder y las elecciones en México*. México: IFE-Plaza.
- Martori, J. C. (2008). Nuevas técnicas de estadística espacial para la detección de clusters residenciales de población inmigrante. *revista electrónica de geografía y ciencias sociales*, 12, 741-798.
- Martori, J. C., Hoberg, K., & Hoberg, K. (2008). *La incorporación del espacio en los métodos estadísticos* . Barcelona: X Coloquio Internacional de Geocrítica.
- Medus, N. (2005). La Geografía Electoral en la historia de la ciencia geográfica. *Anuario* , 7, 15-32.
- Monzón, N. (2013). Aportes a la investigación geográfica desde la geografía electoral. *Revista Geográfica Digital*, 1.

- Moreno, A. (2003). *El votante mexicano: democracia, actitudes políticas y conducta electoral*. México: Fondo de Cultura Económica, 4 (7), 251-252.
- Openshaw. (1984). Ecological Fallacies and the Analysis of Areal Census Data. *Department of Geography, University of Newcastle*, 16 (1), 17-31.
- Robinson, W. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, 15 (3), 351-357.
- Sánchez Rivero, M. (2008). Análisis espacial de datos y turismo: nuevas técnicas para el análisis turístico. Una aplicación al caso extremeño. *Revista de Estudios Empresariales*, 2, 48-66.
- Sánchez-Peña, L. L. (2012). Alcances y límites de los métodos de análisis espacial para el estudio de la pobreza urbana . *El Colegio de México* , 72, 147-179.
- Sendra, J. B., & Ballesteros, A. G. (1985). Evolución y tendencias actuales de la Geografía Política. *Documents d'Analisi Geografica*, 6, 115-132.
- Sendra, J. B., & Buzar, G. D. (2017). Geografía electoral de la Ciudad Autónoma de Buenos Aires, 2015. Elecciones a Jefe de Gobierno municipal y a Presidente de la República Argentina. *Persona y Sociedad*, 31 (1), 48-73.
- Serrano, R. M., & Valcarce, E. V. (2004). Econometría espacial: nuevas técnicas para el análisis regional. Una aplicación a las regiones europeas. *Investigaciones Regionales* , 83-106.
- Sonnleinter, W. (2013). Explorando las dimensiones territoriales del comportamiento político. *Estudios Sociológicos XXXI*, 98.
- Taylor. (1994). *Geografía Política. Economía-mundo. Estado-Nación y Localidad*. Madrid : Trama Editorial .
- Tingsten, H. (1937). Political Behavior; Studies in Election Statistics. . *American Political Science Review*, 1164-1165.
- Tobler, W. (1970). A computer model simulation of urban growth in the Detroit region. *Economic Geography*, 46(2), 234-240.
- Villodres, C. O., & Cerezo, J. M. (2013). El efecto «amigos y vecinos» sobre la conducta electoral. Un estudio de los comicios locales de 2011 en Andalucía. *RES*, 19, 93-115.
- Wing, I. S., & Walker, J. L. (2004). The Geographic Dimensions of Electoral Polarization in the 2004 U.S. Presidential Vote. *Progress in Spatial Analysis* , 253-285.
- Yrigoyen, C. C. (s.f.). Análisis exploratorio de datos espaciales al servicio del geomarketing . *Universidad Autónoma de Madrid*.

2. Páginas web:

Instituto de Estadística y Cartografía de Andalucía. (s.f). *Andalucía pueblo a pueblo fichas municipales*. Consejería de Economía, Conocimiento, Empresas y Universidad.

<https://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/ficha.htm?mun=29006>

Ministerio del Interior (s.f). *Información electoral*.

<http://www.infoelectoral.mir.es/infoelectoral/min/>

ANEXOS

1. Legislación

- 1.1 Decreto 178/2018, de 2 de octubre, por el que se aprueba la creación del municipio de Fuente Carreteros por segregación del término municipal de Fuente Palmera (Córdoba). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 13-27.
- 1.2 Decreto 179/2018, de 2 de octubre, por el que se aprueba la creación del municipio de Fornes por segregación del término municipal de Arenas del Rey (Granada). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 27-41.
- 1.3 Decreto 180/2018, de 2 de octubre, por el que se aprueba la creación del municipio de La Guijarrosa por segregación del término municipal de Santaella (Córdoba). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 41-61.
- 1.4 Decreto 181/2018, de 2 de octubre, por el que se aprueba la creación del municipio de San Martín del Tesorillo, por segregación del término municipal de Jimena de la Frontera (Cádiz). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 61-76.
- 1.5 Decreto 182/2018, de 2 de octubre, por el que se aprueba la creación del municipio de Tharsis, por segregación del término municipal de Alonso (Huelva). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 76-100.
- 1.6 Decreto 183/2018, de 2 de octubre, por el que se aprueba la creación del municipio de El Palmar de Troya por segregación del término municipal de Utrera (Sevilla). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 100-122.
- 1.7 Decreto 184/2018, de 2 de octubre, por el que se aprueba la creación del municipio de Torrenueva costa por segregación del término municipal de Motril (Granada). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 122-143.
- 1.8 Decreto 185/2018, de 2 de octubre, por el que se aprueba la creación del municipio de la Zarza-Perrunal por segregación del término municipal de calañas (Huelva). *Boletín Oficial de la Junta de Andalucía*. Andalucía. 9 de octubre de 2018, núm. 196, pp. 143-170.
- 1.9 Ley 7/1993, de 27 de julio, Reguladora de la Demarcación Municipal de Andalucía. *Boletín Oficial de la Junta de Andalucía*. Andalucía. 7 de agosto de 1993, núm. 86, pp.7582-7592

2. Relativo a conceptos

Concepto	Definición
R²	El R ² es una medida estadística de qué tan cerca están los datos de la línea de regresión ajustada. La definición de R-cuadrado es bastante sencilla: es el porcentaje de la variación en la variable de respuesta que es explicado por un modelo lineal, o, en otras palabras, cuanto de la variable independiente explica a la variable dependiente
Log Likelihood (función de verosimilitud)	Es una medida de bondad de ajuste para cualquier modelo. Cuanto mayor sea el valor, mejor es el modelo, por ejemplo, un valor de -6 es mejor que uno de -12. Al tratarse de productos de probabilidades la función de verosimilitud será siempre menor que 1 y por tanto su logaritmo será negativo.
Breusch-Pagan	Si el estadístico de prueba tiene un valor p por debajo de un umbral apropiado ($p < 0.05$ por ejemplo), se rechaza la hipótesis nula de homocedasticidad y se asume la heterocedasticidad.
Likelihood Ratio Test	Conocido en español como razón de verosimilitud, el Likelihood Ratio Test se define como la razón entre la posibilidad de observar la ocurrencia de un evento frente a la posibilidad de no ocurrencia del mismo. Por ejemplo, si el valor es alto y $p < 0,05$ existirá más probabilidad de encontrar peso de esas variables en los municipios vecinos que de no encontrarla.
W	Refleja la dependencia espacial inherente a nuestros datos de muestra, midiendo la influencia promedio en observaciones por sus observaciones vecinas. Su coeficiente se mueve hasta el valor 1, por eso cuanto mayor sea, mejor explica la dependencia espacial.
LAMBDA	Este es un estadístico que se basa en la reducción proporcional del error de predicción cuando se utilizan los valores de la variable independiente para predecir los de la dependiente. Su coeficiente también va hasta el 1, por lo que cuanto mayor sea, mejor explica la dependencia espacial

Fuente: elaboración propia con información obtenida de textos de la bibliografía

3. Modelos de regresión espacial redefinidos en términos de error y de retardo

3.1 Partido Socialista Obrero Español

Tabla 20. Modelo redefinido en términos de error; (VD: Tasa_PSOE)

```

Data set           : ANDALUCIA
Spatial Weight     : ANDALUCIA
Dependent Variable : Tasa_PSOE  Number of Observations: 777
Mean dependent var : 38,080785  Number of Variables   : 6
S.D. dependent var : 10,752201   Degrees of Freedom    : 771
Lag coeff. (Lambda) : 0,451165

R-squared         : 0,404916  R-squared (BUSE)      : -
Sq. Correlation   : -          Log likelihood        : -2761,576086
Sigma-square      : 68,7976   Akaike info criterion : 5535,15
S.E of regression : 8,29443   Schwarz criterion     : 5563,08
    
```

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	28,7633	5,33151	5,39497	0,00000
porc_extra	-18,8997	4,11931	-4,58806	0,00000
Porc_65y	0,556344	0,113035	4,92188	0,00000
menos20	0,506033	0,156179	3,24007	0,00120
Paro_Mun	0,145293	0,0725876	2,00162	0,04533
rentamedia	-0,00120219	0,000117502	-10,2312	0,00000
LAMBDA	0,451165	0,0451968	9,98223	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	61,9820	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	75,5950	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Tabla 21. Modelo redefinido en términos de retardo; (VD: Tasa_PSOE)

Data set : **ANDALUCIA**
 Spatial Weight : **ANDALUCIA**
 Dependent Variable : Tasa_PSOE Number of Observations: 777
 Mean dependent var : 38,0808 Number of Variables : 7
 S.D. dependent var : 10,7522 Degrees of Freedom : 770
 Lag coeff. (Rho) : 0,397162

R-squared : 0,401325 Log likelihood : -2760,21
 Sq. Correlation : - Akaike info criterion : 5534,42
 Sigma-square : 69,2128 Schwarz criterion : 5567,01
 S.E of regression : 8,31942

Variable	Coefficient	Std.Error	z-value	Probability
W_Tasa_PSOE	0,397162	0,0421651	9,41921	0,00000
CONSTANT	16,1561	5,15644	3,13319	0,00173
porc_extra	-19,3695	3,41482	-5,67218	0,00000
Porc_65y	0,449928	0,105154	4,27874	0,00002
menos20	0,448331	0,144704	3,09826	0,00195
Paro_Mun	0,0956353	0,0603919	1,58358	0,11329*
rentamedia	-0,00102324	0,000107055	-9,55805	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	67,9638	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	78,3291	0,00000

*Nota: en este caso difiere del modelo MCO y del modelo por error en que la variable no es significativa.

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

3.2 Partido Popular

Tabla 22. Modelo redefinido en términos de retardo espacial; (VD: Tasa_PP)

Data set : **ANDALUCIA**
 Spatial Weight : **ANDALUCIA**
 Dependent Variable : **Tasa_PP** Number of Observations: 777
 Mean dependent var : 22,2762 Number of Variables : 7
 S.D. dependent var : 8,77466 Degrees of Freedom : 770
 Lag coeff. (Rho) : 0,392972
 R-squared : 0,265895 Log likelihood : -2681,26
 Sq. Correlation : - Akaike info criterion : 5376,51
 Sigma-square : 56,5221 Schwarz criterion : 5409,1
 S.E of regression : 7,51812

Variable	Coefficient	Std. Error	z-value	Probability
W_Tasa_PP	0,392972	0,0451286	8,70781	0,00000
CONSTANT	22,9686	4,62648	4,96459	0,00000
porc_extra	8,72838	3,06976	2,84335	0,00446
Porc_65y	0,0506854	0,094311	0,537428	0,59097
menos20	-0,343427	0,130979	-2,62201	0,00874
Paro_Mun	-0,246476	0,0553881	-4,44997	0,00001
rentamedia	4,4626e-005	9,34045e-005	0,477772	0,63281

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	88,2317	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	66,2722	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Tabla 23. Modelo redefinido en términos de error; (VD: Tasa_PP)

Data set : ANDALUCIA
 Spatial Weight : ANDALUCIA
 Dependent Variable : Tasa_PP Number of Observations: 777
 Mean dependent var : 22,276229 Number of Variables : 6
 S.D. dependent var : 8,774661 Degrees of Freedom : 771
 Lag coeff. (Lambda) : 0,389769
 R-squared : 0,250457 R-squared (BUSE) : -
 Sq. Correlation : - Log likelihood :-2689,143472
 Sigma-square : 57,7108 Akaike info criterion : 5390,29
 S.E of regression : 7,59676 Schwarz criterion : 5418,22

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	32,3195	4,84261	6,67398	0,00000
porc_extra	8,16799	3,67958	2,21981	0,02643
Porc_65y	0,0840699	0,102653	0,818974	0,41280
menos20	-0,362167	0,141997	-2,55052	0,01076
Paro_Mun	-0,290852	0,0650001	-4,47463	0,00001
rentamedia	4,0114e-005	0,00010603	0,378326	0,70519
LAMBDA	0,389769	0,0476683	8,17669	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	81,8617	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	50,4963	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

3.3 Ciudadanos

Tabla 24. Modelo redefinido en términos de error; (VD: Tasa_Cs)

```

Data set           : ANDALUCIA
Spatial Weight     : ANDALUCIA
Dependent Variable : Tasa_Cs   Number of Observations: 777
Mean dependent var : 13,498005 Number of Variables : 6
S.D. dependent var : 4,958611  Degrees of Freedom : 771
Lag coeff. (Lambda) : 0,272588

R-squared          : 0,340504  R-squared (BUSE) : -
Sq. Correlation    : -          Log likelihood : -2190,059375
Sigma-square       : 16,2156   Akaike info criterion : 4392,12
S.E of regression  : 4,02686   Schwarz criterion : 4420,05
  
```

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	3,17328	2,5196	1,25944	0,20787
porc_extra	5,23876	1,85274	2,82758	0,00469
Porc_65y	-0,00715076	0,0533897	-0,133935	0,89345
menos20	0,124212	0,0739739	1,67913	0,09313
Paro_Mun	0,0186422	0,0328895	0,566813	0,57084
rentamedia	0,00066804	5,44532e-005	12,2681	0,00000
LAMBDA	0,272588	0,0517614	5,26624	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	52,0987	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	28,2452	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Tabla 25. Modelo redefinido en términos de retardo espacial; (VD: Tasa_Cs)

Data set : ANDALUCIA
 Spatial Weight : ANDALUCIA
 Dependent Variable : Tasa_Cs Number of Observations: 777
 Mean dependent var : 13,498 Number of Variables : 7
 S.D. dependent var : 4,95861 Degrees of Freedom : 770
 Lag coeff. (Rho) : 0,221973
 R-squared : 0,333406 Log likelihood : -2192,41
 Sq. Correlation : - Akaike info criterion : 4398,83
 Sigma-square : 16,3901 Schwarz criterion : 4431,41
 S.E of regression : 4,04847

Variable	Coefficient	Std.Error	z-value	Probability
W_Tasa_Cs	0,221973	0,0472395	4,69887	0,00000
CONSTANT	0,0978137	2,43639	0,040147	0,96798
porc_extra	4,57108	1,63892	2,78908	0,00529
Porc_65y	0,0265237	0,0508201	0,521913	0,60173
menos20	0,131317	0,0704966	1,86274	0,06250
Paro_Mun	0,0139142	0,0293847	0,473519	0,63584
rentamedia	0,000609969	5,23968e-005	11,6413	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	54,1952	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	23,5376	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

3.4 Vox

Tabla 26. Modelo redefinido en términos de error (VD: Tasa_Vox)

```

Data set           : ANDALUCIA
Spatial Weight     : ANDALUCIA
Dependent Variable : Tasa_Vox  Number of Observations: 777
Mean dependent var : 8,230940  Number of Variables   : 6
S.D. dependent var : 4,174725  Degrees of Freedom    : 771
Lag coeff. (Lambda) : 0,600899

R-squared          : 0,463687  R-squared (BUSE)      : -
Sq. Correlation    : -          Log likelihood         : -2000,088931
Sigma-square       : 9,34704   Akaike info criterion : 4012,18
S.E of regression  : 3,05729   Schwarz criterion     : 4040,11
  
```

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	8,99641	1,99807	4,50255	0,00001
porc_extra	12,3881	1,60131	7,73624	0,00000
Porc_65y	-0,133792	0,0423376	-3,16013	0,00158
menos20	-0,052315	0,0582855	-0,897564	0,36942
Paro_Mun	-0,0596853	0,0280649	-2,12669	0,03345
rentamedia	0,000301278	4,46335e-005	6,75003	0,00000
LAMBDA	0,600899	0,0380856	15,7776	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	23,9359	0,00022

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	192,0219	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Tabla 27. Modelo redefinido en términos de retardo espacial (VD: Tasa_Vox)

Data set : ANDALUCIA
 Spatial Weight : ANDALUCIA
 Dependent Variable : Tasa_Vox Number of Observations: 777
 Mean dependent var : 8,23094 Number of Variables : 7
 S.D. dependent var : 4,17472 Degrees of Freedom : 770
 Lag coeff. (Rho) : 0,546646

R-squared : 0,448542 Log likelihood : -2005,13
 Sq. Correlation : - Akaike info criterion : 4024,25
 Sigma-square : 9,61099 Schwarz criterion : 4056,84
 S.E of regression : 3,10016

Variable	Coefficient	Std.Error	z-value	Probability
W_Tasa_Vox	0,546646	0,0369696	14,7864	0,00000
CONSTANT	4,7426	1,86207	2,54695	0,01087
porc_extra	10,0714	1,309	7,69395	0,00000
Porc_65y	-0,0983679	0,0390125	-2,52145	0,01169
menos20	-0,0681885	0,0539045	-1,26499	0,20588
Paro_Mun	-0,0510131	0,0225826	-2,25895	0,02389
rentamedia	0,000237427	3,91357e-005	6,06677	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	21,7941	0,00057

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	181,9468	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

3.5 Adelante Andalucía

Tabla 28. Modelo redefinido en términos de error; (VD: Tasa_AA)

```

Data set           : ANDALUCIA
Spatial Weight     : ANDALUCIA
Dependent Variable : Tasa_AA   Number of Observations: 777
Mean dependent var : 13,515457 Number of Variables : 6
S.D. dependent var : 8,496934  Degrees of Freedom : 771
Lag coeff. (Lambda) : 0,440300

R-squared          : 0,252902  R-squared (BUSE) : -
Sq. Correlation    : -          Log likelihood : -2666,251239
Sigma-square       : 53,9389   Akaike info criterion : 5344,5
S.E of regression  : 7,34431   Schwarz criterion : 5372,44
  
```

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	22,13	4,71418	4,69435	0,00000
porc_extra	-9,65922	3,63175	-2,65966	0,00782
Porc_65y	-0,402295	0,0999444	-4,02519	0,00006
menos20	-0,167289	0,138122	-1,21116	0,22583
Paro_Mun	0,1356	0,0640238	2,11796	0,03418
rentamedia	6,70608e-005	0,000103779	0,646189	0,51816
LAMBDA	0,4403	0,0456516	9,6448	0,00000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	99,9316	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	66,2604	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

Tabla 29. Modelo redefinido en términos de retardo espacial; (VD: Tasa_AA)

Data set : ANDALUCIA
 Spatial Weight : ANDALUCIA
 Dependent Variable : Tasa_AA Number of Observations: 777
 Mean dependent var : 13,5155 Number of Variables : 7
 S.D. dependent var : 8,49693 Degrees of Freedom : 770
 Lag coeff. (Rho) : 0,439728

R-squared : 0,267172 Log likelihood : -2658,72
 Sq. Correlation : - Akaike info criterion : 5331,44
 Sigma-square : 52,9086 Schwarz criterion : 5364,02
 S.E of regression : 7,27383

Variable	Coefficient	Std.Error	z-value	Probability
W_Tasa_AA	0,439728	0,0437774	10,0446	0,00000
CONSTANT	13,5541	4,3479	3,11739	0,00182
porc_extra	-7,6573	2,96607	-2,58164	0,00983
Porc_65y	-0,332026	0,0916609	-3,62233	0,00029
menos20	-0,110849	0,126474	-0,876454	0,38078
Paro_Mun	0,135846	0,0531338	2,55668	0,01057
rentamedia	5,39473e-005	9,05195e-005	0,595975	0,55119

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	102,7546	0,00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : ANDALUCIA

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	81,3274	0,00000

Fuente: Elaboración propia a partir de datos obtenidos de GeoDa

