



Concentración empresarial y economías de aglomeración en Aragón

Rafael González Val
Miriam Marcén Pérez



CONSEJO ECONÓMICO Y SOCIAL DE ARAGÓN

COLECCIÓN PREMIOS DE INVESTIGACIÓN

Premio de investigación «Ángela López Jiménez» 2017

Equipo de investigación:

Rafael González Val

Miriam Marcén Pérez

La responsabilidad por las opiniones expresadas en las publicaciones del Consejo Económico y Social de Aragón incumbe exclusivamente a sus autores y su publicación no significa que el Consejo se identifique con las mismas.

2018

© De esta edición digital: Consejo Económico y Social de Aragón.

© Para otras ediciones: los autores.

Esta publicación se edita únicamente en formato digital.

Consejo Económico y Social de Aragón

c/ Joaquín Costa, 18, 1º

50071 Zaragoza (España)

Teléfono: 976 71 38 38 – Fax: 976 71 38 41

cesa@aragon.es

www.aragon.es/cesa

Concentración empresarial y economías de aglomeración en Aragón

Premio de investigación
«Ángela López Jiménez» 2017

Rafael González Val

Miriam Marcén Pérez



Premio de investigación «Ángela López Jiménez» 2017

El Consejo Económico y Social de Aragón, con el fin de promover y divulgar la investigación, convoca anualmente un premio a proyectos de investigación, que desde el año 2007 se denomina Premio de Investigación «Ángela López Jiménez», en reconocimiento a la brillante trayectoria investigadora y a la labor desarrollada por Ángela López, Presidenta del Consejo Económico y Social de Aragón entre mayo de 2000 y marzo de 2007.

El Premio de Investigación «Ángela López Jiménez» correspondiente al año 2017 fue convocado por Resolución de 3 de julio de 2017, de la Presidencia del Consejo Económico y Social de Aragón (BOA nº 134, de 14 de julio de 2017). En la convocatoria pudieron participar los investigadores individuales o equipos de investigadores que presentaran un proyecto sobre materias económicas, sociales o laborales de trascendencia para la Comunidad Autónoma de Aragón.

Por Resolución de 14 de diciembre de 2017, de la Secretaría General Técnica de la Presidencia (BOA nº 244, de 22 de diciembre de 2017), se otorgó el Premio de Investigación «Ángela López Jiménez» 2017, dotado con 10.000 euros, al proyecto «Concentración empresarial y economías de aglomeración en Aragón», del grupo de investigación formado por Rafael González Val y Miriam Marcén Pérez, por su contribución al conocimiento de la realidad empresarial aragonesa, al analizar empíricamente la concentración espacial de empresas por ramas de actividad y estudiar sus efectos sobre el crecimiento del empleo y la población de los municipios de Aragón.

El Jurado encargado de fallar el premio estuvo compuesto por los siguientes miembros del Consejo:

Presidente: D. José Manuel Lasierra Esteban
Secretaria: D^a. Belén López Aldea
Vocales: D^a. M^a Dolores Gadea Rivas
D. José de las Morenas de Toro
D^a. Beatriz Callén Escartín

Índice

1. Introducción	6
2. La distribución espacial de empresas en Aragón	10
2.1.- Análisis preliminar de la distribución espacial de empresas en Aragón	10
2.2.- Metodología	16
2.3.- Resultados	19
2.4.- Conclusiones.....	40
3. Economías de aglomeración en Aragón	41
3.1.- Economías de aglomeración y crecimiento del empleo.....	41
3.1.1.- Metodología.....	43
3.1.2.- Resultados.....	47
3.2.- Economías de aglomeración y crecimiento de la población	54
3.2.1.- Modelo teórico	54
3.2.2.- Metodología.....	58
3.2.3.- Resultados.....	62
3.3.- Conclusiones.....	64
4. Resumen final y conclusiones	67
5. Referencias	70
6. Apéndices	73
Apéndice A: Distribución espacial hasta la máxima distancia	73
Apéndice B: Estimaciones por MCO del efecto de la aglomeración sobre el empleo sectorial local	91

1. Introducción

La concentración creciente de las personas y la producción en el espacio ha sido una de las características del crecimiento económico en las últimas décadas. La población y las empresas tienden a ubicarse en los lugares donde la proximidad física con otras personas y empresas puede aportar algún tipo de ventaja; estas ventajas se conocen en la literatura académica como “economías de aglomeración”. El concepto de las economías de aglomeración es clave en la disciplina conocida como Economía Urbana, ya que la existencia misma de las ciudades se fundamenta en la presencia de estas externalidades que favorecen la aglomeración de empresas y personas. Este concepto se remonta a Marshall (1890), quien observó que la localización de las empresas en áreas con una elevada concentración de empresas del mismo sector conllevaba ventajas como (i) la disponibilidad de mano de obra especializada y (ii) bienes intermedios, y (iii) la facilidad para intercambiar conocimientos sobre productos, procesos e innovaciones. Este planteamiento está en la base de los modelos de geografía económica, junto con la causación circular: los trabajadores acuden a las ciudades con fuertes sectores industriales y a la vez las empresas prefieren situarse cerca de las ciudades más pobladas y con mayor tamaño del mercado.

Recientemente, el concepto inicial de economías de aglomeración de Marshall (1890) fue actualizado por Duranton y Puga (2004), que proponen tres mecanismos basados en micro-fundamentos (modelos teóricos del comportamiento individual de los agentes) para explicar la incidencia positiva de la densidad de empresas y población, a través de la proximidad geográfica: *sharing* (compartir), *matching* (compatibilizar) y *learning* (aprender). El mecanismo *sharing* (compartir) hace referencia a las ganancias que provienen de la gran variedad de factores productivos disponibles y de la especialización industrial; *matching* (compatibilizar) se refiere al mejor emparejamiento entre empleados y empleadores, compradores y oferentes o pares de negocios, y *learning* (aprender) corresponde a la generación, difusión y acumulación de conocimiento aplicado a fines productivos.

La concentración empresarial puede surgir de forma espontánea, o ser algo auspiciado o incentivado por los poderes públicos. En el primer caso, lo que suele ocurrir es que, de manera espontánea, varias empresas pequeñas o de tamaño medio se sitúan cerca de otra empresa de gran éxito en su sector, con el fin de proporcionarle una serie de servicios o factores productivos específicos. Es lo que ocurre, por ejemplo, con General Motors España en Figueruelas (actualmente Opel España SLU) y todas las empresas que le proporcionan bienes y servicios intermedios. Dentro de esta creación espontánea de agrupaciones espaciales de empresas del mismo sector juegan un papel importante las distintas asociaciones empresariales, ya que prestan todo tipo de

servicios a las empresas y emprendedores, fomentando el asociacionismo y la cooperación entre diferentes empresas.

En el segundo caso, son los propios poderes públicos los que, a través de diversos mecanismos de incentivos, persiguen la concentración de empresas en un clúster. En Aragón, existe desde 2016 un Plan de Cooperación Empresarial de Aragón de apoyo a los *clusters* regionales. Entre sus objetivos está tratar de impulsar y promocionar la cooperación empresarial y especialmente a los *clusters* aragoneses, considerados como motor de desarrollo empresarial, capaces de atraer el talento, impulsar la innovación y captar inversiones en sus sectores o mercados, con especial énfasis en la oportunidad que genera que las pymes puedan colaborar para que su reducido tamaño no sea un obstáculo en su competitividad. El resultado de este plan es el apoyo y la promoción a los ocho *clusters* reconocidos en Aragón (que ya existían previamente): Arahealth (Clúster de la Salud de Aragón), Tecnara (Asociación de Empresas de Tecnologías de la Información, Electrónica y Telecomunicaciones de Aragón), Zinae (Clúster Aragonés para el uso eficiente del agua), Clúster Aragonés de Alimentación, AERA (Clúster Aeroespacial de Aragón), ALIA (Clúster Logístico de Aragón), CAAR (Clúster de Automoción de Aragón), y Clúster de la Energía de Aragón. Además de mejorar la competitividad y productividad de las empresas involucradas, el fomento de estos *clusters* puede tener además otros objetivos sociales. Así, se entiende que estas concentraciones empresariales generan crecimiento del empleo y pueden, a su vez, fijar o atraer población, y por lo tanto ayudar a luchar contra la despoblación en algunas zonas de Aragón, que es uno de los grandes problemas demográficos de la Comunidad.

Existe un amplio consenso en la literatura acerca de la importancia de las economías de aglomeración para la productividad y el crecimiento económico. De hecho, la heterogeneidad del crecimiento urbano se interpreta como el resultado de los diferentes niveles de productividad, que, a su vez, son consecuencia del tamaño de las ciudades. En el caso de Aragón, las tres capitales de provincia (y, especialmente, Zaragoza) aglutinan más de la mitad de la población de toda la Comunidad Autónoma (el 57% en 2017), lo que plantea enormes desafíos para las ciudades, dado que cuando emergen las economías de aglomeración emergen también las desigualdades espaciales. A su vez, la distribución de la actividad económica en Aragón es también muy desigual; así, la provincia de Zaragoza cuenta en promedio con el 85% del valor añadido y el 80% del empleo mientras que Huesca tiene el 10% del valor añadido y una proporción algo mayor del empleo, el 13,5%, y Teruel concentra el 5% del valor añadido y el 6,5% de los trabajadores (Informe Económico de Aragón, 2016).

Tradicionalmente se distinguen dos tipos de economías de aglomeración: las economías de localización, relacionadas con la concentración empresarial dentro de un

sector específico, y las economías de urbanización, que operan a través de la concentración general de la actividad económica. Dentro de las economías de localización se incluyen todos aquellos factores que impactan externamente a las empresas y provienen del sector económico en el que se desarrolla la actividad: reducción de costes de transporte, aparición de economías de escala, reducción de costes de transacción, formación de un mercado de trabajo especializado, o creación de una atmósfera industrial (Marshall, 1890) capaz de generar la innovación en estos campos y su rápida difusión. Respecto a las economías de urbanización, incorporan todos aquellos efectos externos que no se generan dentro de la empresa o del sector particular al que pertenece la empresa, sino que provienen de las ventajas que se originan en el lugar, internas a la región o ciudad: economías de diversidad, cualidades de las ciudades o regiones (por ejemplo, las infraestructuras), o el acceso a un mercado laboral eficiente y capacitado. Los beneficios (y costes) de la localización y urbanización son denominadas en conjunto economías (deseconomías) externas, debido a que surgen fuera de la empresa. No obstante, en este estudio nos centramos únicamente en las posibles ventajas de las economías de aglomeración.

A partir de esta clasificación de las economías de aglomeración en economías de localización o de urbanización, los trabajos empíricos más recientes estiman el impacto de las economías externas en los resultados locales de la productividad, los salarios, el empleo y las decisiones de localización de las empresas. Numerosos estudios buscan cuantificar tales efectos en diferentes países y escalas geográficas. Torres Gutiérrez (2017) ofrece un detallado resumen de los principales resultados de los trabajos empíricos. En general, los resultados apoyan la presencia de efectos positivos de la diversidad y la competencia en el crecimiento urbano, mientras que los resultados derivados de la especialización productiva son más ambiguos (Beaudry y Schiffravaurova, 2009; Melo *et al.*, 2009; De Vor y De Groot, 2010; Combes y Gobillon, 2015; Groot *et al.*, 2015). En términos del empleo y la productividad como medidas del crecimiento urbano, la incidencia esperada de las externalidades de localización y urbanización resulta menos previsible para el empleo que para la productividad. En el primer caso, la prevalencia de una u otra externalidad está fuertemente condicionada al contexto y sector de análisis. En el segundo, se espera un impacto positivo de la especialización, antes que de las economías de urbanización.

Otra variable clave en la explicación de la relevancia de las economías de aglomeración es el tamaño de las empresas. En principio, es de esperar un efecto positivo del tamaño promedio de las empresas sobre el empleo y la productividad, dado que son inherentes a las economías internas de escala (Combes, 2000). No obstante, una de las particularidades del caso aragonés (y español) es la importancia de las pymes.

Tabla 1. Empresas según actividad principal y estrato de asalariados. Aragón. 1 de enero de 2018

Ramas de actividad según CNAE-2009	Total	Sin asalariados	De 1 a 9	De 10 a 49	De 50 a 199	De 200 o más
Total de empresas	91.493	48.871	38.135	3.755	591	141
Industria y Energía	6.893	2.199	3.438	1.016	190	50
De las cuales: Industria manufacturera	5.850	1.676	3.003	941	181	49
Industria de la alimentación, fabricación de bebidas e industria del tabaco	1.076	235	609	195	33	4
Industria textil, confección de prendas de vestir e industria del cuero y del calzado	498	214	218	63	3	0
Industria de la madera y del corcho, industria del papel y artes gráficas	697	229	395	59	11	3
Coquerías y refino de petróleo; industria química; fabricación de productos farmacéuticos	156	33	63	43	13	4
Fabricación de productos de caucho y plásticos y de otros productos minerales no metálicos	407	98	194	90	20	5
Metalurgia y fabricación de productos metálicos, excepto maquinaria y equipo	1.306	342	742	187	32	3
Fabricación de productos informáticos, electrónico y ópticos; fabricación de material y equipo técnico; fabricación de maquinaria y equipo n.c.o.p.	530	109	225	151	33	12
Fabricación de material de transporte	155	22	53	42	23	15
Fabricación de muebles; otras industrias manufactureras y reparación e instalación de maquinaria y equipo	1.025	394	504	111	13	3
Construcción	12.228	7.279	4.523	391	33	2
Servicios	72.372	39.393	30.174	2.348	368	89
Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas	19.072	8.439	9.899	655	64	15
Transporte y almacenamiento	5.929	3.589	2.033	255	41	11
Hostelería	7.893	2.462	5.135	261	31	4
Información y comunicaciones	1.283	726	456	84	14	3
Actividades financieras y de seguros	2.148	1.578	546	19	1	4
Actividades inmobiliarias	5.384	4.174	1.194	14	1	1
Actividades profesionales, científicas y técnicas; actividades administrativas y servicios auxiliares	14.708	9.519	4.715	397	56	21
Educación	2.944	1.799	933	146	53	13
Actividades sanitarias y de servicios sociales	4.251	2.581	1.378	216	68	8
Actividades artísticas, recreativas y de entrenamiento; reparación de artículos de uso doméstico y otros servicios	8.760	4.526	3.885	301	39	9

Fuente: Datos básicos de Aragón, 2018 - Instituto Aragonés de Estadística según Directorio Central de Empresas, Instituto Nacional de Estadística.

Unidad: número de empresas.

La Tabla 1 muestra la distribución de empresas en Aragón según la rama de actividad (clasificación CNAE-2009) y el número de empleados a 1 de enero de 2018. Del total de empresas, más de la mitad (53,4%) no tiene ningún asalariado, y el 41,7% tienen de 1 a 9 trabajadores. Ambas categorías suman el 95,1% del total de empresas. Si nos centramos en las empresas que tienen de 50 a 199 trabajadores, estas representan un 0,65% del total, y solo un 0,15% tiene 200 o más empleados. Se trata por tanto de una distribución de empresas muy desigual, con una importancia casi absoluta de las pymes.

La literatura empírica ha establecido que la concentración espacial de la actividad industrial mejora el crecimiento económico, la productividad y la innovación a través de diferentes aproximaciones, en las que el común denominador es el análisis de la dicotomía localización–urbanización. En este estudio se aborda (1) el análisis empírico de la concentración espacial de las empresas aragonesas para las principales ramas de actividad (véase la Tabla 1) a partir de información geográfica de su localización física, y (2) el efecto de la concentración industrial sobre el crecimiento del (i) empleo y (ii) la población en los municipios aragoneses. Para analizar estas cuestiones se estudiará la distribución espacial de las empresas, utilizando información individual de cada empresa y una nueva metodología empírica (Duranton y Overman, 2005, 2008), y posteriormente se analizará el efecto de la concentración empresarial en el crecimiento del empleo y la población a nivel municipal, aplicando rigurosos modelos econométricos. Los datos se obtendrán de la base de datos SABI (más información a continuación) y del padrón municipal.

2. La distribución espacial de empresas en Aragón

2.1.- Análisis preliminar de la distribución espacial de empresas en Aragón

En la introducción se ha indicado cómo la distribución de población y actividad económica es tremendamente desigual en Aragón. Por lo tanto, la distribución en el espacio de empresas también resulta muy atomizada, con un gran número de empresas en la ciudad de Zaragoza. Como una primera aproximación, la Tabla 2 muestra la clasificación de los 20 municipios aragoneses con más empresas, así como la distancia geográfica a Zaragoza, el principal centro de actividad de la Comunidad.

Para realizar este estudio utilizamos datos geográficos individuales de todas las empresas aragonesas, considerando las principales ramas de actividad según la Clasificación Nacional de Actividades Económicas CNAE-2009. La muestra de empresas utilizada es la proporcionada por la base de datos SABI (Sistema de Análisis de Balances Ibéricos), que contiene lo declarado en sus cuentas anuales por las empresas

en el Registro Mercantil (Sociedades Anónimas y Sociedades Limitadas). En promedio, la versión de los datos utilizada (versión software 72.00, actualización de datos 08/08/2017) proporciona información de alrededor de 25.000 empresas activas con sede en Aragón en 2017, aunque el número varía de un año a otro. Esta base de datos proporciona información geográfica muy detallada; no solo se indica la provincia y localidad en las que está situada la sede de la empresa (de acuerdo a la información registral), sino que también se indican las coordenadas geográficas (latitud y longitud), lo que nos permite realizar un análisis considerando el espacio continuo e incluyendo todas las distancias físicas bilaterales entre empresas.

Merece la pena destacar que en este tipo de literatura de economía industrial espacial se suelen considerar los patrones de localización de los establecimientos (es decir, los centros de trabajo) mientras que nuestra muestra está formada por las sedes sociales, que es algo distinto. No obstante, confiamos en que, dado que la inmensa mayoría de las empresas en Aragón son pymes (recuérdese que el 95,1% tiene de 0 a 9 trabajadores, véase la Tabla 1), ambos conceptos (sede social y establecimiento productivo o centro de trabajo) coincidan en la mayoría de los casos y aquellos en que no coincidan sean una anomalía.

Existe una segunda problemática relacionada con el uso de datos societarios; en algunos casos es posible que la localización de la sede sea meramente instrumental, estando el centro de actividad en otro lugar (o incluso en otra Comunidad Autónoma), por ejemplo por motivos fiscales. Un ejemplo de actualidad de esta problemática serían las empresas que antes estaban localizadas en Cataluña y que, tras los acontecimientos políticos de los últimos meses han decidido trasladar la sede social a otras Comunidades (algunas de ellas a Aragón) manteniendo el centro de actividad en el mismo lugar. De nuevo, dado que la fiscalidad en Aragón no es especialmente ventajosa en comparación con la de las Comunidades vecinas y que la mayoría de las empresas aragonesas son pymes, confiamos en que el número de estos casos no sea importante.

La Tabla 2 muestra que aproximadamente un 3,12% de estas empresas están localizadas en la ciudad de Teruel, un 4,34% en Huesca y casi la mitad del total de la muestra (el 46,42%) en Zaragoza. Las siguientes localidades cuentan con un número muy inferior de empresas, representando un porcentaje del total que oscila desde el 2,26% (Cuarte de Huerva) al 0,67% (Tarazona). A nivel provincial, el Informe sobre la situación económica y social de Aragón en 2017 (2018) indica que el 72% del tejido productivo aragonés tiene su sede en la provincia de Zaragoza, el 18% en Huesca y el 10% en Teruel.

Por lo tanto, existe una enorme concentración de empresas en la ciudad de Zaragoza. No obstante, el análisis que se realiza a continuación sigue siendo muy

relevante, ya que la metodología nos permite obtener una estimación continua de la densidad (concentración) de empresas por kilómetro (con independencia de las fronteras municipales), considerando todas las distancias bilaterales geográficas entre empresas dentro de una misma rama de actividad.

Tabla 2. Top 20 Municipios aragoneses con mayor número de empresas

Municipio	Empresas	% del total	Distancia a Zaragoza
1 Zaragoza	11.714	46,42	
2 Huesca	1.096	4,34	66,46
3 Teruel	787	3,12	147,20
4 Cuarte de Huerva	571	2,26	8,56
5 Alcañiz	417	1,65	91,71
6 Fraga	404	1,6	103,23
7 Ejea de los Caballeros	390	1,55	56,74
8 Calatayud	360	1,43	72,05
9 Monzón	360	1,43	93,41
10 La Puebla de Alfindén	334	1,32	11,02
11 Barbastro	333	1,32	93,12
12 Binéfar	327	1,3	99,70
13 Utebo	266	1,05	11,57
14 La Muela	223	0,88	21,52
15 Jaca	222	0,88	105,24
16 Villanueva de Gállego	218	0,86	13,34
17 Cadrete	196	0,78	13,11
18 Zuera	188	0,75	24,86
19 Tauste	172	0,68	42,99
20 Tarazona	169	0,67	75,29
Total	25.233		

Notas: Distancias bilaterales calculadas utilizando la distancia de haversine. Fuente: SABI, Versión software 72.00, actualización de datos 08/08/2017.

Las siguientes Tablas 3 y 4 muestran información sobre la distribución de las empresas aragonesas por actividades económicas. Para ello utilizamos la Clasificación Nacional de Actividades Económicas CNAE-2009. Si bien esta clasificación fue introducida en el año 2009 y puede plantear algún problema metodológico si se utilizan datos de periodos anteriores (trataremos esta cuestión más adelante), el análisis de la distribución espacial de empresas en Aragón de esta sección se refiere en todo momento a los datos del año 2017. Nótese que el número de empresas total en estas Tablas es ligeramente inferior al que se mostró anteriormente; la razón es que para algunas empresas SABI no ofrece la información exacta de su geolocalización (coordenadas latitud y longitud), lo que reduce el tamaño muestral inicial hasta 22.354 (el 88,6% del total).

La Tabla 3 muestra el número de empresas por Sección de actividad, que es el epígrafe más agregado de dicha clasificación, representando grandes grupos generales de actividades. Las secciones están ordenadas por orden alfabético. Se observa que la Sección G (Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas) es la actividad más importante en cuanto al número de empresas con un 22% del total de empresas, seguida por la Construcción (Sección F) y la industria manufacturera (Sección C). Estas tres actividades engloban casi la mitad del número total de empresas de Aragón (el 48,90%). Por detrás aparecen actividades del sector servicios que también cuentan con un importante número de empresas: Actividades profesionales, científicas y técnicas, Actividades inmobiliarias y Hostelería; secciones M, L e I, respectivamente. Por último, hay que destacar la importancia del sector primario ya que la Sección A (Agricultura, ganadería, silvicultura y pesca) cuenta con más de 1.000 empresas en todo Aragón. El resto de actividades presentan un número de empresas muy inferior.

La Tabla 4 nos ofrece una información más detallada, ya que contiene la distribución de empresas según las principales ramas de actividad, descendiendo hasta el nivel de cuatro dígitos de la clasificación. Consideramos un número mínimo de 200 empresas, y en esta ocasión las ramas de actividad están ordenadas según su importancia cuantitativa. De esta forma podemos observar que las ramas concretas de actividad más importantes en cuanto al número de empresas son las relacionadas con la actividad inmobiliaria y la construcción (Alquiler de bienes inmobiliarios por cuenta propia, Construcción de edificios residenciales o Promoción inmobiliaria), y en el resto de los casos nos encontramos mayoritariamente con actividades del sector servicios orientadas tanto al consumidor como a la empresa.

Tabla 3. Número de empresas por Sección de actividad, clasificación CNAE-2009

	Empresas	Porcentaje
SECCIÓN A: Agricultura, ganadería, silvicultura y pesca	1.378	6,16
SECCIÓN B: Industrias extractivas	95	0,42
SECCIÓN C: Industria manufacturera	2.632	11,77
SECCIÓN D: Suministro de energía eléctrica, gas, vapor y aire acondicionado	448	2,00
SECCIÓN E: Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación	66	0,30
SECCIÓN F: Construcción	3.376	15,10
SECCIÓN G: Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas	4.924	22,03
SECCIÓN H: Transporte y almacenamiento	887	3,97
SECCIÓN I: Hostelería	1.315	5,88
SECCIÓN J: Información y comunicaciones	465	2,08
SECCIÓN K: Actividades financieras y de seguros	529	2,37
SECCIÓN L: Actividades inmobiliarias	1.759	7,87
SECCIÓN M: Actividades profesionales, científicas y técnicas	2.053	9,18
SECCIÓN N: Actividades administrativas y servicios auxiliares	699	3,13
SECCIÓN P: Educación	330	1,48
SECCIÓN Q: Actividades sanitarias y de servicios sociales	494	2,21
SECCIÓN R: Actividades artísticas, recreativas y de entretenimiento	453	2,03
SECCIÓN S: Otros servicios	448	2,00
SECCIÓN T: Actividades de los hogares como empleadores de personal doméstico; actividades de los hogares como	1	0,00
SECCIÓN U: Actividades de organizaciones y organismos extraterritoriales	2	0,01
Total	22.354	100

Notas: Fuente: SABI, Versión software 72.00, actualización de datos 08/08/2017.

Tabla 4. Número de empresas por rama de actividad, clasificación CNAE-2009 (cuatro dígitos)

CNAE-2009	Literal CNAE-2009	Empresas	Porcentaje
6820	Alquiler de bienes inmobiliarios por cuenta propia	1.311	5,86
4121	Construcción de edificios residenciales	981	4,39
4110	Promoción inmobiliaria	719	3,22
4941	Transporte de mercancías por carretera	640	2,86
6920	Actividades de contabilidad, teneduría de libros, auditoría y asesoría fiscal	606	2,71
5610	Restaurantes y puestos de comidas	514	2,30
5630	Establecimientos de bebidas	430	1,92
4520	Mantenimiento y reparación de vehículos de motor	400	1,79
4321	Instalaciones eléctricas	393	1,76
4322	Fontanería, instalaciones de sistemas de calefacción y aire acondicionado	351	1,57
3519	Producción de energía eléctrica de otros tipos	331	1,48
7112	Servicios técnicos de ingeniería y otras actividades relacionadas con el asesoramiento	328	1,47
0146	Explotación de ganado porcino	306	1,37
7022	Otras actividades de consultoría de gestión empresarial	249	1,11
5510	Hoteles y alojamientos similares	217	0,97
6832	Gestión y administración de la propiedad inmobiliaria	210	0,94
4399	Otras actividades de construcción especializada n. c. o. p.	209	0,93
6622	Actividades de agentes y corredores de seguros	205	0,92
	Total	22.354	

Notas: Fuente: SABI, Versión software 72.00, actualización de datos 08/08/2017.

2.2.- Metodología

Se propone aplicar la metodología desarrollada por Duranton y Overman (2005, 2008), que ha sido aplicada a algunos sectores específicos con datos de empresas de toda España por Albert *et al.* (2012). Esta técnica permite medir la distancia de cada empresa al centro de su correspondiente grupo industrial, a partir de la estimación del grado de concentración de los puntos analizados mediante la función K de Ripley, que mide la aleatoriedad espacial y muestra cómo cambia la densidad (concentración) de las empresas con la distancia. La función K de Ripley se caracteriza principalmente por determinar los patrones de concentración en las diferentes ramas de empresas; permite, a través de una técnica estadística espacial, hacer comparación industrial, controlar la distribución desigual global, controlar la concentración de los establecimientos y, además, no presenta sesgo de agregación. Específicamente, la función permite determinar si las industrias, o los valores asociados a estas, exhiben una concentración o una dispersión estadísticamente significativa en un rango de distancias (Ripley, 1976).

En primer lugar, este método consiste en estimar la distribución de distancias bilaterales entre todas las empresas a partir de sus coordenadas geográficas (latitud y longitud). Definimos d_{ij} como la distancia entre las empresas i y j . Dadas n empresas de la misma rama de actividad, el estimador de la densidad de distancias bilaterales (denominada K-densidad) en cualquier punto (distancia) d es

$$\hat{K}(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=i+1}^n f\left(\frac{d-d_{ij}}{h}\right),$$

donde f es una función kernel Gaussiana con bandwidth (parámetro de alisado) h . Para simplificar el análisis, consideramos solo el rango de distancias entre cero y 65 kilómetros. Como sugieren Duranton y Overman (2005), este umbral coincide aproximadamente con la media y mediana de la distribución de distancias entre todas las empresas (alrededor de 63 kilómetros en ambos casos). No obstante, hemos repetido el análisis considerando toda la distribución de distancias hasta la distancia máxima (casi 300 kilómetros), y los resultados son similares; estos resultados pueden consultarse en el Apéndice A.

Esta densidad de empresas es no ponderada. Aunque Duranton y Overman (2005) consideran la posibilidad de ponderar las empresas por su número de empleados, en el caso aragonés no tendría sentido ya que la mayoría de las empresas en Aragón son pymes (alrededor del 95,1% tiene de 0 a 9 trabajadores) y, en particular, el 53,4% de ellas tienen cero empleados, lo que impediría la ponderación (véase la Tabla 1).

En segundo lugar, hay que contrastar si la distribución espacial de las empresas aragonesas obtenida exhibe algún patrón espacial significativo; esto es, si no es diferente de una distribución cualquiera aleatoria de empresas en el espacio. Para contrastar esta hipótesis, se pueden construir intervalos de confianza globales basados en la distribución que obtendríamos simulando una distribución aleatoria de empresas en el territorio aragonés.

Para construir estos contrafactuales partimos de la distribución real de todas las empresas. Consideramos que el conjunto de todas las ubicaciones existentes (S), es decir, las coordenadas geográficas de todas las empresas de Aragón, representa el conjunto de todas las ubicaciones posibles para cualquier empresa. No obstante, debido a limitaciones computacionales debemos limitar el conjunto S a una muestra aleatoria de 15.000 empresas (el 67% del total de las empresas para las que disponemos de información geográfica).

A partir de este conjunto S de ubicaciones posibles extraemos de manera aleatoria un número de localizaciones (coordenadas latitud y longitud) igual al número de empresas en la rama o sección de actividad considerada. Esto significa que, por ejemplo, BSH Electrodomesticos España SA (la segunda empresa más importante de Aragón según el número de empleados) podría estar situada en cualquier otro lugar de Aragón donde actualmente exista una empresa. Esta simulación se repite un número elevado de veces (Duranton y Overman (2005) recomiendan 2.000 simulaciones) y para cada simulación la densidad de distancias entre pares de empresas se calcula si el mismo número de empresas dentro de la rama o sección de actividad se asignó a través del conjunto S de todas las ubicaciones posibles. El muestreo se realiza sin reemplazo. Por lo tanto, para cualquiera de las ramas o secciones de actividad generamos nuestros contrafactuales para elementos de muestreo sin reemplazo, de modo que cada simulación es equivalente a una redistribución aleatoria de empresas en todos los sitios posibles.

Finalmente, para cada sector de actividad comparamos las estimaciones reales de la K-densidad con los contrafactuales simulados. Para analizar la significación estadística del patrón de localización de las empresas en comparación con la aleatoriedad, construimos bandas globales de confianza utilizando las distribuciones contrafactuales simuladas, siguiendo la metodología de Duranton y Overman (2005, 2008). Las desviaciones de la aleatoriedad implican un patrón de localización o dispersión si las K-densidades estimadas para un grupo de empresas se encuentran por encima o por debajo, respectivamente, de la banda global de confianza durante al menos una distancia d .

$\bar{K}(d)$ denota el límite superior de la banda global de confianza para las empresas dentro de una rama de actividad. Esta banda es alcanzada por el 5% de nuestras

simulaciones entre 0 y 65 kilómetros. De manera similar, el límite inferior de la banda global de confianza, que llamamos $\underline{K}(d)$, es tal que es alcanzado por el 5% de las K-densidades generadas de manera aleatoria que no están localizadas. Así, cuando las K-densidades estimadas se encuentran entre las dos bandas de confianza para una distancia d , $\bar{K}(d) > \hat{K}(d) > \underline{K}(d)$, la localización espacial de las empresas no es diferente de la aleatoriedad. Desviaciones de este patrón aleatorio implican un patrón de concentración (o localización, son sinónimos) si gráficamente las K-densidades estimadas se encuentran por encima de la banda superior para al menos una distancia d , esto es, cuando $\hat{K}(d) > \bar{K}(d)$. Análogamente, cuando gráficamente $\hat{K}(d) < \underline{K}(d)$ y las K-densidades caen por debajo del límite inferior de las bandas para una distancia d se puede observar un patrón espacial de dispersión.

Duranton y Overman (2005) proporcionan algunas directrices y ejemplos para comprender mejor lo que capturan estas estimaciones que fácilmente podemos adaptar al caso aragonés. Consideremos la rama de actividad con mayor número de empresas, Alquiler de bienes inmobiliarios por cuenta propia (código CNAE-2009 6820). Dado que esta actividad requiere de la existencia de un stock de viviendas para alquiler, es de esperar que a mayor población en el municipio mayor número de viviendas disponibles y por lo tanto mayor número de empresas que se dedicarán a esa actividad. Por lo tanto, sería de esperar que existieran importantes grupos de estas empresas en Zaragoza, Huesca y Teruel, las tres capitales de provincia, pero también en otros municipios con elevada población, como Ejea de los Caballeros o Cuarte de Huerva. Esto implicaría un clúster importante de estas empresas para distancias muy cortas, lo que en términos de nuestras K-densidades estimadas nos conduciría a una alta densidad de empresas para distancias cortas, mostrando así la localización (concentración) entre estas distancias.

Además, Zaragoza y otros municipios como Ejea de los Caballeros o Cuarte de Huerva están bastante cerca unos de otros (véase la distancia a Zaragoza en la Tabla 2), de modo que también podría existir una localización de empresas de esta rama de actividad para distancias a 9 o 56 kilómetros (las distancias desde Zaragoza a Cuarte de Huerva y Ejea de los Caballeros, respectivamente). Una multiplicidad de picos en nuestra densidad estimada indicaría entonces una multiplicidad de grupos de empresas cercanos entre sí. Pero nótese que, por el contrario, los picos de densidad que se generarían por las empresas de esta rama de actividad en Huesca y Teruel no aparecerían en nuestro análisis ya que Zaragoza y las otras dos capitales de provincia están a más de 65 kilómetros de distancia, el umbral que hemos considerado. Es por ello que, cuando consideramos toda la distribución de distancias hasta la distancia máxima en el Apéndice A se observan picos de densidad para algunas actividades a grandes distancias, que en algunos casos coinciden precisamente con la distancia

desde Zaragoza a otros centros de actividad importantes como Huesca, Teruel o Calatayud.

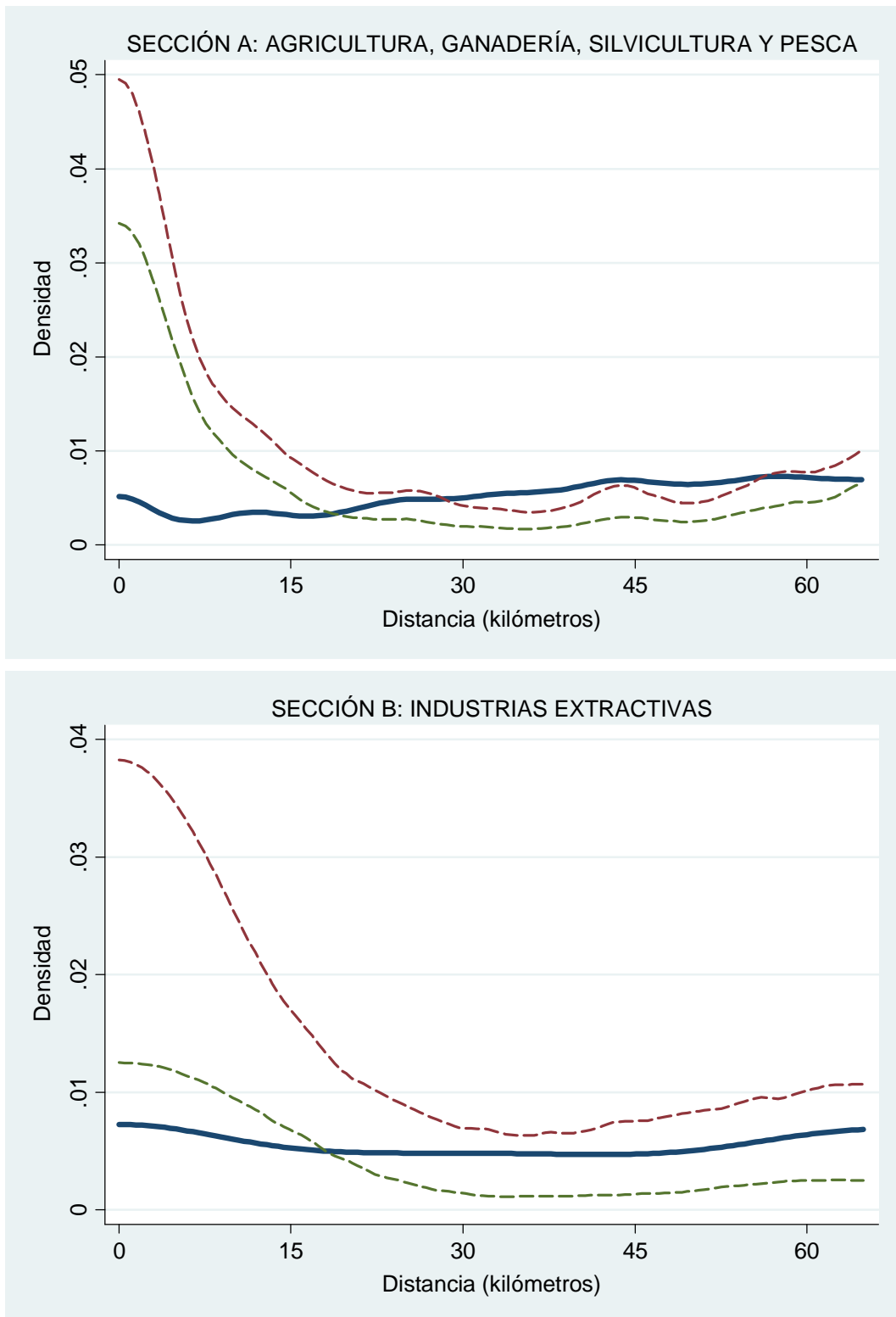
Duranton y Overman (2005) también señalan algunas limitaciones de su enfoque. En particular, no es posible detectar patrones no aleatorios si no implican concentración (localización) o dispersión. Por lo tanto, podríamos aceptar como aleatorio el patrón de algunas actividades cuya ubicación es claramente no aleatoria como, por ejemplo, las empresas ubicadas a lo largo de una costa o una línea ferroviaria. Obviamente en Aragón no existe costa, pero si tenemos otras características geográficas (ríos, montañas, etc.) que podrían influir en la distribución espacial de empresas. En este caso, la probabilidad de un error de este tipo disminuye con el número de empresas. Por ejemplo, una industria con plantas a lo largo de una línea recta se localiza en distancias cortas (el número de vecinos aumenta linealmente con la distancia, mientras que si la ubicación es aleatoria, el aumento es cuadrático) y esto debe detectarse siempre que haya suficientes empresas en la muestra.

Por último, Duranton y Overman (2005) destacan que este tipo de análisis se ocupa únicamente de los patrones espaciales en actividades específicas en relación con la actividad en general, y no sobre los patrones de especialización de determinadas economías locales. El análisis de la especialización es conceptualmente distinto del de la localización y, por tanto, requiere diferentes herramientas. En el apartado 3 de este estudio nos ocuparemos del análisis de los efectos de la especialización y diversidad productiva a nivel local, y es entonces cuando abordaremos la definición de las medidas relativas a ambas características industriales.

2.3.- Resultados

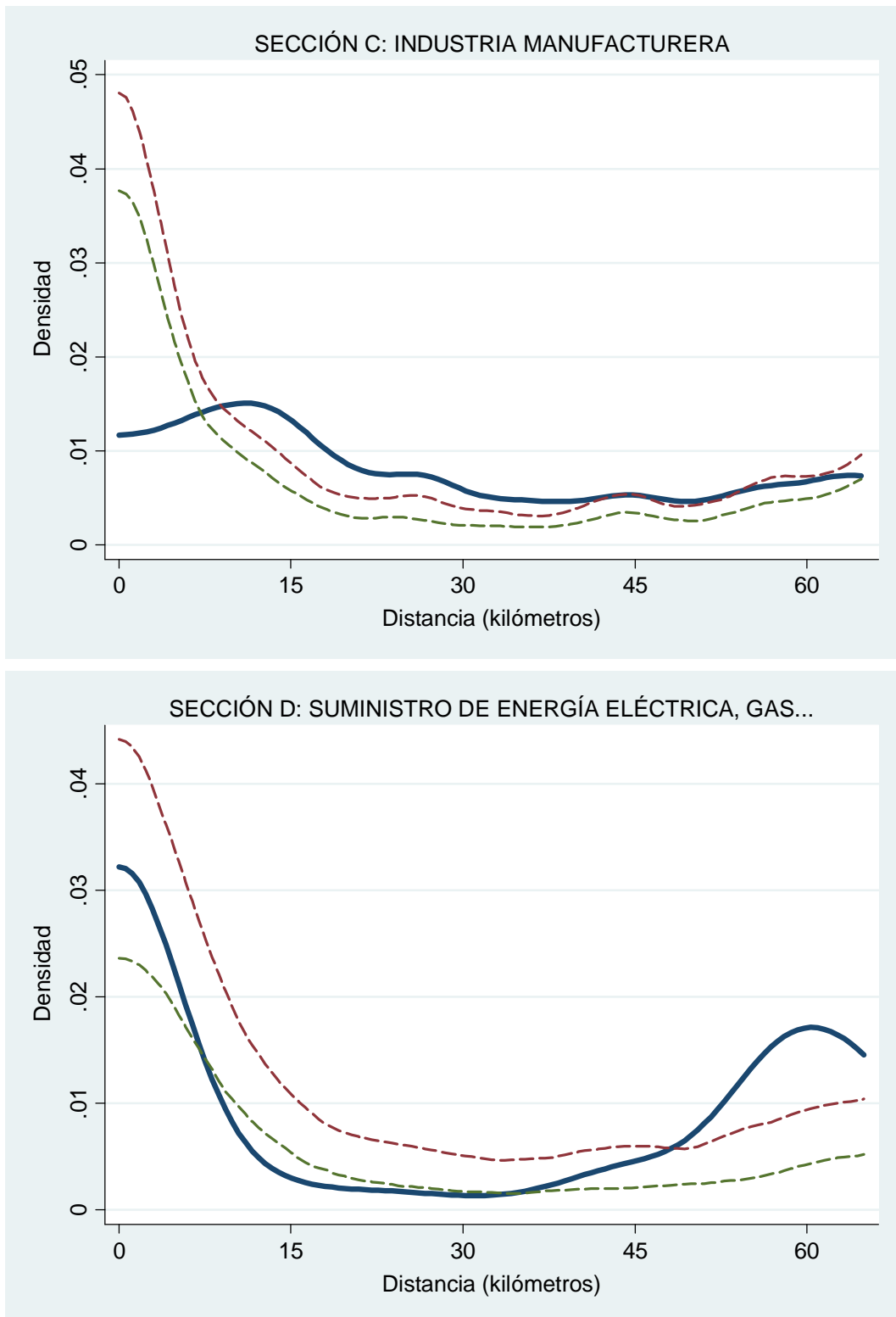
Las Figuras 1 y 2 muestran, respectivamente, los resultados de aplicar la metodología de Duranton y Overman (2005, 2008) a las secciones y principales ramas de actividad (aquellas que cuentan como mínimo con 200 empresas), según la clasificación CNAE-2009 (las Tablas 3 y 4 muestran los tamaños muestrales en cada caso). Como se ha indicado, la metodología permite obtener una estimación de la densidad (concentración) de empresas del mismo sector, representada por la línea continua, que varía con la distancia. Recuérdese además que las bandas de significatividad tienen una interpretación poco usual; las bandas discontinuas representan la distribución aleatoria de empresas en el espacio, por lo que cuando la densidad estimada cae fuera de esas bandas, indica que, para la distancia correspondiente, existe una concentración o dispersión significativa de empresas.

Figura 1. Distribución espacial de empresas por secciones de actividad



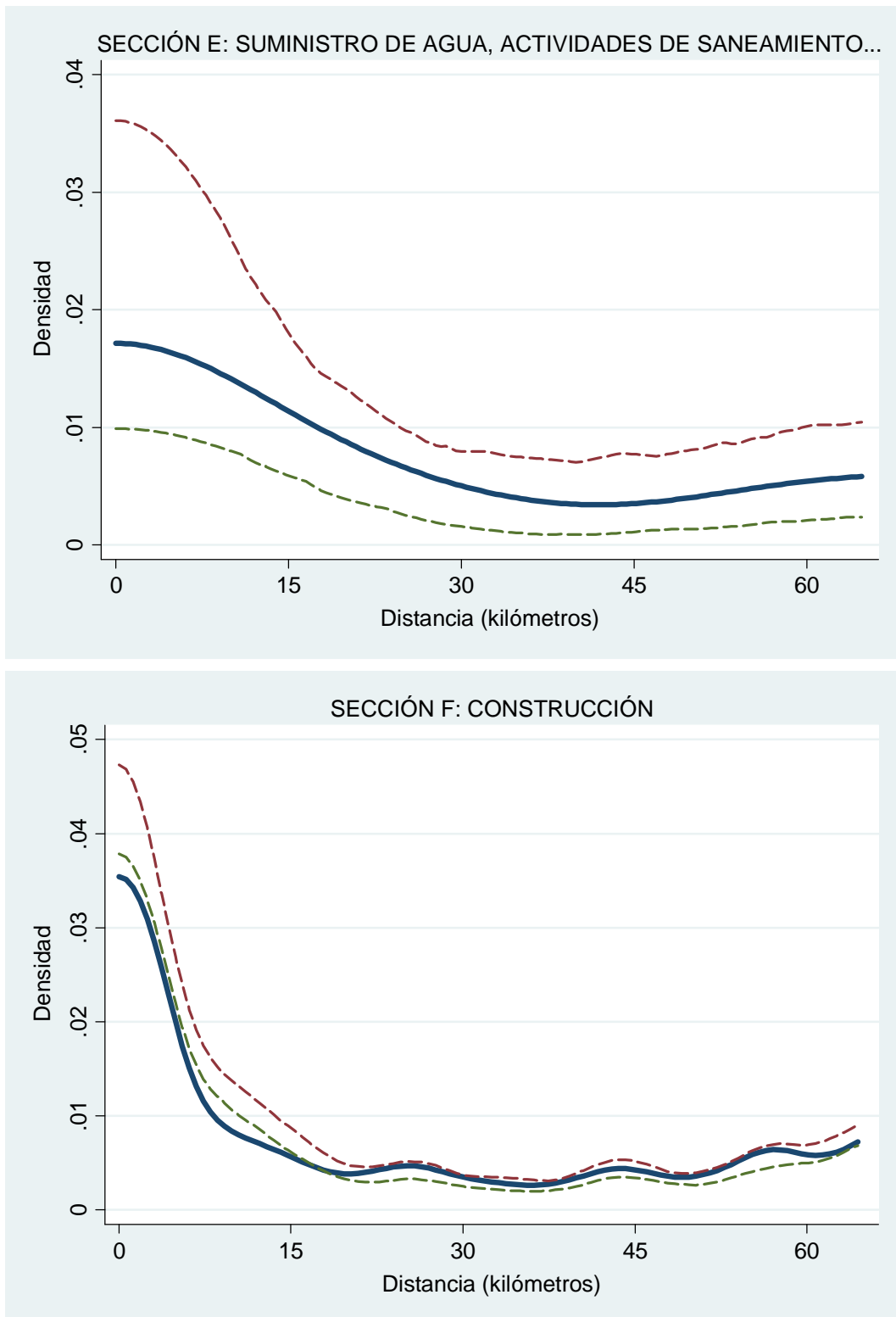
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



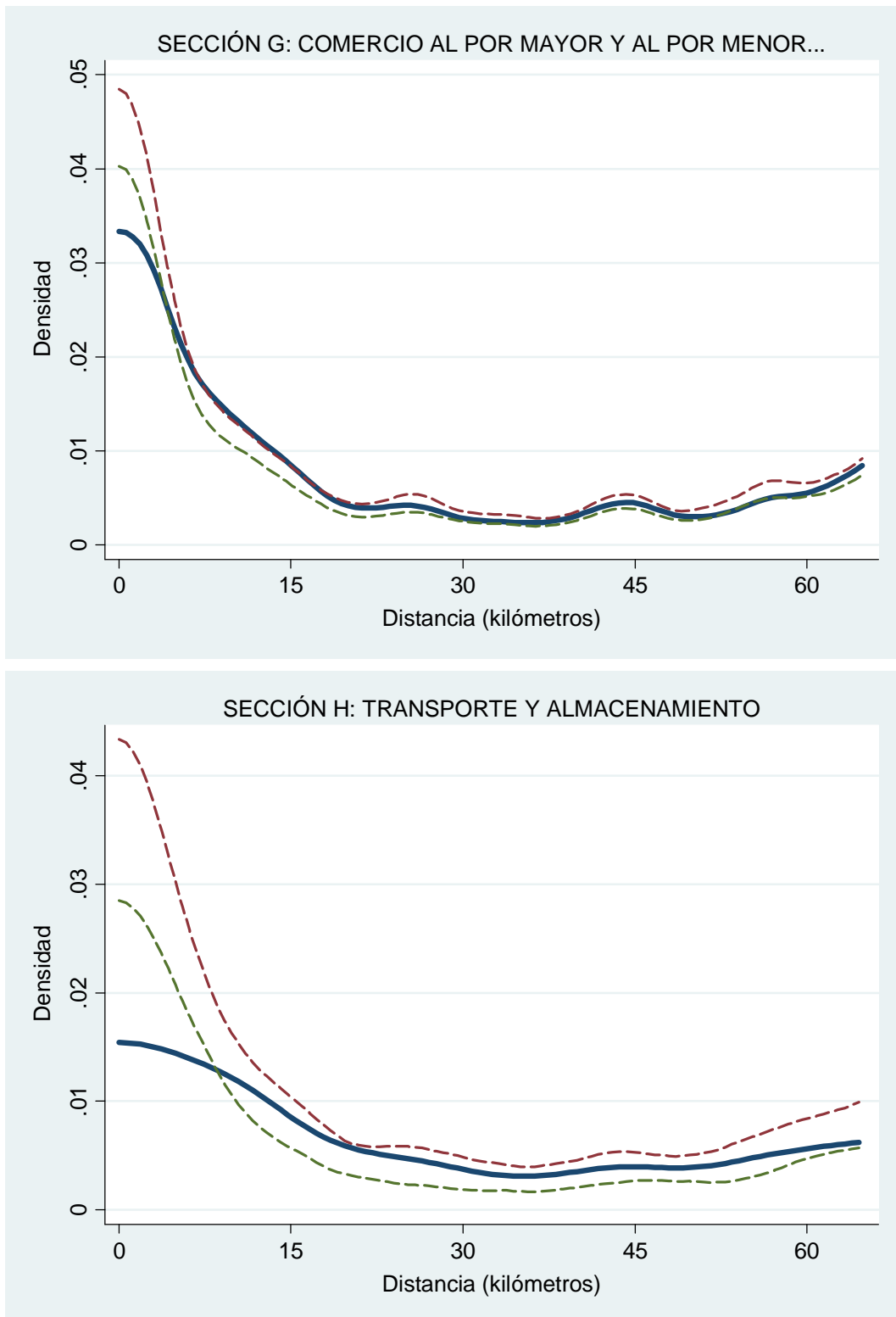
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



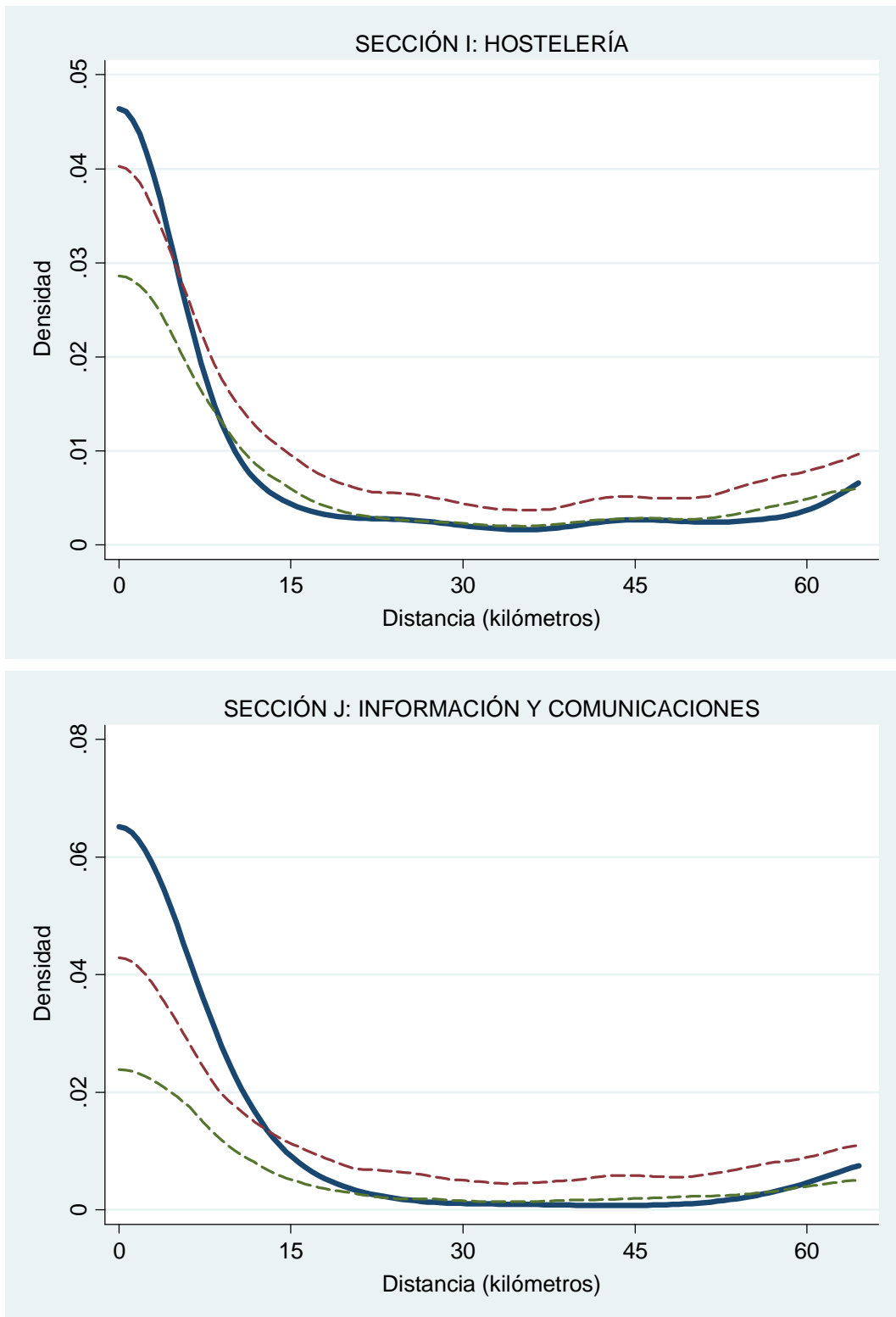
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



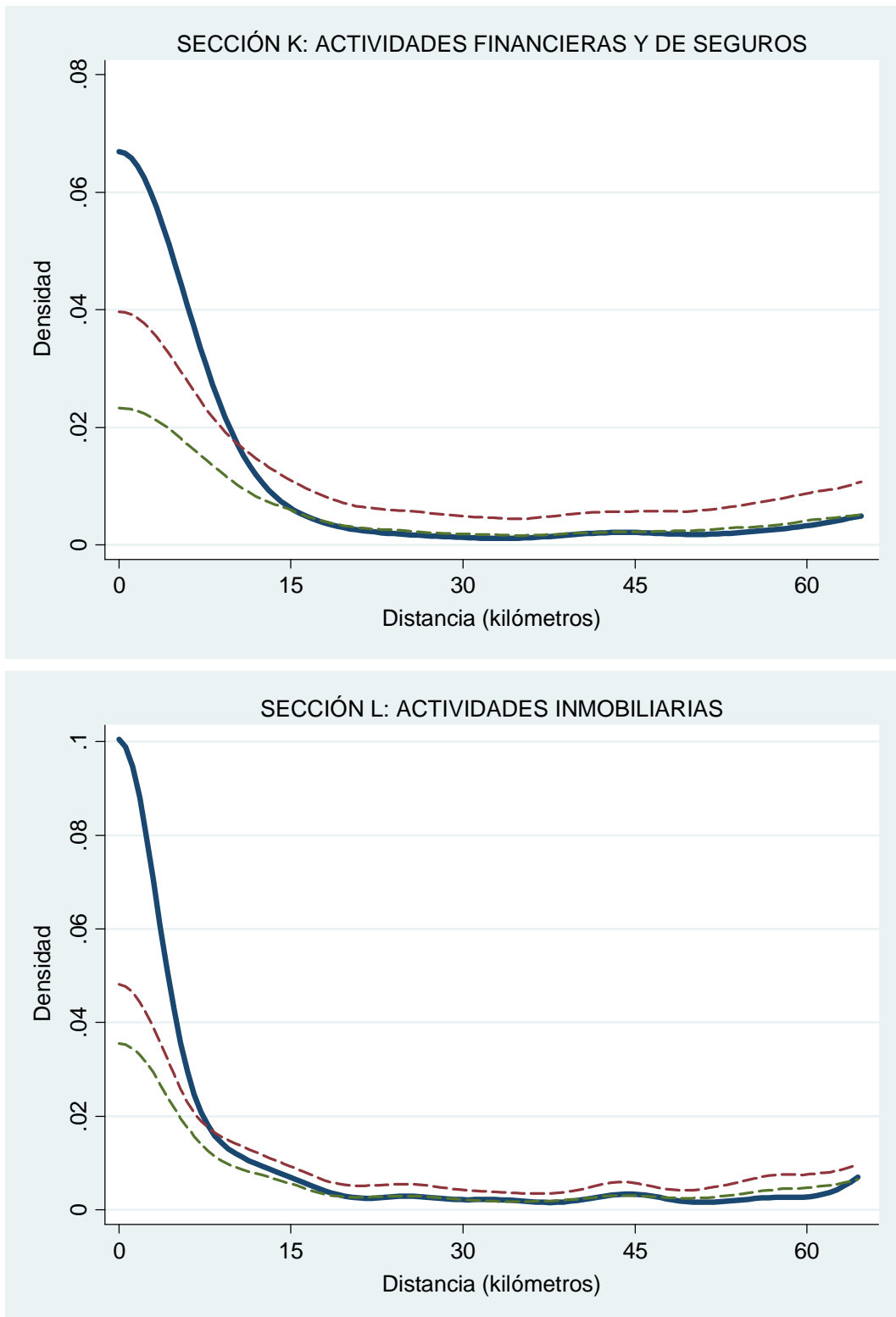
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



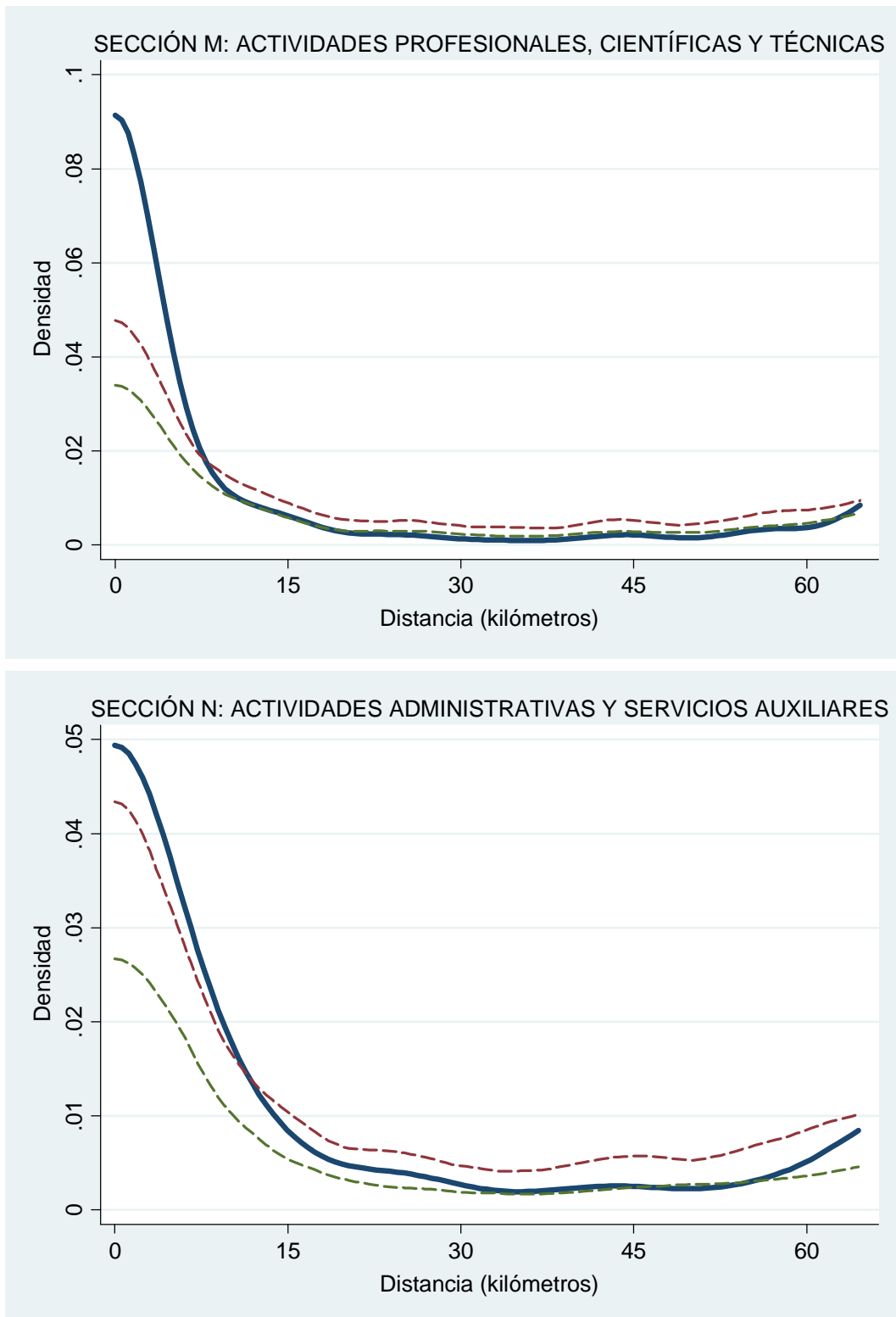
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



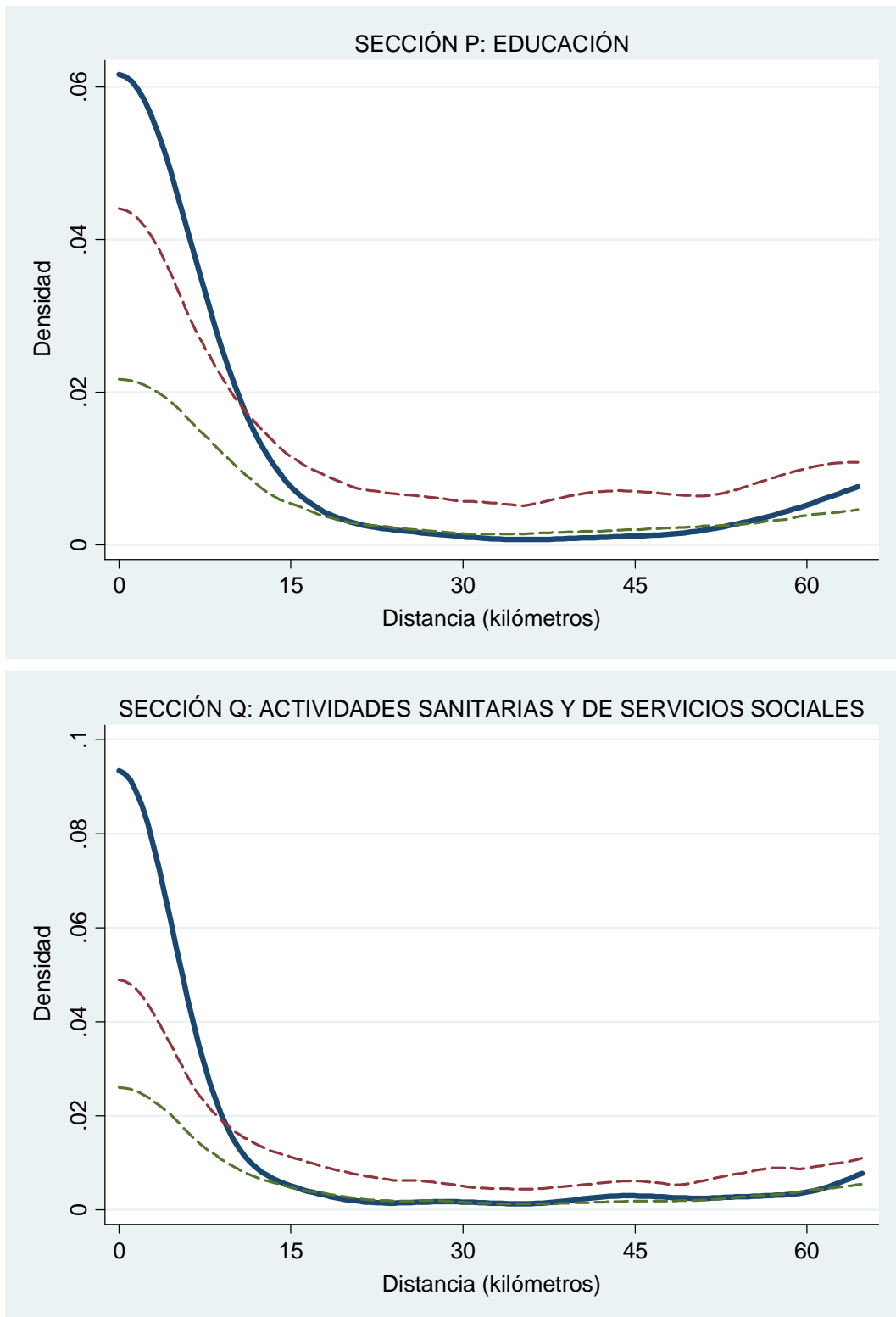
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



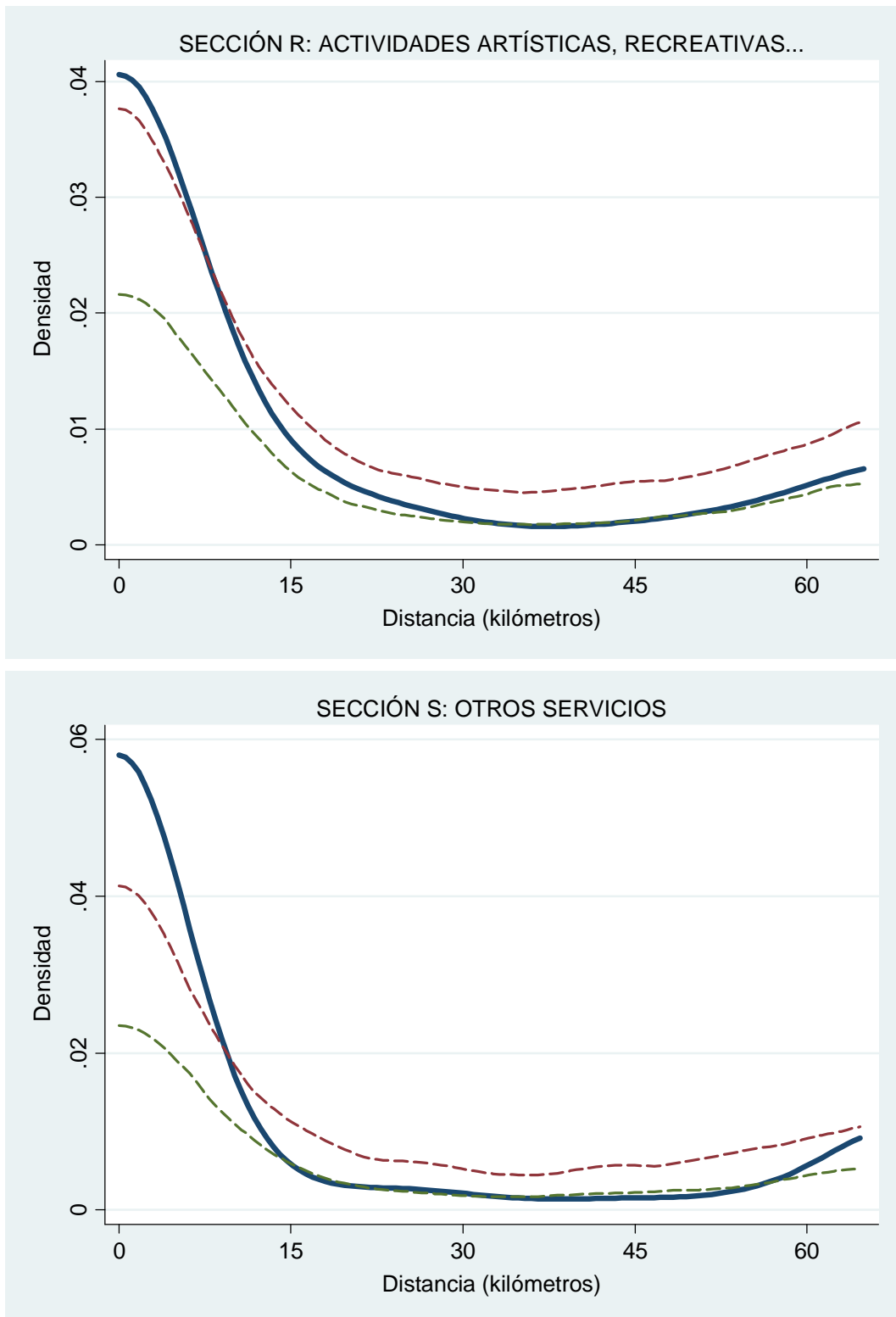
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



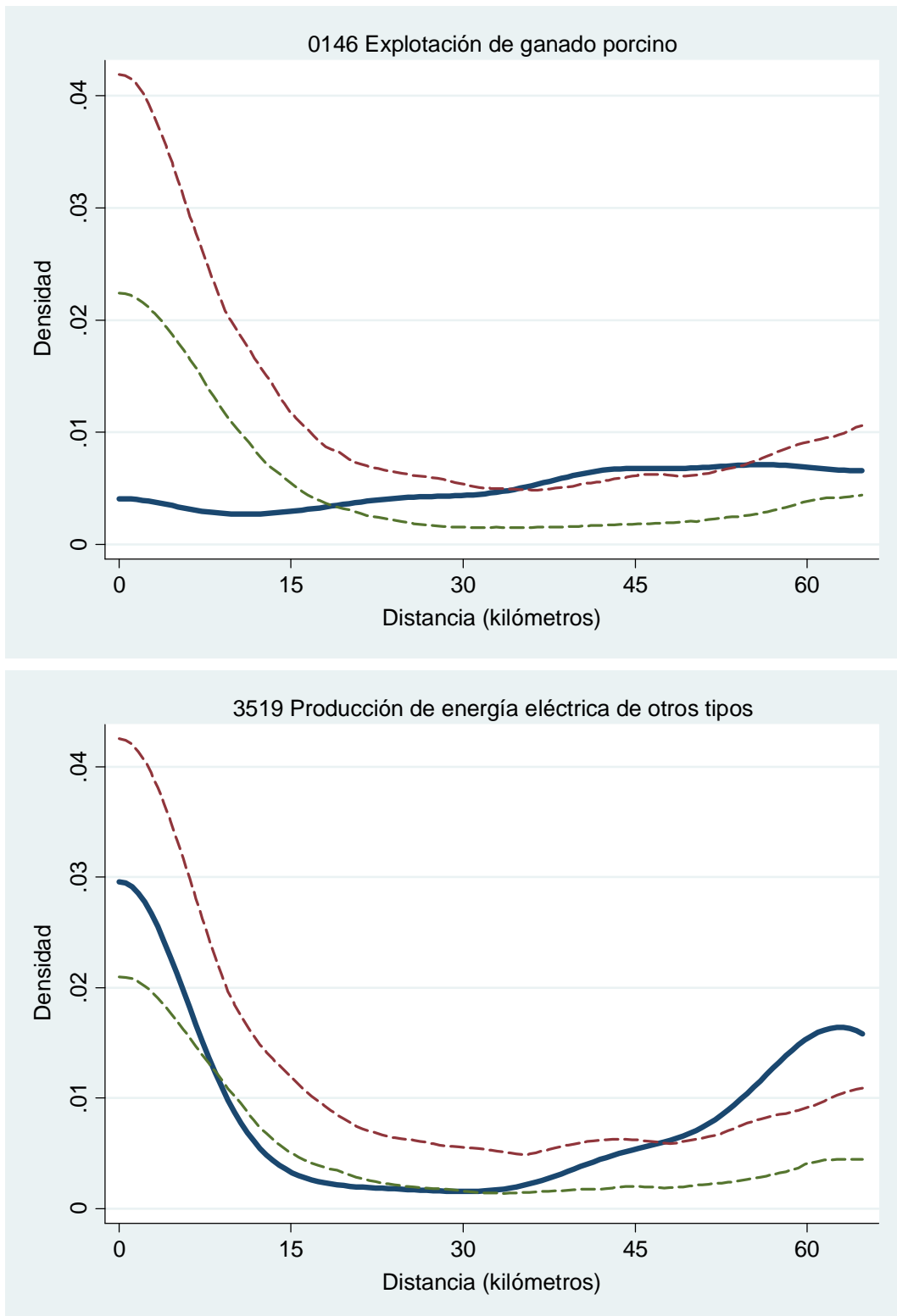
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 1. Distribución espacial de empresas por secciones de actividad (continuación)



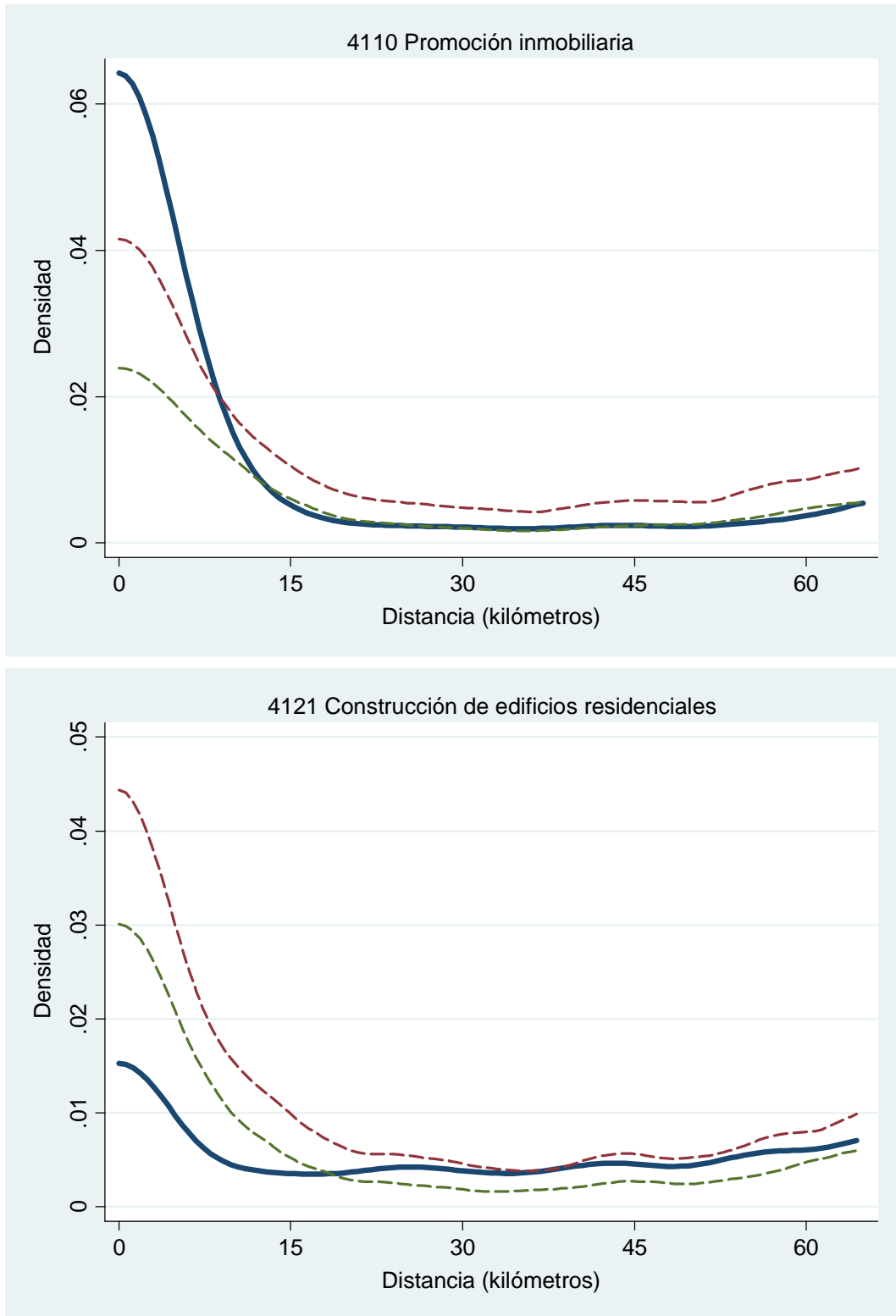
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas



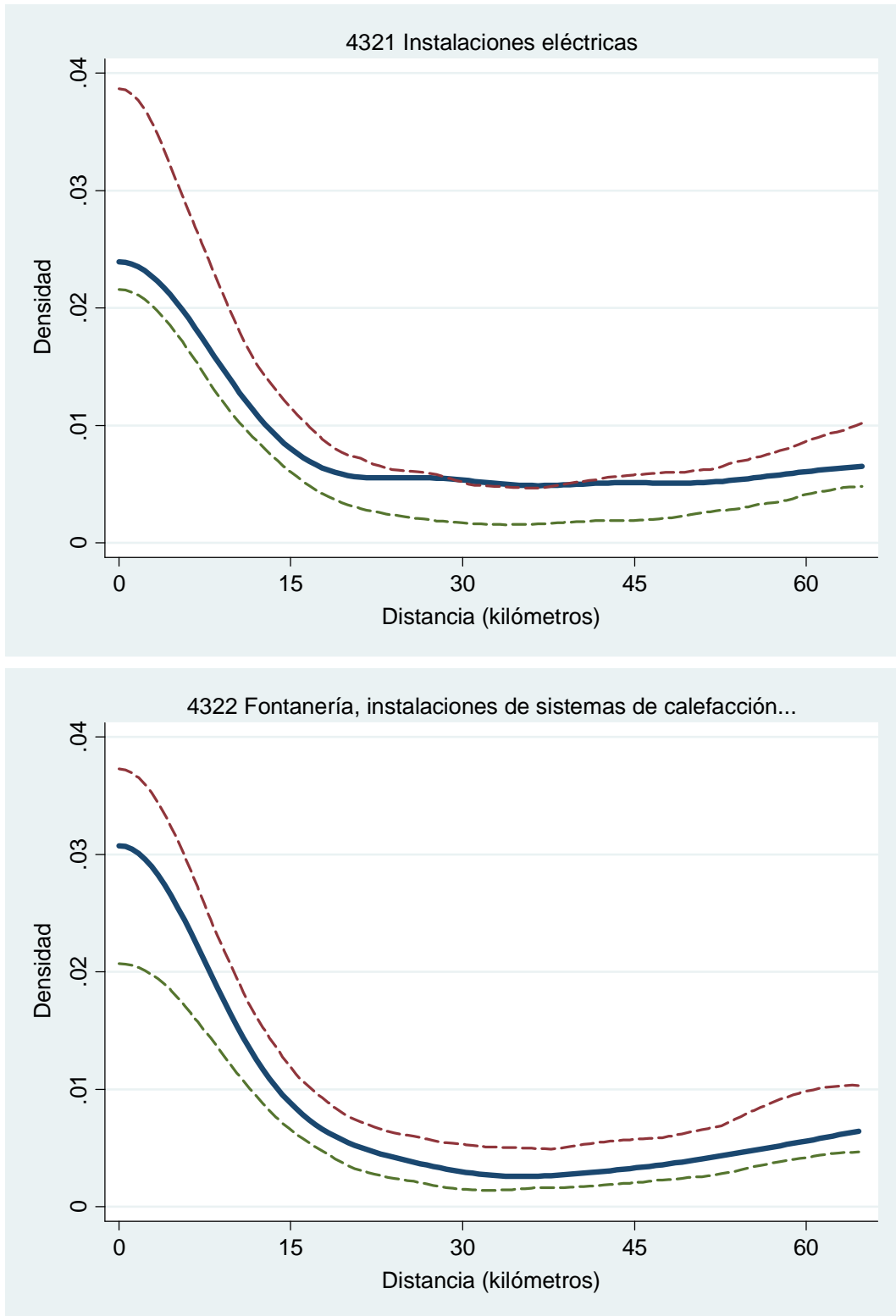
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



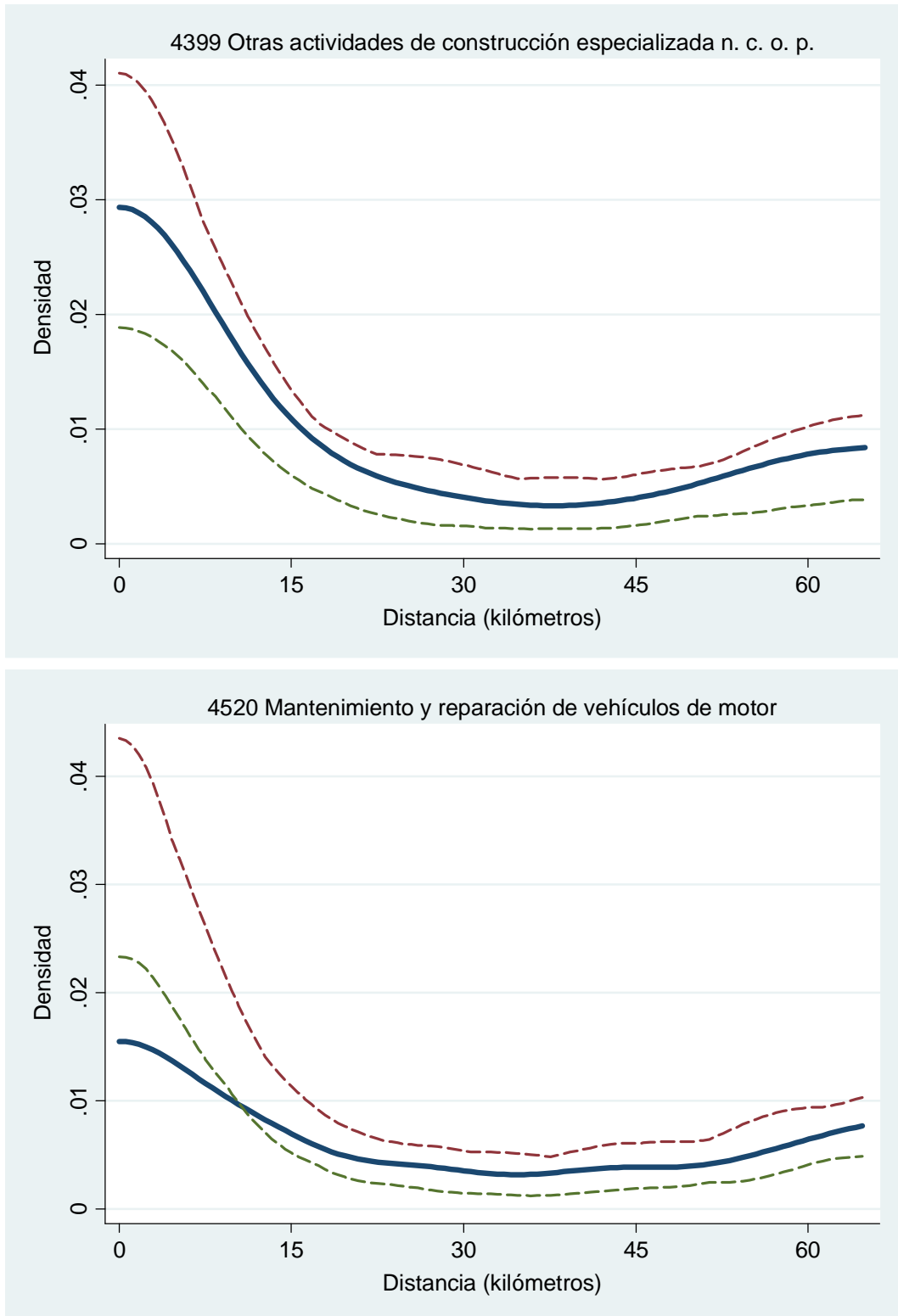
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



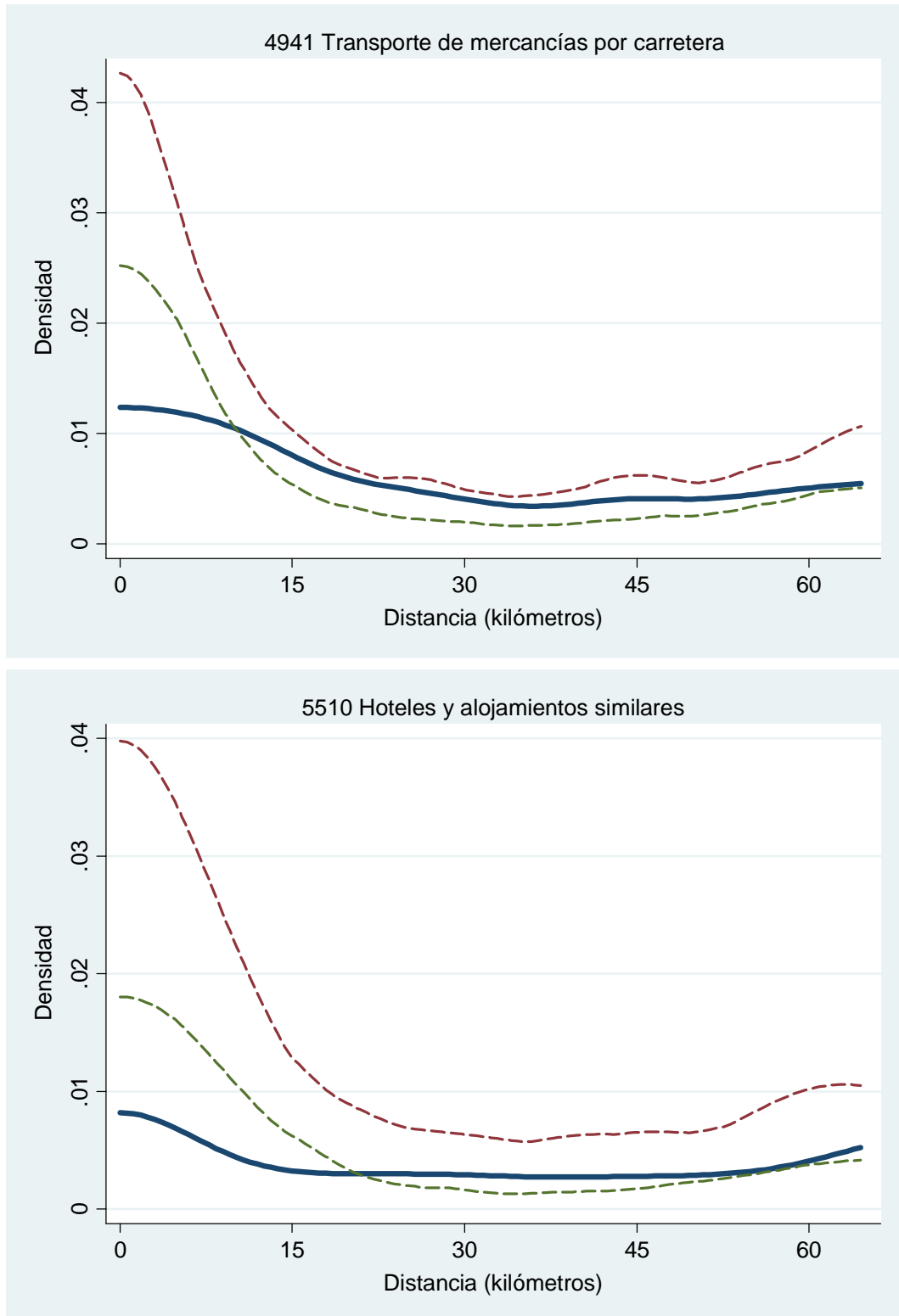
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



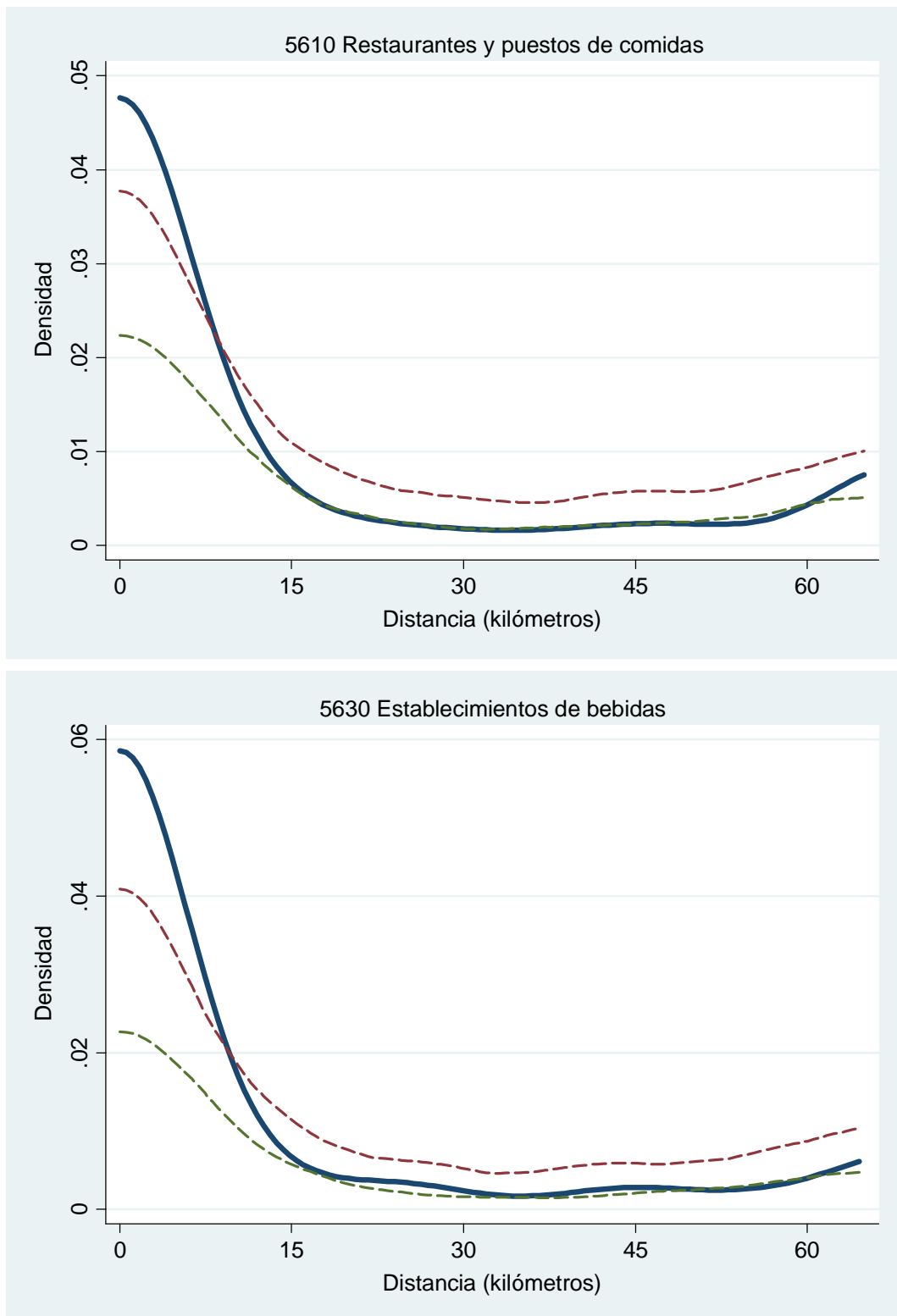
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



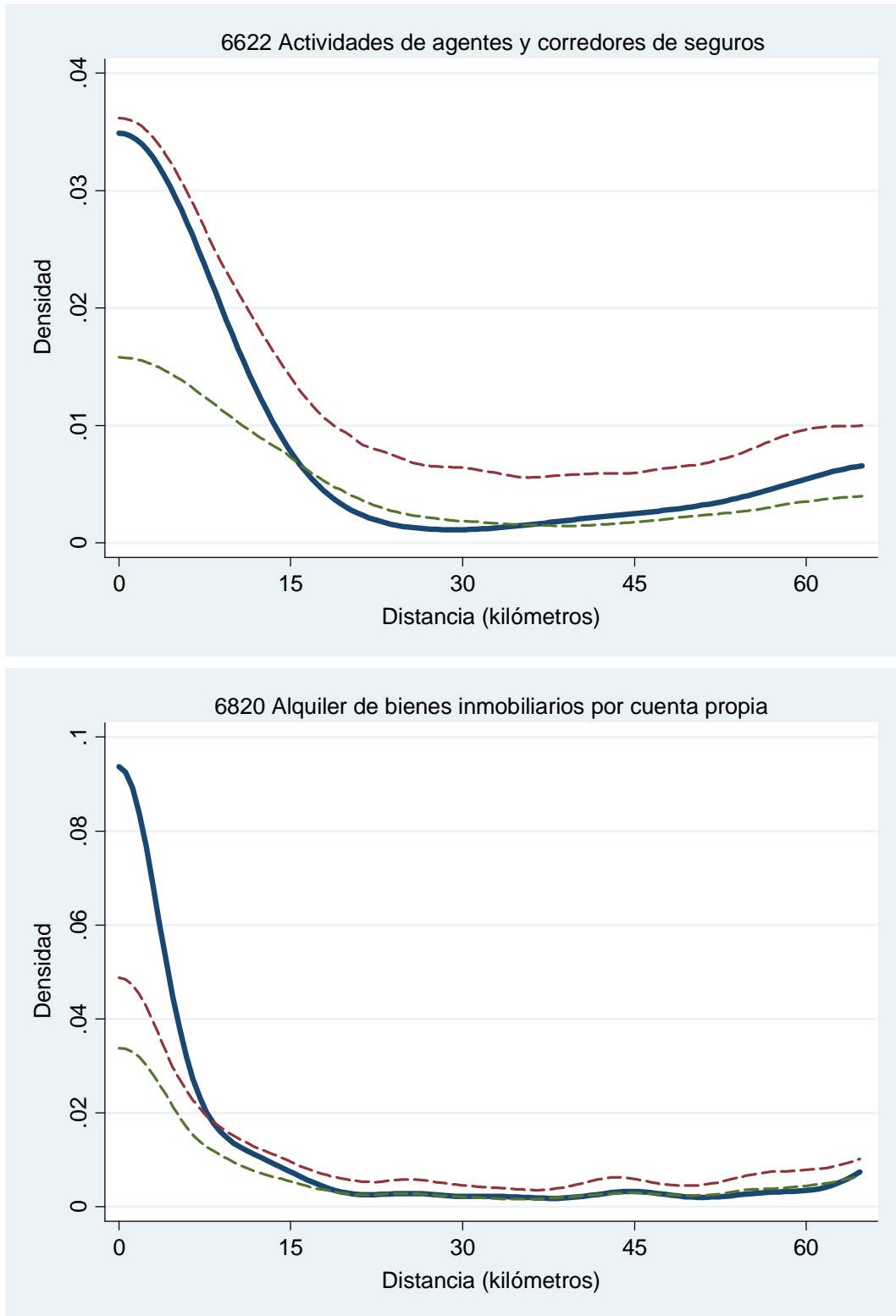
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



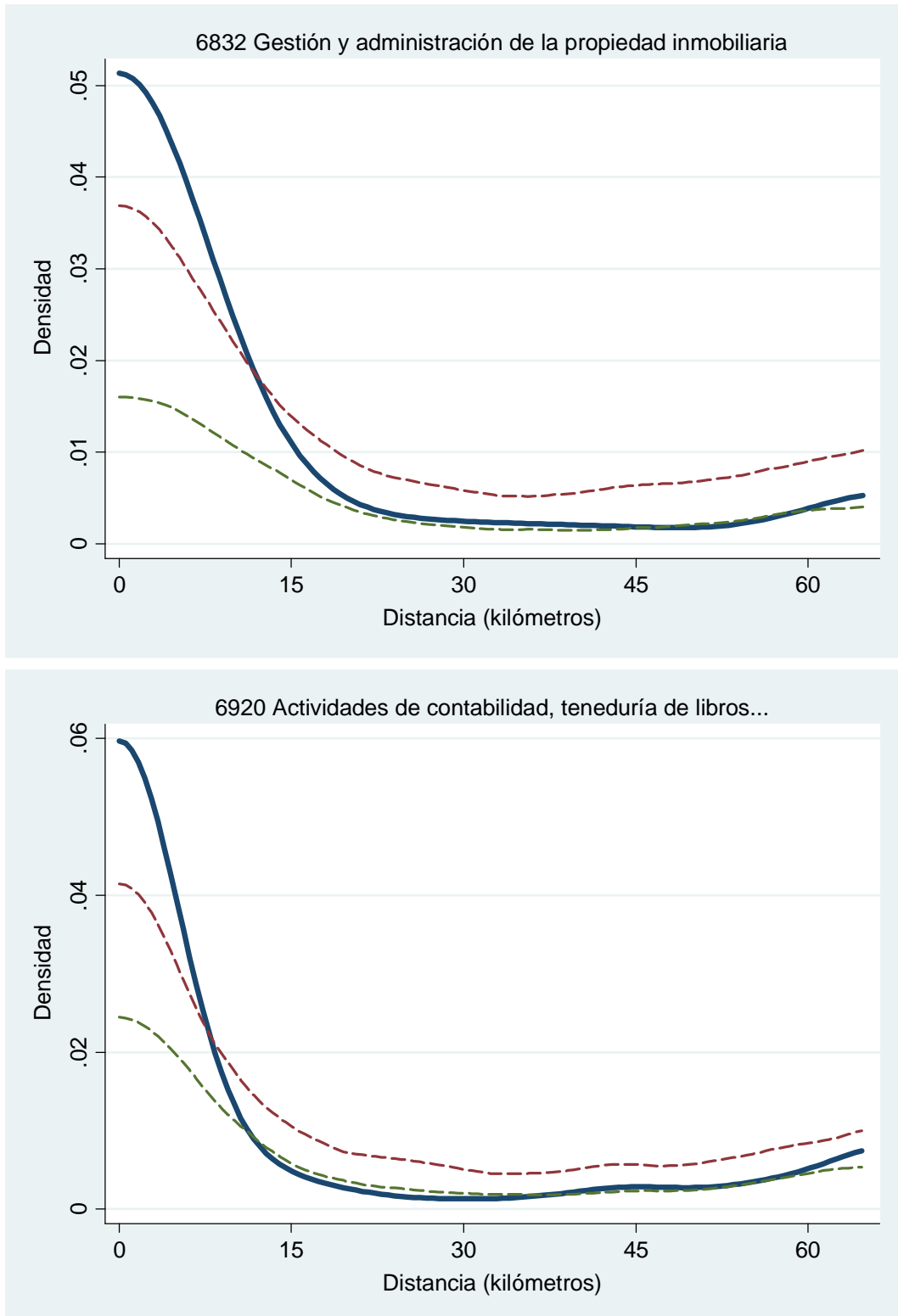
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



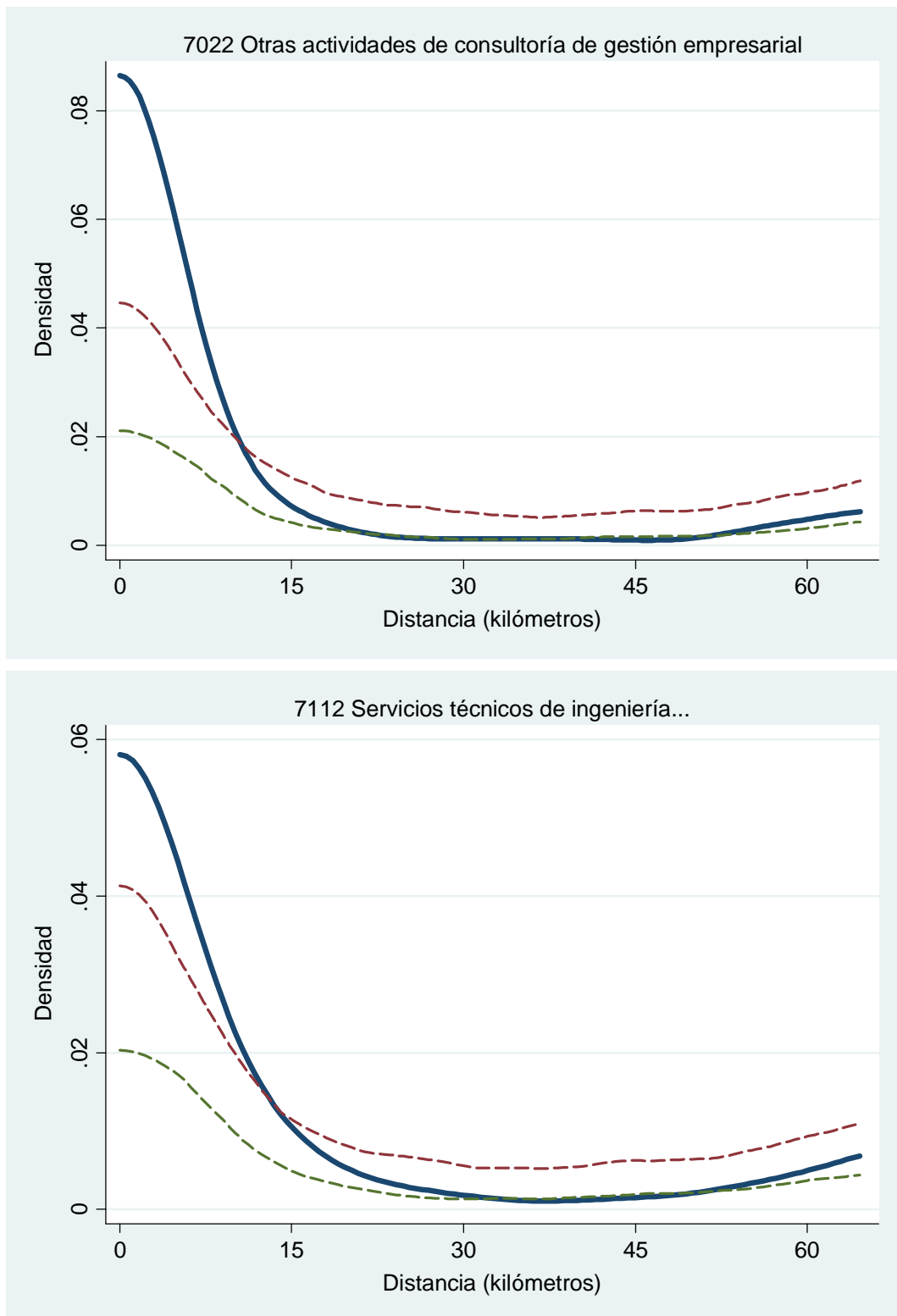
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura 2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Cabe destacar cómo en todos los casos los contrafactuales contruidos a partir de todas las ubicaciones posibles de las empresas, representados por las bandas de confianza, muestran una disminución de la concentración con la distancia, partiendo de una concentración inicial muy elevada de empresas en distancias muy cortas. Sin duda este patrón está muy influido por la elevada concentración de empresas en la ciudad de Zaragoza. Las figuras construidas utilizando todas las distancias posibles se muestran en el Apéndice A, y en ellas se pueden observar picos de densidad en la distribución de los contrafactuales en las distancias 70, 100 y 150 kilómetros, que vendrían a coincidir con las distancias entre Zaragoza y otros municipios con alta densidad de empresas, como Huesca, Calatayud, Alcañiz, Monzón, Fraga, Jaca o Teruel, véase la Tabla 2.

Podemos identificar tres tipos de patrones generales de localización de las empresas: (i) aquellas industrias que comienzan con una dispersión significativa en distancias cortas, pero que terminan exhibiendo un patrón aleatorio a medida que aumenta la distancia, (ii) aquellas industrias que muestran una concentración significativa para distancias cortas, pero que a medida que aumenta la distancia pasan a presentar un patrón no diferente de la aleatoriedad, y (iii) aquellas industrias cuyas K-densidades caen para casi todas las distancias entre las bandas y por lo tanto no muestran ningún patrón significativo de localización.

En el caso de las secciones de actividad, las Secciones A, B, F y H pertenecen al primer caso, las Secciones I, J, K, L, M, N, P, Q, R y S al segundo, y las Secciones D, E y G al tercero. No resulta difícil identificar los factores que generan esta clasificación. Las secciones que contienen las actividades del sector primario (agricultura), construcción y transporte y almacenamiento exhiben una dispersión significativa en distancias cortas (caso (i)), ya que se trata de actividades productivas que requieren de espacio físico (por ejemplo, tierras) para desarrollar su actividad, y por lo tanto la aglomeración resulta más difícil. Por el contrario, las secciones que presentan una localización significativa para distancias cortas (caso (ii)) son actividades del sector servicios, que por su propia naturaleza requieren menos espacio físico, facilitando la aglomeración, y además tienden a localizarse en las zonas más pobladas, cerca de su demanda. Por último, las secciones cuya distribución no es diferente de la aleatoriedad corresponden con industrias de suministro y abastecimiento de agua y energía.

El único caso especial que no podemos encajar en ninguno de los tipos definidos es la industria manufacturera (Sección C), que para distancias cortas inferiores a 8 kilómetros exhibe un patrón de dispersión espacial, ente 8 y 10 kilómetros la estimación de la densidad cruza las bandas, no pudiendo rechazarse la distribución aleatoria de empresas, pero desde 10 hasta 45 kilómetros el patrón es de concentración de empresas. Este resultado vendría dado por las particularidades de

Aragón, con una enorme cantidad de territorio que genera grandes distancias entre las grandes ciudades, donde suelen situarse las empresas industriales. Además, se observan varios picos de densidad a diferentes distancias (15, 25, 45 o 60 kilómetros) que, como se ha indicado anteriormente, pueden interpretarse como diferentes *clusters* de empresas industriales.

Los resultados obtenidos con la clasificación por ramas de actividad (4 dígitos) se muestran en la Figura 2. Si bien los resultados individuales de las ramas son consistentes con los obtenidos para las secciones de actividad, en este caso obtenemos información de aquellas ramas concretas que cuentan con más empresas (como mínimo 200, véase la Tabla 4) y que por su importancia cuantitativa pueden estar influenciando de forma significativa los patrones de toda la sección de actividad. Si mantenemos la misma casuística de tres patrones de distribución espacial generales, se observa que las empresas de las ramas de actividad 0146 (Explotación de ganado porcino), 4121 (Construcción de edificios residenciales), 4520 (Mantenimiento y reparación de vehículos de motor), 4941 (Transporte de mercancías por carretera) y 5510 (Hoteles y alojamientos similares) comienzan exhibiendo una dispersión significativa en distancias cortas (hasta los 15 kilómetros, aproximadamente) que termina cambiando a un patrón de distribución aleatoria a medida que aumenta la distancia, perteneciendo por tanto al caso (i); las ramas de actividad 4110 (Promoción inmobiliaria), 5610 (Restaurantes y puestos de comidas), 5630 (Establecimientos de bebidas), 6820 (Alquiler de bienes inmobiliarios por cuenta propia), 6832 (Gestión y administración de la propiedad inmobiliaria), 6920 (Actividades de contabilidad, teneduría de libros, auditoría y asesoría fiscal), 7022 (Otras actividades de consultoría de gestión empresarial) y 7112 (Servicios técnicos de ingeniería y otras actividades relacionadas con el asesoramiento técnico) se clasificarían dentro del patrón tipo (ii), ya que muestran una concentración significativa para distancias cortas (hasta 10 kilómetros, aproximadamente), pero después pasan a presentar un patrón no diferente de la aleatoriedad a medida que la distancia aumenta la distancia; y, finalmente, las empresas de las ramas 3519 (Producción de energía eléctrica de otros tipos), 4321 (Instalaciones eléctricas), 4322 (Fontanería, instalaciones de sistemas de calefacción y aire acondicionado), 4399 (Otras actividades de construcción especializada n. c. o. p.) y 6622 (Actividades de agentes y corredores de seguros) se incluirían en el caso (iii), con un patrón de distribución espacial aleatorio.

De nuevo, aquellas empresas que muestran más dispersión pertenecen a ramas de actividad de los sectores agrícola o construcción (explotación de ganado porcino, construcción de edificios residenciales, mantenimiento y reparación de vehículos de motor, etc.), mientras que aquellas actividades en las que se genera más aglomeración de empresas se pueden clasificar en actividades relacionadas con el sector inmobiliario (alquiler de bienes inmobiliarios por cuenta propia, promoción inmobiliaria, gestión y

administración de la propiedad inmobiliaria), servicios de consumo (restaurantes y puestos de comidas, establecimientos de bebidas) y actividades de asesoramiento a otras empresas (actividades de contabilidad, teneduría de libros, auditoría y asesoría fiscal, otras actividades de consultoría de gestión empresarial, servicios técnicos de ingeniería y otras actividades relacionadas con el asesoramiento técnico). Finalmente, la mayoría de las ramas de actividad que presentan un patrón espacial aleatorio corresponden a industrias de abastecimiento de algunas materias primas (instalaciones eléctricas, producción de energía eléctrica de otros tipos, fontanería, instalaciones de sistemas de calefacción y aire acondicionado).

2.4.- Conclusiones

El objetivo de este apartado es analizar la concentración espacial de las empresas aragonesas para las principales ramas de actividad. Si bien tradicionalmente en la literatura esta cuestión se suele tratar mediante el cálculo de algunos índices de concentración espacial (índice de Gini espacial o índice de Ellison-Glaeser, por ejemplo), en este estudio se propone aplicar la metodología desarrollada por Duranton y Overman (2005, 2008) que realiza una estimación continua de la concentración de empresas que varía con la distancia. Para ello consideramos las secciones y principales ramas de actividad (aquellas que cuentan como mínimo con 200 empresas) según la clasificación CNAE-2009. La información geográfica de la localización de las empresas proviene de la base de datos SABI (Sistema de Análisis de Balances Ibéricos), que contiene información geográfica de más de 22.000 empresas aragonesas en el año 2017.

Los resultados obtenidos permiten clasificar las secciones o principales ramas de actividad en tres tipos de distribución espacial, según la distribución espacial de sus empresas: (i) aquellas industrias que comienzan con una dispersión significativa en distancias cortas, pero que terminan exhibiendo un patrón aleatorio a medida que aumenta la distancia, (ii) aquellas industrias que muestran una concentración significativa para distancias cortas, pero que a medida que aumenta la distancia pasan a presentar un patrón no diferente de la aleatoriedad, y (iii) aquellas industrias cuyas K-densidades caen para casi todas las distancias entre las bandas y por lo tanto no muestran ningún patrón significativo de localización.

En el primer caso nos encontramos con actividades del sector primario (agricultura), construcción y transporte y almacenamiento. El segundo caso corresponde a las actividades del sector servicios, y en el tercer caso encontramos industrias de suministro y abastecimiento de agua y energía. Ambos tipos de patrones (dispersión o concentración) operan en distancias muy cortas (10–15 kilómetros), pero para

distancias más elevadas el resultado que suele emerger es una distribución de empresas espacial no diferente de la aleatoriedad.

La industria manufacturera es la única actividad económica que no encaja en ninguno de los tipos de esta clasificación. En distancias cortas inferiores a 8 kilómetros muestra un patrón de dispersión espacial, pero para distancias amplias (desde 10 hasta 45 kilómetros) el patrón es de localización (concentración) de empresas. Una posible explicación para este resultado sería la gran extensión de territorio disponible en Aragón, que genera amplias distancias entre las empresas industriales. También se observan diferentes picos de densidad para varias distancias (15, 25, 45 o 60 kilómetros) que pueden interpretarse como diferentes *clusters* de empresas industriales.

El que las empresas del sector industrial presenten dispersión espacial en distancias cortas mientras que aquellas que exhiben un patrón de concentración más fuerte pertenezcan al sector servicios podría cuestionar la fuerza de las economías de aglomeración en Aragón, especialmente cuando la mayoría de estas empresas cuentan con un número de empleados muy reducido. En este contexto, es necesario comprobar si esta distribución de empresas (especialmente en los sectores industriales) es capaz de generar efectos significativos sobre el empleo o la población en la economía aragonesa, cuestión que abordamos en el siguiente apartado.

3. Economías de aglomeración en Aragón

3.1.- Economías de aglomeración y crecimiento del empleo

Como se ha indicado en la Introducción del apartado 1, la literatura distingue dos tipos básicos de economías de aglomeración: las economías de localización, relacionadas con la concentración empresarial dentro de un sector específico, y las economías de urbanización, que operan a través de la concentración general de la actividad económica. Los beneficios de la localización y urbanización son denominadas en conjunto economías externas, debido a que surgen fuera de la empresa. Los fundamentos teóricos de las economías de aglomeración se remontan al concepto de los distritos industriales de Marshall (1890), que observó varias ventajas derivadas de la concentración de empresas, tales como (i) la disponibilidad de mano de obra especializada y (ii) bienes intermedios, y (iii) la facilidad para intercambiar conocimientos sobre productos, procesos e innovaciones. Otros modelos teóricos más recientes desarrollan y actualizan el concepto de economías de aglomeración para adaptarlo al momento actual, como el modelo teórico con micro-fundamentos de Duranton y Puga (2004), que les permite identificar tres mecanismos de transmisión de los efectos de la densidad de empresas: *sharing* (compartir), *matching* (compatibilizar) y *learning* (aprender).

Una vez analizada previamente la distribución espacial de las empresas aragonesas, en esta segunda parte del estudio nos centraremos en los beneficios de las economías de escala externas o economías de aglomeración que dicha concentración sectorial puede generar. A nivel teórico y empírico el foco suele estar centrado en la industria (manufacturas), ya que es el sector de mayor interés por su relación con el crecimiento económico y su capacidad de arrastre del resto de la economía. No obstante, nosotros consideraremos también los servicios que, como hemos analizado anteriormente, se encuentran más concentrados espacialmente (debido a que, por su propia naturaleza, tienen un alto potencial de codependencia y aglomeración, y buscan situarse cerca de su demanda), y la construcción que, hasta el comienzo de la crisis económica en 2008, era uno de los principales motores del crecimiento económico. Según el Informe socioeconómico de la década 2001-2010 en Aragón (2012), la construcción presentó en Aragón hasta 2007 ganancias notables de participación en la producción total, para pasar a importantes pérdidas tras dicho año. El sector de la construcción en Aragón fue un protagonista clave en el largo proceso de expansión económica que se prolongó hasta 2007, incluso más que en España, pero también fue el sector más afectado por la crisis, junto con la industria. Finalmente, dada la importancia del sector primario en la economía aragonesa, también estudiaremos si la concentración de las empresas de este sector genera algún efecto significativo sobre el empleo.

En esta línea, los trabajos pioneros corresponden a Glaeser *et al.* (1992) y Henderson *et al.* (1995), quienes desde una perspectiva dinámica estiman el impacto de un gran número de determinantes locales en el crecimiento del empleo a nivel de industria y ciudad para Estados Unidos. Más recientemente, Combes (2000) establece para Francia que el tamaño del mercado local tiene un efecto positivo en el crecimiento del empleo industrial para las industrias manufactureras pero negativo para los servicios. Viladecans-Marsal (2004) encuentra para España que el efecto en el empleo industrial no es significativo en tres de los seis sectores analizados. Para Alemania, Blien *et al.* (2006) determina que el tamaño del mercado local tiene un impacto positivo en el crecimiento del empleo industrial y en el sector de los servicios. Para Italia, Mameli *et al.* (2008) encuentra que las economías de localización tienen un efecto negativo, y las economías de urbanización un efecto positivo en el crecimiento del empleo; sin embargo sus resultados son sensibles a la desagregación sectorial utilizada. Finalmente, Marrocu *et al.* (2013) encuentran efectos positivos de la diversidad empresarial, particularmente perceptibles en el sector de servicios en las áreas europeas urbanas viejas, mientras que la especialización afecta todavía positivamente los sectores manufactureros de baja tecnología en la nueva Europa.

3.1.1.- Metodología

En este estudio replicamos el trabajo de Combes (2000) sobre las economías de aglomeración en Francia, utilizando datos de empresas y empleo en Aragón. Por tanto, la estrategia empírica y la selección de variables es similar a la de Combes (2000).

Una primera cuestión empírica es cómo cuantificar las economías de aglomeración a nivel local. Nuestro objetivo es examinar si los efectos externos que genera la concentración sectorial son importantes para el crecimiento del empleo. La forma más clara de establecer tales efectos es observando el crecimiento de un mismo sector en diferentes municipios y analizando en cuáles de estos esos sectores crecen más rápidamente. Por tanto, la unidad de observación es cada sector en cada municipio, por lo que definimos la variable a explicar como el crecimiento relativo del empleo:

$$y_{ict} = \ln\left(\frac{emp_{ict}}{emp_{ict-1}}\right) - \ln\left(\frac{emp_{it}}{emp_{it-1}}\right), \quad (1)$$

donde emp_{ict} es el empleo en el sector i en el municipio c en el momento t y emp_{it} es el empleo total de ese sector en Aragón en t . El periodo temporal considerado comienza en el año 2000 y termina en 2015; se trata por lo tanto de un panel de datos anuales desde 2000 a 2015. Como explica Combes (2000), la elección de esta variable significa que no pretendemos explicar por qué el crecimiento de un sector en un municipio determinado es $x\%$, sino por qué es $y\%$ más alto o más bajo en este lugar en comparación con el nivel de ese sector en todo Aragón.

Una vez definida la variable a explicar, procedemos a establecer los índices que nos permitirán medir las economías de localización (especialización productiva) y las economías de urbanización (diversidad productiva) a nivel de sector y municipio. La medida de especialización (esp_{ict}) relacionada con las economías de localización es la siguiente:

$$esp_{ict} = \frac{emp_{ict}/emp_{ct}}{emp_{it}/emp_t}, \quad (2)$$

siendo emp_{ict} el empleo en el sector i en el municipio c , emp_{ct} el empleo total municipal, emp_{it} es el empleo total de ese sector en Aragón, y emp_t el empleo total en Aragón, todos ellos medidos en el año t . Se trata del cociente entre la cuota (o porcentaje) que representa el empleo en el sector i en el municipio c y el porcentaje que representa el empleo en ese mismo sector en el total de la economía aragonesa. Así, si el municipio está más especializado en ese sector concreto que la economía aragonesa en su conjunto, el índice tomará valores superiores a la unidad. Por el contrario, si la cuota de empleo en ese sector en el municipio considerado es inferior al peso total que el sector tiene en la economía aragonesa el índice tendrá un valor inferior a la unidad.

Respecto a la diversidad productiva (div_{ict}), vinculada a las economías de urbanización, habitualmente (Combes, 2000; Henderson *et al.*, 1995) se mide a través del índice de Herfindahl inverso construido a partir de la participación (cuota) de los diferentes sectores dentro del empleo local, exceptuando el sector que se considera. Esta variable se normaliza por el mismo índice construido a nivel del total de Aragón:

$$div_{ict} = \frac{1/\sum_{i^*=1, i^* \neq i}^i (emp_{i^*ct}/(emp_{ct} - emp_{ict}))^2}{1/\sum_{i^*=1, i^* \neq i}^i (emp_{i^*t}/(emp_t - emp_{it}))^2}, \quad (3)$$

donde emp_{ict} es el empleo en el sector i en el municipio c , emp_{ct} el empleo total municipal, emp_{it} es el empleo total de ese sector en Aragón, y emp_t el empleo total en Aragón, todos ellos medidos en el mismo momento t ; i es el número de sectores productivos. El numerador es máximo cuando todos los sectores excepto el que se está considerando (que llamamos i^*) tienen el mismo tamaño en el municipio. Este índice refleja la diversidad sectorial del sector i en el municipio c , y por lo tanto no se halla necesariamente relacionado con el nivel de especialización del sector analizado. Téngase en cuenta que, al estar definido de esta manera, obtenemos una medida de la diversidad distinta para cada sector en cada municipio.

Además de los dos índices que miden la especialización y diversidad productiva a nivel municipal que representan, respectivamente, las economías de localización y urbanización, en la literatura se sugiere la introducción de algunas variables adicionales. Con la intuición de que tradicionalmente las empresas grandes suelen ser más capaces de internalizar algunos de los efectos locales mientras las empresas pequeñas tienen mayor dificultad para hacerlo, Glaeser *et al.* (1992) sugieren incorporar el tamaño promedio de las empresas dentro de la industria local como un determinante adicional de las economías de localización. Al normalizar por el promedio de las empresas en el mismo sector a nivel de todo Aragón, consideramos el siguiente cociente:

$$tamaño_{ict} = \frac{emp_{ict}/n_{ict}}{emp_{it}/n_{it}}, \quad (4)$$

siendo emp_{ict} el empleo en el sector i en el municipio c , emp_{it} es el empleo total de ese mismo sector en todo Aragón, n_{ict} el número de empresas en la industria i en el municipio c en el momento t , y n_{it} el número total de empresas del sector en Aragón en el mismo año. De nuevo, al tratarse de un cociente entre el tamaño medio de las empresas del sector en el municipio c y el promedio del tamaño de las empresas del mismo sector en todo Aragón, valores superiores a la unidad indicarán que las

empresas de ese sector en ese municipio c tienen un tamaño superior (más empleados) que la media del sector en Aragón, mientras que si el cociente es inferior a la unidad representará que las empresas de ese sector en el municipio c son más pequeñas que el promedio total del sector. Como se indicó al comentar la Tabla 1, la gran mayoría de empresas aragonesas son pymes. Por lo tanto, con este ratio se persigue controlar específicamente el efecto del tamaño de la empresa en el caso de las grandes empresas, porque esos valores extremos de empleo (en comparación con la mayoría de la muestra, que son pymes) podrían sesgar el análisis.

Finalmente, Combes (2000) sugiere que, para controlar simultáneamente las diferencias entre ciudades, resulta relevante considerar la densidad del empleo total en estas mediante el siguiente indicador:

$$den_{ct} = \frac{emp_{ct}}{area_c}, \quad (5)$$

donde emp_{ct} es el empleo total municipal en el año t y $area_c$ es el área geográfica del municipio medida en km^2 .

Una vez definidas estas variables e indicadores de la diversidad y concentración empresarial, procedemos a estimar un modelo econométrico en el que la variable a explicar es el crecimiento del empleo en el sector i en el municipio c en el año t (ecuación 1), y las variables explicativas han sido definidas en las ecuaciones (2) a (5). El modelo econométrico básico sería el siguiente:

$$y_{ict} = \beta_0 + \beta_1 \ln esp_{ict} + \beta_2 \ln div_{ict} + \beta_3 \ln tamaño_{ict} + \beta_4 \ln den_{ct} + \delta_i + \eta_t + Prov_c + \epsilon_{ict} \quad (6)$$

donde δ_i denota los efectos fijos (EF) sectoriales, η_t son los efectos fijos temporales (las observaciones son anuales desde el año 2000 al 2015), $Prov_c$ indica los efectos fijos provinciales y ϵ_{ict} es el término de error.

No obstante, el modelo (6) presenta una serie de potenciales problemas econométricos. El más importante de estos problemas es el posible sesgo de selección. En este contexto el sesgo de selección hace referencia a que existen actividades económicas que solo están presentes en ciertas ciudades. Por tanto, para cada municipio es habitual que haya varios sectores productivos sin ninguna empresa, lo que implica que el número de empleados será nulo en esos casos.

Hay dos formas de tratar estas observaciones nulas. Algunos autores superan este inconveniente incluyendo en las regresiones únicamente los casos en que las variables toman valores distintos de cero, pero este proceder puede conducir a estimaciones sesgadas de los parámetros. La segunda alternativa, que es la opción que adoptamos

en este estudio, consiste en estimar un modelo de selección en dos etapas de Heckman (1976, 1979), estimando un modelo Tobit tipo II.

En la primera etapa se estima un modelo Probit de una variable ficticia (*dummy*) que toma el valor 1 si el empleo sectorial es observado en el municipio y cero si no como una función de variables disponibles para todas las ciudades. Es decir, se estima la probabilidad de que un municipio contenga un sector productivo en particular. Formalmente, la probabilidad de que un sector productivo esté presente en el municipio se estima mediante la siguiente ecuación:

$$\text{Prob}(S = 1|Z) = \phi(Z\gamma) \quad (7)$$

S indica el sector ($S = 1$ si $emp_{ict} > 0$ y $S = 0$ en caso contrario), Z es un vector de variables explicativas municipales, ϕ es la función de distribución acumulativa con distribución normal y γ es un vector de parámetros desconocidos a estimar. En nuestro caso las variables explicativas son una constante, la densidad de población y el porcentaje de la población de Aragón que representa la población del municipio (ambas variables en logaritmos), y una serie de efectos fijos sectoriales y provinciales¹. Esta estimación permite construir una nueva variable que se denomina Ratio Inverso de Mills, que captura la magnitud del sesgo y que se incorpora al modelo (6) como una variable explicativa adicional para corregir el mencionado problema de selección. Además, teniendo en cuenta la dimensión temporal de nuestros datos, este modelo Probit se estima para cada uno de los 15 años que incluye el panel de datos, siguiendo las instrucciones de Semykina y Wooldridge (2010) para corregir el sesgo de selección en un panel de datos.

En la segunda etapa del modelo Tobit tipo II se estimaría el siguiente modelo econométrico por Mínimos Cuadrados Ordinarios (MCO):

$$y_{ict} = \beta_0 + \beta_1 \ln esp_{ict} + \beta_2 \ln div_{ict} + \beta_3 \ln tamaño_{ict} + \beta_4 \ln den_{ct} + \beta_5 RIM_{ict} \times \eta_t + \delta_i + \eta_t + Prov_c + \epsilon_{ict} \quad (8)$$

La diferencia entre los modelos (6) y (8) es que en la ecuación (8) se incluye el Ratio Inverso de Mills (RIM) para corregir el sesgo de selección. Además, como recomiendan Semykina y Wooldridge (2010), dada la estructura de panel de datos de nuestros datos el Ratio Inverso de Mills se introduce interactuando con las *dummies* anuales (η_t denotaba los efectos fijos temporales).

¹ Se incluyen efectos fijos provinciales porque, debido a problemas de multicolinealidad, en este caso no es posible incluir efectos fijos a nivel de municipio.

El segundo problema que puede presentar esta estrategia empírica es que las estimaciones por Mínimos Cuadrados Ordinarios son consistentes pero ineficientes. Por tanto, el modelo (8) será estimado por MCO y por máxima verosimilitud (ML), la cual resulta eficiente. En el texto principal presentaremos los resultados obtenidos por ML, mientras que las estimaciones por MCO pueden consultarse en el Apéndice B. Las diferencias entre los resultados obtenidos con ambos estimadores son, en este caso, mínimas.

Por último, dado que los diferentes indicadores se construyen a partir de los mismos valores del empleo sectorial en cada municipio, podría existir un problema de multicolinealidad. Para controlar este potencial problema, a partir de las estimaciones MCO del modelo (8) calculamos el factor de inflación de la varianza (VIF, por sus siglas en inglés, *variance inflation factor*), que cuantifica la intensidad de la multicolinealidad. Los resultados obtenidos revelan que no hay problema de multicolinealidad, ya que el VIF se mantiene dentro de los límites sugeridos en la literatura (véanse los resultados en el Apéndice B).

3.1.2.- Resultados

Estimamos el modelo econométrico (8) de forma separada para los sectores de la agricultura, industria, construcción y servicios (Tablas 5, 6, 7 y 8, respectivamente), utilizando la clasificación de ramas de actividad CNAE-2009. El modelo es estimado por máxima verosimilitud (ML), si bien los resultados de la estimación MCO, muy similares, pueden consultarse en el Apéndice B.

La fuente de datos SABI proporciona información geográfica de las empresas aragonesas y del empleo que agregamos según su localización municipal en casi seiscientos municipios aragoneses distintos. Hay que destacar que estas cifras de empleo corresponden únicamente a los trabajadores asalariados; por lo tanto, los trabajadores autónomos quedan excluidos del estudio por falta de información estadística disponible desagregada por municipio y sector. Dada la importancia cuantitativa de los trabajadores autónomos, que representan los últimos años alrededor de un 20% del total de los trabajadores afiliados a la Seguridad Social, es posible que nuestras estimaciones subestimen los efectos de las distintas economías de aglomeración sobre el empleo, por lo que nuestros resultados deben considerarse como el límite inferior de los posibles efectos sobre el empleo.

El periodo temporal considerado comienza en el año 2000 y termina en 2015. Este periodo temporal nos permite realizar una estimación de largo plazo del efecto de las economías de aglomeración (15 años), y además cubre todo un ciclo económico, incluyendo épocas de bonanza como la primera mitad de la década de los 2000s, así

como la Gran Recesión que comenzó en el año 2008 y se alarga hasta casi el final del periodo considerado.

Los datos de la base de datos SABI nos permiten salvar una dificultad metodológica relacionada con el cambio de la clasificación de actividades económicas que se produjo en el año 2009, al sustituir la clasificación CNAE-2009 a la anterior clasificación CNAE-1993. En la base de datos las empresas están clasificadas según la actual CNAE-2009, pero se ofrecen datos de empleo de periodos anteriores, lo que nos permite retrotraer la clasificación de actividades CNAE-2009 hasta el inicio de nuestro periodo de análisis, siempre que las empresas consideradas ya estuvieran activas entonces.

Todas las tablas tienen una misma estructura común con 5 columnas que representan 5 especificaciones distintas del modelo (8). En las columnas (1) a (3) se utiliza la muestra de todos los municipios disponibles, pero se introducen los controles de forma progresiva: la columna (1) solo incluye las variables explicativas definidas en las ecuaciones (2) a (5), en la columna (2) además se incluyen efectos fijos temporales anuales (es decir, variables ficticias (*dummies*) que toman el valor 1 si la observación pertenece al año correspondiente y 0 si es de otro año) con los que se controlan los cambios que puedan ser debidos a la evolución temporal de las variables (factores como el ciclo económico, la evolución de los patrones migratorios, etc.) y el Ratio Inverso de Mills interactuado con las *dummies* temporales para, como recomiendan Semykina y Wooldridge (2010), corregir el sesgo de selección, y en la columna (3) adicionalmente se introducen en la estimación también efectos fijos provinciales y sectoriales, para controlar factores inobservables a nivel de sector y provincia que puedan influir en el crecimiento del empleo a nivel local.

Las dos columnas restantes corresponden a estimaciones en las que se incluyen todos los controles, pero se utilizan submuestras de municipios para realizar pruebas de robustez. En la columna (4) se realiza la estimación excluyendo Zaragoza capital de la muestra de municipios; como se ha indicado anteriormente, la ciudad de Zaragoza absorbe una gran cantidad de la actividad económica de Aragón (casi la mitad de las empresas totales), por lo que tratamos de comprobar que los resultados obtenidos en las demás estimaciones no corresponden únicamente a la importancia de Zaragoza. Finalmente, en la columna (5) se utiliza una submuestra de 46 municipios con más de 3.000 habitantes². Al centrarnos en las ciudades más pobladas excluimos del análisis los municipios menos poblados que, en muchas ocasiones, corresponden a áreas rurales en las que en teoría las economías de aglomeración no tendrían fuerza suficiente para generar ningún efecto significativo.

² Hemos realizado pruebas con otros umbrales de población, y hemos podido comprobar que a partir de este límite de 3.000 habitantes los resultados no cambian si se consideran otros niveles de población mínima.

Por último, nótese que los elevados tamaños muestrales (el número de observaciones que aparece al pie de cada columna) corresponden al número de municipios (casi seiscientos) multiplicado por el número de ramas de actividad en cada caso para cada actividad y por el número de años (14, ya que se pierde una observación al trabajar con tasas de crecimiento).

La Tabla 5 muestra los resultados correspondientes al sector agrario (ramas de actividad 01 a 03 según la clasificación CNAE-2009). A priori no hay ninguna teoría que prediga ningún efecto de economías de aglomeración en este sector, caracterizado por rendimientos constantes o decrecientes en la producción. Además, el análisis del apartado 2 nos indicó que las actividades de este sector muestran una dispersión significativa en distancias cortas, ya que se trata de actividades que requieren de un factor productivo básico (tierra) que por su propia naturaleza se halla disperso en el espacio. A pesar de ello, también estimamos el modelo para este sector, principalmente por dos razones. En primer lugar, por un afán de exhaustividad, ya que sí que analizamos los efectos de todos los demás sectores productivos. En segundo lugar, porque el sector agrícola tradicionalmente es importante en la economía aragonesa; a modo de ejemplo, el sector agrario es el sector que mejor crecimiento logró en el año 2016 con un 8,2% (Informe Económico de Aragón, 2016), y durante los años de crisis económica el sector agrario fue el único que mostró una evolución favorable con incrementos de su producción final. El Informe sobre la situación económica y social de Aragón en 2017 (2018) también identifica al sector agrario como uno de los que con más fuerza ha tirado de la actividad aragonesa en 2017.

Pues bien, aunque la primera columna muestra un efecto negativo y significativo de la especialización, tamaño y densidad, a medida que introducimos los controles estos coeficientes pierden significatividad. En la columna (3), que incluye todos los controles, solamente el coeficiente del tamaño mantiene su significatividad al 1%. Incluso cuando consideramos las submuestras en las columnas (4) y (5) se mantiene el mismo resultado. Esta variable tamaño medía el tamaño medio de las empresas del sector, normalizado por el tamaño medio de las empresas del sector en todo Aragón. El coeficiente negativo y significativo en todos los casos se interpretaría como evidencia de rendimientos decrecientes a escala en el sector; cuanto más grandes son las empresas de este sector en el municipio, menor es el crecimiento del empleo en las empresas del sector en ese mismo municipio.

Respecto a las dos variables explicativas principales, especialización y diversidad, que representan las economías de localización y urbanización, respectivamente, en los dos casos no encontramos ningún efecto significativo. El coeficiente de la variable diversidad no es significativo en ningún caso, mientras que para la especialización detectamos un cierto efecto negativo y significativo en el modelo de la columna (1)

que a medida que introducimos todos los controles pasa a ser significativo solo al 10% (columna 3), y finalmente se desvanece del todo si excluimos Zaragoza de la muestra (columna 4) o consideramos solo las ciudades más pobladas (columna 5).

Tabla 5. Resultados actividades agrícolas (actividades 01-03 CNAE-2009), ML

Municipios:	(1) Todos	(2) Todos	(3) Todos	(4) Sin Zaragoza	(5) Población >=3000
Especialización	-0,042*** (0,009)	-0,013 (0,009)	-0,017* (0,009)	-0,014 (0,009)	-0,024 (0,021)
Diversidad	-0,004 (0,005)	-0,004 (0,005)	-0,001 (0,006)	-0,000 (0,006)	0,015 (0,015)
Tamaño	-0,060*** (0,011)	-0,082*** (0,011)	-0,086*** (0,012)	-0,097*** (0,012)	-0,061** (0,026)
Densidad	-0,019*** (0,007)	-0,015** (0,006)	-0,004 (0,007)	0,002 (0,007)	-0,011 (0,016)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	3.296	3.296	3.296	3.252	652
Log verosimilitud	-1.979,028	-1.898,248	-1.882,764	-1.825,826	-332,148

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%.

Tabla 6. Resultados actividades industriales (actividades 10-33 CNAE-2009), ML

Municipios:	(1) Todos	(2) Todos	(3) Todos	(4) Sin Zaragoza	(5) Población >=3000
Especialización	-0,020*** (0,004)	-0,026*** (0,005)	-0,030*** (0,005)	-0,031*** (0,005)	-0,030*** (0,007)
Diversidad	-0,001 (0,004)	0,001 (0,004)	0,001 (0,004)	0,000 (0,004)	-0,007 (0,006)
Tamaño	-0,048*** (0,005)	-0,040*** (0,006)	-0,044*** (0,006)	-0,045*** (0,006)	-0,050*** (0,011)
Densidad	-0,008*** (0,002)	-0,006** (0,002)	-0,003 (0,003)	-0,003 (0,003)	-0,005 (0,005)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	12.820	12.820	12.820	12.475	5.829
Log verosimilitud	-7.438,526	-7.376,228	-7.346,534	-7.296,857	-3.441,353

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%.

La Tabla 6 presenta los resultados para el sector industrial (ramas de actividad 10 a 33 según la clasificación CNAE-2009). Este es el sector básico de nuestro análisis, ya que los efectos positivos de las economías de aglomeración están tradicionalmente vinculados a la actividad industrial, desde el trabajo pionero de los distritos industriales de Marshall (1890). El modelo de la columna (1) muestra un efecto negativo y significativo de la especialización, tamaño y densidad. A medida que introducimos los controles el coeficiente de la densidad pierde significatividad (columna 3), pero no los de las otras dos variables que se mantienen negativos y significativos en todos los modelos estimados (columnas 1 a 5). Respecto al coeficiente de la variable que mide la diversidad productiva, no es significativo en ningún caso.

Por lo tanto, obtenemos un efecto negativo de la especialización (economías de localización) sobre el empleo sectorial industrial a nivel municipal y ningún efecto significativo de la diversidad productiva (economías de urbanización). Estos resultados son robustos incluso cuando se consideran submuestras (excluyendo Zaragoza capital o considerando solo los municipios de más de 3.000 habitantes). Aunque no hay ningún modelo teórico que prediga un efecto negativo de las economías de localización sobre el empleo, otros trabajos anteriores también obtienen un coeficiente del mismo signo; por ejemplo Combes (2000) obtiene el mismo resultado en su análisis de las zonas urbanas francesas. Si bien no encontramos explicación para este coeficiente en la teoría de las economías de aglomeración, Combes (2000) sugiere buscar una respuesta en la evolución del ciclo del sector considerado a nivel agregado. Su explicación es que una elevada especialización productiva implica poca flexibilidad y peor adaptabilidad de productos, tecnologías e infraestructuras cuando el sector está en declive, explicación que creemos que se ajusta perfectamente al periodo analizado (2000–2015) que incluye, durante la Gran Recesión que comienza en 2008, una importante destrucción de empleo en el sector industrial. Así, aquellos municipios aragoneses que estaban más especializados en la industria sufrieron una destrucción de empleo superior a la media de Aragón.

Finalmente, el tamaño medio tiene un efecto negativo significativo en todos los casos. Este signo negativo aportaría evidencia en contra de las economías de escala en el sector industrial, aunque también podría ser un reflejo de la evolución en el ciclo de vida de las industrias, o simplemente representar la estructura empresarial aragonesa con un dominio absoluto de las pymes, también en el sector industrial.

El siguiente sector que analizamos es la construcción (ramas de actividad 41 a 43 según la clasificación CNAE-2009), uno de los sectores más castigados con la destrucción de empleo durante la reciente crisis económica; el Informe Económico de Aragón (2016) indica que el empleo en la construcción en Aragón disminuyó desde el comienzo de la crisis en 2008 y no volvió a aumentar hasta 2015. A pesar de ello, según

el Informe sobre la situación económica y social de Aragón en 2017 (2018), Aragón es todavía una comunidad autónoma de marcado carácter industrial, con unos porcentajes de su Producto Interior Bruto y del empleo (23 y 19%, respectivamente), superiores a los valores medios de España.

La Tabla 7 muestra las estimaciones para las ramas de actividad del sector de la construcción. Los resultados son similares a los obtenidos para las actividades industriales: mientras que el coeficiente de la especialización es negativo y significativo, el coeficiente de la variable que mide la diversidad productiva no es significativo en ningún caso. Nuestra interpretación de este resultado es que, de nuevo, las estimaciones están recogiendo el efecto negativo sobre el empleo sectorial de la mala evolución del sector durante el periodo de crisis. No obstante, la excepción es el modelo en el que se utiliza una submuestra de los municipios más poblados (columna 5), en el que el efecto de la especialización tampoco es significativo; lo que indicaría que en las grandes ciudades el sector de la construcción no es capaz de generar economías de aglomeración de ninguno de los dos tipos (localización o urbanización). Además, la variable que mide el tamaño medio de las empresas del sector vuelve a tener un efecto negativo significativo en todos los casos, que podemos volver a interpretar como evidencia contraria a las economías de escala en el sector de la construcción.

Tabla 7. Resultados construcción (actividades 41-43 CNAE-2009), ML

	(1)	(2)	(3)	(4)	(5)
Municipios:	Todos	Todos	Todos	Sin Zaragoza	Población >=3000
Especialización	-0,018** (0,007)	-0,020** (0,009)	-0,031*** (0,009)	-0,031*** (0,009)	-0,044 (0,040)
Diversidad	0,001 (0,004)	0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,016 (0,012)
Tamaño	-0,130*** (0,011)	-0,128*** (0,013)	-0,128*** (0,012)	-0,128*** (0,012)	-0,148*** (0,046)
Densidad	-0,002 (0,004)	-0,002 (0,004)	0,000 (0,005)	0,000 (0,005)	-0,018 (0,014)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	6.164	6.164	6.164	6.119	779
Log verosimilitud	-3.432,856	-3.366,126	-3.344,037	-3.341,081	-307,100

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%.

Tabla 8. Resultados sector servicios (actividades 49-96 CNAE-2009), ML

Municipios:	(1) Todos	(2) Todos	(3) Todos	(4) Sin Zaragoza	(5) Población >=3000
Especialización	-0,011*** (0,004)	-0,016*** (0,004)	-0,025*** (0,005)	-0,037*** (0,006)	-0,040*** (0,008)
Diversidad	0,004 (0,003)	0,004 (0,003)	0,001 (0,003)	0,000 (0,004)	0,013** (0,006)
Tamaño	-0,104*** (0,006)	-0,098*** (0,006)	-0,105*** (0,007)	-0,099*** (0,007)	-0,110*** (0,011)
Densidad	0,001 (0,003)	0,002 (0,003)	0,001 (0,003)	0,000 (0,003)	0,008* (0,004)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	14.374	14.374	14.374	13.753	7.989
Log verosimilitud	-9.280,180	-9.173,532	-9.089,257	-8.915,978	-4.951,565

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%.

Por último, la Tabla 8 presenta los resultados para el sector servicios (ramas de actividad 49 a 96 según la clasificación CNAE-2009), cuyo peso representa en torno al 67% del empleo y la producción totales de la economía aragonesa y que ha liderado la creación de empleo los últimos años (Informe Económico de Aragón, 2016), especialmente desde 2014 (Informe sobre la situación económica y social de Aragón en 2017, 2018). Recordemos que el análisis del apartado 2 reveló una alta concentración espacial en distancias cortas de las actividades del sector servicios, lo que facilitaría la generación de economías de aglomeración.

Obtenemos dos resultados comunes a los obtenidos con los anteriores sectores; por un lado, un efecto negativo y significativo de la especialización productiva, que podemos vincular a la destrucción de empleo durante al Gran Recesión, y por otro lado un coeficiente negativo y significativo del tamaño medio de las empresas, que estaría reflejando otra vez la predominancia de las pymes en la economía aragonesa.

El resultado nuevo que muestra la Tabla 8, y que es específico del sector servicios, es el efecto positivo y significativo de la diversidad productiva cuando utilizamos la muestra de municipios más poblados (columna 5). Este efecto positivo sobre el empleo sectorial puede ser considerado como evidencia favorable de las economías de urbanización sobre el empleo sectorial: un aumento de la diversidad en las distintas ramas del sector servicios estaría generando incrementos en el empleo de otras ramas del sector en el mismo municipio, lo que indicaría vínculos de oferta y demanda entre

las distintas actividades del sector servicios. No obstante, estos resultados indicarían que, para que se genere este efecto positivo, es necesaria una cierta escala o tamaño económico del municipio.

3.2.- Economías de aglomeración y crecimiento de la población

Uno de los mayores problemas sociales y demográficos de Aragón es la despoblación de muchos de sus municipios, normalmente situados en zonas rurales, hasta tal punto que frenar la despoblación es el objetivo primario o secundario de una amplia batería de medidas públicas que actúan sobre ámbitos muy distintos. Como se indicó en la Introducción del apartado 1, incluso la política industrial que favorece la creación de clústeres es justificada en ocasiones con el argumento de crear empleo local y fijar población al territorio. En el apartado 3.1 se ha analizado si la concentración empresarial realmente contribuye a la creación de empleo a nivel local, y este apartado se ocupa de estimar si hay algún tipo de efecto de las economías de aglomeración en el crecimiento de la población de los municipios aragoneses.

Si bien el concepto de economías de aglomeración goza de sólidos fundamentos teóricos (Marshall, 1890; Duranton y Puga, 2004) que nos permiten identificar de forma clara distintos mecanismos de transmisión de sus efectos al empleo, los posibles efectos sobre el crecimiento de la población son más difusos ya que pueden intervenir factores relacionados con la oferta y con la demanda. Es por ello que dedicamos el siguiente epígrafe a la exposición de un modelo teórico del que se puede obtener una ecuación del crecimiento de la población local que posteriormente será estimada.

3.2.1.- Modelo teórico

En este apartado exponemos el modelo de crecimiento urbano de Glaeser *et al.* (1995), posteriormente desarrollado en Glaeser (2000) y Glaeser y Shapiro (2003). Este modelo es un desarrollo teórico del modelo de equilibrio espacial clásico de Roback (1982), y constituye un marco de referencia que incluye factores de oferta y de demanda que pueden influir en el crecimiento de la población local.

Partimos del supuesto de que siempre existe un equilibrio espacial en el que (1) la utilidad individual y (2) los rendimientos del capital se igualan en el espacio. La función de producción del municipio i en el momento t es $A_{it} f(K_{it}, L_{it}) = A_{it} K_{it}^{\alpha} L_{it}^{\beta}$, donde A_{it} es el nivel de productividad, K_{it} es el nivel de capital y L_{it} es el nivel de trabajo de la ciudad, que asumimos que es igual a z veces la población total de la ciudad (N_{it}), con $0 < z \leq 1$. z captura la posibilidad de que parte de la población del municipio no esté trabajando (ya sea porque estén desempleados o porque no formen parte de la población activa), y el precio del output está normalizado y es igual a la unidad.

Asumiendo competencia perfecta, las ganancias del capital son una tasa r exógena (igual a la productividad marginal del capital, $r_{it} = \alpha A_{it} K_{it}^{\alpha-1} L_{it}^{\beta} = \alpha A_{it} K_{it}^{\alpha-1} (z N_{it})^{\beta}$) y el salario será igual también a la productividad marginal del trabajo: $W_{it} = \beta A_{it} K_{it}^{\alpha} L_{it}^{\beta-1} = \beta A_{it} K_{it}^{\alpha} (z N_{it})^{\beta-1}$.

La utilidad de un individuo es $\frac{C_{it} W_{it}}{P_{it}}$, donde C_{it} es un índice de consumo de amenidades locales (*consumption amenity index*), W_{it} representa el salario, y P_{it} representa los precios, todo ello medido a nivel local de municipio. Esta utilidad debe igualarse a un nivel u de utilidad que en equilibrio es constante entre ciudades, $u = \frac{C_{it} W_{it}}{P_{it}}$.

Sustituyendo el salario W_{it} en la función de utilidad por la productividad marginal del trabajo se obtiene: $\frac{C_{it}}{P_{it}} W_{it} = \frac{C_{it}}{P_{it}} \beta A_{it} K_{it}^{\alpha} L_{it}^{\beta-1}$, y aplicando que en el equilibrio espacial debe igualarse la utilidad individual en todos los municipios al nivel constante u :

$$u = \frac{C_{it}}{P_{it}} \beta A_{it} K_{it}^{\alpha} L_{it}^{\beta-1} \quad (9)$$

A partir de esta primera condición de equilibrio podemos obtener los niveles de equilibrio de K_{it} y L_{it} . En particular, nos interesa despejar K_{it} :

$$K_{it}^{\alpha} = u \cdot \left[A_{it} \left(\beta \frac{C_{it}}{P_{it}} \right) L_{it}^{\beta-1} \right]^{-1},$$

$$K_{it} = u^{\frac{1}{\alpha}} A_{it}^{\frac{-1}{\alpha}} \left(\beta \frac{C_{it}}{P_{it}} \right)^{\frac{-1}{\alpha}} \cdot L_{it}^{\frac{1-\beta}{\alpha}}.$$

La segunda condición de equilibrio espacial postula que los rendimientos del capital deben igualarse en el espacio ($r_{it} = r$). Aplicando que en equilibrio, asumiendo competencia perfecta, el rendimiento del capital debe igualarse a su productividad marginal podemos obtener:

$$r = \alpha A_{it} K_{it}^{\alpha-1} L_{it}^{\beta} \quad (10)$$

Sustituyendo en esta condición de equilibrio (10) el nivel de equilibrio de K_{it} obtenido anteriormente:

$$r = \alpha A_{it} K_{it}^{\alpha-1} L_{it}^{\beta} = \alpha A_{it} L_{it}^{\beta} \cdot \left[u^{\frac{1}{\alpha}} A_{it}^{\frac{1}{\alpha}} \left(\beta \frac{C_{it}}{P_{it}} \right)^{\frac{1}{\alpha}} \cdot L_{it}^{\frac{1-\beta}{\alpha}} \right]^{\alpha-1}.$$

Después de algunas cuentas es posible despejar L_{it} :

$$L_{it} = A_{it}^{\frac{1}{1-\alpha-\beta}} \left(\frac{\alpha^{\alpha} \beta^{1-\alpha}}{r^{\alpha} u^{1-\alpha}} \right)^{\frac{1}{1-\alpha-\beta}} \left(\frac{C_{it}}{P_{it}} \right)^{\frac{1-\alpha}{1-\alpha-\beta}}.$$

Sustituyendo L_{it} por (zN_{it}) podemos despejar la población del municipio i en el momento t (N_{it}):

$$(zN_{it}) = A_{it}^{\frac{1}{1-\alpha-\beta}} \left(\frac{\alpha^{\alpha} \beta^{1-\alpha}}{r^{\alpha} u^{1-\alpha}} \right)^{\frac{1}{1-\alpha-\beta}} \left(\frac{C_{it}}{P_{it}} \right)^{\frac{1-\alpha}{1-\alpha-\beta}},$$

$$N_{it} = A_{it}^{\frac{1}{1-\alpha-\beta}} \left(\frac{\alpha^{\alpha} \beta^{1-\alpha}}{r^{\alpha} u^{1-\alpha}} \right)^{\frac{1}{1-\alpha-\beta}} \left(\frac{C_{it}}{P_{it}} \right)^{\frac{1-\alpha}{1-\alpha-\beta}} \cdot \frac{1}{z}.$$

Tomando logaritmos:

$$\begin{aligned} \text{Log}(N_{it}) &= -\text{Log}(z) + \frac{1}{1-\alpha-\beta} \cdot [\text{Log}(\alpha^{\alpha} \beta^{1-\alpha}) - \alpha \text{Log}(r) - (1-\alpha) \text{Log}(u)] + \\ &+ \frac{1}{1-\alpha-\beta} \text{Log}(A_{it}) + \frac{1-\alpha}{1-\alpha-\beta} \text{Log}\left(\frac{C_{it}}{P_{it}}\right) \end{aligned}$$

y agrupando términos:

$$\text{Log}(N_{it}) = \Theta_t + \frac{1}{1-\alpha-\beta} \text{Log}(A_{it}) + \frac{1-\alpha}{1-\alpha-\beta} \text{Log}\left(\frac{C_{it}}{P_{it}}\right),$$

con $\Theta_t = -\text{Log}(z) + \frac{1}{1-\alpha-\beta} [\text{Log}(\alpha^{\alpha} \beta^{1-\alpha}) - \alpha \text{Log}(r) - (1-\alpha) \text{Log}(u)]$, siendo Θ_t un término que es constante entre todos los municipios.

Supongamos que cada municipio i tiene una serie de K características escalares, que podemos denotar como $X_{i1}, \dots, X_{ik}, \dots, X_{iK}$. Estas características pueden incluir todo tipo de factores, desde el clima o el tamaño del sector público local a características de la estructura productiva (especialización o diversidad) a nivel municipal. Por simplicidad, suponemos que estas características no cambian con el tiempo. Si \mathbf{X}_i es el vector de estas características, asumimos que

$\text{Log}(A_{it}) = \mathbf{X}_i' \beta_t + \delta_{it} + \varepsilon_{it}$ y que $\text{Log}\left(\frac{C_{it}}{P_{it}}\right) = \mathbf{X}_i' \gamma_t + \mu_{it}$, donde ε_{it} y μ_{it} son errores

ortogonales en niveles y en tasas de cambio para cada una de las características

observables, y que β_t y γ_t son los vectores de coeficientes correspondientes a las características de cada ciudad. El término δ_{it} es ortogonal en niveles, pero $\delta_{it+1} - \delta_{it} = \mathbf{X}'_i \theta + \xi_{it}$, donde ξ_{it} es otro término de error ortogonal y θ representa la tasa de cambio tecnológico.

Utilizando estos términos y combinando los errores ortogonales, podemos obtener la expresión de los incrementos:

$$\begin{aligned} \text{Log}(A_{it+1}) - \text{Log}(A_{it}) &= \mathbf{X}'_i (\beta_{t+1} - \beta_t) + \mathbf{X}'_i \theta + \xi_{it} + \epsilon_{it+1} - \epsilon_{it}, \\ \text{Log}\left(\frac{C_{it+1}}{P_{it+1}}\right) - \text{Log}\left(\frac{C_{it}}{P_{it}}\right) &= \mathbf{X}'_i (\gamma_{t+1} - \gamma_t) + \mu_{it+1} - \mu_{it}. \end{aligned}$$

Estas expresiones están en términos vectoriales (son vectores $K \times 1$). En términos escalares:

$$\begin{aligned} \text{Log}(A_{it+1}) - \text{Log}(A_{it}) &= \sum_k^K [(\beta_{kt+1} - \beta_{kt}) + \theta_k] X_{ik} + \xi_{it} + \epsilon_{it+1} - \epsilon_{it} \\ \text{Log}\left(\frac{C_{it+1}}{P_{it+1}}\right) - \text{Log}\left(\frac{C_{it}}{P_{it}}\right) &= \sum_k^K (\gamma_{kt+1} - \gamma_{kt}) X_{ik} + \mu_{it+1} - \mu_{it}. \end{aligned}$$

A partir de $\text{Log}(A_{it+1}) - \text{Log}(A_{it})$ y $\text{Log}\left(\frac{C_{it+1}}{P_{it+1}}\right) - \text{Log}\left(\frac{C_{it}}{P_{it}}\right)$ podemos obtener la expresión de la variación en la población, $\text{Log}(N_{it+1}) - \text{Log}(N_{it})$:

$$\text{Log}\left(\frac{N_{it+1}}{N_{it}}\right) = \Delta\Theta + \frac{1}{1 - \alpha - \beta} \cdot \sum_k^K [\beta_{kt+1} - \beta_{kt} + (1 - \alpha)(\gamma_{kt+1} - \gamma_{kt}) + \theta_k] X_{ik} + \zeta_{it}, \quad (11)$$

donde ζ_{it} es el término de error ortogonal.

La ecuación (11) es directamente estimable si se dispone de datos locales de las características individuales de los municipios. Además, en ella se observa que hay tres vías por las que una característica X_k puede influir en el crecimiento del municipio: (1) puede influir en el proceso de producción ($\beta_{kt+1} - \beta_{kt}$), puede afectar al consumo (coste de la vida, amenidades, etc.) ($\gamma_{kt+1} - \gamma_{kt}$), y puede influir en la productividad a través de la tasa de cambio tecnológico (θ_k). Así, si obtenemos que una característica local (por ejemplo, el nivel de especialización productiva) predice un crecimiento positivo de la población municipal, esto podría ocurrir por tres motivos. En primer lugar, porque esta variable X_k se ha vuelto más importante en el proceso de producción, en cuyo caso, $\beta_{kt+1} > \beta_{kt}$ es decir, la característica estaría afectando a la oferta. Segundo, desde el punto de vista de la demanda, esta variable X_k puede

haberse vuelto más importante para los consumidores al reducir el coste de la vida o al aumentar el volumen de amenidades locales disponibles, lo que en términos del modelo implicaría que $\gamma_{kt+1} > \gamma_{kt}$. Por último, podría ser que la variable X_k contribuyera a incrementar la tasa de crecimiento tecnológico, $\theta_k > 0$, otro factor de oferta de la economía. No obstante, en nuestras estimaciones no intentaremos identificar cuál de estas tres posibles explicaciones está detrás de cada coeficiente positivo o negativo, ya que al estimar se estima conjuntamente para cada característica un único coeficiente que es igual a $[\beta_{kt+1} - \beta_{kt} + (1 - \alpha)(\gamma_{kt+1} - \gamma_{kt}) + \theta_k] / (1 - \alpha - \beta)$.

Este modelo ha sido profusamente utilizado para estudiar el crecimiento de las ciudades, especialmente en el caso americano. Glaeser *et al.* (1995) examinan los patrones de crecimiento urbano de las 200 ciudades estadounidenses más pobladas entre 1960 y 1990 en relación con varias características urbanas iniciales de 1960, llegando a la conclusión de que el crecimiento de la renta y de la población de las ciudades estuvo (1) positivamente relacionado con el nivel de educación inicial, (2) negativamente relacionado con el desempleo inicial, y (3) negativamente relacionado con el porcentaje inicial de empleo en el sector de las manufacturas. Posteriormente, Glaeser y Shapiro (2003) analizan el crecimiento durante la década 1990-2000, utilizando un tamaño muestral ligeramente superior (imponen una población mínima de 25.000 habitantes, considerando las 1.000 ciudades más pobladas). Durante esa década las tres variables más relevantes habrían sido el capital humano, el clima y el sistema de transporte de los individuos (público o privado).

3.2.2.- Metodología

A partir de la ecuación (11), que relaciona el crecimiento de la población de los municipios con las características locales, nos proponemos estimar un modelo con el que pretendemos explicar el crecimiento de los municipios aragoneses en función de una serie de variables que, dada la temática de este estudio, estarán relacionadas con las economías de aglomeración (economías de localización o urbanización).

En primer lugar, definimos la variable a explicar como el crecimiento relativo de la población:

$$\begin{aligned} g_{ct} &= \ln Población\ relativa_{ct} - \ln Población\ relativa_{ct-1} = \\ &= \ln\left(\frac{Pob_{ct}}{Pob_t}\right) - \ln\left(\frac{Pob_{ct-1}}{Pob_{t-1}}\right) = \ln\left(\frac{Pob_{ct}}{Pob_{ct-1}}\right) - \ln\left(\frac{Pob_t}{Pob_{t-1}}\right), \end{aligned} \quad (12)$$

donde Pob_{ct} es la población del municipio c en el momento t y Pob_t es la población total de Aragón. El periodo temporal considerado es el mismo que en el epígrafe anterior; se trata de un panel de datos anuales desde el año 2000 al 2015. Los datos de

población municipal se obtienen del padrón municipal elaborado por el Instituto Nacional de Estadística. Al considerar poblaciones relativas (la cuota que representa la población de cada municipio sobre el total de población en Aragón) nuestro análisis versará no sobre el crecimiento total de la población en valor absoluto, sino que trataremos de explicar por qué el crecimiento de un municipio determinado es más alto o más bajo en comparación con el crecimiento total de la población en Aragón.

En segundo lugar, definimos una serie de variables explicativas que nos permitirán obtener una medida de las economías de localización o urbanización a nivel local. Recordemos que, como se ha indicado anteriormente, las economías de localización están relacionadas con la concentración empresarial dentro de un sector específico mientras que las economías de urbanización operan a través de la concentración general de la actividad económica. Es importante destacar que, al contrario que en el análisis anterior del efecto de las economías de aglomeración en el empleo sectorial a nivel local en que teníamos para cada municipio una observación para cada sector (porque el empleo en cada sector dentro del mismo municipio es diferente), la unidad de análisis ahora es la ciudad (no el sector), así que se construyen variables a nivel municipal, por lo que tendremos una única observación para cada municipio en cada año (porque solo hay un dato de población por año y municipio).

La medida de especialización (esp_{ct}) relacionada con las economías de localización a nivel municipal es el índice de especialización de Krugman, que se define como:

$$esp_{ct} = \sum_{i=1}^I \left| \frac{emp_{ict}}{emp_{it}} - \frac{emp_{ct}}{emp_t} \right| \quad (13)$$

donde emp_{ict} es el empleo en el sector i en el municipio c , emp_{ct} el empleo total municipal, emp_{it} es el empleo total de ese sector en Aragón, y emp_t el empleo total en Aragón, todos ellos medidos en el año t ; I es el número total de sectores presentes en el municipio. La fuente de datos del empleo sectorial es, como en los apartados anteriores, SABI. El índice es la suma en valor absoluto de las desviaciones en la estructura productiva del municipio, medida a través de las cuotas de empleo de cada sector respecto al empleo total del municipio, respecto a la referencia de la estructura de empleo sectorial de Aragón. La interpretación es sencilla; cuanto mayor sea el índice más se desvía la estructura económica del municipio de la de Aragón (es decir, más especializado está).

Respecto a las economías de urbanización (div_{ct}), estas se miden a través del índice de Herfindahl inverso normalizado:

$$div_{ct} = \frac{1/\sum_{i=1}^I (emp_{ict}/(emp_{ct} - emp_{ict}))^2}{1/\sum_{i=1}^I (emp_{it}/(emp_t - emp_{it}))^2}, \quad (14)$$

siendo emp_{ict} el empleo en el sector i en el municipio c , emp_{ct} el empleo total municipal, emp_{it} es el empleo total de ese sector en Aragón, y emp_t el empleo total en Aragón, todos ellos medidos en el mismo momento t . Nótese que, a diferencia de la ecuación (3), en esta ocasión no se excluye ningún sector del cálculo, por lo que se obtiene un único índice para todo el municipio en cada año.

Además, se incluye también el tamaño promedio de las empresas, considerando todas las empresas dentro de todos los sectores del municipio (normalizado con el tamaño medio en todo Aragón):

$$tamaño_{ct} = \frac{1}{I} \sum_{i=1}^I \left(\frac{emp_{ict}/n_{ict}}{emp_{it}/n_{it}} \right), \quad (15)$$

donde emp_{ict} es el empleo en el sector i en el municipio c , emp_{it} es el empleo total de ese mismo sector en todo Aragón, n_{ict} el número de empresas en la industria i en el municipio c en el momento t , y n_{it} el número total de empresas del sector en Aragón en el mismo año; I es el número total de sectores presentes en el municipio.

Por último, también añadimos la densidad del empleo total en cada municipio mediante el indicador definido en la ecuación (5) para controlar las diferencias en tamaño geográfico (área) entre ciudades:

$$den_{ct} = \frac{emp_{ct}}{area_c},$$

siendo emp_{ct} el empleo total municipal en el año t y $area_c$ el área geográfica del municipio medida en km^2 .

Después de definir estas variables e indicadores de la diversidad y concentración productiva a nivel local, procedemos a estimar un modelo econométrico en el que la variable a explicar es el crecimiento de la población relativa en el municipio c en el año t (ecuación 12), y las variables explicativas han sido definidas en las ecuaciones (5) y (13) a (15). El modelo econométrico básico sería el siguiente:

$$g_{ct} = \beta_0 + \beta_1 \ln esp_{ct} + \beta_2 \ln div_{ct} + \beta_3 \ln tamaño_{ct} + \beta_4 \ln den_{ct} + \varphi_c + \eta_t + v_{ct}, \quad (16)$$

donde φ_c denota los efectos fijos municipales, η_t son los efectos fijos anuales y v_{ct} es el término de error. Esta ecuación lineal (16) encuentra soporte teórico en el modelo de crecimiento urbano de Glaeser *et al.* (1995) explicado en el epígrafe anterior, y, en concreto, es una forma de especificar al ecuación (11) del modelo.

Si bien es cierto que hay muchas características locales que pueden influir en el crecimiento de la población (desde el clima, la red de transporte disponible o el precio de la vivienda, por ejemplo), las variables explicativas escogidas recogen únicamente factores relacionados con las economías de aglomeración; el resto de las características locales que podrían tener efecto en el crecimiento local de la población se entiende incluidas en los efectos fijos municipales (variables ficticias (*dummies*) que toman el valor 1 si la observación pertenece al municipio correspondiente y 0 si es de cualquier otro municipio).

Respecto a los efectos fijos temporales, (*dummies* que toman el valor 1 si la observación pertenece al año correspondiente y 0 si es de cualquier otro año), nos sirven para controlar los cambios en la evolución de la población que puedan ser debidos a la evolución temporal de las variables. En concreto, dentro de nuestro periodo de estudio (2000–2015) González Pampillón (2018) documenta la existencia de dos fases claramente diferenciadas en la evolución de la población inmigrante, con distintos efectos en la población de las principales ciudades españolas y, dentro de estas, de la distribución de esta población inmigrante entre los distintos vecindarios: el periodo del *boom* de los inmigrantes (2001–2009) y el posterior periodo de congelación de la población inmigrante (2010–2015). Como explica González Pampillón (2018), España experimentó un ingreso masivo de inmigrantes entre finales de los años noventa hasta 2009, causado por un crecimiento económico sostenido; durante estos años España fue el segundo receptor de inmigrantes en términos absolutos (detrás de EE. UU.) y el más alto en relación con su población. Sin embargo, la recesión, que terminó en España aproximadamente en 2013, detuvo estos flujos de inmigrantes y, de hecho, la participación de inmigrantes disminuyó ligeramente entre 2010 y 2015. Es de esperar que estas dinámicas temporales tuvieran alguna consecuencia sobre la población de los municipios aragoneses, pero intentamos aislar estos efectos a través de los efectos fijos anuales.

Sin embargo, el modelo (16) presenta una serie de potenciales problemas econométricos. Estos son, básicamente, la posible persistencia en la tendencia de la variable dependiente (crecimiento de la población), relacionada con cuestiones dinámicas de la propia variable, y problemas de endogeneidad y causalidad inversa. Para resolver ambos problemas potenciales, en algunos de los modelos que estimamos también incluimos como variable explicativa el valor inicial de la población relativa del municipio (lo que convierte el modelo (16) en un modelo dinámico de panel de datos),

y además también utilizamos el estimador GMM en diferencias (incrementos) de Arellano y Bond (1991). Con este estimador se eliminan los efectos individuales no observados específicos (los efectos fijos municipales) y también se controla el efecto de la posible tendencia temporal en nuestras principales variables de interés al utilizar primeras diferencias. Además, las variables independientes se instrumentan con valores retardados de las variables dependientes e independientes³. Por último, para despejar cualquier duda de posible multicolinealidad calculamos el indicador de inflación de la varianza (VIF), que arroja unos valores dentro de los límites sugeridos en la literatura, por lo que concluimos que no hay problema de multicolinealidad.

3.2.3.- Resultados

El modelo (16) es estimado para un panel de datos desde el año 2000 al 2015 con 577 municipios aragoneses distintos⁴, incluyendo efectos fijos (EF) municipales y temporales. El número de observaciones temporales es 14, ya que se pierde un año al calcular la tasa de crecimiento. Se estima con desviaciones típicas robustas aplicándoles clúster a nivel de ciudad.

La Tabla 9 muestra los resultados de la estimación del efecto de las economías de aglomeración sobre el crecimiento de la población (que es la variable a explicar, ecuación 12). La tabla tiene 5 columnas que representan 5 especificaciones distintas del modelo (16). En las columnas (1) y (3) se estima el modelo incluyendo todos los controles, pero utilizando diferentes muestras de municipios, mientras que en los modelos de las columnas (2), (4) y (5) se incluye, además de todos los controles, la población inicial del municipio para controlar la persistencia en el crecimiento de la población, lo que convierte a estos modelos en modelos dinámicos de panel de datos. Dentro de los modelos dinámicos, la estimación de la columna (5) se realiza aplicando el estimador GMM en diferencias de Arellano y Bond (1991), para controlar posibles problemas de endogeneidad y causalidad inversa.

Consideramos dos muestras de municipios; las columnas (1) y (2) contienen los resultados cuando todos los municipios disponibles son incluidos en la muestra, y en los modelos de las columnas (3), (4) y (5) se utiliza una submuestra de 46 municipios con más de 3.000 habitantes⁵. El argumento es el mismo que en el apartado anterior

³ En nuestro caso, utilizamos el tercer retardo de todas las variables independientes, ya que de esa forma los distintos tests arrojan buenos resultados.

⁴ Según el padrón municipal del año 2017, Aragón cuenta con 731 municipios distintos: 202 situados en Huesca, 236 en Teruel y 293 en Zaragoza. Nuestro tamaño muestral es ligeramente inferior porque hay algunos municipios (154) para los que no tenemos información de empleo de ninguna empresa, y por lo tanto no podemos calcular las variables (5) y (13) a (15).

⁵ Este es el mismo límite de población que utilizábamos en el análisis anterior. Al igual que entonces, hemos comprobado que a partir de este límite de 3.000 habitantes los resultados no cambian si se consideran otros umbrales de población.

en que estudiábamos el efecto de las economías de aglomeración sobre el empleo sectorial local: con esta submuestra nos centramos en las ciudades más pobladas, excluyendo del análisis los municipios menos poblados que, en la mayoría de las ocasiones, corresponden a áreas rurales en las que en teoría las economías de aglomeración no tendrían fuerza suficiente para generar ningún efecto significativo⁶.

Cuando consideramos todos los municipios (columnas (1) y (2)) ninguna de las variables introducidas (ecuaciones (5) y (13) a (15)) es significativa en ningún caso. Esto indicaría que, si incluimos todos los municipios aragoneses, no se encuentra evidencia de ningún efecto significativo sobre el crecimiento de la población ni de las economías de localización (especialización) ni de las economías de urbanización (diversidad). No solo ambos coeficientes son no significativos, sino que el valor del coeficiente está muy próximo a cero en los dos casos.

Si nos centramos en los resultados obtenidos con la muestra de municipios más poblados (columnas (3), (4) y (5)), los resultados cambian ya que obtenemos un efecto positivo de ambas variables, especialización y diversidad. Además, dada la magnitud de los coeficientes, se verifica que el efecto de las economías de localización (especialización) es más del doble que el de las economías de urbanización (diversidad). Respecto al resto de variables (tamaño promedio de las empresas y densidad de empleo), no obtenemos ningún coeficiente significativo para ninguna de ellas en ningún caso, excepto el coeficiente de la densidad de empleo en el modelo estimado en la columna (4).

En aquellos casos en que se incluye como variable explicativa adicional la población inicial del municipio (columnas (2), (4) y (5)) el coeficiente de dicha variable es siempre significativo, lo que aportaría evidencia de una estructura dinámica de nuestra variable dependiente, respaldando el uso del modelo dinámico de panel de datos. Dicho coeficiente es negativo, lo que indicaría que en el periodo de tiempo considerado (2000–2015) el crecimiento de la población entre los municipios habría sido convergente. Es decir, un vez que los efectos fijos municipales controlan la tendencia individual del crecimiento de cada ciudad, el patrón general que podemos extraer es que aquellos municipios con mayor población inicial crecieron menos que aquellos menos poblados.

Por último, las columnas (4) y (5) muestran que los resultados no cambian sustantivamente si la estimación se realiza aplicando el estimador GMM en diferencias

⁶ En las estimaciones del apartado 3.1.2. incluimos los resultados para una submuestra adicional excluyendo Zaragoza capital de la muestra de municipios. En términos de empleo sectorial local para Zaragoza teníamos varias observaciones (una para cada rama de actividad dentro de cada sector considerado), pero en términos de población Zaragoza solo representa una única observación. Además, téngase en cuenta que el modelo (16) incluye efectos fijos a nivel de municipio, por lo que la exclusión de solamente uno de ellos no cambia los resultados.

(columna 5), ya que los coeficientes de las dos variables relacionadas con las economías de aglomeración (especialización y diversidad) se mantienen positivos y significativos, aunque es cierto que su magnitud aumenta al estimar por GMM. Así, incluso cuando utilizamos el estimador GMM para controlar posibles problemas de endogeneidad y causalidad inversa los resultados son robustos.

Arellano y Bond (1991) destacan la importancia de pasar algunas pruebas para validar el uso del método GMM. En particular, la prueba del retardo de segundo orden (AR (2)) para contrastar la correlación serial en los residuos diferenciados y las pruebas de Hansen y Sargan para chequear validez de los instrumentos utilizados. Las últimas filas de la columna (5) reportan los p-valores de estos tres tests, mostrando que los instrumentos son válidos y no se encuentra la correlación serial.

Tabla 9. Crecimiento de la población, 2000–2015

	Panel EF (1)	Panel EF (2)	Panel EF (3)	Panel EF (4)	GMM en diferencias (5)
Municipios:	Todos	Todos	Población >=3000	Población >=3000	Población >=3000
Pob. Relativa _{t-1}		-0,132*** (0,014)		-0,142*** (0,010)	-0,098* (0,050)
Especialización	-0,001 (0,004)	-0,001 (0,004)	0,046** (0,017)	0,021* (0,011)	0,151** (0,063)
Diversidad	-0,000 (0,002)	-0,001 (0,002)	0,018*** (0,006)	0,006* (0,004)	0,059** (0,028)
Tamaño	0,003 (0,002)	-0,000 (0,002)	0,005 (0,008)	0,003 (0,005)	-0,029 (0,040)
Densidad	-0,002 (0,002)	0,001 (0,003)	0,014 (0,009)	0,016*** (0,004)	0,005 (0,032)
EF municipales	Sí	Sí	Sí	Sí	No
EF temporales	Sí	Sí	Sí	Sí	Sí
Observaciones	7.196	7.196	593	593	561
Municipios	577	577	46	46	46
R ²	0,016	0,109	0,187	0,447	
AR(2), p-valor					0,482
Test de Sargan, p-valor					0,488
Test de Hansen, p-valor					0,967

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por ciudad. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%.

3.3.- Conclusiones

El objetivo de este apartado es analizar el efecto de la concentración espacial de las empresas aragonesas sobre el crecimiento del empleo sectorial a nivel local y la población del municipio. En particular, analizamos la dicotomía localización-

urbanización. Las economías de localización son ventajas consideradas internas al sector al que pertenece la empresa y que suponen un beneficio por el hecho de estar localizada cerca de otras empresas del mismo sector (especialización), mientras que las economías de urbanización son ventajas que se generan de forma externa al sector y representan un beneficio para todas las empresas localizadas en la misma área independientemente del sector al que pertenecen (diversidad).

En primer lugar, estimamos el efecto de estas economías de aglomeración sobre el empleo sectorial a nivel local para las ramas de actividad de cada uno de los sectores principales de actividad (agricultura, industria, construcción y servicios). En el caso de la agricultura no encontramos ningún efecto, ya que espacialmente las empresas del sector están distribuidas de una manera muy dispersa. En los otros tres grandes sectores (industria, construcción y servicios) la variable que mide la especialización presenta un signo negativo y significativo, lo que indicaría que a mayor especialización de los municipios en las ramas de actividad de estos sectores menor crecimiento del empleo.

Anteriores estudios como el de Combes (2000) también encuentran un efecto negativo de la especialización, que es lo contrario a lo que predicen los modelos teóricos. La explicación a este resultado es que una elevada especialización productiva implica poca flexibilidad y peor adaptabilidad de productos, tecnologías e infraestructuras cuando el sector está en declive. Si tenemos en cuenta que durante el periodo de análisis, 2000–2015, que incluye la profunda crisis económica que comienza en 2008, se produjo una importante destrucción de empleo en estos sectores, nuestros resultados apuntan a que los municipios aragoneses que estaban más especializados en estos sectores sufrieron una destrucción de empleo superior a la media de Aragón. Es decir, aunque se daban las circunstancias para que se generaran economías de localización internas al sector ya que las empresas tanto en la industria como en el sector servicios están espacialmente concentradas (como corroboramos en el análisis del apartado 2), el devenir de la economía general y el impacto de la Gran Recesión provocó que la especialización restara en lugar de sumar al crecimiento del empleo sectorial a nivel local.

Respecto a la variable que mide la diversidad, y que por lo tanto está relacionada con las economías de urbanización, no encontramos ningún efecto significativo en ninguno de los sectores productivos. La única excepción es el sector servicios cuando restringimos el análisis a las ciudades más grandes (con más de 3.000 habitantes), ya que en ese caso encontramos un efecto positivo de la diversidad en la creación de empleo. En consecuencia, para que se genere este efecto positivo es necesaria una cierta escala o tamaño económico del municipio. Este resultado indicaría evidencia favorable a las economías de urbanización: a mayor diversidad en las distintas ramas

del sector servicios mayor crecimiento en el empleo de otras ramas del sector en el mismo municipio, lo que estaría señalando la existencia de vínculos de oferta y demanda entre las distintas actividades del sector servicios.

En segundo lugar, estimamos el efecto de las economías de aglomeración sobre el crecimiento de la población, encontrando evidencia de un efecto positivo significativo de las economías de localización y urbanización en el crecimiento de los municipios aragoneses, pero solo en las grandes ciudades. Si consideramos todos los municipios no se encuentra ningún efecto significativo, lo que implica que las economías de aglomeración requieren una escala mínima, que en este caso cuantificamos en 3.000 habitantes (46 de los 731 municipios de Aragón). En estos grandes municipios, el efecto de las economías de localización (especialización) es más del doble que el de las economías de urbanización (diversidad).

El modelo de crecimiento urbano de Glaeser *et al.* (1995), explicado en el apartado 3.2.1, nos ayuda a comprender los factores de oferta o demanda que pueden estar detrás de este efecto positivo de las economías de aglomeración. De acuerdo a este modelo teórico, hay tres vías por las que una característica local (en este caso, la especialización o la diversidad productiva, que representan las economías de aglomeración) puede influir en el crecimiento del municipio. Primero, porque esta variable se haya vuelto más importante en el proceso de producción. Teniendo en cuenta los resultados obtenidos en el análisis del crecimiento sectorial del empleo a nivel local, que indicaban que la especialización tenía un efecto negativo en el crecimiento del empleo y la diversidad ningún efecto significativo excepto para los servicios en los municipios grandes, esta vía quedaría descartada en el caso de los municipios aragoneses.

Segundo, desde el punto de vista de la demanda, la especialización y la diversidad pueden haberse vuelto más importantes para los consumidores al reducir el coste de la vida o al aumentar el volumen de amenidades locales disponibles. Una mayor especialización a nivel municipal implica más empresas de un mismo sector produciendo bienes similares; desde el punto de vista de la demanda, eso puede incrementar la utilidad de los individuos debido a su estructura de preferencias y a lo que Paul Krugman denominó amor por la variedad (*“love-of-variety effect”*), es decir, que los consumidores prefieren consumir distintas variedades de un mismo bien diferenciado (por ejemplo, los individuos compran ropa de distintas marcas). Respecto a la diversidad productiva, una elevada presencia de empresas de diferentes sectores que ofrecen bienes y servicios distintos puede atraer tanto a consumidores como a otras empresas que prefieren tener a su disposición dentro de su mismo municipio un amplio elenco de bienes y servicios disponibles.

Tercero, podría ser que la especialización o la diversidad productiva contribuyera a incrementar la tasa de crecimiento tecnológico. A pesar de que nuestras estimaciones del crecimiento sectorial del empleo a nivel local indicaran un efecto negativo de la especialización y un efecto no significativo de la diversidad, no podemos descartar esta tercera vía. La concentración empresarial, ya sea de empresas del mismo o diferentes sectores, contribuye a facilitar el intercambio de conocimientos sobre productos, procesos e innovaciones (los denominados *knowledge spillovers* o desbordamientos del conocimiento). De hecho, una de las razones por las que el nivel del empleo podría reducirse es que, debido al cambio tecnológico, se están produciendo innovaciones técnicas que permitan reducir el trabajo necesario en el proceso de producción. Es decir, aunque la especialización tuviera un efecto negativo sobre el empleo, como revelan nuestras estimaciones, puede influir sobre la oferta a través del cambio tecnológico, y por lo tanto afectar al crecimiento de la población. Esta transferencia de conocimiento suele ser más fácil a través de la especialización (es más sencillo compartir procesos productivos o conocimientos de los trabajadores dentro del mismo sector), lo que implicaría que el efecto sería de mayor magnitud en el caso de la especialización que de la diversidad.

En resumen, tanto las economías de localización (especialización) como las de urbanización (diversidad) pueden influir en la oferta y en la demanda y, a través de estas, en el crecimiento de la población local. Aunque en nuestras estimaciones ambos efectos son estimados de forma conjunta con un único coeficiente para cada variable, de acuerdo con nuestros resultados ambos efectos sobre el crecimiento de la población serían positivos, pero el impacto de la especialización sería superior al de la diversidad.

4. Resumen final y conclusiones

Se denomina “economías de aglomeración” a las ventajas que aporta la proximidad física con otras personas y empresas. Concretamente, en este estudio nos centramos en el análisis de la dicotomía localización-urbanización. Las economías de localización están relacionadas con la concentración empresarial dentro de un sector específico y suponen un beneficio por el hecho de estar localizado cerca de otras empresas del mismo sector (especialización), y las economías de urbanización operan a través de la concentración general de la actividad económica y representan un beneficio para todas las empresas localizadas en la misma área independientemente del sector al que pertenecen (diversidad).

En este estudio se aborda (1) el análisis empírico de la concentración espacial de las empresas aragonesas para las principales ramas de actividad a partir de información

geográfica de su localización física, y (2) el efecto de la concentración empresarial sobre el crecimiento de (i) el empleo y (ii) la población en los municipios aragoneses.

A partir de la información que proporciona la base de datos SABI (Sistema de Análisis de Balances Ibéricos) de más de 20.000 empresas activas con sede en Aragón, se realiza un análisis de la concentración espacial de las empresas aplicando la metodología de Duranton y Overman (2005, 2008), que considera el espacio continuo para estimar la densidad (concentración) de empresas a partir de todas las distancias físicas bilaterales (latitud y longitud) entre las empresas de una misma rama de actividad.

Considerando los datos geográficos de las empresas aragonesas en 2017 encontramos distintos patrones en la distribución espacial de empresas dependiendo del tipo de actividad. En general, los resultados obtenidos permiten clasificar las secciones o principales ramas de actividad (según la CNAE-2009) en tres tipos de distribución espacial: (i) aquellas industrias que comienzan con una dispersión significativa en distancias cortas, pero que terminan exhibiendo un patrón aleatorio a medida que aumenta la distancia, (ii) aquellas industrias que muestran una concentración significativa para distancias cortas, pero que a medida que aumenta la distancia pasan a presentar un patrón no diferente de la aleatoriedad, y (iii) aquellas industrias que no muestran ningún patrón significativo de localización.

En el primer caso nos encontramos con actividades del sector primario (agricultura), construcción y transporte y almacenamiento. El segundo caso corresponde a las actividades del sector servicios, y en el tercer caso encontramos industrias de suministro y abastecimiento de agua y energía. Ambos tipos de patrones (dispersión o concentración) operan en distancias muy cortas (10–15 kilómetros), pero para distancias más elevadas el resultado que suele emerger es una distribución de empresas espacial no diferente de la aleatoriedad.

La industria manufacturera es la única actividad económica que no encaja en ninguno de los tipos de esta clasificación. En distancias cortas inferiores a 8 kilómetros muestra un patrón de dispersión espacial, pero para distancias amplias (desde 10 hasta 45 kilómetros) el patrón es de localización (concentración) de empresas, resultado que relacionamos con la geografía particular de Aragón que cuenta con un extenso territorio y grandes distancias entre los municipios más poblados donde se sitúan las principales industrias.

Posteriormente se analiza el efecto de la concentración y diversidad industrial en el crecimiento del empleo y la población a nivel municipal, aplicando rigurosos modelos econométricos y un panel de datos desde el año 2000 al 2015. Primero estimamos el efecto de las economías de aglomeración sobre el empleo sectorial a nivel local para las ramas de actividad de cada uno de los sectores principales de actividad,

encontrando un efecto negativo y significativo de la especialización en los tres grandes sectores de actividad (industria, construcción y servicios). Teniendo en cuenta que el periodo de análisis incluye la Gran Recesión la cual comenzó en 2008 y supuso una gran destrucción de empleo en estos sectores, nuestros resultados pueden indicar que los municipios aragoneses que estaban más especializados en estos sectores sufrieron una destrucción de empleo superior a la media de Aragón. Respecto a la variable que mide la diversidad, relacionada con las economías de urbanización, no encontramos ningún efecto significativo en ninguno de los sectores productivos, a excepción de un efecto positivo sobre la creación de empleo en las actividades del sector servicios cuando restringimos el análisis a los municipios de más de 3.000 habitantes. Por tanto, a mayor diversidad en las distintas ramas del sector servicios mayor crecimiento en el empleo de otras ramas del sector en el mismo municipio siempre que se trate de un municipio grande, lo que implicaría la existencia de vínculos de oferta y demanda entre las distintas actividades del sector servicios.

Por último, partiendo de un modelo teórico de crecimiento urbano estimamos el efecto de las economías de aglomeración sobre el crecimiento de la población, encontrando evidencia de un efecto positivo significativo de las economías de localización y urbanización en el crecimiento de los municipios aragoneses, pero solo en los municipios de más de 3.000 habitantes (46 del total de 731 municipios en Aragón). Además, la magnitud del efecto de la especialización sería superior al de la diversidad. Desde un punto de vista teórico, aunque el efecto previamente estimado de las economías de aglomeración sobre el empleo sectorial local sea o negativo (especialización) o no significativo (diversidad), todavía quedan dos vías adicionales por las que ambas variables pueden influir sobre el crecimiento de la población local a través de la demanda y la oferta (cambio tecnológico), consiguiendo un efecto neto positivo como el que obtenemos con nuestras estimaciones.

En definitiva, si bien encontramos patrones espaciales de localización (concentración) en la distribución de las empresas dentro de algunos sectores específicos (básicamente, industria y servicios) que podrían generar economías de aglomeración, la evidencia que finalmente obtenemos a favor de las economías de localización y urbanización es limitada, quedando solamente restringida a los municipios más grandes de Aragón. Por un lado, esto podría deberse a que las economías de aglomeración requieren una escala de tamaño mínima, que en este caso cuantificamos en 3.000 habitantes; es decir, cualquier política pública industrial con el ánimo de generar empleo o atraer población puede resultar eficaz, pero solo si los municipios a los que va dirigida tienen un tamaño superior a este límite de población. Para los municipios de menor tamaño no encontramos ningún efecto significativo durante el periodo considerado; sin embargo, nuestros resultados se limitan al periodo de análisis 2000–2015, lo que nos impide pronosticar cuál hubiera sido el resultado de

estas políticas en los municipios menos poblados si las condiciones económicas del periodo no hubieran sido tan negativas. Es de esperar que si el entorno económico fuera de creación de empleo en lugar de destrucción la fuerza de las economías de aglomeración se reforzaría, en lugar de atenuarse.

Por otro lado, si incluso en un periodo tan adverso en términos económicos hemos sido capaces de encontrar algún efecto positivo de las economías de aglomeración, es posible que en periodos de bonanza económica las economías de aglomeración puedan jugar un papel clave en la creación de empleo. Esto es relevante para el caso de Aragón ya que esta región cuenta con una elevada concentración de la actividad económica y de la población en la capital, Zaragoza, y esa tendencia a la aglomeración dura ya varias décadas. Tampoco hay que descartar los posibles efectos externos que genera dicha aglomeración (positivos o negativos), los cuales sería necesario cuantificar, ya que, en caso de persistir, pueden conducirnos a una distribución cada vez más desequilibrada de la actividad económica.

5. Referencias

- Albert, J. M., M. R. Casanova, y V. Orts, (2012). Spatial location patterns of Spanish manufacturing firms. *Papers in Regional Science*, 91(1): 107–136.
- Arellano, M., y S. Bond, (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2): 277–297.
- Beaudry, C., y A. Schiffauerova, (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2): 318–337.
- Blien, U., J. Suedekum, y K. Wolf, (2006). Local employment growth in West Germany: A dynamic panel approach. *Labour Economics*, 13(4): 445–458.
- Combes, P. P., (2000). Economic structure and local growth: France, 1984–1993. *Journal of Urban Economics*, 47(3): 329–355.
- Combes, P. P. y L. Gobillon, (2015). The empirics of agglomeration economies. En G. Duranton, V. Henderson, y W. Strange, (editores), *Handbook of Regional and Urban Economics*, 5: 247–348.
- de Vor, F., y H. L. de Groot, (2010). Agglomeration externalities and localized employment growth: the performance of industrial sites in Amsterdam. *The Annals of Regional Science*, 44(3): 409–431.
- Duranton, G., y H. G. Overman, (2005). Testing for Localization Using Microgeographic Data. *Review of Economic Studies*, 72: 1077–1106.

- Duranton, G., y H. G. Overman, (2008). Exploring the detailed location patterns of U.K. manufacturing industries using microgeographic data. *Journal of Regional Science*, 48(1): 213–243.
- Duranton, G., y D. Puga, (2004). Micro-foundations of urban agglomeration economies. En Henderson, Vernon y Thisse, Jacques-Francois (editores), *Handbook of Regional and Urban Economics*, 4: 2063–2117. North-Holland, Amsterdam.
- Glaeser, E. L., (2000). The New Economics of Urban and Regional Growth, En G. L. Clark, M. P. Feldman, y M. Gertler, (editores), *The Oxford Handbook of Economic Geography*, Oxford: Oxford University Press, 83–98.
- Glaeser, E. L., H. Kallal, J. Scheinkman, y A. Shleifer, (1992). Growth in cities. *Journal of Political Economy*, 100(6): 1126–1152.
- Glaeser, E. L., J. A. Scheinkman, y A. Shleifer, (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, 36: 117–143.
- Glaeser, E. L. y J. Shapiro, (2003). Urban Growth in the 1990s: Is city living back? *Journal of Regional Science*, 43(1): 139–165.
- González Pampillón, N., (2018). *Essays on Urban Economics*. Tesis Doctoral, Universidad de Barcelona.
- Groot, H. L., J. Poot, y M. J. Smit, (2015). Which Agglomeration Externalities Matter Most and Why? *Journal of Economic Surveys*, 0(0): 1–27.
- Heckman, J. J., (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*, 5(4): 475–492.
- Heckman, J. J., (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1): 153–161.
- Henderson, J., A. Kuncoro, y M. Turner, (1995). Industrial development in cities. *Journal of Political Economy*, 103(5): 1067–1090.
- Informe Económico de Aragón, (2016). Fundación Basilio Paraíso de la Cámara de Comercio, Industria y Servicios de Zaragoza.
- Informe sobre la situación económica y social de Aragón en 2017, (2018). Consejo Económico y Social de Aragón.
- Informe socioeconómico de la década 2001-2010 en Aragón, (2012). Consejo Económico y Social de Aragón.
- Mameli, F., A. Faggian, y P. McCann, (2008). Employment growth in Italian local labour systems: Issues of model specification and sectoral aggregation. *Spatial Economic Analysis*, 3(3): 343–360.

- Marrocu, E., R. Paci, y S. Usai, (2013). Productivity growth in the old and new Europe: the role of agglomeration externalities. *Journal of Regional Science*, 53(3): 418–442.
- Marshall, A., (1890). *Principles of Economic*. Mc Millan, London.
- Melo, P. C., D. J. Graham, y R. B. Noland, (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics*, 39(3): 332–342.
- Ripley, B. D., (1976). The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability*, 13(2): 255–266.
- Roback, J., (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6): 1257–1278.
- Semykina, A., y J. Wooldridge, (2010). Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics*, 157(2): 375–380.
- Torres Gutiérrez, T., (2017). *Ensayos sobre Economías de Aglomeración*. Tesis Doctoral, Universidad de Barcelona.
- Viladecans-Marsal, E., (2004). Agglomeration economies and industrial location: city-level evidence. *Journal of Economic Geography*, 4(5): 565–582.

6. Apéndices

Apéndice A: Distribución espacial hasta la máxima distancia

Figura A1. Distribución espacial de empresas por secciones de actividad

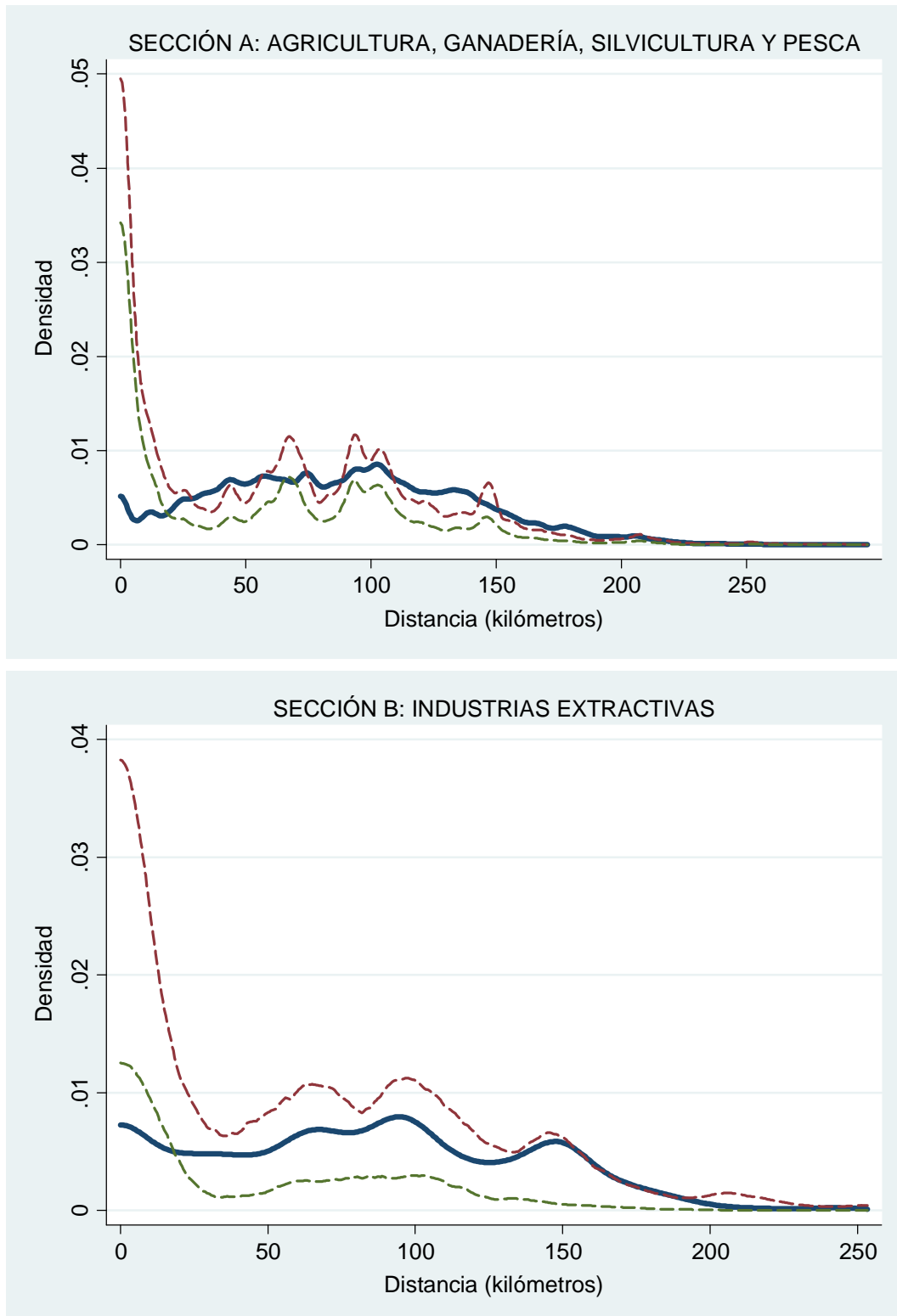
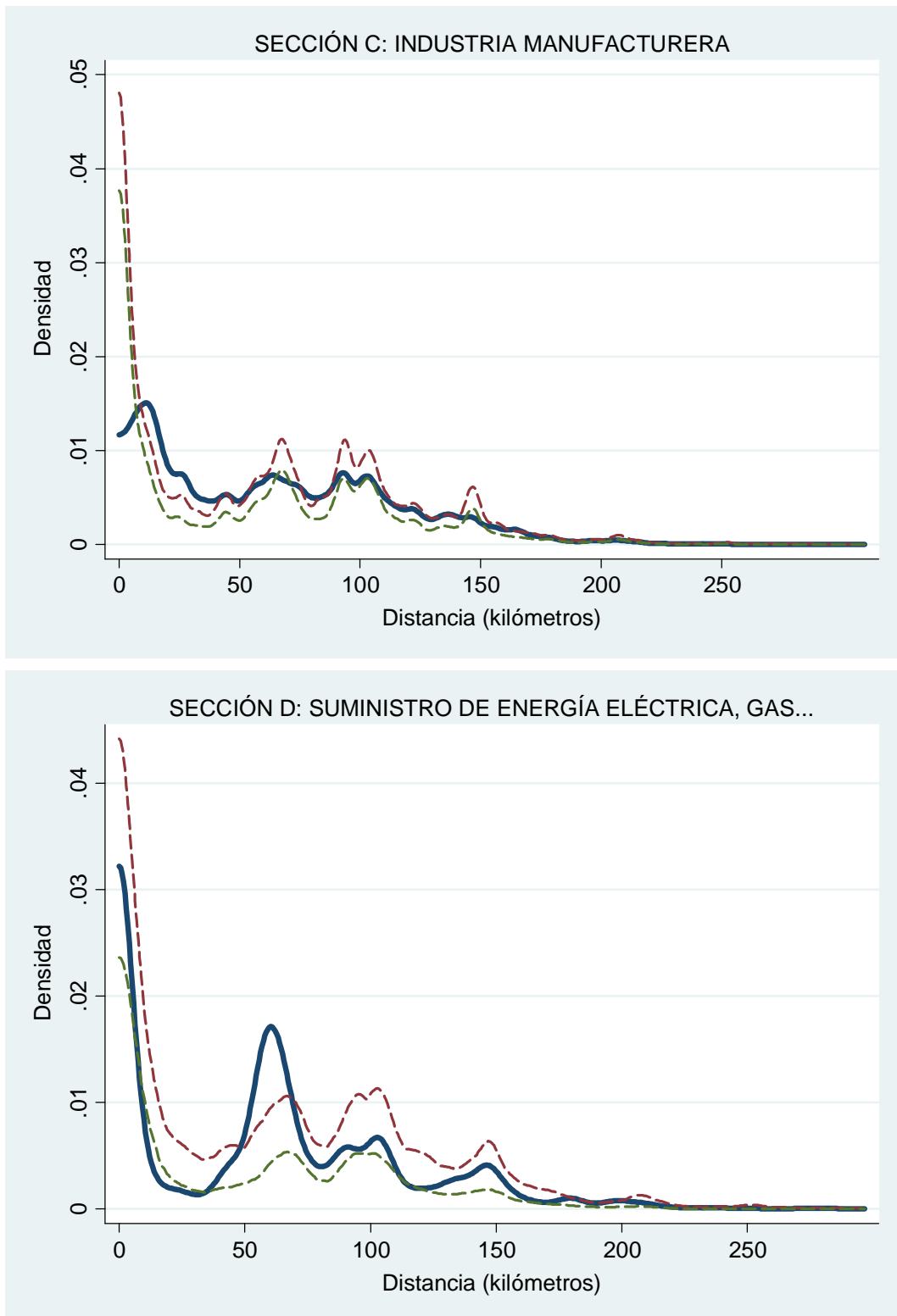
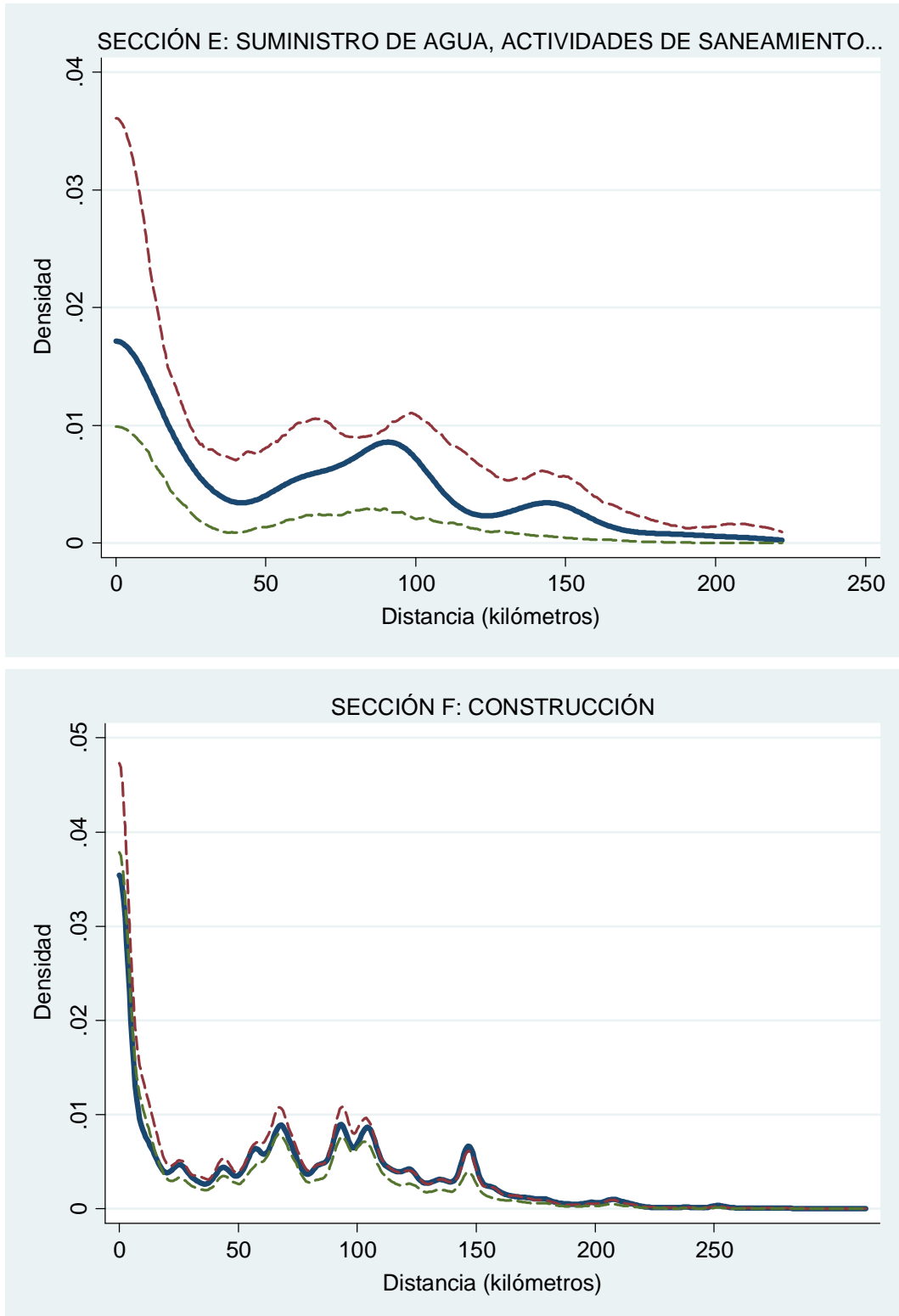


Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



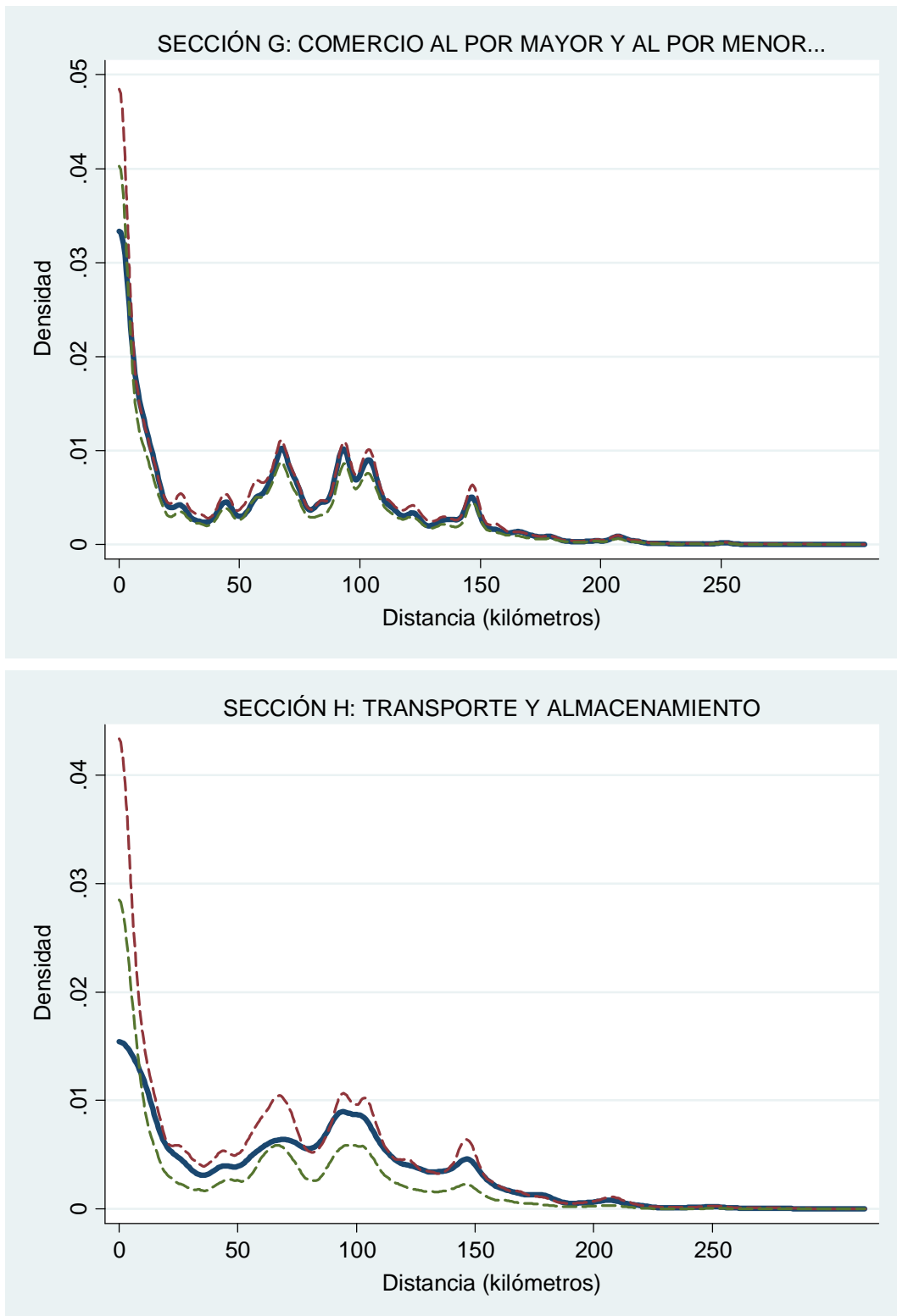
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



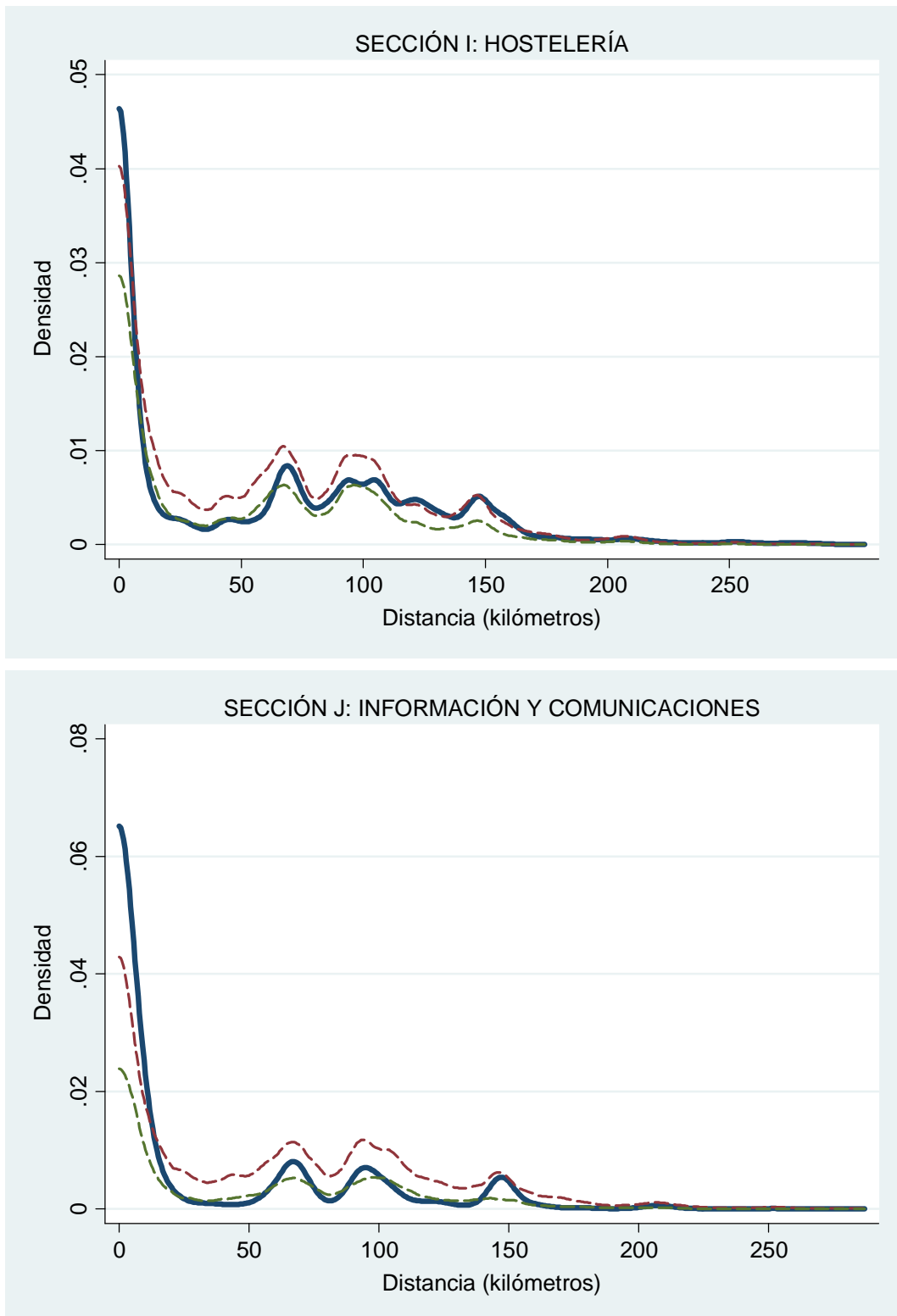
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



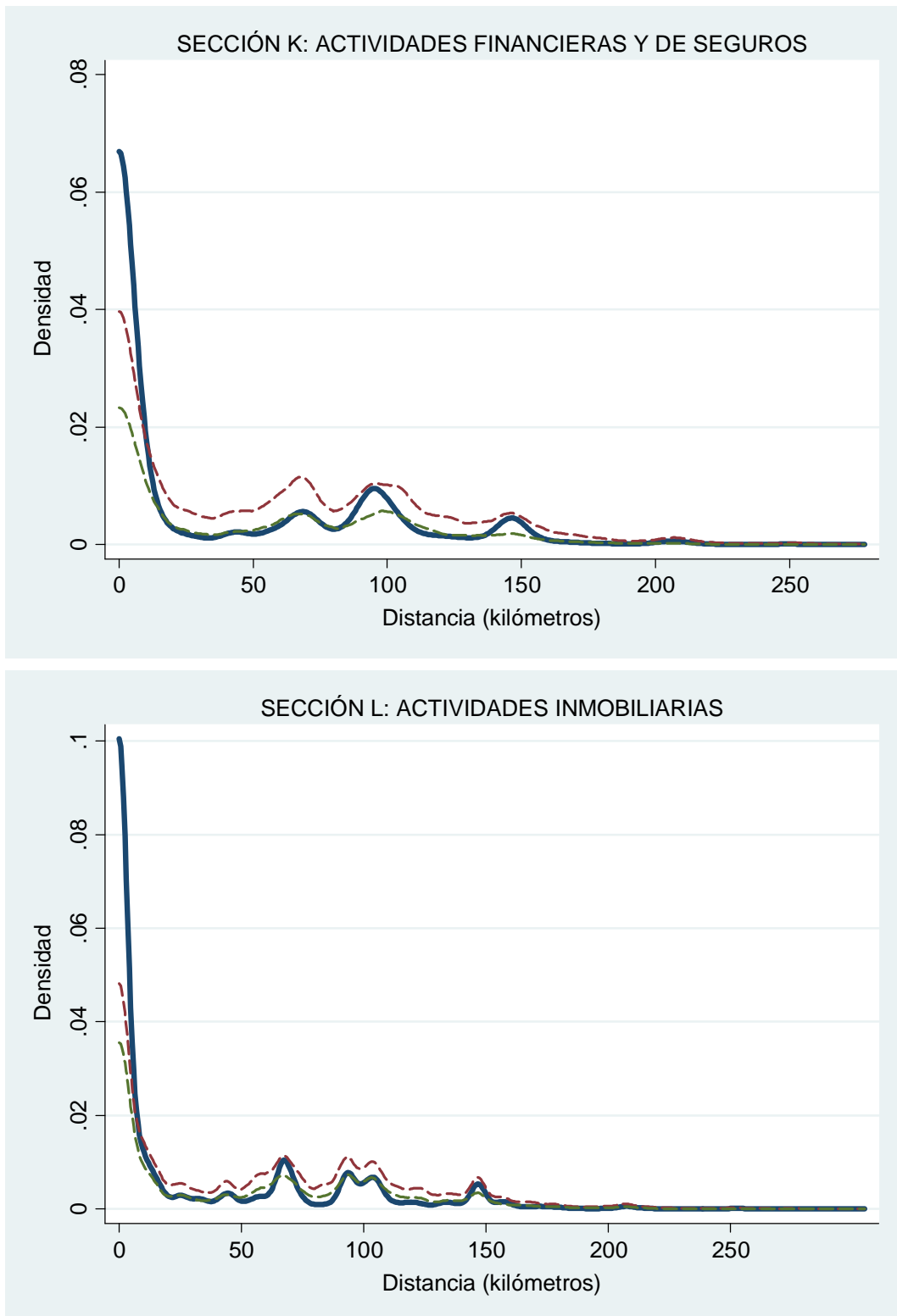
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



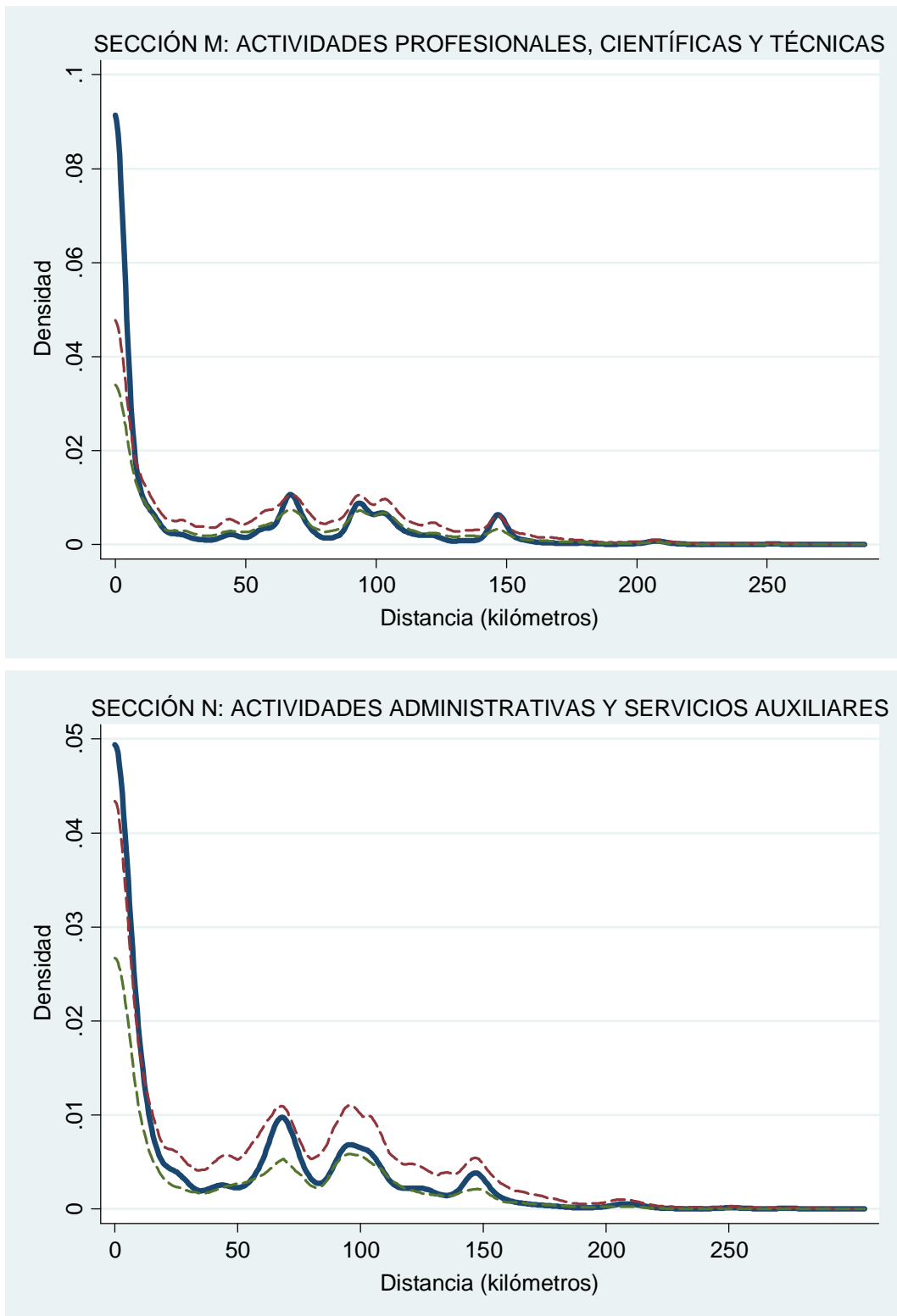
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



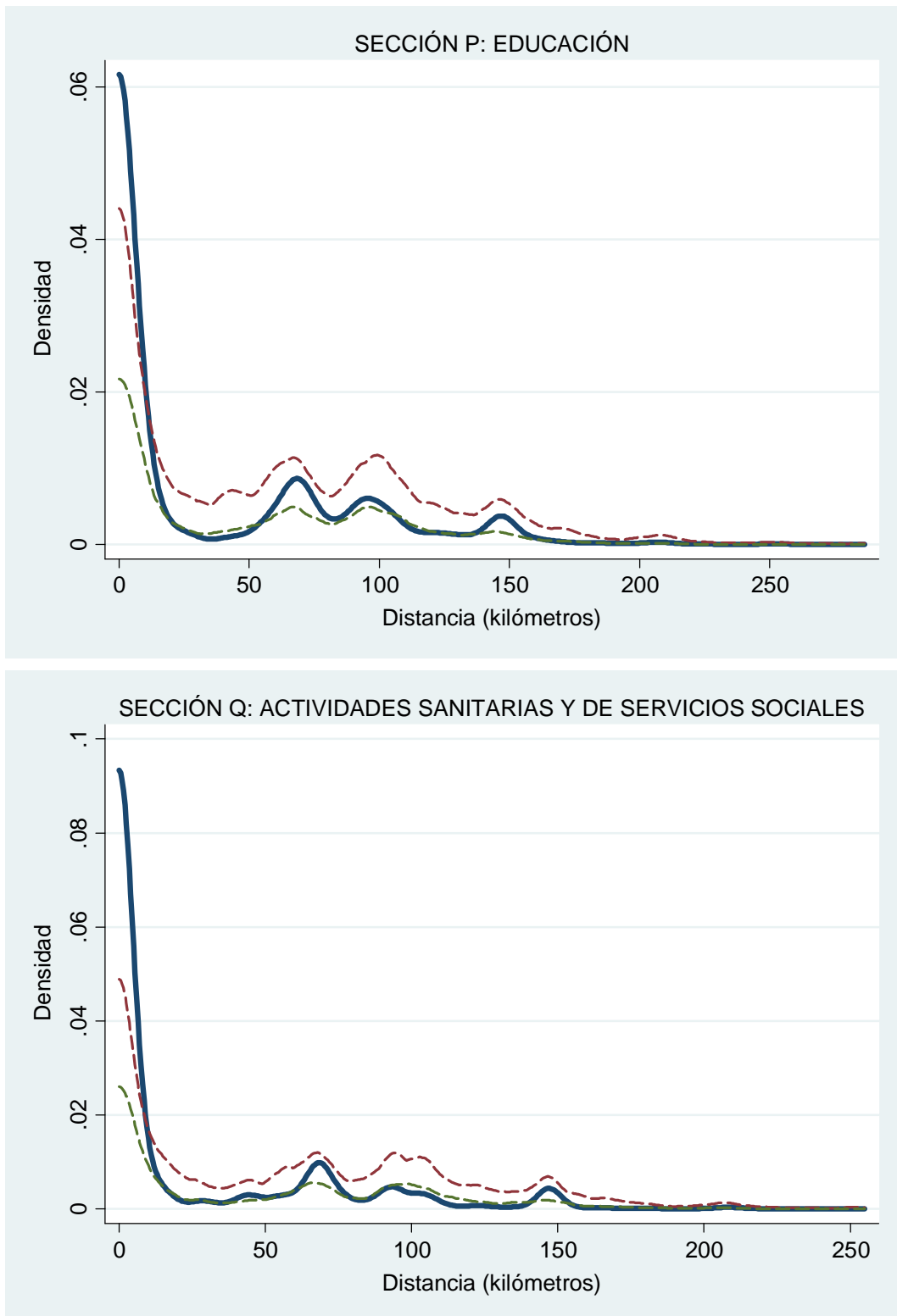
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



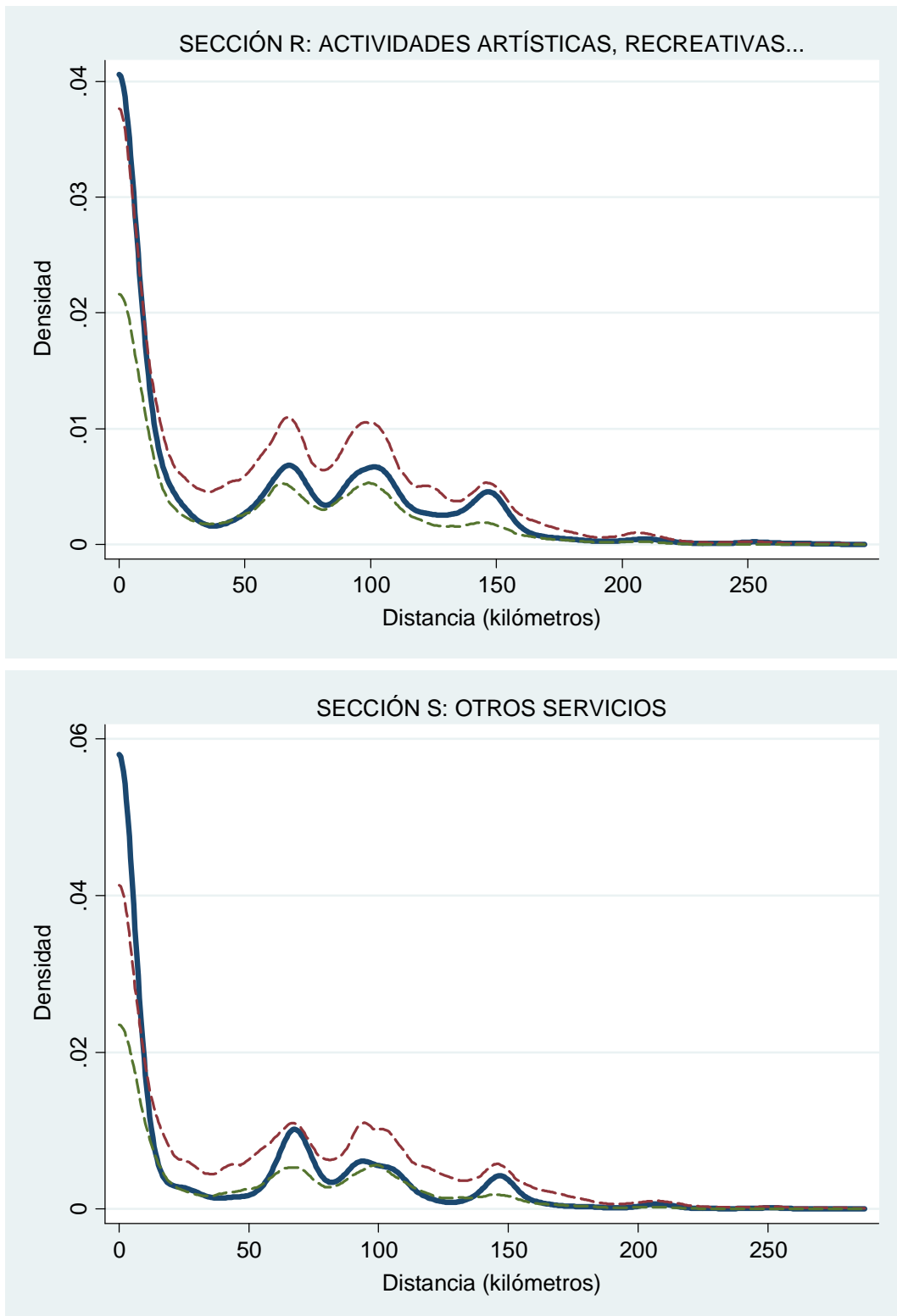
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



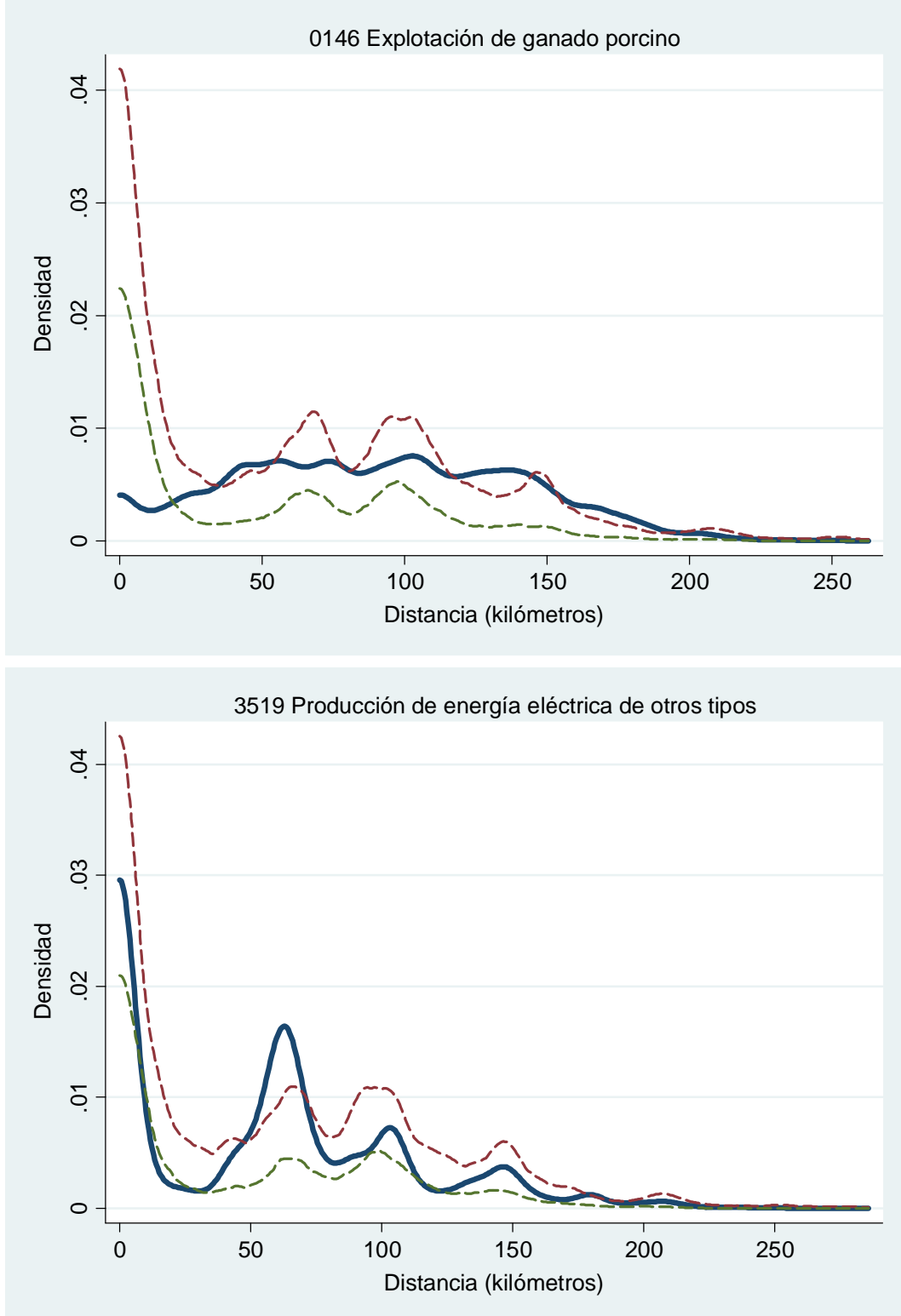
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A1. Distribución espacial de empresas por secciones de actividad (continuación)



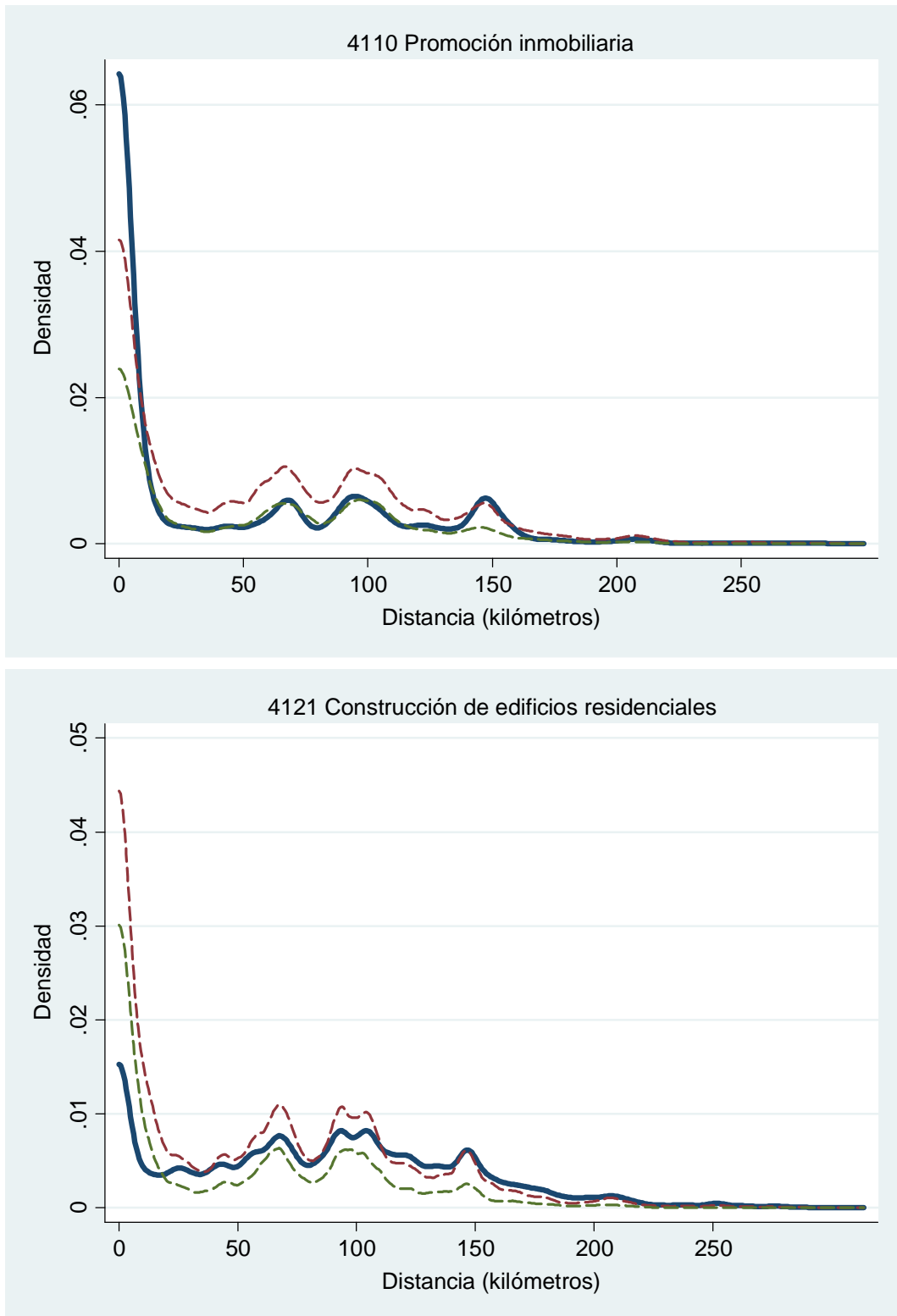
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas



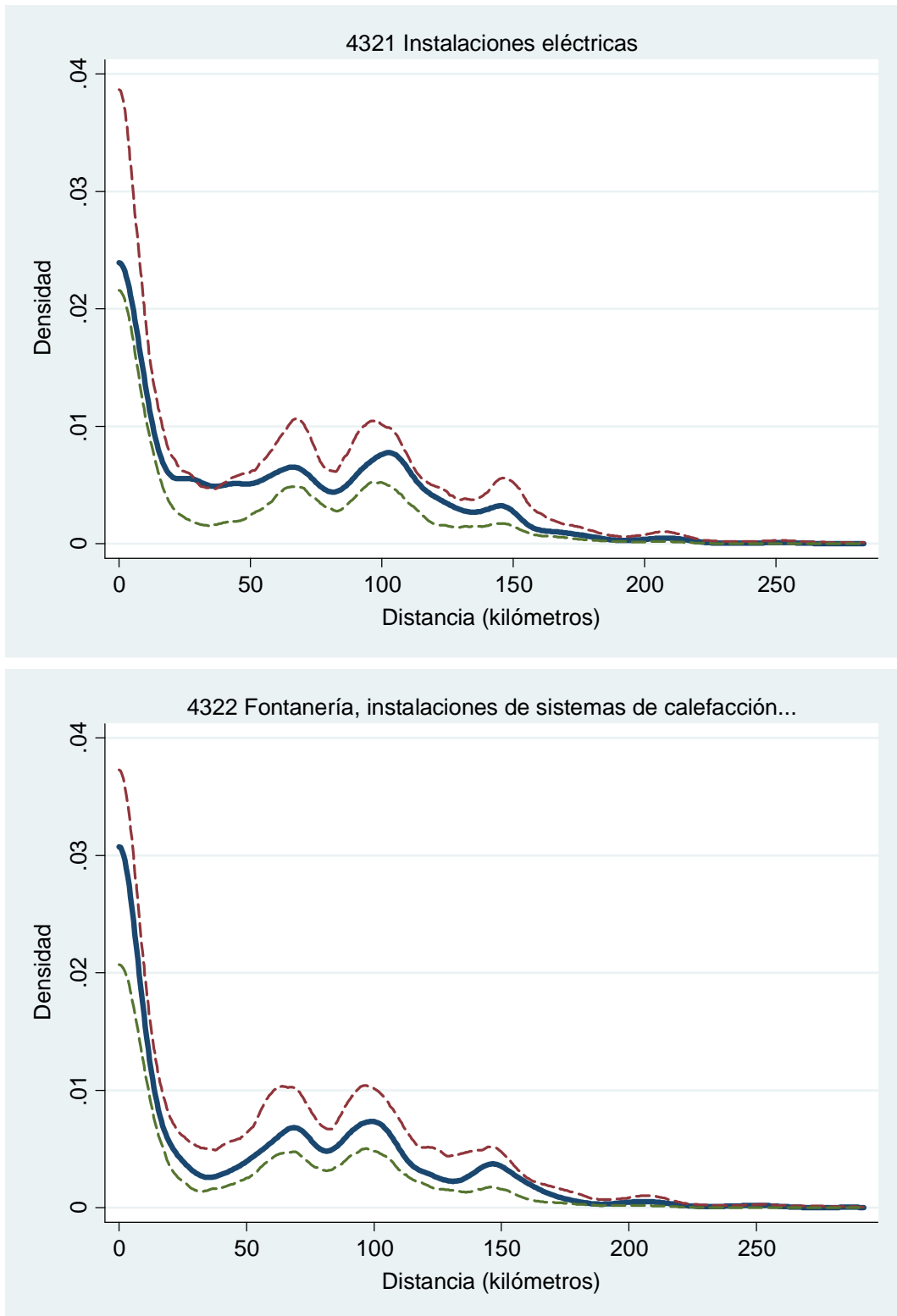
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



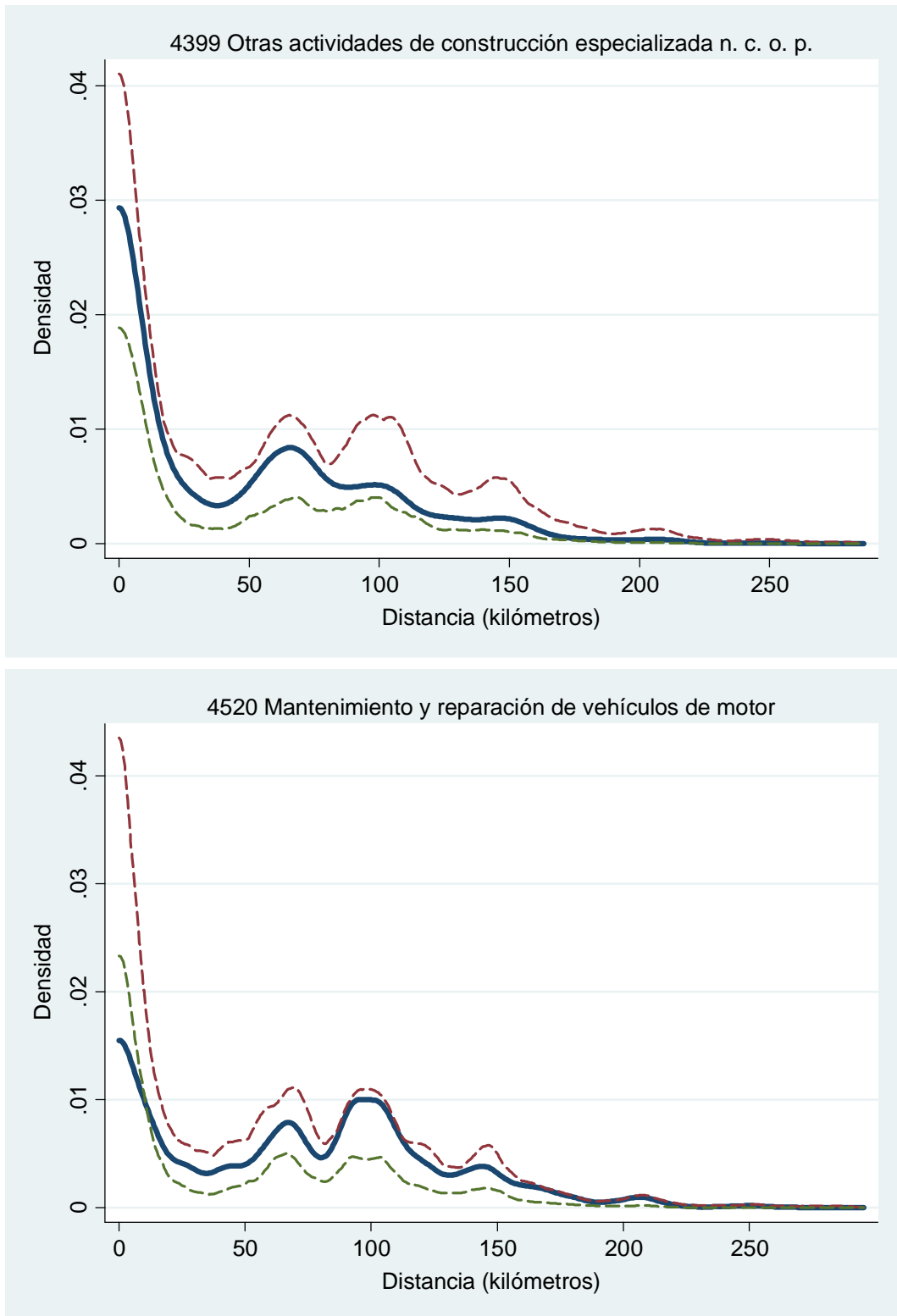
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



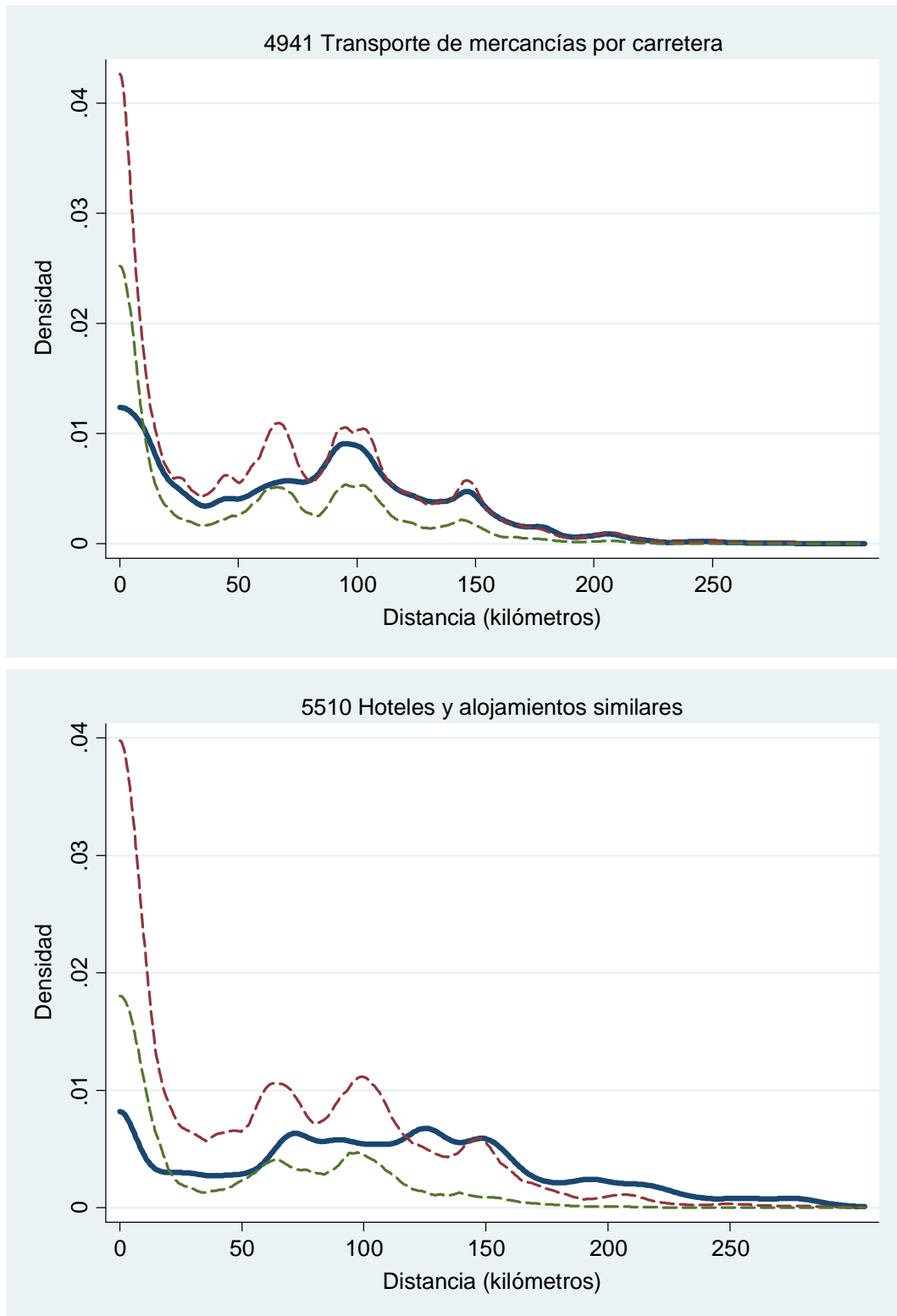
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



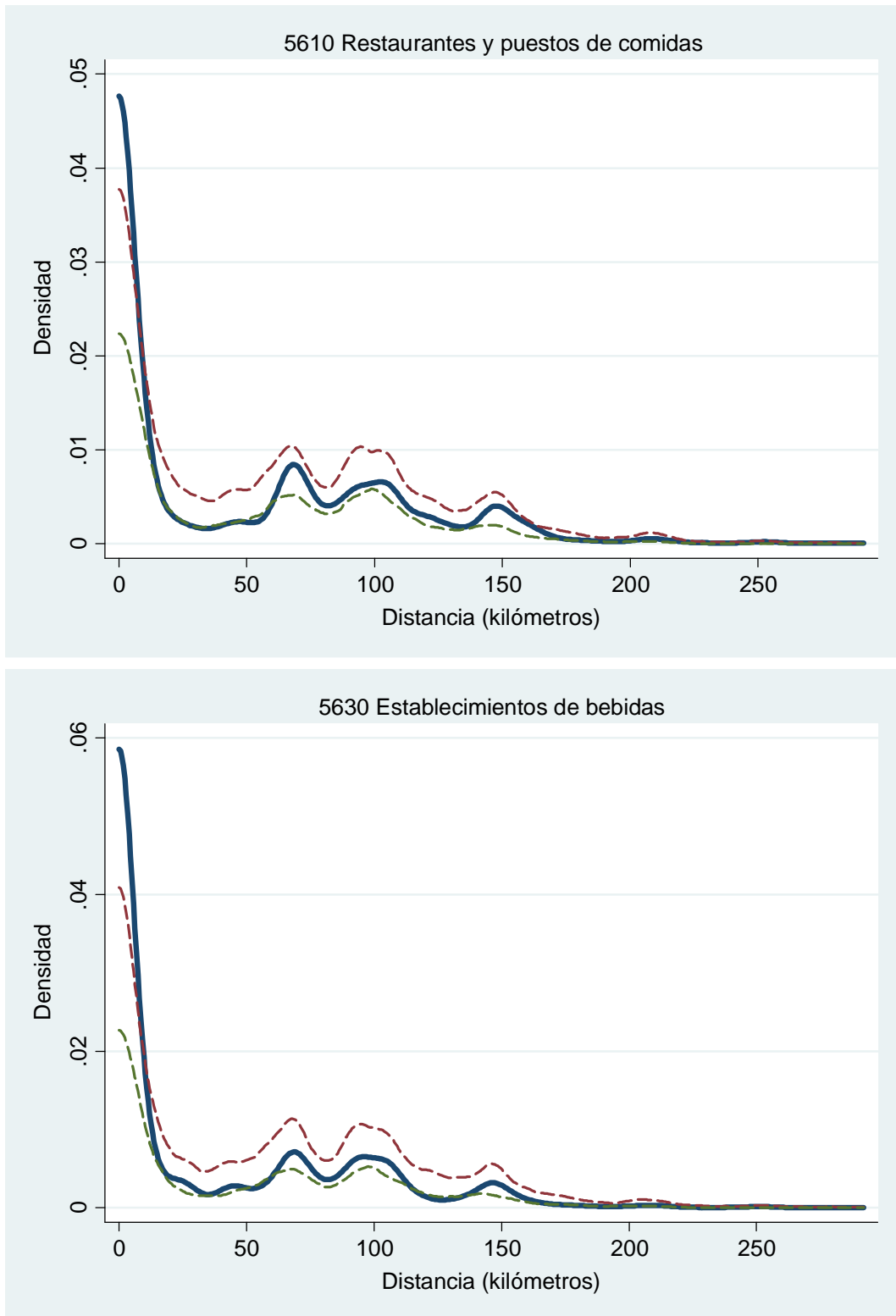
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



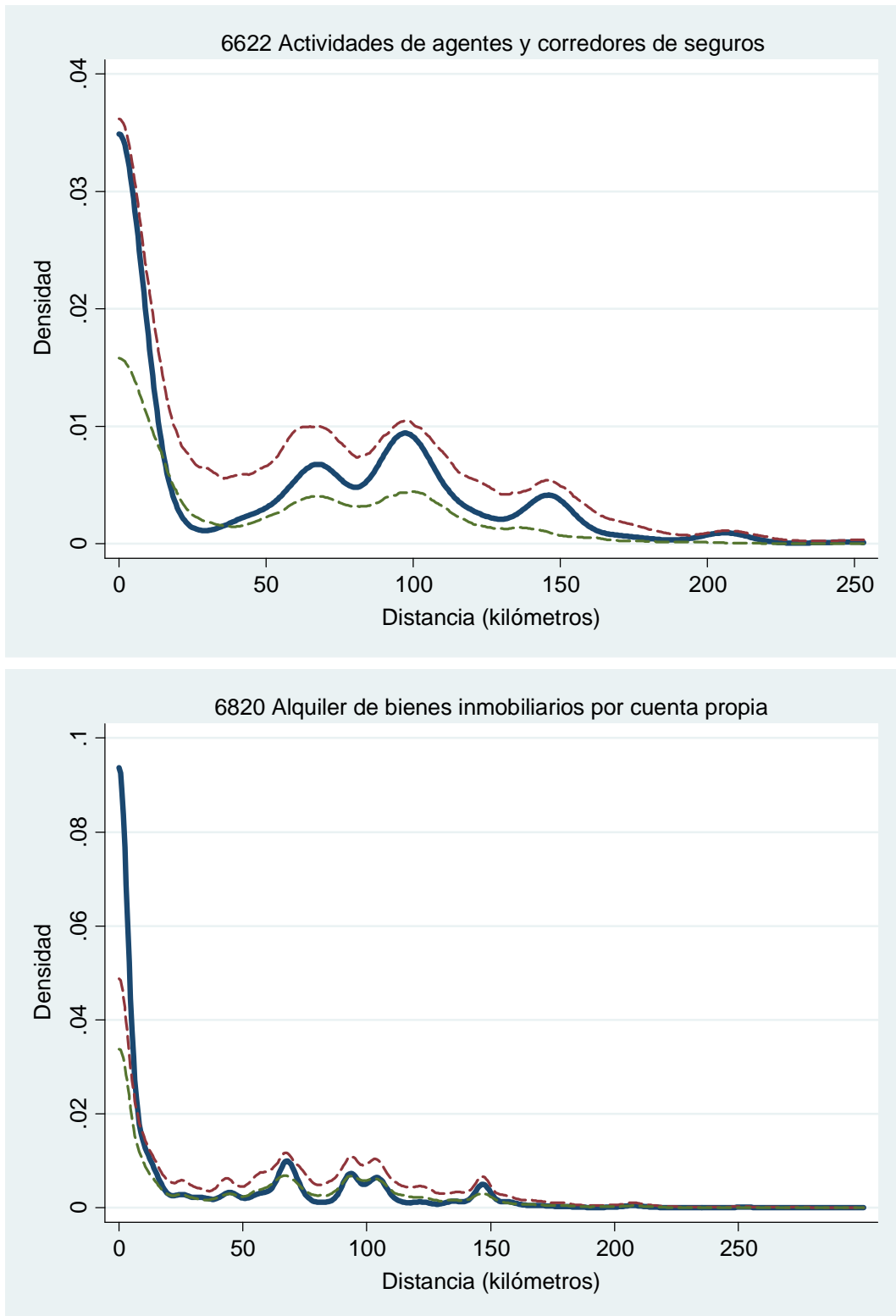
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



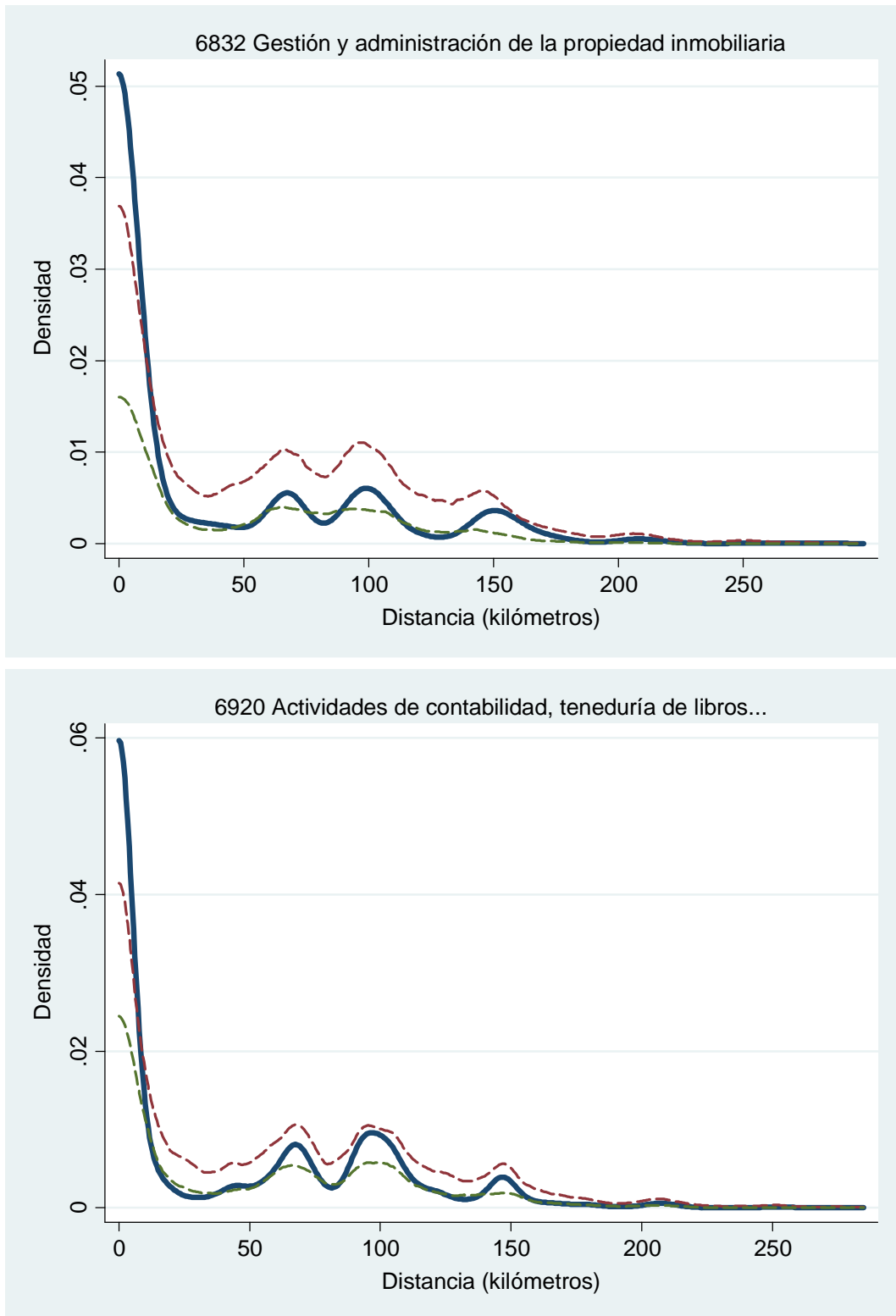
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



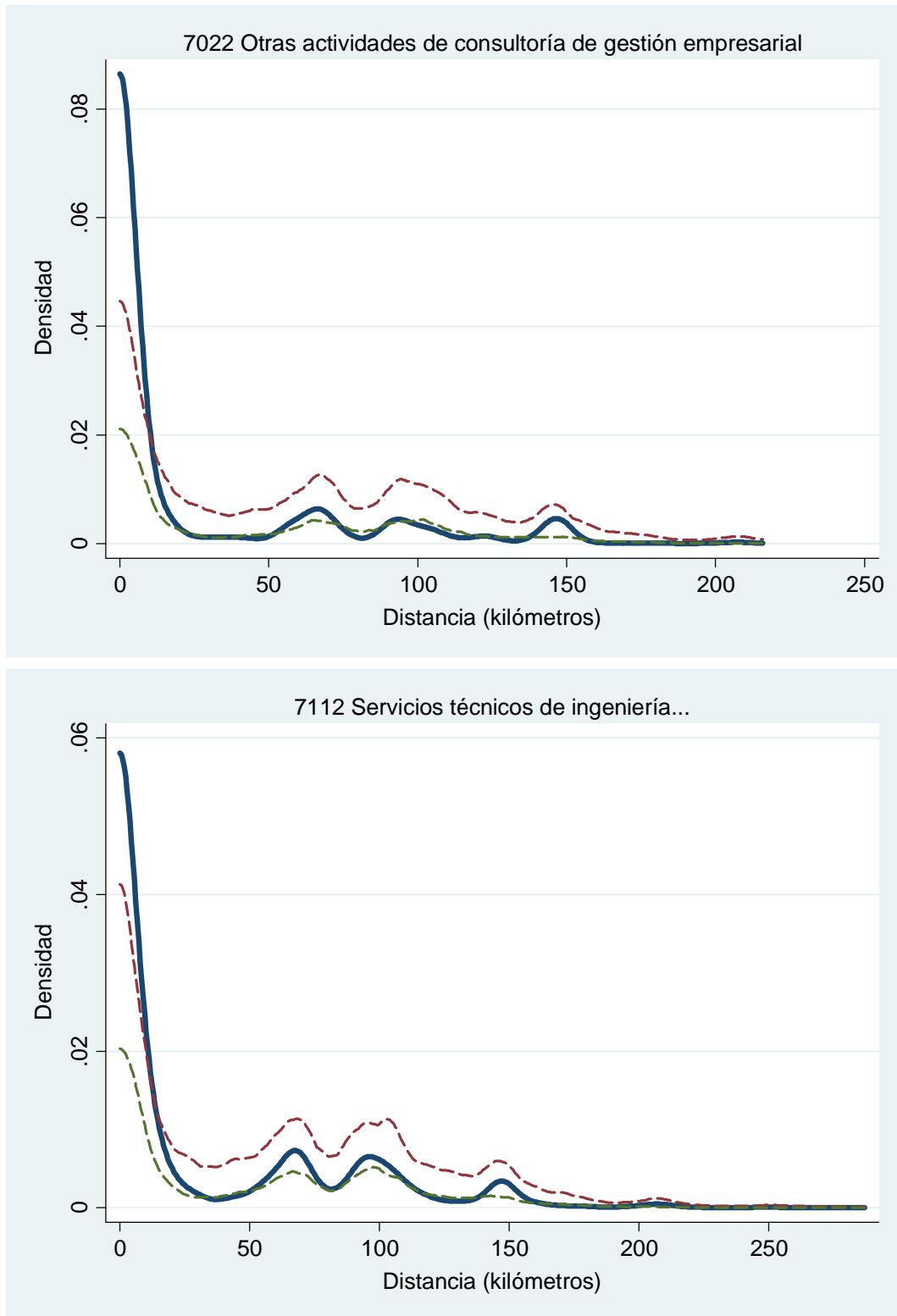
Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Figura A2. Distribución espacial de empresas en las ramas de actividad con más de 200 empresas (continuación)



Notas: K-densidades estimadas utilizando la metodología de Duranton y Overman (2005). Las líneas discontinuas representan las bandas globales de confianza al 95%, basadas en 2.000 simulaciones.

Apéndice B: Estimaciones por MCO del efecto de la aglomeración sobre el empleo sectorial local

Tabla B1. Resultados actividades agrícolas (actividades 01-03 CNAE-2009), MCO

	(1)	(2)	(3)	(4)	(5)
Municipios:	Todos	Todos	Todos	Sin Zaragoza	Población >=3000
Especialización	-0,042*** (0,010)	-0,013 (0,008)	-0,017* (0,009)	-0,014 (0,009)	-0,024 (0,024)
Diversidad	-0,004 (0,005)	-0,004 (0,006)	-0,001 (0,006)	-0,000 (0,006)	0,015 (0,015)
Tamaño	-0,060*** (0,014)	-0,082*** (0,013)	-0,086*** (0,016)	-0,097*** (0,014)	-0,061* (0,034)
Densidad	-0,019** (0,009)	-0,015** (0,007)	-0,004 (0,010)	0,002 (0,008)	-0,011 (0,019)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	3.296	3.296	3.296	3.252	652
R ²	0,053	0,098	0,107	0,115	0,152

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%. VIF (modelo 3, que incluye todos los controles) = 10,74.

Tabla B2. Resultados actividades industriales (actividades 10-33 CNAE-2009), MCO

	(1)	(2)	(3)	(4)	(5)
Municipios:	Todos	Todos	Todos	Sin Zaragoza	Población >=3000
Especialización	-0,020*** (0,004)	-0,026*** (0,004)	-0,030*** (0,005)	-0,031*** (0,005)	-0,030*** (0,008)
Diversidad	-0,001 (0,003)	0,001 (0,003)	0,001 (0,004)	0,000 (0,004)	-0,007 (0,006)
Tamaño	-0,048*** (0,005)	-0,040*** (0,006)	-0,044*** (0,007)	-0,045*** (0,007)	-0,050*** (0,012)
Densidad	-0,008*** (0,002)	-0,006*** (0,002)	-0,003 (0,003)	-0,003 (0,003)	-0,005 (0,005)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	12.820	12.820	12.820	12.475	5.829
R ²	0,032	0,041	0,046	0,046	0,046

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%. VIF (modelo 3, que incluye todos los controles) = 5,24.

Tabla B3. Resultados construcción (actividades 41-43 CNAE-2009), MCO

Municipios:	(1) Todos	(2) Todos	(3) Todos	(4) Sin Zaragoza	(5) Población >=3000
Especialización	-0,018** (0,008)	-0,020** (0,009)	-0,031*** (0,010)	-0,031*** (0,010)	-0,044 (0,033)
Diversidad	0,001 (0,005)	0,001 (0,005)	-0,001 (0,005)	-0,001 (0,005)	-0,016 (0,012)
Tamaño	-0,130*** (0,013)	-0,128*** (0,014)	-0,128*** (0,015)	-0,128*** (0,015)	-0,148** (0,058)
Densidad	-0,002 (0,004)	-0,002 (0,004)	0,000 (0,005)	0,000 (0,005)	-0,018 (0,014)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	6.164	6.164	6.164	6.119	779
R ²	0,068	0,088	0,095	0,095	0,139

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%. VIF (modelo 3, que incluye todos los controles) = 4,99.

Tabla B4. Resultados sector servicios (actividades 49-96 CNAE-2009), MCO

Municipios:	(1) Todos	(2) Todos	(3) Todos	(4) Sin Zaragoza	(5) Población >=3000
Especialización	-0,011** (0,004)	-0,016*** (0,004)	-0,025*** (0,005)	-0,037*** (0,006)	-0,040*** (0,009)
Diversidad	0,004 (0,003)	0,004 (0,003)	0,001 (0,004)	0,000 (0,004)	0,013* (0,007)
Tamaño	-0,104*** (0,006)	-0,098*** (0,007)	-0,105*** (0,008)	-0,099*** (0,008)	-0,110*** (0,011)
Densidad	0,001 (0,003)	0,002 (0,003)	0,001 (0,004)	0,000 (0,004)	0,008* (0,005)
RIM×Año	No	Sí	Sí	Sí	Sí
EF temporales	No	Sí	Sí	Sí	Sí
EF provinciales	No	No	Sí	Sí	Sí
EF sectoriales	No	No	Sí	Sí	Sí
Observaciones	14.374	14.374	14.374	13.753	7.989
R ²	0,063	0,077	0,088	0,092	0,108

Notas: Todas las variables en logaritmos. Desviaciones típicas robustas entre paréntesis, clúster por sector y municipio. Todos los modelos incluyen constante. Significativo al *10%, **5%, ***1%. VIF (modelo 3, que incluye todos los controles) = 4,61.