

Una manera de hacer Europa

EVALUACIÓN DE EFECTOS SOBRE LA LOCALIZACIÓN DE LAS EMPRESAS DE LAS INVERSIONES EN BANDA ULTRARRÁPIDA FINANCIADA CON FONDOS EUROPEOS



EVALUACIÓN DE EFECTOS SOBRE LA LOCALIZACIÓN DE LAS EMPRESAS DE LAS INVERSIONES EN BANDA ULTRARRÁPIDA FINANCIADA CON FONDOS EUROPEOS

Autores:

Ignacio Moral Arce, Miguel Gómez de Antonio, María Gorriti Gutiérrez-Cortines y Daniel Fernández Romero.

Instituto de Estudios Fiscales

Contenido

1. RESUMEN DEL IMPACTO DE LAS INVERSIONES EN INFRAESTRUCTURAS DE BANDA ANCHA SOBRE LA LOCALIZACIÓN DE LAS EMPRESAS CON TECNICAS DE ECONOMETRIA ESPACIAL..... 5

2. OBJETIVO DEL TRABAJO	6
3. PREGUNTAS DE LA EVALUACIÓN SOBRE EL IMPACTO DE LA BANDA ANCHA EN LA LOCALIZACIÓN DE LAS EMPRESAS	8
4. TEORÍA ECONÓMICA SOBRE LOS EFECTOS DE LA BANDA ANCHA SOBRE LA LOCALIZACIÓN DE LAS EMPRESAS	9
5. TEORÍAS DE ECONÓMICA ESPACIAL.....	10
5.1. FUERZAS CENTRALIZADORAS Y DESCENTRALIZADORAS.....	10
5.2. ECONOMÍAS DE AGLOMERACIÓN.....	10
6. ESTUDIOS PREVIOS.....	12
6.1. ESTUDIOS SOBRE PRODUCTIVIDAD DE LAS EMPRESAS	12
6.2. ESTUDIOS SOBRE LOCALIZACIÓN DE LAS EMPRESAS.....	14
7. OTRAS VARIABLES QUE INFLUYEN EN LA LOCALIZACIÓN DE LAS EMPRESAS.....	17
7.1. TIPO DE MUNICIPIO- RURAL FRENTE A URBANO-.....	17
7.2. VELOCIDAD DE LA BANDA ANCHA	17
7.3. ENDOGENEIDAD	17
7.4. SECTOR DE ACTIVIDAD	17
7.5. TAMAÑO DE LA EMPRESA.....	18
8. ECONOMETRÍA ESPACIAL.....	19
9. HETEROGENEIDAD ESPACIAL	19
10. DEPENDENCIA ESPACIAL. DEFINICIÓN, FUENTES, TIPOS	20
a. FUENTES DE LA DEPENDENCIA ESPACIAL	20
b. MATRICES DE PONDERACIÓN O DE ORDENACIÓN ESPACIAL.....	20
2. ANÁLISIS EXPLORATORIO/DESCRIPTIVO DE LOS DATOS.....	21
a. I DE MORAN	21
b. DIAGRAMA DE DISPERSIÓN DE LA I DE MORAN (MORAN SCATTERPLOT)	22
3. RESULTADOS DEL ANÁLISIS EXPLORATORIO DE LOS DATOS	23
a. RESULTADOS DE LA I DE MORAN Y DE LA I DE MORAN LOCAL (LISA) EN LA VARIABLE DEPENDIENTE: NÚMERO DE EMPRESAS Y CRECIMIENTO DEL NÚMERO DE EMPRESAS	24
b. RESULTADO DE LA I DE MORAN Y DE LA I DE MORAN LOCAL EN LA VARIABLE EXPLICATIVA: INFRAESTRUCTURA DE BANDA ANCHA- MEDIDA A TRAVÉS DE COBERTURA DE 100 MB DE VELOCIDAD	25
4. ESTRATEGIAS PARA LA ESTIMACIÓN DE LOS MODELOS DE REGRESIÓN.....	27
• PLANTEAMIENTO ESPECIFICO-GENERAL.....	27
• PLANTEAMIENTO GENERAL- ESPECÍFICO.....	27
5. VARIABLES DEL MODELO A ESTIMAR.....	28
i. VARIABLE DEPENDIENTE (número de empresas versus tasa de variación).....	28
ii. VARIABLES EXPLICATIVAS.....	29

6.	RESULTADOS DE LAS ESTIMACIONES DE LOS MODELOS	30
7.	VALIDACIÓN DEL MODELO.....	32
8.	CONCLUSIONES	33
9.	CONSIDERACIONES FINALES.....	34
10.	RECOMENDACIONES. CONTESTACIÓN A LAS PREGUNTAS DE EVALUACIÓN	35
	ANEXO I: AREAS URBANAS ESPAÑA	37
	ANEXO II: DENSIDAD DE POBLACIÓN DE LOS MUNICIPIOS ESPAÑOLES, 2019 (Habitantes/km2)	38
	ANEXO III: VARIABLES DE CONTROL ANALIZADAS	39
	ANEXO IV: ESTIMACIÓN DE LOS MODELOS	41

1. RESUMEN DEL IMPACTO DE LAS INVERSIONES EN INFRAESTRUCTURAS DE BANDA ANCHA SOBRE LA LOCALIZACIÓN DE LAS EMPRESAS CON TÉCNICAS DE ECONOMETRÍA ESPACIAL

Ley de Tobler: "Everything is related to everything else, but near things are more related than distant things"

El trabajo analiza si las inversiones en Banda Ancha financiadas con Fondos FEDER durante el periodo 2013-2019 han tenido efectos sobre la localización de las empresas.

Para el estudio solo se han tenido en cuenta municipios en zonas rurales, dado que la mayoría de las zonas donde se han llevado a cabo inversiones en banda ancha financiadas con Fondos Europeos son áreas no urbanas, donde ya existía este tipo de cobertura o donde los operadores privados estaban ejecutando o tenían previsto ejecutar este tipo de inversiones sin necesidad de apoyo público.

Los resultados de los modelos indican que la variable explicativa incremento en la infraestructura de banda ancha, medida a través del indicador de cobertura, **es estadísticamente significativa para explicar el número de empresas en cada municipio, teniendo en cuenta que solo se han analizado municipios que no forman parte de áreas urbanas metropolitanas.**

Algunas variables de control, como las características del mercado de trabajo o la proximidad a un aeropuerto, son estadísticamente significativas a la hora de influir en el número de empresas del municipio.

Mientras que otras variables que se han incluido como variables de control, como por ejemplo accesibilidad por carretera o localización en zona montañosa, no han resultado significativas para el caso de España en el periodo de análisis.

Se utilizan técnicas de econometría espacial ya que se observa que la variable de número de empresas en un municipio se ve influida por el número de empresas que existen en los municipios cercanos, es decir, existe autocorrelación espacial.

Palabras clave: autocorrelación espacial, localización de empresas, impacto, econometría espacial, crecimiento económico, banda ancha, TIC, tecnología de la información.

2. OBJETIVO DEL TRABAJO¹

Este trabajo forma parte de la Evaluación encargada por la DG de Fondos Europeos al Instituto de Estudios Fiscales sobre el impacto de las inversiones en banda ancha financiadas con fondos FEDER.

Partiendo de que existe una fuerte interacción entre la tecnología, las empresas y el mercado de trabajo, se considera importante evaluar si los escasos recursos gubernamentales- en este caso procedentes de los Fondos FEDER de la Unión Europea- deberían de invertirse en mejorar las infraestructuras de banda ancha con el objetivo de favorecer el crecimiento económico local de las zonas rurales.

En esta evaluación se analiza **si el incremento en la cobertura de banda ancha de redes de nueva generación (RNG) de 100 Mbps es un factor de localización de las empresas a nivel municipal.**

El objetivo de este análisis es determinar el efecto del crecimiento en la infraestructura de banda ancha, durante en el periodo 2012-2019 en el que FEDER financió las inversiones a través de las Convocatorias PEBA, en el número de empresas que se han establecido durante ese periodo de tiempo en los municipios receptores de las inversiones. Se asume que la gran mayoría de los municipios que forman parte de áreas urbanas cuentan con infraestructura de banda ancha, por lo que el análisis se lleva a cabo para los municipios que no forman parte de áreas urbanas metropolitanas.

Para el análisis de los fenómenos económicos que se producen en las áreas urbanas se necesitan divisiones que reflejen adecuadamente la realidad urbana con mayor precisión que las delimitaciones administrativas, como por ejemplo las ciudades, las regiones o las provincias. La delimitación de las áreas urbanas no es sencilla debido a que no se ajustan bien a los límites administrativos y pueden cambiar a lo largo del espacio y del tiempo como consecuencia de la evolución de la economía y la sociedad.

Los países utilizan diferentes metodologías para definir las áreas urbanas o zonas metropolitanas, por lo que la comparación de las unidades de áreas urbanas identificadas en diferentes países tampoco es sencilla.

En España se han utilizado diversas alternativas para identificar las áreas urbanas. Por un lado, el Ministerio de Fomento a la hora de definir las áreas urbanas utiliza umbrales de población con datos de 2019. Según esta clasificación existen 52 áreas metropolitanas en España.

Otra alternativa consiste en utilizar la metodología propuesta por *Boix et al. (2012)* para definir las áreas metropolitanas funcionales, según la cual se pueden identificar 62 áreas urbanas. *Boix et al. (2012)* propusieron una metodología general para identificar **áreas funcionales** con fines comparativos.

Este enfoque funcional resulta ser el más adecuado para este análisis, ya que tiene en cuenta las **relaciones socioeconómicas** entre los municipios que forman el área urbana.

¹ https://www.dgfc.sepg.hacienda.gob.es/sitios/dgfc/es-ES/ipr/fcp1420/e/PEF1420/Documents/Plan_de_Evaluacion_Comun_FEDER_v14.pdf

De acuerdo con esta definición, en el mapa del anexo 1 se pueden identificar 67 áreas urbanas en España, que difieren claramente de los límites administrativos de las regiones o las provincias, y que son las que se han excluido de esta evaluación.

Una de las razones por las que se ha elegido analizar el efecto sobre las zonas rurales es que el objetivo de las convocatorias, tal y como establecía el Ministerio de Asuntos Económicos y Transformación Digital era:

“Acelerar la extensión de la cobertura de las redes públicas de comunicaciones electrónicas capaces de proporcionar servicios de banda ancha de muy alta velocidad a las zonas sin cobertura actual ni prevista en los próximos años (zonas blancas) así como a aquellas que la tienen de un sólo operador (zonas grises) , lo que permitirá mejorar la funcionalidad y la calidad de los servicios digitales necesarios para la transformación digital y con ello aumentar el bienestar y la calidad de vida de los ciudadanos, contribuyendo al éxito de un modelo de crecimiento sostenible basado en la sociedad del Gigabit”.

El Ministerio hizo una clasificación de las zonas en negras, grises y blancas según la situación que presentan en cuanto al despliegue de infraestructuras de banda ancha RNG o a la que se les prevé para el futuro, que define la situación del mercado tanto en relación con las redes básicas como en cuanto a las redes de banda ancha RNG. Las zonas se clasifican en:

- Zonas blancas, si no disponen de red de banda ancha RNG y no está previsto que la tengan en los tres próximos años; en este caso, se pueden permitir ayudas estatales.
- Zonas grises, si solo cuentan con una red de banda ancha RNG operativa (o está previsto que la tengan en los tres próximos años); en este caso, la zona puede optar a ayudas estatales, pero será preciso realizar un análisis detallado.
- Zonas negras, si poseen dos o más redes de banda ancha RNG operativas o está previsto que las tengan en los tres próximos años; en este caso, en ausencia de un fallo del mercado claramente demostrado, no se permite la intervención estatal.

Las zonas donde se han llevado las inversiones cofinanciadas con fondos son las zonas blancas y grises. Y tal y como señala [Fabritz \(2013\)](#) estas zonas se encuentran predominantemente en áreas rurales.

3. PREGUNTAS DE LA EVALUACIÓN SOBRE EL IMPACTO DE LA BANDA ANCHA EN LA LOCALIZACIÓN DE LAS EMPRESAS

Algunas de las preguntas que se plantean los investigadores sobre el acceso a la banda ancha y el uso que realizan las empresas incluyen:

- ¿Cuáles son los enfoques más relevantes para estimar los efectos de la banda ancha sobre el establecimiento de empresas? ¿Se pueden utilizar modelos de econometría convencional o es más preciso utilizar modelos de econometría espacial?
- ¿Afecta el despliegue de la infraestructura de banda ancha ultrarrápida a la localización de las empresas en las zonas rurales?
- ¿En qué contexto o cuales son los factores que influyen en que la banda ancha ultrarrápida genere efectos positivos para el desarrollo económico local (áreas con buena accesibilidad por carretera, bien dotadas de mano de obra calificada, próximas a aeropuertos, etc.)?
- Con inversiones realizadas en el último periodo de programación 2007-2014, ¿existe suficiente tiempo para percibir el impacto de las mismas en el establecimiento de empresas en las zonas con conexión?
- ¿Influye la velocidad- es decir la calidad de la banda ancha- en las decisiones de localización de las empresas? ¿tiene un mayor impacto la banda ultrarrápida que la banda rápida? En este sentido, teniendo en cuenta que la mayor parte de España tiene ya acceso a banda ancha a velocidad 30 Mbps ¿es adecuado seguir invirtiendo en tener acceso a conexiones de 100 Mbps? ¿generará importantes beneficios económicos adicionales, en particular, se establecerá un mayor número de empresas en zonas rurales con velocidad de banda ultrarrápida?
- ¿Existen efectos mayores en las zonas situadas en la periferia de las grandes urbes que en las zonas rurales remotas?

4. TEORÍA ECONÓMICA SOBRE LOS EFECTOS DE LA BANDA ANCHA SOBRE LA LOCALIZACIÓN DE LAS EMPRESAS

Los efectos positivos de las inversiones en banda ancha a nivel macroeconómico parecen claros – como se ha visto en el meta análisis llevado a cabo sobre el impacto de dichas inversiones sobre el crecimiento económico- pero los cauces a través de los cuales dichos efectos requieren que se lleve a cabo un análisis más detallado.

Los estudios recientes vienen a demostrar que **la banda ancha se ha convertido en una condición necesaria pero no suficiente para el crecimiento económico local.**

La teoría económica señala que, en mercados de competencia perfecta, las empresas tendrían una mayor probabilidad de establecerse en mercados donde la esperanza de obtener beneficios es más alta. De hecho, el acceso de banda ancha puede mejorar la rentabilidad específica de la ubicación de las empresas de varias formas:

- Reduce los costes de producción de las empresas ya que pueden acceder a proveedores que estaban más distantes y acceder de manera más fácil a servicios de subcontratación (Lamie & al., 2011).
- Permite a las empresas aumentar sus ventas al expandir su mercado.
- Conduce a una reducción de los costes de comunicación, ya que permite un acceso más fácil a la información, el conocimiento y las ideas y facilita la coordinación con los socios.
- Fomenta el espíritu empresarial local (Fairlie, 2006).
- Permite un mejor acercamiento entre empresas y trabajadores y acorta el proceso de contratación (Autor, 2001).
- Las empresas pueden recopilar, compartir y difundir información independientemente de distancia y tiempo, así como ofrecer soluciones innovadoras (Angelides, 1997; Cappelli, 2001; Sharma, 2002).
- Se crean nuevos modelos de negocio virtuales y las empresas se ven obligadas a cambiar sus modelos de negocio para actuar online (Bakos, 1998; Griffith y Krampf, 1998).

Como internet es una tecnología de propósito general (GPT) en terminología de Schumpeter (1942), se puede considerar una infraestructura que afecta al entorno empresarial actual (Bresnahan y Trajtenberg, 1995; David y Wright, 2003; Lipsey, Carlaw y Bekar, 2005; Ruttan, 2008). El entorno colaborativo ayuda a generar nuevas capacidades, mercados y estrategias que son importantes para el crecimiento (Austin y Bradley, 2005).

5. TEORÍAS DE ECONOMÍA ESPACIAL

Sin embargo, las teorías de economía espacial, indican que existen una serie de fuerzas y de factores que influyen en las decisiones de localización de las empresas, entre las que destacan:

5.1. FUERZAS CENTRALIZADORAS Y DESCENTRALIZADORAS

El desarrollo de las TICs ha afectado a las fuerzas centralizadoras y descentralizadoras que impulsan las actividades económicas. Las TICs han estado impulsando nuevas actividades económicas en una serie de ciudades y regiones con infraestructuras de información, mientras que dentro de estas ciudades y regiones algunas industrias tienden a dispersarse como resultado de las comunicaciones seguras y suficientes a distancia que permiten las nuevas TIC.

Liu, Dicken and Yeung (2004) sugieren que la aplicación de nuevas TIC es una fuerza descentralizadora en el ubicación a escala local. La aplicación de nuevas TIC tiende a conducir a un “agrupamiento virtual” de empresas, ya que se trata de una paso esencial de la integración de las cadenas de suministro, mientras que los proveedores eligen o no ubicarse en la proximidad de otros proveedores o de los ensambladores dependiendo de otros factores.

5.2. ECONOMÍAS DE AGLOMERACIÓN

Las **economías de aglomeración** hacen referencia a los beneficios obtenidos por las empresas por localizarse en las cercanías de otras ('aglomeración'). Este concepto está relacionado con las ideas de economías de escala y efectos de red. Mientras más empresas en campos relacionados se establezcan en áreas contiguas, sus costes de producción pueden disminuir significativamente (las empresas tienen que competir por proveedores múltiples; un resultado de aglomerarse sería una mayor división y especialización en el trabajo). Incluso puede tener ventajas porque el grupo atrae a más proveedores y más clientes que en la situación en la que las empresas estuvieran ubicadas en sitios distintos. Las ciudades conforman y expanden economías de aglomeración.

Las economías de aglomeración están asociadas a una serie de factores:

La aglomeración de empresas genera **la puesta en común del mercado de trabajo**. Las grandes poblaciones de trabajadores calificados entran en el área y son capaces de intercambiar conocimientos, ideas e información. Cuantas más empresas estén ubicadas en una determinada área, mayor será la competencia en la obtención de los trabajadores y por lo tanto se traduce en mayores salarios para los trabajadores. Sin embargo, mientras menos empresas haya y más trabajadores estén, menor será el salario para los trabajadores.

Otro tema relevante es **el acceso a bienes y servicios especializados**, insumos intermedios, que permiten que las empresas ubicadas dentro de ese área de trabajo tengan rendimientos crecientes a escala. Si los insumos intermedios son relevantes para la producción se puede producir una inversión centro-periferia que tendrá muchas empresas localizadas unas cerca de otras y a su vez cerca de sus fuentes necesarias.

Otro tema importante es **la “difusión tecnológica”**. Otra ventaja de los clusters de empresas es que, en campos específicos, se produce una difusión más rápida de las ideas o de la adopción de las ideas. Las nuevas formas de la tecnología, al permitir un flujo de información más rápido

y ágil, pueden reducir la cantidad de incertidumbre y las complicaciones involucradas con los procesos de creación y puesta en marcha de nuevos productos o servicios. . Además, la difusión tecnológica puede ser más beneficiosa para ciudades más pequeñas debido a las redes de información existentes en las grandes ciudades que ya han ayudado a formar y crecer.

Para que la producción sea llevada a su nivel máximo y se pueda vender sus productos, las empresas requieren de algún tipo de acceso viable a los mercados de capitales. Los negocios bancarios y la tecnología se concentran en áreas específicas y por lo tanto se beneficia la empresa localizada cerca de estas zonas.

No obstante, las aglomeraciones también generan efectos negativos: congestión del tráfico, precios del suelo y de la vivienda, contaminación ...

6. ESTUDIOS PREVIOS

En esta evaluación solo nos vamos a centrar en los estudios que analizan directamente el impacto de las inversiones en banda ancha sobre las empresas: tanto sobre su productividad y competitividad como sobre sus decisiones de localización.

Sin embargo, existen numerosos estudios previos, que hemos analizado en las otras evaluaciones realizadas sobre banda ancha que se centran en los efectos sobre el crecimiento económico, sobre el mercado de trabajo o sobre el valor de las propiedades (entre otros, Atasoy, 2013; Bai, 2017; Briglauer & al., 2019; Crandall et al. (2007), Czernich, 2014; De Stefano & al., 2014; Deller & Whitacre, Forthcoming; Forman & al., 2012; Guidry & al., 2012; Gurney, 2012; Holt & Jamison (2009); Ivus & Boland, 2015; Katz et al. (2010); Kandilov & Renkow, 2010; Kolko, 2012; Lapointe, 2015; Lehr & al., 2006; Mahasuweerachai & al., 2010; Mack, 2014; 2015; Mack & Rey, 2014; Mack & Wentz, 2017; Stiroh et al (2007); Van Gaasbeck, 2008; Whitacre & al., 2014-a; 2014).

Los efectos espaciales son otra cuestión clave a la hora de investigar empíricamente los efectos económicos de la banda ancha. Los estudios existentes han probado y contabilizado los efectos espaciales de varias formas.

La mayoría de ellos ha utilizado modelos de retraso espacial o error espacial, dependiendo de la naturaleza de la autocorrelación espacial, y han demostrado que no tener en cuenta los efectos espaciales puede conducir a estimaciones fuertemente sesgadas (p. Ej., Gurney, 2012; Grubenic, 2003; Mack & Rey, 2014; Whitacre & al., 2014-a). Otro enfoque, utilizado en Mack & Wentz (2017), comprende tener en cuenta los efectos espaciales controlando el acceso de banda ancha de los vecinos (modelo WX).

6.1. ESTUDIOS SOBRE PRODUCTIVIDAD DE LAS EMPRESAS

Las TIC permiten una difusión eficiente de la información, disminuyendo los costes de producción porque es posible acceder al conocimiento generado al mínimo coste.

También mejora la competitividad de las empresas, ya que las TIC superan las limitaciones de tiempo y espacio. Como resultado, la transmisión de datos entre compradores y vendedores aumenta, el mercado se vuelve más amplio y existe un mayor acceso al suministro internacional de bienes.

Esta eficiencia organizativa se traduce en mejoras de la productividad tanto de las empresas productoras de TIC como de las empresas que las utilizan.

- Por lo general, el análisis de la literatura sobre difusión de tecnología se centra en empresas individuales que adoptan nueva tecnología y contempla los efectos que emanan de la adopción de nueva tecnología dentro de la empresa que la adopta.
- Los estudios sobre organizaciones industriales aplicadas también se centra en las empresas que adoptan nuevas tecnologías, pero trata de capturar no solo los efectos sobre el empleo dentro de la empresa, sino también los efectos secundarios a otras industrias y las posibles interdependencias intraindustriales e interindustriales.

- Los estudios basados en el análisis input-output y centrados en los efectos multiplicadores valoran conceptualmente la inversión de los operadores de telecomunicaciones y analizan los efectos de esa inversión en diferentes sectores de la economía.

Finalmente, destacar que las empresas que adoptan la fibra tienen más ventajas competitivas en comparación con las que no la adoptan. Disponer de Fibra en casa (conocido por sus iniciales en inglés como Fiber To The Home – FTTH) permite la innovación y nuevas oportunidades comerciales en la economía del conocimiento, impulsando a las empresas y organizaciones a adoptar nuevos modelos comerciales y estrategias de marketing.

Autores como [Bresnahan y Trajtenberg \(1995\)](#) sugiere que el despliegue de banda ancha y la adopción de servicios y aplicaciones conducen a innovaciones y aumentos de productividad en muchos sectores importantes de la economía, lo que en última instancia da lugar a un crecimiento total del factor productividad.

Los efectos sobre las empresas dependen de una serie de factores, como:

- Inversiones complementarias que las mismas realicen (por ejemplo, capacitar a los trabajadores o reorganizar la estrategia de ventas o las cadenas de suministro para aprovechar las conexiones a Internet más rápidas).
- Del tipo de industria y trabajadores, siendo las industrias tecnológicas y de servicios las que posiblemente se beneficien más, en comparación con las industrias manufactureras.
- De la cualificación de los trabajadores, ya que la banda ancha mejora la eficiencia organizativa y la productividad marginal de la mano de obra calificada ([Biagi, 2013](#)).

Tal y como señala [Moshiri \(2016\)](#) el uso de las TIC aumenta la productividad internamente al elevar la calidad y la productividad de los insumos, y externamente al facilitar la difusión del conocimiento de una empresa, industria o país a otro.

Existen una serie de estudios que examinan el impacto de la banda ancha en las economías locales ([Holt y Jamison, 2009](#); [Kolko, 2012](#); [Majumdar, Carare y Chang, 2010](#); [Van Gaasbeck, 2008](#)).

[Stiroh \(2005\)](#) dijo que: “la evidencia apunta claramente a un efecto positivo de productividad de la TIC”. [Jorgeson, Ho y Stiroh \(2005\)](#) encontraron que las inversiones en TIC eran responsables de un 25% del crecimiento del factor productividad y un 22% del crecimiento de la productividad laboral en el periodo 1984-2002. [Grimes et al. \(2012\)](#) indican para el caso de Nueva Zelanda que una conexión más rápida a internet a través de banda ancha incrementa la productividad de las empresas frente a las que no la tienen.

La adopción de aplicaciones de TI habilitadas para banda ancha puede afectar la economía al cambiar el comportamiento y la productividad tanto de las empresas como de las personas. Estudios como los de [Forman, Goldfarb y Greenstein \(2005\)](#), [Bresnahan, Brynjolfsson y Hitt \(2002\)](#) y [Brynjolfsson y Hitt \(1996\)](#) se han centrado en los cambios en el comportamiento de las empresas, llegando a la conclusión de que las mejoras en la productividad se producen fundamentalmente en las empresas que se comprometen más intensamente con la integración de TI en los nuevos procesos comerciales.

En la misma línea, [Forman et al \(2005\)](#) distinguen entre empresas "que utilizan TI" y "empresas que mejoran las TI". Las primeras simplemente adoptan las aplicaciones de Internet existentes

para hacer que los procesos comerciales actuales sean más productivos: por ejemplo, utilizan el correo electrónico y la navegación web para aumentar la calidad y reducir los costos de recopilar información de mercado y comunicarse con proveedores y clientes. Las segundas, las empresas que mejoran las TI, desarrollan e integran aplicaciones de “comercio electrónico” más complejas, como CRM y ERP, que pueden habilitar procesos y modelos comerciales completamente nuevos, como la gestión automatizada de la cadena de suministro en línea y las ventas en línea en mercados geográficamente distantes.

Por otro lado, [Colombo et al \(2012\)](#) toman una muestra de 799 PYMEs y observan que el impacto de la adopción de aplicaciones básicas de banda ancha es insignificante (o incluso negativo) para el periodo 1998-2004. Por el contrario, las PYME se benefician de la adopción de aplicaciones de banda ancha avanzadas seleccionadas en función de varios factores: (i) tipo de industria (servicios frente a fabricación); (ii) la relevancia de las aplicaciones de software de y (iii) la realización de cambios estratégicos y organizativos complementarios.

En la medida en que la disponibilidad y el uso de banda ancha fomente cualquier tipo de adopción y uso de TI por parte de las empresas, se puede esperar que se produzcan mejoras de productividad y otros impactos económicos asociados.

6.2. ESTUDIOS SOBRE LOCALIZACIÓN DE LAS EMPRESAS

Existen una serie de trabajos que analizan los efectos de las inversiones en banda ancha sobre la localización de las empresas. Muchas son las preguntas que se plantean y muy diversos los resultados que obtienen.

Autores como [Briglauer y Gugler \(2018\)](#) utilizando un conjunto de datos de panel completo de los Estados miembros de la EU27 para el período 2003-2015, estiman un efecto pequeño pero significativo de la banda ancha ultrarrápida basada en fibra superior a los efectos de la banda ancha básica en el PIB. Este mayor impacto se debe a que los servicios que se pueden prestar o utilizar con banda ancha ultrarrápida son mayores por el aumento en la velocidad de la misma.

Algunos autores consideran que el uso de internet reduce la importancia de la geografía en la localización de las empresas ([Cairncross, 1997](#)). Otros, por el contrario, consideran que el uso de las TIC no reduce la importancia de los beneficios de aglomeración de los centros urbanos ([Graham,1999; Learner&Storper, 2001; Zook, 2002](#)).

Además, [Grubestic y Murray \(2004\)](#) y [Strover \(2001\)](#) demuestran que la desigualdad de la disponibilidad de banda ancha favorece a los grandes centros urbanos en comparación con las ciudades más pequeñas y las áreas rurales, ya que han sido éstos los primeros en llevar a cabo las inversiones de infraestructuras de banda ancha. En el estudio que hicieron [Kandilov y Renkow \(2010\)](#) utilizando las metodologías de evaluación de impacto, tanto la de DIF in DIF como el Propensity Score Matching Method concluyen que los préstamos entregados en 2002 y 2003 tuvieron un impacto positivo sustancial en el empleo, los salarios anuales y el número de establecimientos comerciales en las zonas beneficiarias. Sin embargo, de manera desagregada los resultados positivos se producen fundamentalmente en las zonas próximas a las áreas urbanas.

Por su parte, [Mack \(2012\)](#) concluyó que el impacto positivo de la banda ancha en la ubicación de las empresas no es capaz de superar las externalidades negativas que tienen las zonas remotas, alejadas de los centros urbanos. **En cambio, cuando se hace referencia a zonas ubicadas en la periferia de las principales áreas metropolitanas, sí que la banda ancha genera una ventaja competitiva para retener y atraer empresas en sectores intensivos en conocimiento.**

Algunos de los modelos, como el de [Mack y Rey \(2011\)](#) se centran en un determinado tipo de empresa, al examinar el vínculo entre los cambios en el número de empresas intensivas en conocimiento con la banda ancha. Concluyen que la existencia de banda ancha tiene un efecto positivo y estadísticamente significativo en las decisiones de localización de las empresas intensivas en conocimiento en las 54 áreas metropolitanas en estudio- excepto en cinco- aunque con variaciones en función al tamaño de la zona y a su especialización.

El estudio de [Lehr, Osorio, Gillett y Sirbu \(2006\)](#) antes mencionado utiliza un conjunto de datos de panel transversal de áreas en EEUU (a dos niveles: por estado y por código postal) para analizar los efectos de la banda ancha para el periodo 1998 y 2002. El análisis se basa en zonas con disponibilidad – no pudieron utilizar datos de uso porque no estaban disponibles- de banda ancha en diciembre de 1999 y concluye que existe un crecimiento mayor en el número de establecimientos que en las zonas donde no hay banda ancha. Este impacto positivo (0.5% mayor tasa de crecimiento) fue mayor para los establecimientos de mayor tamaño y para los sectores intensivos en IT. Solamente para las empresas pequeñas el efecto es negativo (reducción del 1%).

El trabajo de [Kim y Orazem \(2012\)](#) utilizan la metodología de evaluación de impacto, de Diferencias en Diferencias, para comparar la diferencia en la probabilidad estimada de que las empresas se establecieran en un determinado municipio en el periodo 1990-1992 antes de que la banda ancha estuviera disponible con el mismo municipio después del despliegue de la banda ancha- periodo 2000-2002-. Encuentran un efecto positivo y significativo en la entrada de empresas en el código postal, pero solo en los mercados rurales adyacentes a un área metropolitana o con una población urbana más grande. El acceso de banda ancha no afecta la entrada de nuevas empresas en mercados rurales más remotos.

El trabajo de [Falk and Hangsten \(2021\)](#) analiza la relación entre el acceso de banda ancha de fibra de alta velocidad y la dinámica de establecimiento a nivel municipal para 290 municipios en Suecia durante el periodo 2010-2018. Los resultados de las estimaciones del modelo de efectos fijo y espacial de Durbin revelan un efecto directo significativo, pero bastante pequeño, del acceso de banda ancha de alta velocidad retrasado, en el número de micro establecimientos. Sin embargo, cuando la proporción de establecimientos con acceso de banda ancha de alta velocidad se combina con la presencia local de trabajadores e investigadores con titulación universitaria, aparece una relación indirecta más fuerte. También hay evidencia de que la disponibilidad de banda ancha en los municipios vecinos tiene spillovers. Se encuentra un alto grado de heterogeneidad en los resultados.

[Kolko \(2012\)](#) concluye que el impacto sobre el desarrollo local de la inversión en banda ancha se produce fundamentalmente para los sectores que dependen más de la tecnología de la información y en áreas con menor densidad de población.

[Deller, Whitacre y Conroy \(2021\)](#) analizan el impacto de la calidad del acceso, distinguiendo entre cuatro tipos de velocidad de subida y de descarga de datos, para concluir que la velocidad para descargar información es más relevante que la velocidad para subir datos. La banda

ancha es importante para el desarrollo rural, pero sus efectos dependen del tipo de industria que se esté analizando.

Hasbi (2017) utilizando datos de panel a nivel micro que cubren casi 5000 municipios en la Francia metropolitana, de 2010 a 2015, llega a la conclusión de que los municipios con una red de banda ancha de muy alta velocidad tienden a ser más atractivos para las empresas, con un efecto positivo en la creación de establecimientos dentro del sector terciario y el sector de la construcción y también favorecen el espíritu empresarial. Sin embargo, los beneficios no se distribuyen por igual entre los municipios y dependen del nivel educativo de la población.

McCoy et al (2018) indican que para el caso irlandés, la disponibilidad de banda ancha es un determinante significativo importante para el establecimiento de empresas en una zona, pero sus efectos se pueden ver afectados por la presencia de suficiente capital humano en la misma.

Parajuli y Haynes (2017) han llegado a la conclusión de que la banda ancha tiene efectos positivos sobre el espíritu empresarial en la región, pero no observan cambios económicos significativos en el corto plazo. Los autores han comparado los resultados a través de un modelo OLS estándar frente a un modelo de regresión ponderado geográficamente, concluyendo que éste último daba mejor resultados, dada la heterogeneidad entre los condados de un mismo Estado en EEUU.

Atasoy (2013) encuentra que si bien la banda ancha conduce a un aumento del empleo, no tiene ningún efecto sobre el número de empresas porque el efecto sobre el empleo se debe a un aumento en el tamaño de las empresas que ya están establecidas y no a la creación de un mayor número de empresas

Dinterman (2016) y Mack & Wentz (2017) tampoco encuentran una correlación significativa entre la banda ancha y la ubicación de la empresa.

Van Gaasbeck (2008) concluye que la banda ancha conduce a un aumento de la productividad y, por lo tanto, a una mayor nómina y empleo, pero a un menor número de establecimientos.

7. OTRAS VARIABLES QUE INFLUYEN EN LA LOCALIZACIÓN DE LAS EMPRESAS

7.1. TIPO DE MUNICIPIO- RURAL FRENTE A URBANO-

El acceso a internet puede permitir a las empresas escapar de los efectos negativos de la aglomeración, que se han indicado previamente: congestión de tráfico, precios elevados del suelo y de la vivienda, contaminación, ausencia de zonas verdes.

Sin embargo, algunos autores consideran que los beneficios en las zonas rurales serán menores porque la adopción de la banda ancha es mucho menor en dichas zonas (Forman, 2005; Salemink & al., 2017), por lo que las oportunidades económicas generadas por el acceso a la banda ancha podrían permanecer desaprovechadas. En segundo lugar, si bien la banda ancha puede compensar en parte la lejanía, no puede remediar todas las deficiencias (falta de mano de obra calificada, espíritu empresarial, etc.) y puede resultar difícil para las áreas estructuralmente débiles atraer empresas incluso si están conectadas (Galloway, 2007). Las empresas de uso intensivo de tecnología que emplean mano de obra calificada, y que son las más receptivas a banda ancha ultrarrápida, se encuentran poco representadas en las zonas urbanas.

Las empresas locales pueden incluso verse afectadas negativamente, en el caso de que las empresas online, con productos más diversificados, empiecen a abarcar y competir con los minoristas establecidos en las zonas rurales (Cumming y Johan, 2010), lo que puede llevar a la quiebra de algunas empresas locales (Kandilov y Renkow, 2010).

7.2. VELOCIDAD DE LA BANDA ANCHA

Junto a la posibilidad de tener acceso a banda ancha, medido mediante la cobertura, también la **variable de velocidad de internet permite hacer un uso mayor de la misma**, y por lo tanto, se debe considerar, ya que la mejora en las características técnicas permite acceder a nuevos contenidos, aplicaciones y servicios, o suministrarlos.

7.3. ENDOGENEIDAD

La banda ancha experimenta endogeneidad. Autores como Holt & Jamison (2009) encuentran que existe endogeneidad al menos con dos otras variables: el crecimiento económico- como se observó en la meta evaluación- y la demanda de banda ancha.

7.4. SECTOR DE ACTIVIDAD

Existe **heterogeneidad en el uso de TIC** por parte de las empresas por una serie de razones que han sido señaladas en estudios previos:

- La propensión a la adopción de nuevas tecnologías depende del tipo de sector que se esté considerando ya que están relacionados con cuatro características de los mismos: el uso

previo de TIC , los costes laborales, el crecimiento del sector y la proximidad geográfica (Forman et al., 2003).

- Es difícil establecer cuál ha sido el efecto de la banda ancha sobre cada sector, pero en base a la revisión sistemática de la literatura previa en algunos sectores industriales, como los intensivos en conocimiento -KIBS knowledge intensive business sectors- y sectores creativos, la concentración de empresas incluso ha aumentado (Polèse & Shearmur, 2004) ya que el efecto es mayor que en los sectores primarios o en los sectores de finanzas, inversiones e inmobiliario- FIRE finance, investment and real state- (Duvivier, 2020). Así, en algunas industrias, Internet aparece mucho más como un complemento, más que como un sustituto, de la aglomeración.

7.5. TAMAÑO DE LA EMPRESA

El tamaño de la industria también es una razón por la que existe heterogeneidad en el uso de la banda ancha (Burgess, 2009; Forman, 2005; Gibbs&Tanner, 1997; Karshenas&Stoneman, 1993; Pontikakis&Collins,2010). Según algunos de los estudios realizados previamente, suelen ser las empresas de más tamaño- que son las que hacen un uso mayor de las TIC las que más se benefician de las inversiones en banda ancha. Las PYMEs hacen un menor uso porque tienen incertidumbre sobre cuáles serían los retornos que pueden esperar, desconocen las aplicaciones y los beneficios que les reportan, no disponen de personal con experiencia y cualificaciones en TI y les falta, en consecuencia, la capacidad técnica para implementar nuevas soluciones TI (Gibbs, 2001).

8. ECONOMETRÍA ESPACIAL

El análisis econométrico convencional considera que las unidades de observación son independientes. Sin embargo las unidades territoriales (ciudades, regiones, países...) están interrelacionadas. Existe simultaneidad porque las decisiones de una empresa afectan a las vecinas y las decisiones de éstas afectan a la primera.

Cuando la localización importa, hay que llevar a cabo un análisis espacial.

Varios problemas pueden hacer que los resultados obtenidos, mediante los modelos de regresión tradicionales en dos etapas, puedan ser sesgados:

Se violan los supuestos básicos del modelo de regresión clásico, en el que se indica que los residuos son independientes, por dos razones:

- **Heterogeneidad espacial:** no existe una única regresión lineal cuyos errores presenten una varianza constante para toda la muestra (heteroscedasticidad en los residuos). Existe una desigual distribución de los recursos. **No se puede asumir que el proceso que estamos estudiando opera bajo las mismas reglas en todo el territorio).**
- **Dependencia espacial:** los regresores no son fijos. Los valores de las observaciones están relacionados.

No siempre es fácil distinguir entre dependencia y heterogeneidad espacial y una especificación incorrecta de uno de ellos, puede conllevar el que el otro también sea erróneo.

La **autocorrelación espacial**, que supone la existencia de una relación funcional entre lo que ocurre en un punto determinado del espacio y lo que ocurre en otro lugar, lo cual ha sido debidamente señalado en estudios de distribución espacial de banda ancha (Grubestic, 2006, 2008b) y establecimientos (Maoh & Kanaroglou, 2007; Carroll, Reid, & Smith, 2008; Banasick, Ge Y Hanham, 2009; Mack y Grubestic, 2009).

9. HETEROGENEIDAD ESPACIAL

La **heterogeneidad espacial**, es decir, desigual distribución de los recursos. Se refiere a la falta de estabilidad en el espacio del comportamiento de la variable analizada, lo que hace que la forma funcional y los parámetros de la regresión puedan variar en función a la localización y no sean homogéneos en toda la muestra².

La Heterogeneidad espacial aparece cuando no podemos asumir que el proceso que queremos estudiar operaba bajo las mismas reglas en todo el territorio:

- Desde una perspectiva global: existe un único equilibrio. Es decir existe estacionariedad: una única forma funcional y coeficientes fijos.
- Desde una perspectiva local: existen múltiples equilibrios. No existe estacionariedad: la forma funcional cambio y/o los parámetros del modelo cambian (incidental parameter problema)

² Econometría espacial: nuevas técnicas para el análisis regional. Una aplicación a las regiones europeas. Rosina Moreno Serrano y Esther Vayá Valcarce.

10. DEPENDENCIA ESPACIAL. DEFINICIÓN, FUENTES, TIPOS

Por un lado, no tiene que ver con las características inherentes al territorio sino con la distribución espacial de las observaciones. La existencia de **dependencia espacial** supone que los valores de las observaciones están relacionados a través de la distancia.

La heterogeneidad espacial y la dependencia espacial no son sinónimos.

a. FUENTES DE LA DEPENDENCIA ESPACIAL³

- Datos. Pueden producirse errores sistemáticos en la medición o ser consecuencia de la propia delimitación administrativa.
- Fenómenos sociales. Algunos fenómenos sociales pueden explicar por qué existe dependencia espacial como, por ejemplo:
 - Existencia de fenómenos socioeconómicos de interacción, de difusión o transferencia, de dispersión, de imitación etc entre distintas localizaciones.
 - Existencia externalidades espaciales: la actividad de consumo o de producción afecta realizada afecta a los demás.
 - Multidireccional
 - Otros: normas sociales, efecto compañero, contaminación, knowledge spillover, tax competition, yardstick competition...

La presencia de cualquiera de las formas de dependencia espacial invalida los supuestos subyacentes al modelado de regresión lineal clásico y **necesita estimadores econométricos espaciales alternativos (Anselin, 1988).**

b. MATRICES DE PONDERACIÓN O DE ORDENACIÓN ESPACIAL

Para comprobar si existe autocorrelación espacial se utilizan las **matrices de ponderación espacial**, que son la representación de la estructura espacial de los datos. Es una cuantificación de las relaciones espaciales que existen entre las entidades del dataset (o, al menos, una cuantificación del modo en el que se conceptualizan esas relaciones). Se trata de matrices cuadradas, representadas generalmente por W , en las que cada elemento, w_{ij} , recoge la interacción entre las observaciones que representa. Debido a que la matriz de ponderaciones espaciales impone una estructura en los datos, debe seleccionar una conceptualización que refleje lo mejor posible el modo en que las entidades realmente interactúan entre sí (teniendo en cuenta, por supuesto, qué es lo que intenta medir).

Existen distintos tipos de matrices de ponderación espacial:

Los elementos de la matriz pueden ser de diversa naturaleza.

- **Basadas en la geografía:** entre las que se distinguen:

³ Anselin, 1988

- **matrices de contigüidad binaria y simétrica.** Los criterios de contigüidad pueden ser varios: torre, alfil o reina. En las llamadas matrices de contactos cada elemento está representado por un 1 o por un 0, dependiendo de si se supone que hay o no proximidad entre cada par de observaciones.
 - **matrices de distancia-** pueden ser inversa de la distancia o umbral de distancia. En el caso de las matrices con ponderación espacial inversa de distancias: las w_{ij} suelen representar la inversa de la distancia entre las observaciones i y j , ya que la tesis que se trata de corroborar en la mayoría de los procesos espaciales es que la interacción es función inversa de la distancia.
 - **matriz de los k vecinos más cercanos** (nearest neighbours).
- **Basadas en factores socioeconómicos:** comercio, migraciones, capacidad tecnológica, redes sociales...

2. ANÁLISIS EXPLORATORIO/DESCRIPTIVO DE LOS DATOS

Antes de la estimación de los modelos, se va a realizar un análisis exploratorio de los datos, en el que se calcule el valor de la I de Moran global (1948) y de la I de Moran Local (Anselin, Ibnu y Smirnov, 2002) para examinar la cantidad de agrupamiento en cada una de las variables y evaluar si la autocorrelación puede ser un problema que sería necesario abordar para la estimación. El objetivo es **conocer si existen patrones espaciales en la variable dependiente-** en nuestro caso el número de empresas y la variación en el número de empresas- , en las variables explicativas o en el error.

Para llevar a cabo un análisis exploratorio de los datos se utilizan:

- La I de Moran
- Diagrama de dispersión de la I de Moran (Moran Scatterplot)

a. I DE MORAN

La I de Moran mide la autocorrelación espacial basada en las ubicaciones y los valores de las entidades simultáneamente. Dado un conjunto de entidades y un atributo asociado, evalúa si el patrón expresado está agrupado, disperso o es aleatorio.

La I de Moran se calcula a través de la siguiente fórmula:

$$I = \frac{N \sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^N (x_i - \bar{x})^2}$$

$$S_0 = \sum_i \sum_j w_{ij}$$

Si W estandarizada por filas $S_0 = N$

N : tamaño muestral

w_{ij} es el elemento de la matriz W que relaciona i, j

$|I| < 1$ como un coeficiente de correlación

La herramienta calcula el valor medio y la varianza para el atributo que se evalúa. A continuación, resta el valor medio en cada valor de la entidad, lo que crea una *desviación del valor medio*. Los valores de desviación para todas las entidades vecinas (por ejemplo, las

entidades dentro de la banda de distancia especificada) se multiplican de forma conjunta para crear un *producto cruzado*.

Cuando los valores para las entidades vecinas son mayores que el valor medio o menores que el valor medio, el producto cruzado será positivo. Cuando un valor es menor que el valor medio y el otro es mayor que el valor medio, el producto cruzado será negativo. En todos los casos, mientras mayor sea la desviación del valor medio, mayor será el resultado del producto cruzado. Si los valores en el dataset tienden a agruparse espacialmente (los valores altos se agrupan cerca de otros valores altos; los valores bajos se agrupan cerca de otros valores bajos), el Índice de Moran será positivo. Cuando los valores altos rechazan otros valores altos y tienden a estar cerca de valores bajos, el Índice será negativo. Si los valores positivos de los productos cruzados equilibran los valores negativos de los productos cruzados, el Índice será cercano a cero. El numerador está normalizado por la varianza de modo que los valores del Índice oscilan entre -1,0 y +1,0.

Después de calcular el valor del Índice, la I de Moran global calcula el valor del Índice esperado. Después se comparan los valores del Índice esperado y del Índice observado. Dado el número de entidades en el dataset y la varianza del total de los valores de datos, la herramienta calcula [una puntuación z y un valor P](#) que indican si esta diferencia es estadísticamente significativa. Los valores del índice no se pueden interpretar directamente; sólo se pueden interpretar dentro del contexto de la hipótesis nula.

La I de Moran depende de la W que utilicemos y no es comparable entre distintas muestras ni para la misma muestra si se utilizan distintas.

La hipótesis nula establece que el atributo que se analiza está distribuido en forma aleatoria entre las entidades del área de estudio; es decir, los procesos espaciales que promueven el patrón de valores observado constituyen una opción aleatoria.

Si el p-valor > 0,05 no hay dependencia espacial, ya que no es significativo, independientemente del signo.

Si la I de Moran está sin estandarizar :

Si $p - valor < 0,05$ y $I > \frac{-1}{N-1}$ | Autocorrelación espacial positiva

Si $p - valor < 0,05$ y $I < \frac{-1}{N-1}$ | Autocorrelación espacial negativa

Si la I de Moran está estandarizada :

Si $p - valor < 0,05$ y $z - value > 0$ | Autocorrelación espacial positiva

Si $p - valor < 0,05$ y $z - value < 0$ | Autocorrelación espacial negativa

b. DIAGRAMA DE DISPERSIÓN DE LA I DE MORAN (MORAN SCATTERPLOT)

Representa la I de Moran como la pendiente de una regresión:

$$x = I_M Wx.$$

En el eje X se representa la variable x y en el eje Y su retardo espacial.

Es un instrumento gráfico que nos permite:

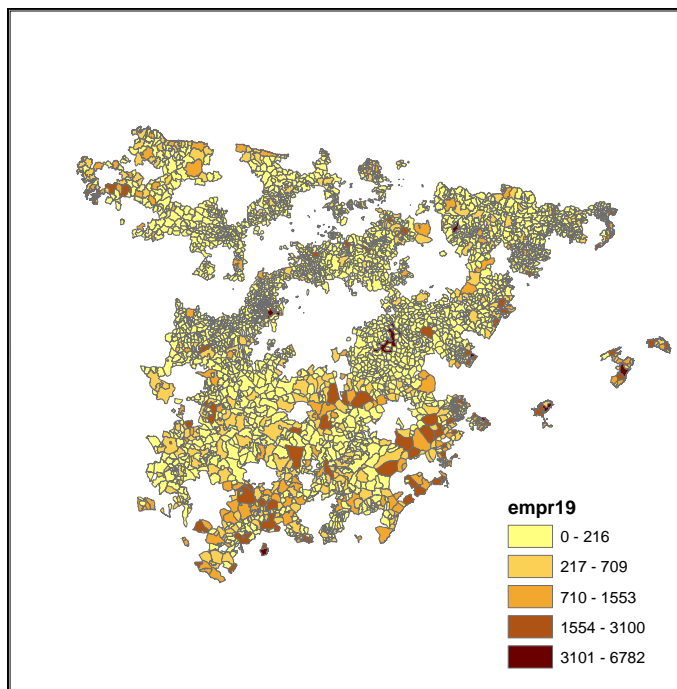
- Determinar la magnitud de la autocorrelación espacial: podemos ver si la pendiente es alta, baja, positiva o negativa
- Descomponer los tipos de dependencia espacial. Puede ser positiva (cuando los valores semejantes de la variable tienden a localizarse cerca de valores semejantes) o negativa (los valores cercanos de la variable que observamos son más disímiles que los que observaríamos si la variable tuviera una distribución espacial aleatoria).
- Detectar outliers y eliminarlos
- Detectar outliers espaciales

3. RESULTADOS DEL ANÁLISIS EXPLORATORIO DE LOS DATOS

Para comprobar si la banda ancha y el establecimiento de empresas presenta alguna forma de dependencia espacial se observa el siguiente gráfico:

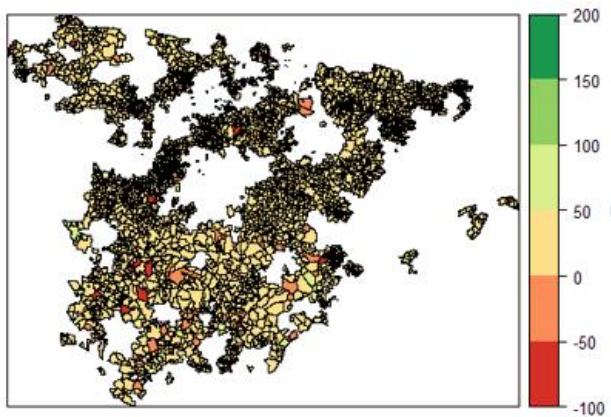
En la figura 1 se presenta la distribución por cuartiles del número de empresas en 2019 en los **municipios que no pertenecen a áreas urbanas**. Como se observa valores similares de la variable dependiente se encuentran agrupados en el espacio, indicando **la presencia de autocorrelación espacial positiva**.

Figura 1



En la figura 2 se representa la tasa de crecimiento de la variable dependiente en el periodo 2012-2019. Como se puede observar los municipios donde aumenta (disminuye) el número de empresas también se agrupan en el espacio.

Figura 2: TASA DE CRECIMIENTO DE LA VARIABLE DEPENDIENTE



Fuente: elaboración propia

a. RESULTADOS DE LA I DE MORAN Y DE LA I DE MORAN LOCAL (LISA) EN LA VARIABLE DEPENDIENTE: NÚMERO DE EMPRESAS Y CRECIMIENTO DEL NÚMERO DE EMPRESAS

En el periodo de análisis existe **dependencia espacial positiva tanto en la variable dependiente como en su evolución a lo largo del periodo**, tal y como confirma el contraste de la I de Moran.

Cuando se aplica este contraste sobre la variable número de empresas en cada municipio en el año 2019 el estadístico de la I de Moran toma un valor estandarizado igual a 0.3159 y su p-valor es $2.2e-16$. Mientras que el valor del estadístico para la tasa de crecimiento del número de empresas es 0.0616 y su p-valor es $1.071e-08$.

Ambos contrastes **rechazan la hipótesis nula de ausencia de autocorrelación espacial en las variables**⁴.

Los resultados son robustos cuando se utiliza **una matriz de ponderación espacial de contigüidad binaria, una matriz de ponderación espacial inversa de distancias, o una matriz de los k=4 vecinos más cercanos**. Una matriz de ponderaciones del vecino más cercano garantiza que todas las áreas de código postal tengan el mismo número de vecinos.

A continuación, en el gráfico 1 representamos el diagrama de dispersión de la I de Moran para la variable crecimiento del número de empresas, calculado para la matriz de ponderación espacial binaria.

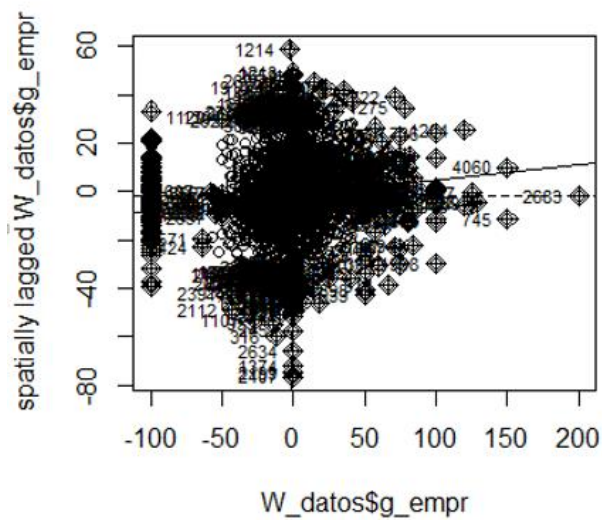
Este gráfico permite determinar la intensidad de la dependencia espacial, así como la existencia de municipios que se desvían del patrón general de dependencia espacial positiva. Como se puede observar la mayoría de las observaciones se sitúan en el cuadrante superior derecho y en el cuadrante inferior izquierdo.

Existe autocorrelación espacial positiva cuando:

⁴ Tanto cuando la hipótesis nula se contrasta de forma analítica como por simulaciones de Monte Carlo se contrasta que estamos en presencia de dependencia espacial positiva.

- En el cuadrante superior derecho tenemos localizaciones donde la X es superior a la media y están rodeados de vecinos cuya variable x también es superior a la media, es decir, valores estandarizados superiores a cero en unidades de desviación típica.
- Y cuando en el cuadrante inferior izquierdo tenemos localizaciones donde la X es inferior a la media y están rodeados por vecinos que también tienen la X por debajo de la media. Valores estandarizados inferiores a cero en unidades de desviación típica.

Es decir, **aquellos municipios donde el número de empresas ha crecido (disminuido) están rodeados por municipios donde ha crecido (disminuido) el número de empresas.**



El análisis exploratorio confirma **la presencia de dependencia espacial en la variable dependiente** y en su evolución a lo largo del periodo de análisis, por lo que es necesario estimar modelos espaciales que permitan su correcta modelización.

Tal y como se indicó antes, la dependencia espacial puede surgir de las interacciones entre las unidades espaciales, ya sea porque las jurisdicciones cercanas se afectan directamente entre sí, porque se ven afectadas por los mismos factores no observados o porque algunas de las variables utilizadas en los modelos estén medidas con error, debido a que la escala del fenómeno que se mide puede no coincidir con la escala del proceso espacial subyacente, dando lugar al problema de la unidad de área modificable (MAUP). Las inversiones en banda ancha se llevan a cabo a través de las unidades singulares y no de municipios.

b. RESULTADO DE LA I DE MORAN Y DE LA I DE MORAN LOCAL EN LA VARIABLE EXPLICATIVA: INFRAESTRUCTURA DE BANDA ANCHA- MEDIDA A TRAVÉS DE COBERTURA DE 100 MB DE VELOCIDAD

A continuación, se contrasta la existencia de dependencia espacial en la variable explicativa cuyo impacto se quiere evaluar, a saber, la tasa de crecimiento en el periodo 2012-2019 de la infraestructura de banda ancha a 100MB de velocidad.

El valor del estadístico de la I de Moran estandarizado es -0.00155 y el p valor 0.8583- si el p valor es mayor a 0,05 no hay dependencia espacial- por lo que **se acepta la hipótesis nula de**

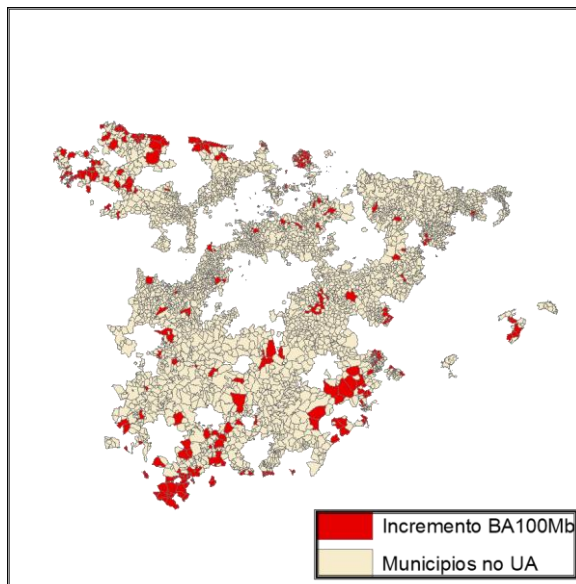
ausencia de autocorrelación espacial en esta variable, es decir, esta variable no presenta autocorrelación espacial.

Este resultado contradice algunos de los estudios previos que indican que existe autocorrelación en la variable explicativa (Grubestic, 2003)⁵. La falta de autocorrelación de este estudio podría deberse a la manera en la que se han lanzado las convocatorias para las inversiones en banda ancha en los municipios rurales:

- Se trata de inversiones financiadas a través de los Fondos FEDER en los que la SEAD establecía aquellas zonas donde se puede llevar a cabo la inversión. Por lo tanto, las zonas donde se iban a llevar a cabo las inversiones se justificaban por objetivos del Ministerio- incluyendo aspectos de viabilidad económica- que eran más amplios que los que hubiera determinado el sector privado en exclusividad.
- La variable explicativa es una variable de inversión en infraestructura- ya que no se ha utilizado como variable explicativa la cobertura a través del sistema móvil (3G, 4G y 5G)
- La forma en la que se han determinado los municipios, al analizarse exclusivamente los rurales. Si se hubieran elegido todos los municipios, incluidos los urbanos, es posible que sí que hubiera resultado que había autocorrelación espacial, dado que las empresas privadas tienen en cuenta los costes de establecer la banda ancha, y estos son menores si se encuentran cerca municipios donde ya han realizado el cableado.

En la figura 3 se muestran los municipios que han experimentado un aumento en la infraestructura de banda ancha de 100Mb- es preciso tener en cuenta que hay muchos municipios donde no se ha invertido-.

Figura 3



⁵ Si no se tiene en cuenta esta autocorrelación, los resultados salen sesgados y los estimadores son inconsistentes.

4. ESTRATEGIAS PARA LA ESTIMACIÓN DE LOS MODELOS DE REGRESIÓN

Cuando no existen argumentos teóricos sólidos para defender un modelo espacial tipo Lag frente a un modelo espacial tipo Error es necesario recurrir a contrastes estadísticos que permitan determinar el modelo que mejor se adapta al proceso generador de los datos (PGD).

En la literatura de econometría espacial existen dos estrategias para identificar el modelo correcto que nos permita interpretar los parámetros estimados.

- **PLANTEAMIENTO ESPECIFICO-GENERAL**

El primer enfoque sigue un **planteamiento Específico-General** y consiste en estimar un modelo de regresión lineal no espacial y, a continuación, se comprueba si es necesario ampliar el modelo con efectos de interacción espacial.

- **PLANTEAMIENTO GENERAL- ESPECÍFICO**

El segundo enfoque, introducido por [Elhorst \(2010\)](#), plantea un **enfoque General-Específico** que comienza con una especificación del modelo lo más general posible, que incluye los retardos espaciales de la variable dependiente, de las variables explicativas y de los residuos. Posteriormente, mediante contrastes estadísticos, se determina si los parámetros estimados de los distintos retardos espaciales son distintos de cero desde un punto de vista estadísticamente significativo.

MODELOS ESPACIALES

El que exista dependencia espacial global no ofrece ninguna explicación de cuál es la razón por la que los valores similares aparecen agrupados. Distintos procesos pueden dar lugar a una misma distribución espacial de la variable. Para ver qué variables explican la evolución de nuestra variable dependiente, es preciso estimar modelos de regresión espaciales.

Para entender los modelos espaciales, es preciso introducir el **concepto de spillovers espaciales**, que suponen que cambios en una unidad espacial producen efectos en otras unidades espaciales. Es decir, cambios en la variable explicativa en la región j tiene un impacto en la variable dependiente de la región i .

Existen dos **tipos de spillovers**:

- **Spillover espaciales globales:** aparecen cuando los cambios en las características de una región tienen efectos en las demás regiones.
- **Spillover espaciales locales.** El impacto del cambio en las características de una región solo se produce en las regiones más cercanas.

Los spillover varían en función a una serie de factores, entre otros, varían con el tamaño de la plantilla de las empresas ([Barff, 1987](#); [Sweeney y Feser, 1998](#); [Duranton y Overman, 2005](#); [Arbia et al., 2012](#)).

El trabajo de Mack and Rey, que sigue la siguiente estructura básica, que está compuesta por tres modelos (no espacial, lag y error):

$$y = X\beta + \mu$$

$$y = \rho Wy + XB + \mu$$

$$y = X\beta + \lambda W\epsilon + \mu$$

donde X es una matriz de variables independientes- que se especificarán en el apartado 8-, β es una matriz de estimaciones de coeficientes para cada una de las variables independientes – son los valores de interés-, μ es la parte del término de error sin dependencia espacial, ρ es la dependencia espacial en la variable dependiente, λ es la dependencia espacial en el término de error, y W es una matriz de ponderaciones espaciales.

La ecuación 1 recoge un **modelo de regresión lineal estándar** estimado por mínimos cuadrados ordinarios.

La Ecuación 2 recoge las especificaciones generales para los **modelos de retardo espacial**. El retardo espacial en (2) incorpora a través de un promedio ponderado de los valores de la variable dependiente en áreas vecinas. Esto es apropiado cuando las interacciones espaciales son operativas en el funcionamiento del mercado, como grupos de trabajo compartidos y vínculos entre empresas que dan como resultado una dependencia espacial en la variable dependiente. **Este tipo de interacción espacial es común cuando la variable dependiente es la localización de empresas.**

La ecuación 3 recoge el error espacial que modela la **autocorrelación espacial** en los términos de error, respectivamente. Alternativamente, el modelo de error espacial en (3) especifica la dependencia espacial a través del error, que puede surgir debido a que no coinciden los límites de la actividad empresarial con los límites administrativos – en nuestro modelo los municipios- para los cuales se recopilan los datos y otras formas de error de medición.

Los modelos donde las pruebas de diagnóstico no revelen dependencia espacial en la variable dependiente o el término de error, se estimarán mediante mínimos cuadrados de dos etapas con variables instrumentales- para hacer frente a la endogeneidad que existe entre inversiones en banda ancha y localización de empresas-.

5. VARIABLES DEL MODELO A ESTIMAR

i. VARIABLE DEPENDIENTE (número de empresas versus tasa de variación)

El número de empresas establecidas en el municipio después de la inversión en banda ancha:

- g_emp: tasa de variación anual entre 2019 y 2012 del número de empresas totales del municipio. En porcentaje.
- finance_seg: número de empresas del municipio dedicadas a las actividades financieras y de seguros de los años 2012-2019.
- prof_tecnicas: número de empresas del municipio dedicadas a las actividades profesionales y técnicas 2012-2019.

- **inform_comunic:** número de empresas del municipio dedicadas a las actividades de información y comunicación entre los años 2012-2019.

De acuerdo con la información antes explicada, se puede ver si los resultados difieren en función del sector en el que trabajan las empresas.

También se podría tomar como variables dependiente: los datos de población por municipio para ver si la banda ancha ha afectado al tema de la migración- y es una buena medida frente al problema de la España despoblada-.

ii. VARIABLES EXPLICATIVAS

VARIABLES ASOCIADAS A LA INVERSIÓN DE BANDA ANCHA: es la principal variable que se desea analizar

- **Cobertura.** En este trabajo, se analiza a través de la variable de cobertura. Sin embargo en algunos de los estudios previos, se utilizó el número de operadores u otras variables relacionadas con el nivel de penetración- número de líneas....-.
- **Penetración: estimada a través de líneas contratadas.** En el momento de analizar el impacto de las inversiones en banda ancha sobre variables económicas de resultado, es preciso tener en cuenta que en determinados países, y en particular en la UE, existen bajas tasas de adopción de la banda ancha para determinados consumidores que disponiendo de acceso a velocidad baja, decidan no contratar servicios a velocidades altas. Por lo tanto, si los ejercicios se hacen teniendo en cuenta la variables de cobertura y no de penetración- número de líneas contratadas, puede ser que los coeficientes estimados representan un efecto proporcional sobre los resultados económicos menor que el efecto a través de la adopción de banda ancha (Czernich, 2014).
- **Uso de internet⁶:** Utilización de aparatos electrónicos con acceso a la red de comunicación de datos Internet para navegar (buscar o consultar información desplazándose por las páginas web), realizado a través de cualquier equipo que permita el acceso, desde cualquier lugar y durante cualquier periodo de tiempo.

Para este estudio nos quedamos con la variable de cobertura medida a través de red fija de banda ancha de distintas velocidades- cada velocidad permite llevar a cabo un tipo de actividad diferente-:

- **g_redfija2mb:** tasa de variación anual entre 2019 y 2013 de la variable de cobertura redfija2mb. En porcentaje.
- **g_redfija10mb:** tasa de variación anual entre 2019 y 2013 de la variable de cobertura redfija10mb. En porcentaje.
- **g_redfija30mb:** tasa de variación anual entre 2019 y 2013 de la variable de cobertura redfija30mb. En porcentaje.

- 6

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608

- **g_redfija100mb**: tasa de variación anual entre 2019 y 2013 de la variable de cobertura redfija100mb. En porcentaje.
- **g_cob3g**: tasa de variación anual entre 2019 y 2013 de la variable de cobertura cob3g. En porcentaje.
- **g_cob4g**: tasa de variación anual entre 2019 y 2013 de la variable de cobertura cob4g. En porcentaje.

Aunque esa es la información disponible en los Ficheros de la SEAD, el estudio no analiza las implicaciones de las líneas móviles- 3G y 4G- y se centra únicamente en la banda ancha ultrarrápida.

Esta decisión está relacionada con los servicios que pueden realizar las empresas dependiendo de la velocidad de la banda ancha.

Además de la principal variable independiente de interés (banda ancha), se incluyeron en los modelos otras variables de control (Anexo II):

- **Características del empleo en el municipio**: nivel educativo (porcentaje de población con estudios superiores) y tamaño del mercado de trabajo (incremento en el periodo de la población en edad de trabajar).
- **Características geográficas**: porcentajes de hogares en alquiler en 2011 (último censo disponible) y el rango medio de elevación para capturar la presencia de montañas en el municipio.
- **Accesibilidad**: kilómetros de carreteras en cada municipio y distancia del centroide del municipio al aeropuerto más cercano. Para evitar problemas de endogeneidad debidos a la causalidad inversa entre la creación de empresas y las mejoras en la red de transporte por carretera se ha utilizado el mapa carreteras principales y secundarias construidas antes de finales del siglo XVIII como fuente de variación exógena para la definición de esta variable.

García-López (2012) encuentra que el uso del mapa histórico del siglo XVIII es un buen instrumento para el mapa actual de carreteras en España. Las dos variables de accesibilidad y la variable rango de elevación del municipio se han construido utilizando el Sistema de Información Geográfica Arcmap.

- **Economías de aglomeración urbana**: densidad de población (hogares/superficie 2011). A diferencia de otros estudios analizados previamente, en este estudio no se tienen en cuenta las variables impositivas dado que en el impuesto de sociedades todas las comunidades en régimen general tributan lo mismo, es decir, se paga lo mismo en todas: un 25%. Todas las Comunidades Autónomas tendrían el mismo impuesto de sociedades, excepto el País Vasco y Navarra.

6. RESULTADOS DE LAS ESTIMACIONES DE LOS MODELOS

Los resultados de las estimaciones de los modelos aparecen detallados en el Anexo IV. Se han utilizado las dos posibilidades: estimar los modelos de específico a general y de general a específico.

El siguiente cuadro recoge los resultados de forma sintetizada:

Siguiendo **el modelo de específico a general** los resultados que se obtienen son los siguientes:

Específico-General			
	Modelo Lag		Modelo Error
	Parámetro estimado	Efecto directo	Parámetro estimado
Variables			
Wy	0.13464*** (4,8516)	
Wu	0.31009***
Constante			
Incremento Banda ancha	0.024419** (2,7160)	0.024528** (0,008911665)	0.022219*** (0,0033443)
Estudios superiores	0.49522*** (19,8933)	0.04974430*** (0,02479335)	0.049689*** (0,0029276)
Población edad trabajar	0.17666** (2,6742)	0.1774548** (0,06779202)	0.039735** (0,013943)
% hogares en alquiler	2,6623*** (6,0882)	2.674264*** (0,430487)	2.7618*** (0,31531)
Rango de elevación	(-0.0010554) (-0,2638)	-0.001060118 (0,004055737)	0.00091735 (0,0059199)
Densidad de carreteras	0.018245 (0,0205)	0.0143274 (0,6827562)	-0.36317 (0,28596)
Distancia al aeropuerto	(-0.000018245*) (-2,3213)	-0.0001832711* (0,00008140228)	-0.00051853*** (0,000089471)
Economías de aglomeración urbana	0.00000064876 (0,8018)	0.0000006516723 (0,0000008179830)	-0.0000036541 (0,000023801)
		Efecto indirecto	
Incremento Banda ancha		0.003689** (0.001796)	

Siguiendo el modelo **de general a específico** se obtienen los siguientes resultados:

Variables	General-Específico			
	Modelo SAC		Modelo LAG	
	Parámetro estimado	Efecto directo	Parámetro estimado	Efecto directo
Wy	0.11754*** (0.025468)	0.13464*** (0.025506)
Wu	0.18989*** (0.050930)
Constante				
Incremento Banda ancha	0.024088** (0,0092089)	0.02416991** (0,009081917)	0.024419** (0,0089036)	0.0245286** (0,008774523)
Estudios superiores	0.49657*** (0,024341)	0.04982552*** (0,0092089)	0.49522*** (0,025282)	0.042476908*** (20.05755601)
Población edad trabajar	0.022940*** (0,068607)	0.024088*** (0,0092089)	0.17666** (0,066347)	0.1774548** (0,06718719)
% hogares en alquiler	2.7336*** (0,46720)	0.024088*** (0,0092089)	2.6623*** (0,45013)	2.674264*** (0,441004)
Rango de elevación	0.0016971 (0,0046054)	0.024088 (0,0092089)	-0.0010554 (0,004073)	-0.001060118 (0,004059282)
Densidad de carreteras	0.054969 (0,72176)	0.024088 (0,0092089)	0.014263 (0,71823)	0.0143274 (0,7171906)
Distancia al aeropuerto	-0.00021233* (0,000095674)	0.024088* (0,0092089)	-0.00018245* (0,000082373)	-0.0001832711* (0,00008316146)
Economías de aglomeración urbana	-0.00000036738 (0,00000081525)	0.024088 (0,0092089)	0.00000064876 (0,00000081752)	0.0000006516723 (0,0000008112212)
		Efecto indirecto		Efecto indirecto
Incremento Banda ancha		0.003126530* (0.001596857)		0.003689846** (0.001692686)

7. VALIDACIÓN DEL MODELO

Para verificar la robustez de los resultados, se realiza un análisis de sensibilidad con respecto a la especificación de la matriz de ponderaciones espaciales. Se han utilizado diferentes matrices de ponderaciones espaciales: una matriz inversa de distancias al cuadrado, una matriz de los k=4 vecinos más cercanos y las matrices de contigüidad binaria como matrices de ponderación espacial.

8. CONCLUSIONES

Los resultados de todos los modelos espaciales son muy estables. En todos los modelos se encuentra una relación positiva entre el incremento de infraestructura de banda ancha y el número de empresas que existen en un municipio.

Los resultados de los modelos indican que la variable explicativa incremento en la infraestructura de banda ancha **es estadísticamente significativa para explicar el número de empresas en cada municipio.**

El resto de variables explicativas presentan el signo esperado. Las variables que recogen las características del mercado de trabajo son estadísticamente significativas.

La variable que captura la existencia de economías de aglomeración no es significativa, reflejando el hecho de que en estos municipios no **existen economías de aglomeración urbana.**

La variable de accesibilidad por carretera tampoco es estadísticamente significativa, mientras que estar cerca de un aeropuerto afecta positivamente al número de empresas en el municipio. El hecho de que el municipio esté situado en una zona geográfica montañosa tampoco tiene un impacto en el número de empresas del municipio.

El que la mayoría de las variables geográficas no sean significativas- salvo distancia al aeropuerto- va en línea con estudios previos que indican que la banda ancha reduce la importancia de la geografía en la localización de las empresas ya que internet permite reducir el número de desplazamientos- trámites y gestiones electrónicas, videoconferencias que sustituyen a reuniones presenciales, compras online frente a desplazamientos a centros comerciales...

Estos resultados son robustos cuando, para estimar los modelos, se utilizan una matriz inversa de distancias al cuadrado, una matriz de los $k=4$ vecinos más cercanos y las matrices de contigüidad binaria como matrices de ponderación espacial.

Los resultados de la evaluación realizada coinciden con las conclusiones del meta análisis de Duvivier en el sentido de que, de acuerdo con los datos disponibles para las inversiones FEDER en Banda ancha en España, se observa un efecto positivo sobre el número de empresas establecidas en los municipios beneficiarios de las inversiones.

9. CONSIDERACIONES FINALES

De los estudios recientes que se han señalado en el trabajo, se concluye que existen varias líneas de análisis que podrían ser interesantes para el futuro:

Por un lado, comprobar la heterogeneidad en los resultados dependiendo de la tipología de la industria y en las diferentes fases de la producción (investigación, producción, distribución...)

También sería relevante analizar si las inversiones realizadas se ajustarían a algún patrón espacial y si se han ido realizando desde el centro de las zonas metropolitanas hacia las zonas rurales, en base al criterio de viabilidad económica de los operadores.

Por último, tal y como señalan algunos de los autores con estudios más recientes, la velocidad tanto de subida como de descarga de datos, influye en las decisiones de localización de las empresas, por lo que se puede considerar como otro posible tema a desarrollar. De hecho, existen pocos estudios (Guidry & al, 2012; Hasbi, 2017). que comparan los efectos de pasar a banda ancha ultrarrápida desde la velocidad de 30 Mbps- que se ha ido logrando prácticamente en la totalidad de España-. Este tipo de trabajo permitiría comparar los beneficios de estas inversiones frente a los costes de incrementar la velocidad, especialmente para las zonas rurales en las que el uso de dichas tecnologías serían sustitutos de las reuniones presenciales.

Teniendo en cuenta los planes para invertir en la tecnología 5 G, un estudio que podría ser de interés sería evaluar si se producen mayores beneficios en las tecnologías móviles dado que ofrecen una mayor flexibilidad para los usuarios. La evaluación de los efectos de las inversiones en banda ancha frente a cobertura a través de líneas móviles es de enorme importancia para las posibles sugerencias de política económica. Especialmente relevante es el análisis en las zonas rurales porque cada vez más individuos utilizan las líneas móviles (Manlove&Whitacre, 2019).

10. RECOMENDACIONES. CONTESTACIÓN A LAS PREGUNTAS DE EVALUACIÓN

Las recomendaciones o sugerencias para la programación de Fondos provienen de las contestaciones que se obtienen a las preguntas que se han planteado en la evaluación. Durante el ejercicio, la no disponibilidad de datos y la falta de tiempo no han permitido que se pueda contestar a todas las preguntas- que quedarían pendientes para desarrollos posteriores de este estudio-:

- ¿Cuáles son los enfoques más relevantes para estimar los efectos de la banda ancha sobre el establecimiento de empresas? ¿Se pueden utilizar modelos de econometría convencional o es más preciso utilizar modelos de econometría espacial?

Es preciso realizar un análisis exploratorio de los datos previo para analizar si existe autocorrelación espacial en la variable de resultado que se pretende analizar. Sin embargo, como se ha visto durante el estudio, las evaluaciones más recientes encuentran dicha autocorrelación, en muchas ocasiones en ambas- variables dependientes y explicativas- y utilizan, cada vez con más frecuencia, metodologías de econometría espacial para analizar el impacto de cualquier medida de política económica en la localización de las empresas.

Muchos son los estudios que han tenido en cuenta dichos efectos espaciales, de muy diferentes maneras. La mayoría de ellos ha utilizado modelos de retraso espacial o error espacial, dependiendo de la naturaleza de la autocorrelación espacial, y han demostrado que **no tener en cuenta los efectos espaciales puede conducir a estimaciones fuertemente sesgadas** (Gurney, 2012; Mack & Rey, 2014; Whitacre & al., 2014-a).

En la evaluación realizada, el análisis exploratorio confirma **la presencia de dependencia espacial en la variable dependiente**- número de empresas establecidas en un municipio rural- y en su evolución a lo largo del periodo de 2012-2019.

Por lo que nuestro trabajo confirma que es necesario estimar modelos espaciales que permitan su correcta modelización del impacto de la banda ancha sobre la variable de resultado seleccionada.

- ¿Afecta el despliegue de la infraestructura de banda ancha ultrarrápida a la localización de las empresas en las zonas rurales?

El despliegue de la infraestructura de banda ancha, medida a través del nivel de cobertura logrado en los municipios rurales, según la definición utilizada por *Boix et al. (2012)* tiene efectos positivos sobre el número de empresas establecidas en los mismos, por lo que se considera una buena medida de política económica para favorecer el crecimiento económico en dichas zonas.

- ¿En qué contexto o cuales son los factores que influyen en que la banda ancha ultrarrápida genere efectos positivos para el desarrollo económico local (áreas con buena accesibilidad por carretera, bien dotadas de mano de obra calificada, próximas a aeropuertos, etc.)?

Las otras variables que influyen de manera estadísticamente significativa en que las empresas se establezcan en un municipio son las variables relacionadas con el mercado de trabajo, es decir, la existencia de población en edad de trabajar en el municipio rural

analizado y el porcentaje de población que tiene estudios superiores. Este resultado indica que la mano de obra cualificada es complementaria al uso de la banda ancha, por lo que las empresas a la hora de tomar la decisión de establecerse en zonas con cobertura tienen en cuenta si podrán contratar trabajadores que puedan hacer uso de las nuevas tecnologías. La otra variable que también influye es la distancia a los aeropuertos.

- Con inversiones realizadas en el último periodo de programación 2007-2014, ¿existe suficiente tiempo para percibir el impacto de las mismas en el establecimiento de empresas en las zonas con conexión?

El tiempo es importante a la hora de determinar el efecto que tienen las inversiones en banda ancha. En el estudio actual se ha tenido en cuenta el periodo 2012-2019. Sin embargo teniendo en cuenta que existe un lapso de tiempo entre que se realizan las inversiones en infraestructura, se mejora la cobertura, las empresas empiezan a utilizar dichas telecomunicaciones.... Es posible que el impacto sea diferente cuando se deje más tiempo, dado que la última convocatoria PEBA tuvo lugar en noviembre de 2020.

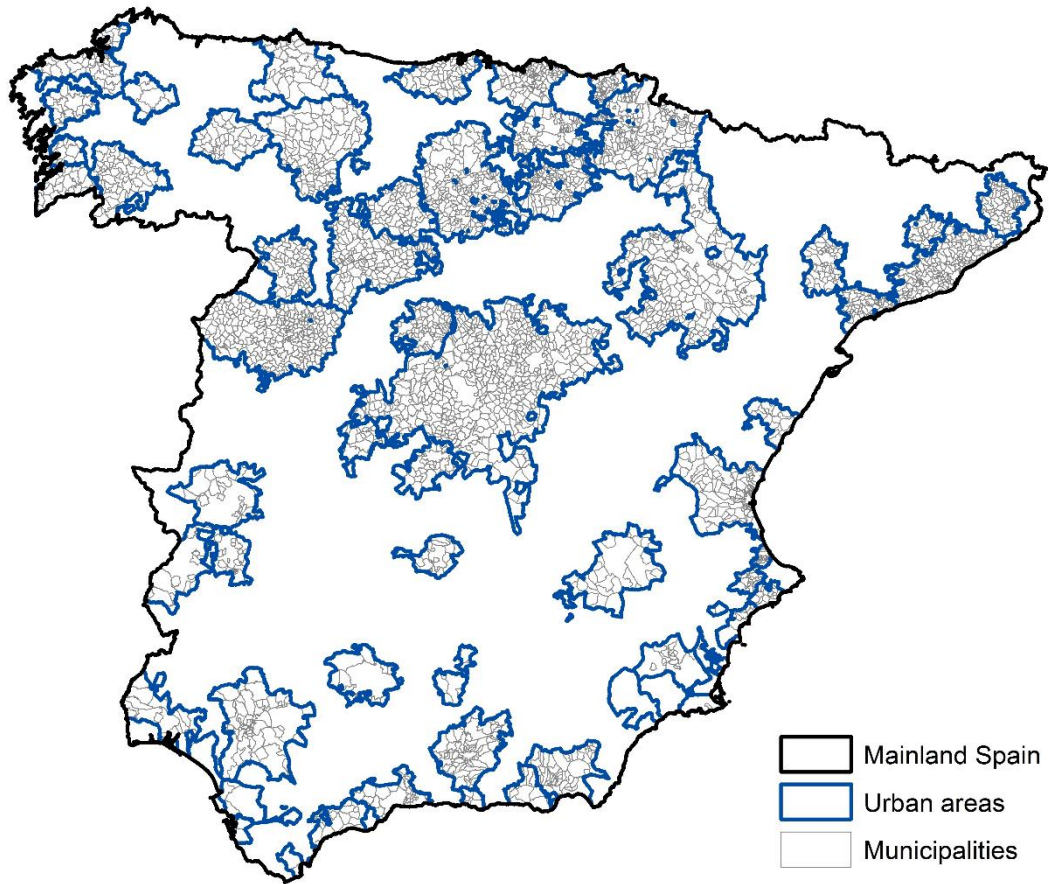
De hecho algunos autores, como [Holt & Jamison \(2009\)](#) indicaron que a corto plazo el efecto de este tipo de inversión puede ser contraproducente, durante el tiempo en el que las empresas están experimentando con las nuevas tecnologías.

Sería, en consecuencia, interesante volver a repetir este ejercicio, en dos o tres años, para ver si el valor del impacto ha cambiado. Además, los datos utilizados no tienen en cuenta los cambios estructurales que ha generado la pandemia, que ha modificado la manera de gestionar los negocios y que es previsible que haya tenido fuerte repercusión en las decisiones de localización de las empresas.

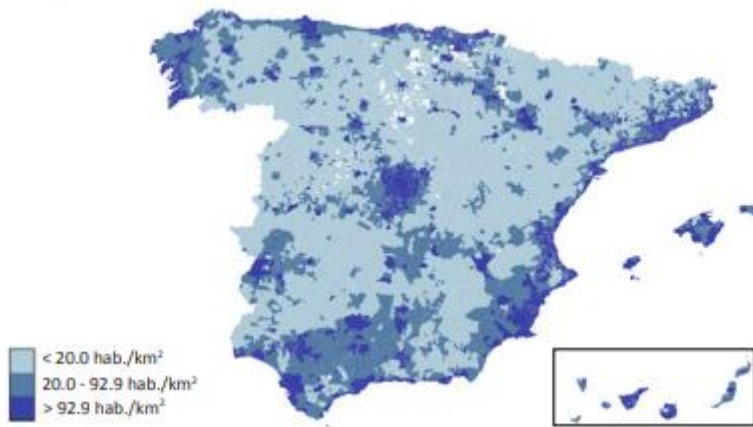
- ¿Influye la velocidad- es decir la calidad de la banda ancha- en las decisiones de localización de las empresas? ¿tiene un mayor impacto la banda ultrarrápida que la banda rápida? En este sentido, teniendo en cuenta que la mayor parte de España tiene ya acceso a banda ancha a velocidad 30 Mbps ¿es adecuado seguir invirtiendo en tener acceso a conexiones de 100 Mbps? ¿generará importantes beneficios económicos adicionales, en particular, se establecerá un mayor número de empresas en zonas rurales con velocidad de banda ultrarrápida?
- ¿Existen efectos más relevantes en zonas urbanas o rurales? Si se analizan solo las zonas rurales, ¿Existen efectos mayores en las zonas situadas en la periferia de las grandes urbes que en las zonas rurales remotas?

En esta evaluación, solo se han tenido en cuenta los municipios rurales, dado que las inversiones de los PEBA's estaban centrados fundamentalmente en los mismos. Algunos previos indican que los efectos son mayores en zonas urbanas (Duvivier, 2019).

ANEXO I: AREAS URBANAS ESPAÑA



ANEXO II: DENSIDAD DE POBLACIÓN DE LOS MUNICIPIOS ESPAÑOLES, 2019
(Habitantes/km²)



Fuente: FUNCAS. La despoblación de la España interior Eduardo Bandrés, Vanessa Azón

ANEXO III: VARIABLES DE CONTROL ANALIZADAS

VARIABLES GEOGRÁFICAS

- **superficie:** área del municipio, medida en km².
- **distance_coast:** distancia de la costa. Proximidad al mar/al océano/los humedales. Distancia de cada centroide a la costa más cercana (unidad de medida: metros).
- **coast_dummy:** variable ficticia igual a 1 si la distancia del centroide del municipio a la costa es inferior a 10km, 0 en caso contrario (unidad de medida: metros).
- **distance_roads:** distancia (en metros) a vías primarias o secundarias del siglo XVIII. Igual para todos los años de la muestra 2010-2019.
- **km_roads:** kilómetros de carreteras en cada municipio (vías principales y secundarias del mapa del s.XVIII). Igual para todos los años de la muestra 2010-2019.
- **rugged_terrain:** modelo digital del terreno español (MDT200) con una malla de elevación de 200 metros. Proporcionado por el Instituto Geográfico Nacional. Igual para todos los años de la muestra 2010-2019.
- **elevation_range:** diferencia (en metros) entre el terreno elevado máximo y el terreno elevado mínimo. Igual para todos los años de la muestra 2010-2019.
- **elevation_mean:** diferencia media (en metros) entre el terreno elevado máximo y el terreno elevado mínimo. Igual para todos los años de la muestra 2010-2019.
- **elevation_median:** diferencia mediana (en metros) entre el terreno elevado máximo y el terreno elevado mínimo. Igual para todos los años de la muestra 2010-2019.
- **urbanarea_code:** identifica a qué área urbana concreta pertenece el municipio. Datos obtenidos de Rafael Boix. Igual para todos los años de la muestra 2010-2019.
- **urban_area:** variable ficticia igual a 1 si el municipio pertenece a una zona urbana, 0 en caso contrario. Datos obtenidos de Rafael Boix. Igual para todos los años de la muestra 2010-2019.

VARIABLES SOCIO ECONOMICAS

Respecto a las variables de carácter económico cuando se analizan los beneficios económicos y locales:

Nombre y descripción de las variables municipales empleadas:

- **densidad:** número de población total del municipio dividido entre la superficie (km²) del municipio, concretamente $densidad = total/superficie$.
- **densidad_hog:** número de hogares residentes en el municipio dividido entre la superficie (km²) del municipio, concretamente $densidad_hog = hogares/superficie$.
- **tamahog:** tamaño medio (unidad de medida: número medio de personas) residentes en el hogar del municipio en el año 2011.
- **hog_alquiler:** porcentaje de número de hogares (unidad de medida: al menos un miembro es arrendatario de toda o parte de la vivienda en alquiler) con respecto al número de hogares totales en el municipio del año 2011. En porcentaje.
- **hogares:** tamaño del hogar (unidad de medida: número de hogares) residentes en el hogar del municipio en el año 2011.
- **est_secundaria:** número de personas con estudios de segundo grado en el año 2011 (se incluyen personas que llegaron al último curso de ESO, EGB o Bachiller Elemental o tiene el

Certificado de Escolaridad o de Estudios Primarios; también se incluye Bachiller, BUP, Bachiller Superior, COU, PREU; FP grado medio, FP I, Oficialía Industrial o equivalente, Grado Medio de Música y Danza, Certificados de Escuelas Oficiales de Idiomas; y FP grado superior, FP II, Maestría industrial o equivalente).

- **est_superior:** número de personas con estudios de tercer grado en el año 2011 (en educación de tercer grado se incluye: Diplomatura universitaria, Arquitectura Técnica, Ingeniería Técnica o equivalente; Grado Universitario o equivalente; Licenciatura, Arquitectura, Ingeniería o equivalente; Máster oficial universitario (a partir de 2006), Especialidades Médicas o análogas; y Doctorado).

ANEXO IV: ESTIMACIÓN DE LOS MODELOS

ESTIMACIÓN DE LOS MODELOS ESPACIALES DE ESPECÍFICO A GENERAL

Esta estrategia consiste en estimar un modelo no espacial y utilizar los contrastes clásicos de los multiplicadores de Lagrange, y sus versiones robustas, para determinar, desde un punto de vista empírico, si el modelo correcto para tratar la dependencia espacial es un modelo espacial tipo Lag o un modelo espacial tipo Error.

Estos contrastes se aplican sobre los residuos del modelo estimado por Mínimos Cuadrados Ordinarios (MCO).

LMerr = 345.22, df = 1, p-value < 2.2e-16

LMlag = 436.09, df = 1, p-value < 2.2e-16

RLMerr = 138.65, df = 1, p-value < 2.2e-16

RLMlag = 229.51, df = 1, p-value < 2.2e-16

Aunque las versiones robustas de ambos contrastes rechazan la hipótesis nula de ausencia de autocorrelación espacial, los valores de los estadísticos de los contrastes son más elevados cuando la hipótesis alternativa es modelizar la dependencia con un modelo espacial tipo Lag. Es decir, los contrastes indican que el modelo correcto es un modelo espacial tipo Lag. No obstante, como los contrastes LMerror también confirman la hipótesis alternativa, que la dependencia espacial puede ser modelizada con un modelo espacial tipo Error, también se presentan los resultados de este modelo.

El modelo espacial tipo Lag presenta la siguiente forma funcional:

$$y = \alpha \tau_N + \rho W y + X \beta + \epsilon \quad (1)$$

donde y representa el vector de N dimensiones que consiste en una observación sobre la variable dependiente para cada unidad de la muestra ($i=1, \dots, N$), τ_N es un vector $N \times 1$ de unos asociados al parámetro del término constante α , X es una matriz $N \times K$ de variables explicativas, β son los parámetros asociados contenidos en un vector $K \times 1$, y $\epsilon = (\epsilon_1, \dots, \epsilon_N)'$ es un vector de términos de perturbación distribuidos de forma independiente e idéntica (i.i.d.) con media cero y varianza σ^2 . W es la matriz de ponderación espacial $N \times N$ que describe la estructura de dependencia entre las unidades espaciales de la muestra. Los elementos de la diagonal principal son cero (ya que ninguna unidad espacial puede considerarse como su propia vecina) y los elementos de fuera de la diagonal (w_{ij}) recogen los patrones espaciales de interacción entre la región i y la región j . La variable Wy es el retardo espacial de la variable dependiente y el coeficiente ρ mide la intensidad de la interacción.

ESTIMACION DEL MODELO ESPACIAL TIPO LAG

En la tabla X se muestran los resultados de estimar el **modelo espacial tipo Lag por el método de Mínimos Cuadrados en dos etapas (2SLS)** y, debido a que los diagnósticos de especificación indican que la varianza de los residuos no es constante (heteroscedasticidad), para hacer inferencia los errores estándar se corrigen por el método de White.

all:stsls(formula = W_datos\$empr19 ~ (g_rf100 + est_sup + g_ptrab + h_alquil + elrange + km_road + d_airp + h_dens), data = W_datos, listw = lwW, zero.policy = TRUE, na.action = na.omit, robust = T)

Coefficients:

	Estimate	Std.error	Z value	Pr(> z) ⁷
Rho	1,35E+03	2,78E+02	48.516	1,23E-03
(Intercept)	9,14E+04	6,57E+04	13.908	0.164279
g_rf100	2,44E+02	8,99E+01	27.160	0.006608
est_sup	4,95E+03	2,49E+02	198.933	< 2.2e-16
g_ptrab	1,77E+03	6,61E+02	26.742	0.007490
h_alquil	2,66E+04	4,37E+03	60.882	1,14E-06
elrange	-1,06E+01	4,00E+01	-0.2638	0.791897
km_road	1,43E+02	6,94E+03	0.0205	0.983612
d_airp	-1,82E+00	7,86E-01	-23.213	0.020271

Los resultados del modelo indican que **la variable explicativa incremento en la infraestructura de banda ancha es estadísticamente significativa para explicar el número de empresas en cada municipio.**

El resto de variables explicativas presentan el signo esperado. Las variables que recogen las características del mercado de trabajo son estadísticamente significativas. La variable que captura la existencia de economías de aglomeración no es significativa, **reflejando el hecho de que en estos municipios no existen economías de aglomeración urbana.** La variable de accesibilidad por carretera tampoco es estadísticamente significativa mientras que estar cerca de un aeropuerto afecta positivamente al número de empresas en el municipio. El hecho de que el municipio esté situado en una zona geográfica montañosa tampoco tiene un impacto en el número de empresas del municipio.

En los modelos de regresión lineal no espaciales los impactos marginales de las variables explicativas sobre la variable dependiente se identifican con la derivada parcial de la variable explicativa respecto de la variable dependiente. [LeSage y Pace \(2009\)](#) señalan que los modelos espaciales que contienen el retardo espacial de la variable dependiente requieren de una interpretación especial de los parámetros. Las estimaciones de los parámetros contienen una gran cantidad de información sobre las relaciones entre las observaciones o jurisdicciones.

[LeSage y Pace \(2009, p. 38\)](#) definen y calculan el promedio de los efectos de directos e indirectos. Un cambio en una variable explicativa afecta a la propia jurisdicción (un impacto directo), pero también afecta a otras jurisdicciones de forma indirecta (un impacto indirecto o spillover). Es decir, en los modelos espaciales que incluyen el retardo espacial de la variable dependiente para determinar el efecto de las variables explicativas es necesario interpretar los efectos totales de cada variable explicativa, que son el resultado de sumar los efectos directos más los efectos indirectos. Los efectos más relevantes en términos absolutos suelen ser los efectos directos, no obstante, pueden verse compensados o reforzados por efectos indirectos (spillovers) del mismo

⁷ Si el p-valor ≤ 0.01 el parámetro estimado de la variable explicativa es estadísticamente significativo al 99%. Si es menor o igual a 0.05 explicativa es estadísticamente significativo al 95%, etc.

signo o de signo contrario, que pueden aumentar o disminuir el efecto total de las variables explicativas sobre la variable dependiente.

Impact measures (lag, trace):

	Direct	Indirect	Total
g_rf100	2.452860e-02	3.689846e-03	2.821845e-02
est_sup	4.974430e-01	7.483051e-02	5.722735e-01
g_ptrab	1.774548e-01	2.669459e-02	2.041494e-01
h_alquil	2.674264e+00	4.022904e-01	3.076554e+00
elrange	-1.060118e-03	-1.594740e-04	-1.219592e-03
km_road	1.432740e-02	2.155276e-03	1.648268e-02
d_airp	-1.832711e-04	-2.756953e-05	-2.108406e-04
h_dens	6.516723e-07	9.803128e-08	7.497036e-07

Simulation results (HCO IV variance matrix):

Simulated standard errors

	Direct	Indirect	Total
g_rf100	8.911665e-03	1.796729e-03	1.053185e-02
est_sup	2.479335e-02	1.453977e-02	1.989501e-02
g_ptrab	6.779202e-02	1.259401e-02	7.877423e-02
h_alquil	4.304870e-01	1.409559e-01	5.526326e-01
elrange	4.055737e-03	6.211113e-04	4.659079e-03
km_road	6.827562e-01	1.025183e-01	7.828576e-01
d_airp	8.140228e-05	1.609502e-05	9.626228e-05
h_dens	8.179830e-07	1.338480e-07	9.474327e-07

Simulated z-values:

	Direct	Indirect	Total
g_rf100	2.734451105	2.09238282	2.67075181
est_sup	20.052223844	5.09254720	28.71102943
g_ptrab	2.630849179	2.14298241	2.60668159

h_alquil 6.239129406 2.91960711 5.60480833
 elrange -0.277288595 -0.24806475 -0.27445023
 km_road -0.009184437 -0.06142823 -0.01605432
 d_airp -2.275809489 -1.79329946 -2.22433214
 h_dens 0.776676219 0.78749862 0.78181068

Simulated p-values:⁸

	Direct	Indirect	Total
g_rf100	0.0062484	0.0364043	0.0075682
est_sup	< 2.22e-16	3.5328e-07	< 2.22e-16
g_ptrab	0.0085172	0.0321145	0.0091424
h_alquil	4.4001e-10	0.0035047	2.0849e-08
elrange	0.7815585	0.8040843	0.7837386
km_road	0.9926720	0.9510182	0.9871911
d_airp	0.0228574	0.0729251	0.0261261
h_dens	0.4373498	0.4309900	0.4343258

ESTIMACION DEL MODELO ESPACIAL TIPO ERROR

Como los contrastes de los multiplicadores de Lagrange también indicaban la posibilidad de modelizar la dependencia espacial mediante un modelo espacial tipo Error, a continuación, se presentan los resultados de estimar el modelo error.

La especificación del modelo espacial tipo Error es:

$$y = \alpha \tau_N + \rho W y + \chi \beta + u \quad (2)$$

$$u = \lambda W u + \epsilon$$

La notación es la misma que en la ecuación (1). El término adicional Wu recoge los efectos de interacción entre los términos de perturbación de las diferentes unidades espaciales, λ es el coeficiente de autocorrelación espacial, u sigue un proceso autorregresivo y ϵ es un ruido blanco, $\epsilon \sim N(0, \sigma_\epsilon^2)$.

⁸ Si el p-valor ≤ 0.01 el parámetro estimado de la variable explicativa es estadísticamente significativo al 99% (***) . Si es menor o igual a 0.05 explicativa es estadísticamente significativo al 95% (**), si es menor que 0.1 al 90% (*).

Call: GMerrorsar(formula = W_datos\$empr19 ~ (g_rf100 + est_sup + g_ptrab + tamahog + h_alquil + elrange + km_road + d_airp + h_dens), data = W_datos, listw = lwW, na.action = na.omit, zero.policy = TRUE, method = "nlminb", verbose = TRUE, returnHcov = TRUE)

Type: GM SAR estimator

Coefficients: (GM standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.4709e+01	1.2047e+01	-2.0510	0.040265
g_rf100	2.2219e-02	3.3443e-03	6.6438	3.057e-11
est_sup	4.9689e-01	2.9276e-03	169.7243	< 2.2e-16
g_ptrab	3.9735e-01	1.3943e-01	2.8498	0.004375
tamahog	3.4051e+01	4.8019e+00	7.0911	1.330e-12
h_alquil	2.7618e+00	3.1531e-01	8.7592	< 2.2e-16
elrange	9.1735e-04	5.9199e-03	0.1550	0.876853
km_road	-3.6317e-01	2.8596e-01	-1.2700	0.204081
d_airp	-5.1853e-04	8.9471e-05	-5.7955	6.813e-09
h_dens	-3.6541e-06	2.3801e-05	-0.1535	0.877981

Lambda: 0.31009 (standard error): 0.034741 (z-value): 8.9256 p-value=2.1e-16

Residual variance (sigma squared): 15284, (sigma: 123.63)

El contraste de Hausman evalúa si los parámetros estimados por MCO coinciden como los parámetros estimados por el modelo espacial tipo error. Cuando se acepta la hipótesis nula los parámetros estimados son insesgados. El estadístico de Hausman rechaza la hipótesis nula con un p-valor de 2.2e-16 indicando que la especificación del modelo error no es correcta. No obstante, las variables explicativas que presentan parámetros estimados estadísticamente significativos distintos de cero son las mismas que con el modelo espacial tipo Lag.

ESTIMACIÓN DE LOS MODELOS ESPACIALES DE GENERAL A ESPECÍFICO

El enfoque de general a específico comienza con la especificación del modelo más general:

$$y = \alpha \tau_N + \rho W y + X \beta + W X \theta + u \quad (3)$$

$$u = \lambda W u + \epsilon$$

La notación es la misma que en las ecuaciones 1 y 2, con un término adicional WX que recoge la interacción espacial entre las variables independientes, θ es un vector $K \times 1$ que recoge los parámetros estimados cuando las variables explicativas presentan autocorrelación espacial.

No obstante, en esta estrategia es necesario excluir al menos uno de los efectos de interacción porque de lo contrario los parámetros del modelo no están identificados (*Manski, 1993*). En este punto se plantean dos alternativas, estimar el modelo Durbin espacial (SDM), que excluye el término de autocorrelación del error (es decir, imponiendo la restricción $\lambda=0$), o estimar el modelo SAC, excluyendo los retardos espaciales de las variables explicativas WX (es decir, imponiendo la restricción $\theta=0$). *LeSage y Pace (2009)* sostienen que se debe comenzar con el SDM:

$$y = \alpha \tau_N + \rho W y + X \beta + WX \theta + \epsilon; \epsilon \sim N(0, \sigma_i^2) \quad (4)$$

Este modelo únicamente se puede estimar por el método de estimación de la Máxima Verosimilitud (ML). Debido a que las perturbaciones no siguen una distribución normal no es posible utilizar el método de estimación de la Máxima verosimilitud y, como no es posible implementar métodos de variables instrumentales (IV) para estimar el SDM, el modelo más general que se puede estimar con un método de estimación robusto es el modelo SAC, que consiste en asumir $\theta=0$. *Piras (2010)* y *Kelejian y Prucha (1998)* argumentan que los patrones espaciales implicados en un modelo SAC son más ricos que los implicados en los modelos espaciales tipo Error o tipo Lag.

Como se mencionó anteriormente, nuestra aplicación empírica requiere el uso de estimadores espaciales de heteroscedasticidad y autocorrelación consistentes (HAC). *Piras (2010)* permite la estimación de dos modelos diferentes. El primer modelo supone que el proceso de perturbación ϵ sigue un proceso autorregresivo espacial de primer orden:

$$y = \alpha \tau_N + \rho W y + X \beta + \epsilon;$$

$$\epsilon = \lambda W \epsilon + \varphi \quad (5)$$

donde las innovaciones (ϕ_1, \dots, ϕ_n) se suponen independientes con media cero y varianza σ_i^2 no constante. El procedimiento de estimación consiste en dos pasos en los que se alternan los estimadores de Momentos Generalizados (GMM) y IV (*Piras, 2010*). *Kelejian y Prucha (2010)* ofrecen resultados sobre la distribución asintótica conjunta de los estimadores IV y GMM para el modelo SAC. Sus resultados permiten contrastar la hipótesis (conjunta) de ausencia de spillovers espaciales originados por las variables endógenas o las perturbaciones.

Modelo SAC

Generalized stsls

```
Call: gstslshet(formula = W_datos$empr19 ~ (g_rf100 + est_sup + g_ptrab + h_alquil + elrange + km_road + d_airp + h_dens), data = W_datos, listw = lwW, na.action = na.omit, zero.policy = TRUE)
```

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	1.2378e+01	8.0606e+00	1.5356	0.1246362
g_rf100	2.4088e-02	9.2089e-03	2.6157	0.0089036 **
est_sup	4.9657e-01	2.4341e-02	20.4004	< 2.2e-16 ***
g_ptrab	2.2940e-01	6.8607e-02	3.3437	0.0008266 ***

```

h_alquil  2.7336e+00 4.6720e-01 5.8511 4.882e-09 ***
elrange   1.6971e-03 4.6054e-03 0.3685 0.7125013
km_road   -5.4969e-02 7.2176e-01 -0.0762 0.9392923
d_airp    -2.1233e-04 9.5674e-05 -2.2193 0.0264664 *
h_dens    -3.6738e-07 8.1525e-07 -0.4506 0.6522565
rho       1.1754e-01 2.5468e-02 4.6153 3.925e-06 ***
lambda    1.8989e-01 5.0930e-02 3.7285 0.0001926 ***

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Impact measures (lag, evalues):

	Direct	Indirect	Total
g_rf100	2.416991e-02	3.126530e-03	2.729644e-02
est_sup	4.982552e-01	6.445243e-02	5.627076e-01
g_ptrab	2.301833e-01	2.977566e-02	2.599590e-01
h_alquil	2.742942e+00	3.548167e-01	3.097758e+00
elrange	1.702872e-03	2.202771e-04	1.923149e-03
km_road	-5.515547e-02	-7.134706e-03	-6.229018e-02
d_airp	-2.130518e-04	-2.755959e-05	-2.406114e-04
h_dens	-3.686273e-07	-4.768426e-08	-4.163116e-07

=====

Simulation results (GSTSLS variance matrix):

=====

Simulated standard errors

	Direct	Indirect	Total
g_rf100	9.081917e-03	1.596857e-03	1.049375e-02
est_sup	2.366390e-02	1.400047e-02	2.371787e-02
g_ptrab	6.953517e-02	1.160187e-02	7.889144e-02
h_alquil	4.593472e-01	1.218405e-01	5.580262e-01
elrange	4.451178e-03	6.116980e-04	5.048420e-03
km_road	7.094672e-01	9.573165e-02	8.028318e-01

d_airp 9.593327e-05 1.560980e-05 1.100760e-04
h_dens 8.256534e-07 1.100904e-07 9.324732e-07

Simulated z-values:

	Direct	Indirect	Total
g_rf100	2.7010154	2.0537395	2.6501419
est_sup	20.9919945	4.6255209	23.6746298
g_ptrab	3.3404024	2.6237247	3.3300896
h_alquil	6.0025539	3.0113925	5.5986011
elrange	0.3711542	0.3885777	0.3743282
km_road	-0.1089876	-0.1427250	-0.1133318
d_airp	-2.2654090	-1.8609749	-2.2382485
h_dens	-0.4901452	-0.4403918	-0.4859904

Simulated p-values:

	Direct	Indirect	Total
g_rf100	0.00691281	0.0400009	0.00804580
est_sup	< 2.22e-16	3.7366e-06	< 2.22e-16
g_ptrab	0.00083657	0.0086974	0.00086818
h_alquil	1.9424e-09	0.0026005	2.1609e-08
elrange	0.71052268	0.6975886	0.70816019
km_road	0.91321234	0.8865073	0.90976747
d_airp	0.02348760	0.0627477	0.02520485
h_dens	0.62403112	0.6596534	0.62697400

Los resultados de las variables explicativas coinciden con los obtenidos al estimar un modelo espacial tipo Lag en la estrategia particular a general. Como vimos el modelo espacial tipo Error, aunque el contraste de Hausman rechazase que era el modelo correcto, también ofrecía resultados similares. Como se observa en la tabla X los parámetros ρ y λ son estadísticamente significativos al 99%.

Modelo LAG

La segunda especificación del modelo consiste en asumir que el vector de perturbaciones no necesariamente es generado por un proceso autorregresivo sino por un proceso muy general. A

continuación, se estima el **modelo SAC** mediante el método de estimación GS2SLS asumiendo que el vector de perturbaciones es generado por un proceso muy general (Piras, 2010)⁹:

$$\mathbf{y} = \alpha\tau_N + \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}; \quad \boldsymbol{\varepsilon} = \mathbf{R}\boldsymbol{\theta} \quad (6)$$

Donde $\boldsymbol{\theta}$ es un vector de innovations y R es una matriz no estocástica de $N \times N$. La estimación de esta forma funcional permite contrastar la robustez de los resultados obtenidos en la estrategia Particular a General. Con esta especificación del modelo los parámetros estimados son los mismos que los obtenidos al estimar un modelo espacial tipo Lag, pero el estimador espacial SAC es robusto frente a una posible especificación errónea de las perturbaciones y permite formas (desconocidas) de heteroscedasticidad y autocorrelación espacial.

Stsls with Spatial HAC standard errors

```
Call: stslshac(formula = W_datos$empr19 ~ (g_rf100 + est_sup + g_ptrab + h_alquil + elrange +
km_road + d_airp + h_dens), data = W_datos, listw = lwW, na.action = na.omit, zero.policy =
TRUE, distance = dists, type = "Triangular")
```

Coefficients:

	Estimate	SHAC	St.Er.	t-value	Pr(> t)
Wy	1.3464e-01	2.5506e-02	5.2790	1.299e-07	***
(Intercept)	9.1393e+00	6.7049e+00	1.3631	0.172861	
g_rf100	2.4419e-02	8.9036e-03	2.7426	0.006095	**
est_sup	4.9522e-01	2.5282e-02	19.5880	< 2.2e-16	***
g_ptrab	1.7666e-01	6.6347e-02	2.6627	0.007752	**
h_alquil	2.6623e+00	4.5013e-01	5.9146	3.328e-09	***
elrange	-1.0554e-03	4.0730e-03	-0.2591	0.795545	
km_road	1.4263e-02	7.1823e-01	0.0199	0.984156	
d_airp	-1.8245e-04	8.2373e-05	-2.2150	0.026763	*
h_dens	6.4876e-07	8.1752e-07	0.7936	0.427444	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Impact measures (lag, trace):

	Direct	Indirect	Total
g_rf100	2.452860e-02	3.689846e-03	2.821845e-02
est_sup	4.974430e-01	7.483051e-02	5.722735e-01

⁹ Se contrasta la robustez de la especificación del modelo a diferentes funciones Kernel.

g_ptrab 1.774548e-01 2.669459e-02 2.041494e-01
 h_alquil 2.674264e+00 4.022904e-01 3.076554e+00
 elrange -1.060118e-03 -1.594740e-04 -1.219592e-03
 km_road 1.432740e-02 2.155276e-03 1.648268e-02
 d_airp -1.832711e-04 -2.756953e-05 -2.108406e-04
 h_dens 6.516723e-07 9.803128e-08 7.497036e-07

=====

Simulation results (IV HAC variance matrix):

=====

Simulated standard errors

	Direct	Indirect	Total
g_rf100	8.774523e-03	1.692686e-03	1.029758e-02
est_sup	2.476908e-02	1.325857e-02	2.108188e-02
g_ptrab	6.718719e-02	1.215998e-02	7.811855e-02
h_alquil	4.410040e-01	1.281248e-01	5.486209e-01
elrange	4.059282e-03	6.165669e-04	4.664282e-03
km_road	7.171906e-01	1.109928e-01	8.259293e-01
d_airp	8.316146e-05	1.569038e-05	9.785525e-05
h_dens	8.112212e-07	1.324113e-07	9.400093e-07

Simulated z-values:

	Direct	Indirect	Total
g_rf100	2.78374341	2.21009475	2.73530442
est_sup	20.05755601	5.61662839	27.09795433
g_ptrab	2.63491638	2.21633831	2.61120163
h_alquil	6.05900294	3.19546172	5.61674336
elrange	-0.24762551	-0.19693963	-0.24153947
km_road	0.08303186	0.03374538	0.07663509
d_airp	-2.19099948	-1.81978772	-2.15379220
h_dens	0.78739884	0.79441545	0.79142219

Simulated p-values:

	Direct	Indirect	Total
g_rf100	0.0053736	0.0270986	0.0062323
est_sup	< 2.22e-16	1.9472e-08	< 2.22e-16
g_ptrab	0.0084158	0.0266683	0.0090225
h_alquil	1.3697e-09	0.0013961	1.9459e-08
elrange	0.8044242	0.8438748	0.8091370
km_road	0.9338262	0.9730802	0.9389138
d_airp	0.0284518	0.0687913	0.0312565
h_dens	0.4310484	0.4269536	0.4286977

Como se puede observar los niveles de significación de las variables explicativas están en línea con los obtenidos cuando se estima el modelo espacial tipo Lag corrigiendo la heteroscedasticidad por el método de White en la estrategia Particular a General.