

DIFERENCIAS EN LAS CARACTERÍSTICAS DE LOS TRABAJADORES Y LA GEOGRAFÍA DEL DESEMPLEO EN ESPAÑA

Enrique LÓPEZ-BAZO

Universitat de Barcelona

Elisabet MOTELLÓN

Universitat Oberta de Catalunya

Resumen

El mercado de trabajo español se caracteriza por marcadas diferencias territoriales en desempleo. Trabajos previos han analizado los rasgos distintivos de la distribución regional de las tasas de paro mediante factores macroeconómicos, utilizando datos agregados regionales. Este estudio analiza en qué medida los resultados en esos trabajos conducen a interpretaciones incorrectas al no tener en cuenta que las regiones difieren en cuanto a la distribución de las características de los trabajadores, que determinan su propensión a estar desempleados y, por agregación, influyen sobre la tasa de paro de cada región. Se propone un método para calcular las tasas de paro de las provincias españolas neto del efecto composición, que se comparan con las tasas realmente observadas.

Palabras clave: tasa de paro, efecto composición, dependencia espacial, reponderación y *matching*.

Abstract

The Spanish labor market is characterized by marked territorial differences in unemployment. Previous work has analyzed the distinctive features of the regional distribution of unemployment rates through macroeconomic factors, using regional aggregate data. This study analyzes the extent to which the results in these studies lead to incorrect interpretations by not taking into account that the regions differ in the distribution of the characteristics of the workers, which determine their propensity to be unemployed and, by aggregation, influence the unemployment rate of each region. We propose a method to calculate the unemployment rates of the Spanish provinces net of the composition effect, which are compared with the rates actually observed.

Key words: unemployment rate, composition effect, spatial dependence, reweighting and matching.

JEL classification: C21, J64, R12.

I. INTRODUCCIÓN

DEBIDO a sus evidentes implicaciones sociales y económicas, el desempleo es uno de los problemas habituales en la agenda de académicos y gestores políticos, especialmente en períodos de recesión y en economías como la española, caracterizada por un elevado número de personas desempleadas entre la población activa. Entre los diversos hechos distintivos del desempleo se encuentra el que hace referencia a su desigual incidencia en el territorio. Se observan sustanciales diferencias no solo entre países, sino también entre regiones dentro de un mismo país. Además, las disparidades regionales en las tasas de paro tienden a persistir en el tiempo, lo que supone que haya regiones que presentan tasas muy superiores a otras durante décadas sin que se observe ninguna tendencia a converger a tasas de desempleo similares ni, por supuesto, a que sea habitual el intercambio de posiciones en el *ranking* regional de esa magnitud del mercado laboral.

España es un país con elevados niveles de desempleo desde principios de la década de los ochenta presentando, a su vez, amplias y persistentes diferencias regionales. Así, se habla habitualmente de la existencia de una división geográfica entre territorios del suroeste, con tasas muy por encima de la media del país, y regiones en el noreste con tasas por debajo de las medias. En algunos casos, la diferencia en tasas de paro resulta abismal, especialmente si se tiene en cuenta que corresponde a provincias o comunidades autónomas (CC.AA.) separadas únicamente por algunos cientos de kilómetros (véase, por ejemplo, López-Bazo, Del Barrio y Artis, 2002). La magnitud de las disparidades territoriales ha fluctuado con el ciclo económico, decreciendo en términos absolutos durante el período de crecimiento desde finales de los noventa hasta el inicio de la recesión en 2008, e incrementando de manera abrupta y persistente desde ese momento. Además, y como resultado de las diferencias geográficas en el impacto de la burbuja inmobiliaria y su posterior efecto en el sector de la construcción (y actividades relacionadas) y en el bancario, el des-

empleo ha aumentado su incidencia en territorios en los que históricamente las tasas habían presentado valores moderados (por ejemplo, las provincias del litoral mediterráneo). Todo ello hace que el mercado de trabajo español y, en particular, su tasa de desempleo sea un caso de estudio interesante.

Trabajos previos han analizado el desempleo regional en España desde diversas perspectivas, aunque en su práctica totalidad lo han hecho adoptando una aproximación macroeconómica. Es decir, han tratado de explicar los rasgos distintivos de la distribución regional de las tasas de paro mediante factores macroeconómicos (relacionados con la oferta y demanda de empleo a nivel regional) y utilizando datos agregados para CC.AA., provincias y, menos frecuentemente, municipios. Conviene indicar que este tipo de estudios se han llevado a cabo también en otros países (Blanchard y Katz, 1992; Cracolici, Cuffaro y Nijkamp, 2007; Filiztekin, 2009), e incluso para el conjunto de la Unión Europea (Decressin y Fatas, 1995; Overman y Puga, 2002; Ríos, 2016). De hecho, diversos trabajos recientes (por ejemplo OCDE, 2011) confirman que el caso de España no es una anomalía dado que el nivel de disparidades regionales en las tasas de paro en otros países (incluyendo Alemania, Italia, México y Turquía) es incluso más elevado que el existente en España.

Desde un punto de vista teórico no se deberían observar diferencias persistentes en las tasas de paro entre territorios en ausencia de costes de ajuste y rigideces en el mercado laboral y, en general, en la economía. En ese caso el exceso de oferta de trabajo en una región se desvanecería rápidamente, al trasladarse los trabajadores sin empleo a lugares con mayores oportunidades de empleo (con menores tasas de paro), a la vez que la disponibilidad de trabajadores y salarios más bajos atraerían a empresas. Sin embargo, la evidencia existente parece contradecir esa predicción teórica dada la elevada persistencia que caracteriza a la distribución geográfica de las tasas de paro (Lazar, 1997; Evans y McCormick, 1994; Martin, 1997; Martin y Sunley, 1999; Overman y Puga, 2002; López-Bazo, Del Barrio y Artis, 2005). En este sentido, se ha argumentado que un lento e imperfecto ajuste salarial junto a elevados costes migratorios para trabajadores y de cambio de localización para las empresas, podría explicar porqué los *shocks* que afectan de manera desigual a las regiones provocarían tasas de paro divergentes durante dilatados períodos de tiempo. En ese caso, las diferencias regionales en el desempleo corresponderían a un fenómeno de desequilibrio (Marston, 1985). Por el contrario, una

explicación alternativa parte de la idea de que existe una relación estable entre las tasas regionales de desempleo y la dotación de determinados factores en las regiones. Como sugiere Marston (1985), a partir de los argumentos más generales de Hall (1972) y Rosen (1974), la distribución geográfica del desempleo no cambiará sustancialmente durante períodos de tiempo largos dado que, por sus propias características, la dotación de esos factores permanece estable. Ello constituye lo que se ha dado en denominar como la hipótesis de equilibrio de los diferenciales regionales del desempleo. Según la misma, los trabajadores desempleados no tendrán suficientes incentivos para migrar a otras regiones porque valoran la dotación de esos factores. Por su parte, al seleccionar su localización, además del nivel salarial y la tasa de desempleo, las empresas toman en consideración las dotaciones regionales de determinados factores (Partridge y Rickman, 1997). La evidencia acerca de la coexistencia de salarios elevados y altas tasas de paro en algunas regiones en determinados períodos, junto a la preferencia por ciertas *amenities*, apoyan esta hipótesis de equilibrio.

La práctica totalidad de los trabajos empíricos existentes han tratado de analizar el efecto de los determinantes del desempleo regional mediante una especificación en la que la tasa de paro en cada región se relaciona con magnitudes regionales que tratan de aproximar los factores de desequilibrio y equilibrio (Partridge y Rickman, 1997; Taylor y Bradley, 1997; López-Bazo, Del Barrio y Artis, 2002, 2005; Elhorst, 2003; Filiztekin, 2009; Bande y Karanassou, 2009). Algunas de las magnitudes incluidas en esas especificaciones persiguen controlar por diferencias entre regiones en la composición de la población total en general y/o de la población activa en particular. Por ejemplo, la inclusión del porcentaje de población joven permite tener en cuenta el hecho de que en muchas economías la tasa de desempleo de ese segmento de la población es bastante mayor a la de los trabajadores de edad superior. Lo mismo sucede en el caso de variables que tratan de controlar por el nivel educativo medio de la población de cada región, dada la evidente heterogeneidad en las tasas de paro de individuos según la educación acumulada.

Por otra parte, algunos de esos trabajos (López-Bazo, Del Barrio y Artis, 2002; Overman y Puga, 2002; Ríos, 2016) han destacado un rasgo característico de la distribución del desempleo, como es un elevado nivel de dependencia espacial en la misma. Es decir, que las tasas de paro no se distri-

buyen de manera independiente en el espacio, sino que valores elevados (bajos) en una región suelen coincidir con valores también elevados (bajos) en las regiones próximas geográficamente. Para explicar este fenómeno, se ha recurrido al efecto de las externalidades espaciales en el mercado laboral. Por ejemplo, en Ríos (2016) se amplía el modelo de Blanchard y Katz (1992) al considerar que tanto la oferta y demanda de trabajo como el salario de una región están afectados por factores internos a la región y también por esos factores en otras regiones. Asimismo, las decisiones de migración de los trabajadores tomarían en consideración la tasa de paro y otras variables representativas de la situación del mercado laboral de la región y las de otras regiones. En todo ello se asume que las externalidades espaciales son más intensas cuando proceden de regiones próximas geográficamente que cuando su origen es más remoto. La consideración de externalidades que desbordan los mercados de trabajo regionales ha conducido a ampliar la especificación empírica utilizada para analizar el desempleo regional, incluyendo entre los factores explicativos la tasa de desempleo medio en el conjunto de regiones próximas geográficamente (según un determinado criterio de proximidad), los valores en ese grupo de regiones de los determinantes del desempleo y/o de los *shocks* aleatorios sufridos por las mismas.

Por tanto, las contribuciones empíricas previas reconocen, de manera más o menos explícita, que las regiones difieren en la composición de su población activa, y que esas diferencias pueden causar al menos parte de las observadas en las tasas de desempleo. De manera más clara, podemos decir que la tasa de paro de una región es el resultado de, por una parte, la propensión al desempleo de sus trabajadores y, por otra, de la composición de su población activa, bajo el supuesto de que no todos los tipos de trabajadores tienen la misma probabilidad de estar parados. En ese contexto, se puede argumentar que las diferencias en la composición de la población activa podría explicar buena parte de las diferencias regionales en las tasas de paro. A su vez, los trabajos más recientes consideran que la elevada dependencia espacial existente en la distribución de las tasas de paro tiene su origen en mecanismos ligados al funcionamiento de los mercados de trabajo regionales. Obvian, por tanto, el hecho de que si las características de los trabajadores que se localizan próximos en el espacio son más parecidas que las de aquellos en localidades más alejadas, la dependencia espacial observada en las tasas de paro podría ser debida en gran parte a que las regiones difieren en las características de los trabajadores.

En consecuencia, si las especificaciones empíricas utilizadas para modelizar el desempleo regional con datos agregados no son capaces de controlar de manera adecuada por las diferencias regionales en la distribución de las características de los trabajadores, la dependencia espacial observada podría ser debida, al menos en parte, al efecto composición y no tanto a mecanismos intrínsecamente espaciales que operan en el mercado de trabajo.

Tomando en consideración estas circunstancias, el objetivo esencial de este trabajo es analizar la distribución geográfica de la tasa de paro en España neta del efecto composición, es decir, descontando el efecto de las diferencias territoriales en las características de los trabajadores relacionadas con la probabilidad de estar desempleado. Para ello se utilizan métodos de reponderación y de emparejamiento (*matching*) de las muestras provinciales proporcionadas por la *Encuesta de Población Activa (EPA)* para 2015, para obtener unas tasas de paro contrafactuales en cada una de las provincias. Dichas tasas contrafactuales aproximan las tasas de paro que se hubiesen observado en el caso en el que todas las provincias hubiesen tenido trabajadores con similares características observadas. La comparación entre la distribución espacial de la tasa de paro real y la contrafactual nos permite valorar el efecto de las diferencias regionales en las características de los trabajadores sobre las observadas en las tasas de paro. Para ello, se utilizan métodos gráficos, mapas y estadísticos estándar de asociación espacial. Conviene destacar que para la realización del ejercicio empírico se utilizan los microdatos de las muestras provinciales que, por el diseño de la *EPA*, representan a la población de las respectivas provincias. Hasta donde conocemos, la utilización de este tipo de información para analizar las diferencias regionales en desempleo se limita a los trabajos de López-Bazo y Motellón (2013a, b). Asimismo, también se debe indicar que la metodología empleada resulta novedosa en el contexto del estudio de las diferencias regionales aunque, a su vez, debemos reconocer que puede presentar algunas limitaciones. Por ejemplo, no se puede descartar *a priori* que existan características inobservables de los trabajadores que afectan a la propensión de estar desempleado y, a su vez, que estén relacionadas con la localización de éstos en una provincia concreta. Esta circunstancia cuestionaría la validez de los resultados, por lo que será discutida en la sección correspondiente.

El resto del artículo se organiza como sigue. La base de datos y la selección del período analizado

se discuten en la siguiente sección. En esa misma sección también se definen y comentan las principales variables empleadas en el estudio. En la tercera sección se realiza un análisis descriptivo a modo de evidencia preliminar. En ella se presentan los rasgos principales de la distribución provincial del desempleo y se muestra evidencia que confirma la desigual distribución geográfica de las características de los trabajadores, que han sido señaladas habitualmente como determinantes de la propensión al desempleo. La metodología para obtener las tasas de paro contrafactuales y la evidencia derivada de las mismas se presenta en la sección cuarta, mientras que en la quinta se valoran las diferencias existentes entre las recientes fases expansiva y recesiva de la economía española. Finalmente, la sección sexta sintetiza las conclusiones alcanzadas.

II. BASE DE DATOS Y VARIABLES

Un elemento diferencial de este estudio respecto a la práctica totalidad de análisis precedentes del desempleo regional es el de la utilización de información estadística referida a los individuos participantes en el mercado laboral. En concreto, el ejercicio empírico que se presenta en los siguientes apartados utiliza los datos individuales de la *EPA*, elaborada por el Instituto Nacional de Estadística (INE) siguiendo estándares internacionales y de manera consistente con las guías y protocolos establecidos por EUROSTAT para todos los estados miembros de la Unión Europea. El grueso del análisis se realiza utilizando los datos para el año 2015, que es el más próximo para el que se dispone de información en el momento de realizar el ejercicio empírico. No obstante, para valorar las posibles diferencias entre las fases expansivas, de bajo nivel de paro, y de las recesivas, de elevado desempleo, también se utiliza la información de la *EPA* referida a 2006 y 2013.

Los microdatos de la *EPA* proporcionan información sobre la situación de los individuos en el mercado de trabajo, distinguiendo entre los que participan y los que no lo hacen y, para los primeros, entre los que se encuentran empleados y los que están desempleados. Además, incluye información sobre un número considerable de características de los individuos, como la edad, nacionalidad, género y el nivel educativo; y del hogar, como el número total de miembros, hijos según edad, y la situación laboral del cónyuge, si lo hay, y de los otros miembros. Dado que el objetivo esencial se articula en torno a la tasa de paro, la muestra uti-

lizada se limita a los individuos entre 16 y 65 años que participan en el mercado laboral, localizados en cada una de las provincias españolas, exceptuando Ceuta y Melilla. En este sentido, cabe mencionar dos aspectos importantes. El primero tiene que ver con el número y delimitación de las unidades espaciales consideradas. El tipo de técnicas utilizadas para analizar la distribución espacial del desempleo aconseja emplear un número suficientemente elevado de unidades espaciales, por lo que un ámbito inferior al provincial hubiese sido preferible. Sin embargo, esto no ha sido posible al no disponer de información referida a muestras representativas para un nivel inferior al provincial. El segundo aspecto se refiere al hecho de que la muestra disponible de microdatos de la *EPA* garantiza la representatividad para la población de cada una de las provincias, lo que permite derivar conclusiones sobre la tasa de desempleo de cada provincia a partir de la información contenida en las respectivas muestras.

Como es bien sabido, la *EPA* proporciona información sobre la situación del mercado de trabajo para cada uno de los trimestres del año. En el trabajo se presentan únicamente los resultados obtenidos al utilizar la información correspondiente al segundo trimestre dado que, como se ha indicado en otros estudios, el segundo es el menos afectado por variaciones estacionales (causadas en gran medida por la sensibilidad del mercado de trabajo español al turismo y a otras actividades de servicios y del sector primario). En todo caso, cabe señalar que los principales resultados obtenidos no dependen del trimestre utilizado. Por otra parte, se descartó la utilización de todos los trimestres de los años considerados para evitar incluir varias observaciones de buena parte de los individuos en la muestra, lo que podría distorsionar los resultados del análisis a realizar. Esto es así dado que la *EPA* incluye a un mismo hogar a lo largo de seis trimestres consecutivos, por lo que la muestra para cada año incluiría una única observación para algunos individuos y entre dos y cuatro para otros, dependiendo de en qué trimestre se incluyese al hogar correspondiente en la encuesta.

Como se ha indicado anteriormente, los microdatos de la *EPA* incluyen la información que permite clasificar a los individuos en activos e inactivos, laboralmente hablando, y para los primeros en ocupados y parados. En este sentido, cabe indicar que siguiendo el criterio de la *EPA*, este estudio considera a un individuo como ocupado cuando trabajó como mínimo una hora durante la semana anterior a la que fue encuestado. Por su parte, se considera a un trabajador como parado cuando estuvo sin tra-

bajo en la semana referenciada, a pesar de buscar activamente un empleo y estar disponible para trabajar. Por último, indicar que la ratio entre parados y activos en la muestra de cada provincia da lugar a su tasa de desempleo. En el caso de utilizar los factores de reponderación incorporados en la EPA es posible reproducir con exactitud los valores publicados por el INE para las tasas de paro de cada provincia y de España en su conjunto. No obstante, dichas reponderaciones muestrales no han sido empleadas en los cálculos para obtener los resultados de los siguientes apartados. Cabe indicar, en todo caso, que la incidencia de esta decisión sobre el cálculo de las tasas de paro provinciales es marginal y que de ninguna manera altera los principales resultados acerca del efecto de las características de los individuos sobre las citadas tasas.

Finalmente, indicar que la definición de las variables utilizadas en el ejercicio empírico referidas a las características del individuo y el hogar se recoge en el cuadro n.º 1. La última columna del cuadro muestra la proporción de individuos en cada una de las categorías cuando se trata de características dicotómicas y la media para el caso de variables continuas. En todos los casos, esos datos hacen referencia únicamente a la muestra de individuos activos, es decir, a los que participaban en el mercado laboral en el segundo trimestre de 2015, bien como empleados o como parados. A la información contenida en el cuadro n.º 1 cabe únicamente añadir que el nivel educativo más bajo, «sin estudios», incluye tanto a los individuos que se declararon como analfabetos como a aquellos sin los estudios primarios completos. A su vez, el nivel más elevado corresponde a «estudios superiores» que

CUADRO N.º 1

DEFINICIÓN Y DESCRIPCIÓN DE LAS CARACTERÍSTICAS DE INDIVIDUOS Y HOGARES, 2015

MEDIDAS DE PARTICIPACIÓN Y DESEMPLEO		
Activo	= 1 si el individuo participaba en el mercado laboral según criterio EPA = 0 en caso contrario	75,3 (0,14)
Parado	= 1 si el individuo se encontraba desempleado según criterio EPA = 0 en caso contrario	22,5 (0,15)
CARACTERÍSTICAS DEL INDIVIDUO		
Edad	= 1 si el individuo se encontraba en el correspondiente tramo de edad = 0 en caso contrario	
	Tramos:	
	16-24	6,9 (0,09)
	25-34	19,1 (0,14)
	35-44	28,5 (0,16)
	45-54	28,7 (0,16)
	55-65	16,7 (0,14)
Género	= 1 si el individuo es hombre = 0 si es mujer	52,8 (0,18)
Nacionalidad	= 1 si el individuo tenía esa situación respecto a la nacionalidad = 0 en caso contrario	
	Española	91,2 (0,10)
	Doble nacionalidad	2,3 (0,05)
	Diferente a española	6,5 (0,09)
Casado	= 1 si el individuo tenía ese estado civil = 0 en caso contrario	56,1 (0,18)

CUADRO N.º 1 (Continuación)

DEFINICIÓN Y DESCRIPCIÓN DE LAS CARACTERÍSTICAS DE INDIVIDUOS Y HOGARES, 2015

Nivel educativo	= 1 si el individuo se encontraba en el correspondiente nivel educativo = 0 en caso contrario	
Niveles:		
Sin estudios		1,6 (0,04)
Primaria		6,8 (0,09)
Secundaria inferior		32,0 (0,17)
Secundaria superior		22,8 (0,15)
Superior		36,8 (0,18)
<i>CARACTERÍSTICAS DEL HOGAR</i>		
Cabeza familia	= 1 si el individuo se declara como cabeza de familia del hogar = 0 en caso contrario	46,6 (0,18)
Número de hijos pequeños	Número de hijos menores de 10 años	0,4 (0,68)
Miembros de edad mayor de 65 años	= 1 si hay miembros del hogar de edad superior a 65 años e inactivos laboralmente = 0 en caso contrario	12,7 (0,12)
Mismo municipio	= 1 si el individuo no ha variado su municipio de residencia = 0 en caso contrario	1,1 (0,04)

Notas: Los datos de la última columna corresponden a la proporción de individuos con la correspondiente característica para las variables binarias y a la media para las variables continuas. Desviación estándar entre paréntesis. Resultados derivados de la muestra correspondiente a los individuos activos en la totalidad del territorio español excluyendo Ceuta y Melilla.

incluye tanto la enseñanza de formación profesional de grado superior como la enseñanza universitaria.

III. EVIDENCIA PRELIMINAR

Como se ha indicado en la sección introductoria, una característica bien conocida del desempleo en España es que la incidencia del mismo varía de manera muy notable entre los distintos territorios, sean estos municipios, CC.AA. o, como en el caso de las unidades espaciales consideradas en este trabajo, provincias. Para el segundo trimestre de 2015, la primera columna del cuadro n.º 2 reproduce las tasas de paro de las provincias españolas obtenidas a partir de la muestra de la EPA. La comparación de las provincias con valores extremos nos proporciona un primer indicio de la magnitud de las diferencias espaciales en el desempleo. Frente a una tasa de paro del 11,6 por 100 en Navarra y del 13,6 por 100 en Segovia, nos encontramos con una cifra de casi el 39 por 100 en Cádiz y más del 34 por 100 en Jaén. Es decir, tasas de desempleo en unas provincias que triplican las observadas en otras.

Para facilitar la visualización de las discrepancias provinciales en el desempleo, en el gráfico 1 se representa la distribución de las tasas de paro provinciales. La línea continua corresponde a la estimación de la función de densidad de las tasas de paro observadas en el período analizado (las restantes funciones de densidad serán comentadas en la sección siguiente). Se aprecia como la moda se sitúa ligeramente por debajo del 20 por 100 y como hay una importante masa de probabilidad en valores en el entorno del 30 por 100. Esto nos indica una cierta polarización en dos grupos de provincias, con tasas de desempleo significativamente diferenciadas. La visualización de la función de densidad estimada nos confirma, por tanto, que las diferencias territoriales en desempleo no se limitan a unas pocas provincias con tasas bajas y elevadas (en términos relativos), sino que se trata de una característica esencial de todo el sistema territorial.

Otra circunstancia bien conocida del desempleo en España es su particular distribución espacial. Con ello nos referimos a una cierta división territorial dependiendo de la incidencia del desempleo.

CUADRO N.º 2

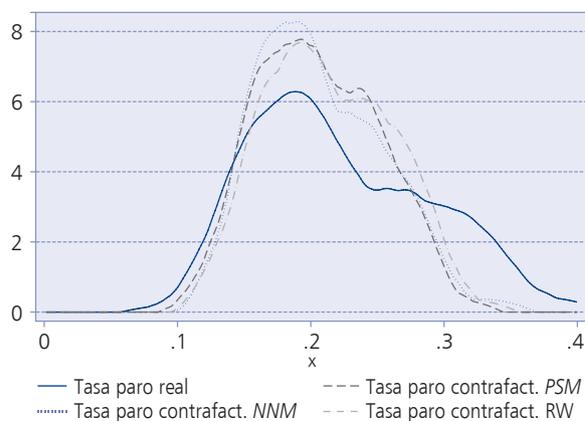
TASAS DE PARO REALES Y CONTRAFACUALES EN LAS PROVINCIAS ESPAÑOLAS, 2015

	REALES	CONTRAFACUAL		
		RW	PSM	NNM
Álava	0,1533	0,1809	0,1700	0,1514
Albacete	0,2488	0,2294	0,2299	0,2163
Alicante	0,2397	0,2174	0,1872	0,2023
Almería	0,2870	0,2708	0,2536	0,2647
Ávila	0,2260	0,2415	0,2276	0,2091
Badajoz.....	0,3113	0,2849	0,2694	0,2579
Islas Baleares.....	0,1603	0,1383	0,1244	0,1422
Barcelona.....	0,1909	0,1732	0,1708	0,1719
Burgos.....	0,1872	0,1930	0,1891	0,1760
Cáceres	0,2656	0,2422	0,2164	0,2182
Cádiz	0,3891	0,3142	0,3037	0,3272
Castellón.....	0,2324	0,2164	0,2215	0,2245
Ciudad Real.....	0,3234	0,2819	0,2812	0,2812
Córdoba	0,3029	0,2378	0,2591	0,2621
La Coruña	0,1812	0,1827	0,1791	0,1791
Cuenca	0,2296	0,2548	0,2254	0,2277
Gerona	0,1874	0,1570	0,1588	0,1590
Granada.....	0,3090	0,2730	0,2647	0,2738
Guadalajara	0,1817	0,1773	0,1786	0,1784
Guipúzcoa	0,1468	0,1676	0,1580	0,1670
Huelva	0,2826	0,2452	0,2228	0,2375
Huesca.....	0,1521	0,1529	0,1542	0,1687
Jaén	0,3447	0,2838	0,2697	0,2750
León	0,2061	0,1853	0,1669	0,1765
Lérida	0,1564	0,1431	0,1368	0,1472
La Rioja.....	0,1577	0,1634	0,1571	0,1620
Lugo.....	0,1589	0,1806	0,1756	0,1725
Madrid.....	0,1782	0,1782	0,1782	0,1782
Málaga.....	0,3027	0,2594	0,2461	0,2504
Murcia	0,2456	0,2152	0,2103	0,2142
Navarra.....	0,1155	0,1349	0,1487	0,1364
Orense.....	0,2011	0,2005	0,2151	0,2087
Asturias	0,2037	0,2146	0,2092	0,2249
Palencia	0,1794	0,2602	0,2359	0,1976
Las Palmas	0,3255	0,2704	0,2626	0,2568
Pontevedra.....	0,2350	0,2250	0,2235	0,2250
Salamanca	0,1818	0,1966	0,1744	0,1773
Sta. Cruz de Ten.....	0,3065	0,2624	0,2679	0,2794
Cantabria.....	0,1787	0,2119	0,2068	0,1921
Segovia.....	0,1359	0,1484	0,1340	0,1514
Sevilla	0,2985	0,2633	0,2536	0,2655
Soria	0,1426	0,1628	0,1449	0,1440
Tarragona	0,2261	0,2093	0,1825	0,1911
Teruel	0,1655	0,1851	0,1692	0,1723
Toledo	0,2899	0,2571	0,2381	0,2483
Valencia.....	0,2219	0,2158	0,2143	0,2134
Valladolid.....	0,1711	0,1916	0,1933	0,1779
Vizcaya	0,1733	0,1824	0,1879	0,1841
Zamora	0,2228	0,2159	0,2067	0,2131
Zaragoza.....	0,1677	0,1740	0,1708	0,1676

Notas: Madrid es la provincia utilizada como referencia. RW denota que se ha utilizado el método de reponderación para aislar el efecto de las diferencias en las características observables, mientras que PSM y NNM que el método utilizado ha sido, respectivamente, el de propensity score matching y nearest-neighbor matching.

La misma se refleja en el mapa de la parte superior izquierda del gráfico 2, que presenta a las provincias coloreadas en función de su pertenencia a uno de los cuatro cuartiles en los que se divide la distribu-

GRÁFICO 1
FUNCIONES DE DENSIDAD DE LAS TASAS DE PARO REALES Y CONTRAFACUALES, 2015

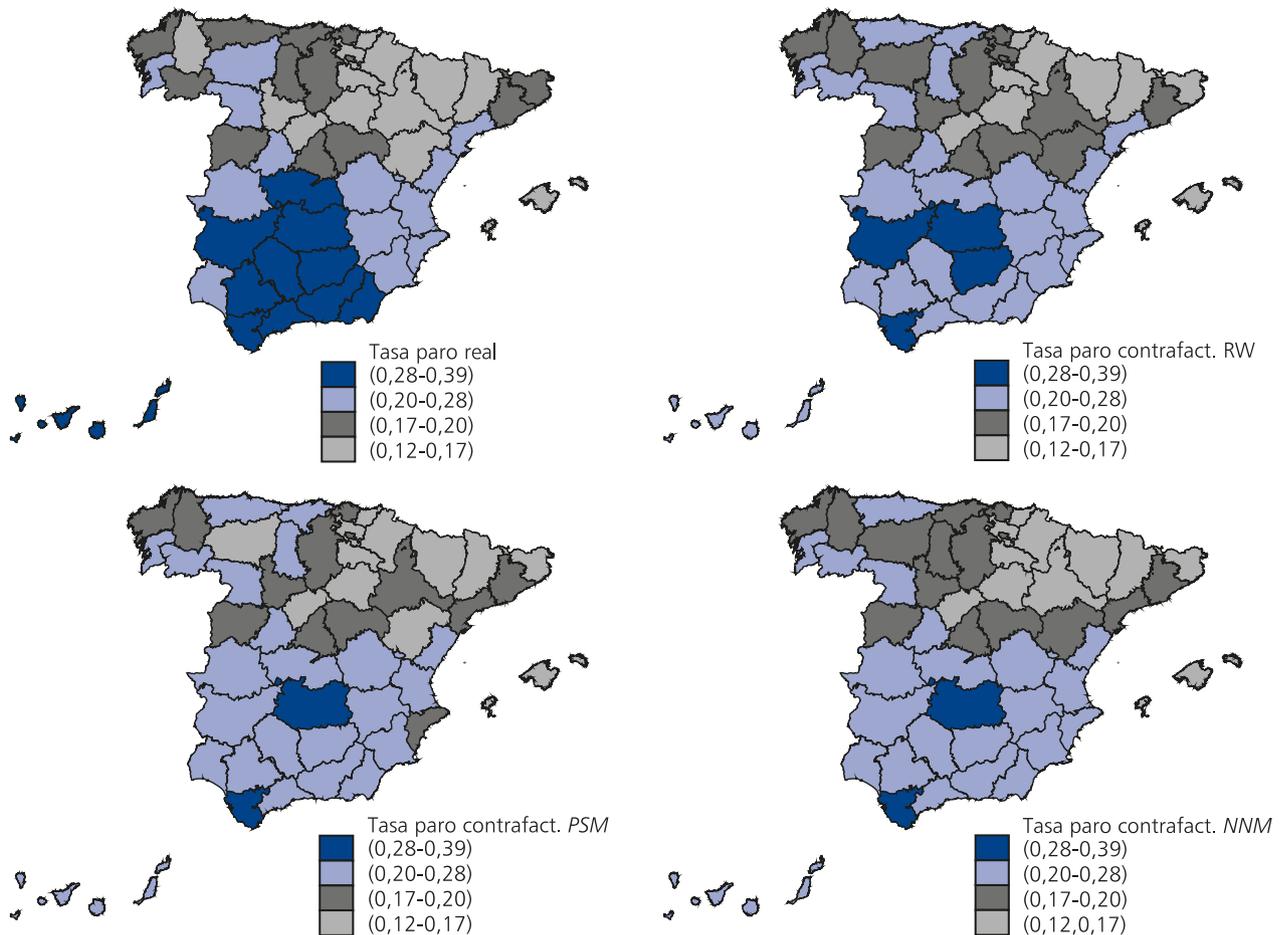


ción de las tasas provinciales de paro mostradas en la primera columna del cuadro n.º 2. De tonalidad menos a más intensa se representan las tasas de paro de valores menos a más elevados. El mapa permite distinguir claramente una España con tasas de paro secularmente muy por encima de la media (provincias en el suroeste) de la España con una incidencia moderada, en términos relativos, del desempleo (noreste). En todo caso, conviene también indicar que desde el comienzo de la crisis, la mayor parte de las provincias mediterráneas han visto aumentar sustancialmente sus tasas de paro, alcanzando valores que conducen a excluirlas del grupo de provincias del noreste caracterizado por las menores tasas de desempleo. En cualquier caso, la imagen general que se deriva del mapa que representa las tasas provinciales de desempleo es la de una intensa relación entre el valor observado en una provincia y el correspondiente al de las provincias próximas geográficamente. O en palabras más técnicas, que la distribución geográfica del desempleo en España muestra claros síntomas de dependencia espacial.

Es posible sintetizar el grado de dependencia espacial en el desempleo, y a su vez contrastar su significación desde un punto de vista estadístico, con algunas de las medidas propuestas en la literatura (véase, por ejemplo, Moreno y Vayá, 2000). En este trabajo, vamos a utilizar dos de las más habituales, la I de Moran y la c de Geary. Ambas consideran la relación existente entre el valor de la variable de interés en una unidad territorial y los valores en los territorios vecinos a ésta. En el caso de este estudio,

GRÁFICO 2

TASAS DE PARO REALES Y CONTRAFACUALES (MISMA COMPOSICIÓN QUE MADRID), 2015



entre la tasa de paro de cada provincia y la de sus vecinas. Para ello es esencial la delimitación del concepto de vecindad, existiendo diversas opciones sugeridas en la literatura, aunque para no introducir complejidades innecesarias en este estudio nos vamos a limitar a las más inmediatas y, a la vez, intuitivas. En este sentido, cabe decir, en primer lugar, que excluimos todas aquellas opciones basadas en un criterio no exclusivamente geográfico. Es decir, que no vamos a considerar definiciones de vecindad basadas en conceptos de distancia distinta a la geográfica, como por ejemplo de tipo económico (como las basadas en flujos de agentes o de bienes y servicios) o social (por ejemplo, basadas en similitudes culturales o empresariales). En segundo lugar, cabe señalar que, entre las diversas opciones basadas en criterios exclusivamente geográficos,

nos limitamos a considerar las correspondientes a la contigüidad física y a la distancia entre territorios.

Así, en el primer caso consideramos que dos provincias son vecinas si son contiguas geográficamente (p. ej., si comparten frontera). Conviene indicar que en el caso de las provincias insulares se ha asignado artificialmente la vecindad con la provincia más cercana. De esta forma es posible definir lo que en la literatura del ámbito de la estadística y la econometría espacial se conoce como la matriz de pesos espaciales, W . Esta matriz tiene como elemento característico w_{ij} , que toma valor 1 cuando las provincias i y j son contiguas, y valor 0 en caso contrario. Por su parte, a los elementos de la diagonal de W , w_{ii} , se les asigna valor 0. Esta matriz de contigüidad espacial puede ser normalizada por filas, de manera que los

pesos para cada provincia suman uno. Por su parte, en el caso de la matriz W basada en las distancias, utilizamos como pesos espaciales la inversa de la distancia al cuadrado entre cada una de las provincias. Para ser más precisos, el elemento característico w_{ij} se define como $1/d_{ij}^2$, siendo d_{ij} la distancia más corta por carretera entre la capital de la provincia i y la j . Para el caso de las provincias insulares, la distancia incorpora una penalización asociada a la obvia imposibilidad del transporte exclusivo a través de ese medio. También en este caso, la matriz es normalizada por filas, de manera que para cada provincia i se considera la distancia relativa a cada una de las restantes provincias, en lugar de la distancia absoluta. Conviene indicar que es frecuente considerar un umbral de distancia a partir del que se asigna un valor nulo a los pesos espaciales y/o que se pueden considerar funciones de la distancia alternativas a la indicada anteriormente. También que existe la posibilidad de definir la matriz W con elementos binarios (0 y 1) a partir de un criterio de distancia (por ejemplo, en el caso de la matriz de pesos espaciales basada en los «*k* vecinos más cercanos»). Como se ha apuntado anteriormente, los resultados que mostramos en este trabajo son aquéllos obtenidos con las matrices de contacto espacial más inmediatas y posiblemente más intuitivas. En todo caso, es importante señalar que las mismas conclusiones se obtuvieron a partir de los resultados obtenidos con algunas de las otras opciones.

Los resultados de los contrastes I de Moran y c de Geary para las tasas de paro provinciales observadas en 2015 se muestran en la primera fila correspondiente a cada uno de los bloques de las

dos matrices de pesos espaciales. Se puede apreciar, claramente, como ambos contrastes rechazan sus respectivas hipótesis nula de ausencia de dependencia espacial, tanto en el caso de la matriz de contigüidad como en la de distancia. Por tanto, los contrastes confirman la evidencia informal que habíamos derivado a partir de la representación de las tasas de paro provinciales en el mapa del gráfico 2. De hecho, la información referida al valor de los estadísticos de Moran y Geary reafirman lo que fácilmente se deduce de ese mapa, en el sentido que la distribución provincial del desempleo está caracterizada por correlación espacial positiva, lo que implica una tendencia a que los valores elevados y bajos de la tasa de paro se encuentren agrupados espacialmente.

Como se ha mencionado en las secciones anteriores, la clara dependencia espacial observada en la distribución geográfica del desempleo ha motivado que la literatura reciente haya considerado modelos empíricos que incorporan efectos espaciales, a través de la inclusión de retardos espaciales (valores del conjunto de regiones vecinas) del desempleo y/o de sus determinantes observados e inobservados (capturados por la componente estocástica). Sin embargo, no podemos, *a priori*, descartar que la distribución geográfica de las tasas de paro y, por tanto, la dependencia espacial observada en esa variable, esté en buena medida causada por el hecho de que las provincias difieran en cuanto a las características de la población activa, que afectan a la probabilidad de estar desempleado. Y a que los individuos de provincias próximas geográficamente tengan mayor tendencia a compartir esas caracte-

CUADRO N.º 3

ESTADÍSTICOS DE DEPENDENCIA ESPACIAL PARA LAS TASAS DE PARO REALES Y CONTRAFACUALES, 2015

	MORAN'S I			GEARY'S C		
	<i>I</i>	<i>z</i>	<i>p-val</i>	<i>c</i>	<i>z</i>	<i>p-val</i>
W contigüidad						
Real	0,72	7,887	0,0000	0,216	-7,395	0,0000
Contrafact. <i>PSM</i>	0,595	6,539	0,0000	0,304	-6,681	0,0000
Contrafact. <i>NNM</i>	0,66	7,257	0,0000	0,265	-6,855	0,0000
Contrafact. <i>RW</i>	0,622	6,817	0,0000	0,303	-6,738	0,0000
W inversa distancia²						
Real	0,532	9,902	0,0000	0,403	-6,87	0,0000
Contrafact. <i>PSM</i>	0,395	7,435	0,0000	0,492	-6,145	0,0000
Contrafact. <i>NNM</i>	0,472	8,835	0,0000	0,435	-6,329	0,0000
Contrafact. <i>RW</i>	0,41	7,693	0,0000	0,503	-6,142	0,0000

Nota: Ver cuadro n.º 2.

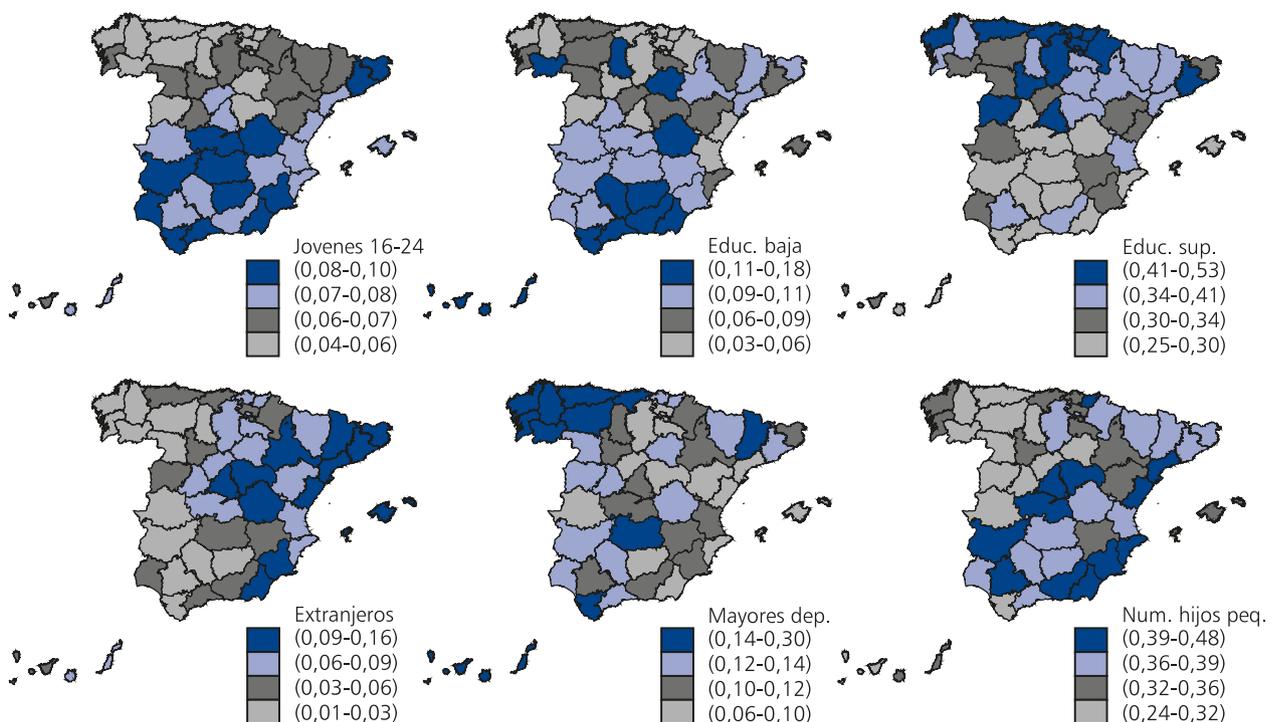
rísticas de la que tienen los individuos de provincias más distantes. A modo de ejemplo, imaginemos que la situación de desempleo para cada individuo viene determinada únicamente por un factor. Supongamos que ese factor es el nivel educativo, de manera que un individuo con nivel educativo bajo estaría desempleado mientras que si el nivel educativo es suficientemente elevado, el individuo tendría empleo. Ahora imaginemos que la distribución espacial de los individuos según esa característica, su nivel educativo, no es aleatoria ni uniforme. Al contrario, imaginemos que las unidades espaciales en el norte de un territorio concentran muchos más individuos con nivel educativo elevado que aquellas en el sur, que concentran mayoritariamente a individuos con bajo nivel educativo. En ese caso, dado que suponemos que el desempleo depende únicamente del nivel educativo de los individuos, el origen de la dependencia espacial en las tasas de paro estaría en la distribución geográfica de esa característica individual.

Evidentemente, la realidad es más compleja y cabe pensar que múltiples características intervienen en la determinación del desempleo de los trabajadores. Por ello, y sin ánimo de ser exhaustivos,

hemos analizado la distribución provincial de algunas características individuales de cuya información disponemos en la EPA. En concreto, a continuación se presentan y comentan brevemente los resultados obtenidos para i) los activos entre 16 y 24 años; ii) los que acreditaban un nivel de estudios bajo (inferior a secundarios); iii) los de nivel de estudios elevado (superiores a los secundarios); iv) los que no tenían la nacionalidad española; v) los que pertenecían a hogares con algún miembro de 65 o más años no activo; y vi) el número de hijos pequeños (edad menor a 10 años). Los mapas correspondientes a la proporción de individuos en cada una de esas categorías en cada provincia, y la media provincial en el caso del número de hijos pequeños, se presentan en el gráfico 3, mientras que los correspondientes estadísticos de dependencia espacial se muestran en el cuadro n.º 4. De los mapas se deduce que el valor promedio de esas características difiere notablemente entre provincias y que su distribución geográfica muestra patrones bastante marcados, especialmente en algunos casos. Así, a modo de ejemplo, se aprecia una división entre provincias en la mitad norte y en la sur para el porcentaje de población con educación superior. En algunas provincias de la mitad norte ese porcentaje

GRÁFICO 3

DISTRIBUCIÓN GEOGRÁFICA DE LA MEDIA DE ALGUNAS CARACTERÍSTICAS DE LA POBLACIÓN ACTIVA, 2015



CUADRO N.º 4

ESTADÍSTICOS DE DEPENDENCIA ESPACIAL PARA ALGUNAS CARACTERÍSTICAS DE LA POBLACIÓN ACTIVA, 2015

	MORAN'S I			GEARY'S C		
	I	z	p-val	c	z	p-val
W contigüidad						
Jóvenes 16-24.....	0,700	7,647	0,0000	0,318	-6,629	0,0000
Educ. baja.....	0,384	4,314	0,0000	0,555	-4,128	0,0000
Educ. sup.....	0,428	4,778	0,0000	0,546	-4,268	0,0000
Extranj.	0,610	6,743	0,0000	0,290	-6,475	0,0000
Mayores dep.....	0,646	7,718	0,0000	0,462	-3,458	0,0000
Núm. hijos peq.	0,382	4,290	0,0000	0,633	-3,419	0,0000
W inversa distancia²						
Jóvenes 16-24.....	0,467	8,705	0,0000	0,494	-6,340	0,0000
Educ. baja.....	0,350	6,642	0,0000	0,692	-3,402	0,0000
Educ. sup.....	0,486	9,068	0,0000	0,737	-3,000	0,0010
Extranj.	0,324	6,197	0,0000	0,480	-5,503	0,0000
Mayores dep.....	0,325	6,698	0,0000	0,560	-2,524	0,0060
Núm. hijos peq.	0,203	4,006	0,0000	0,802	-2,204	0,0140

duplica al existente en otras de la mitad sur. Por su parte, también se aprecia una cierta división nortesur para el porcentaje de población activa menor de 24 años. En cualquier caso, en los mapas de todas esas características se aprecian agrupaciones de provincias próximas espacialmente con valores similares, lo que apunta a la existencia de correlación espacial en las mismas. Esta circunstancia es corroborada por los contrastes de dependencia espacial que se muestran en el cuadro n.º 3. Para todas las características consideradas y para las dos matrices de pesos espaciales utilizadas, los contrastes *I* de Moran y *c* de Geary rechazan la hipótesis nula de ausencia de dependencia espacial a los niveles de significación habituales. De hecho, los valores de los contrastes confirman la existencia de correlación espacial positiva, es decir, la agrupación espacial de provincias con valores similares en el porcentaje de población con esas características (en el valor medio en el caso del número de hijos pequeños).

Por otra parte, contrastes formales de igualdad de proporciones, y de medias para el caso del número de hijos pequeños, confirman la existencia de notables diferencias entre provincias en las características de los trabajadores. A modo de ejemplo, el cuadro n.º 5 recoge los resultados de dichos contrastes para las características enumeradas anteriormente y cuando los valores en cada provincias son comparados con los de los mostrados por los trabajadores en Madrid, provincia que es utilizada como referencia. Se puede apreciar como, para casi

todas las provincias, se rechaza la hipótesis de igualdad de la proporción de trabajadores con educación baja y elevada, y de extranjeros, mientras que se rechaza la de igualdad de medias en el número de hijos pequeños. El número de provincias con valores significativamente distintos a los de Madrid es algo menor en el caso de la proporción de trabajadores jóvenes y de individuos en hogares con mayores dependientes. Pero incluso en esos casos, los resultados avalan la necesidad de controlar por las diferencias provinciales en las características de los individuos en los modelos diseñados para explicar el desempleo regional.

Además del control por esas diferencias, la marcada dependencia espacial en la tasa de paro provincial y en los valores medios de las características que supuestamente determinan el desempleo justificaría la inclusión de los respectivos retardos espaciales en los modelos empíricos especificados para analizar el desempleo regional. Sin embargo, hay dos consideraciones que deben ser tenidas en cuenta. La primera, es que los coeficientes asociados al impacto de los determinantes del desempleo regional medidos en la regiones vecinas podrían estar absorbiendo el efecto de factores distintos al de las externalidades espaciales en el mercado de trabajo. La segunda, que la media provincial de cada una de las características individuales podría estar siendo una medida bastante imperfecta de aquellos factores que determinan la probabilidad de estar desempleado en cada territorio y, por agregación, de la tasa de paro en el mismo. Combinando ambas obtenemos la motivación prin-

CUADRO N.º 5

CONTRASTES DE IGUALDAD DE CARACTERÍSTICAS ENTRE CADA PROVINCIAL Y MADRID, 2015

	JÓVENES 16-24	EDUC. BAJA	EDUC. SUP	EXTRANJ.	MAYORES DEP.	NÚM. HIJOS PEQ.
Álava	-0,76	-0,08	0,31	-4,68***	0,36	-4,03***
Albacete	4,86***	7,89***	-9,80***	-6,32***	0,19	-4,81***
Alicante	1,54	5,36***	-13,68***	-2,21**	0,54	-1,85**
Almería	1,01	10,92***	-10,87***	-0,92	-0,57	1,22
Ávila	1,17	5,51***	-10,92***	-3,35***	2,40**	-5,78***
Badajoz.....	3,90***	9,80***	-15,24***	-12,70***	3,43***	-2,22**
Islas Baleares.....	0,81	2,22**	-14,60***	4,28***	-1,55	-3,83***
Barcelona.....	3,03***	8,07***	-7,00***	1,23	1,44	-3,94***
Burgos.....	0,32	0,69	-0,90	-5,04***	-1,28	-2,40**
Cáceres.....	1,67*	7,20***	-13,10***	-11,00***	-1,12	-7,16***
Cádiz.....	2,51**	18,02***	-14,29***	-12,32***	3,59***	-6,65***
Castellón.....	1,11	0,35	-10,51***	1,00	-1,76*	0,65
Ciudad Real.....	3,57***	10,47***	-14,93***	-8,01***	4,45***	-2,76***
Córdoba.....	3,37***	14,60***	-14,35***	-12,95***	1,12	-4,08***
La Coruña.....	-1,14	2,73***	-5,87***	-15,80***	13,71***	-6,87***
Cuenca.....	1,70*	10,73***	-11,97***	-3,12***	1,32	-2,30**
Gerona.....	2,17**	5,71***	-12,47***	7,53***	1,13	-0,69
Granada.....	2,63***	14,22***	-9,94***	-6,81***	0,48	-1,76*
Guadalajara.....	-1,51	3,34***	-4,86***	-1,01	-2,97***	0,65
Guipúzcoa.....	-0,93	-0,49	2,17**	-4,64***	2,73***	0,17
Huelva.....	4,23***	4,05***	-10,61***	-5,79***	1,93*	-3,13***
Huesca.....	-1,30	4,26***	-4,73***	-2,60***	3,42***	-2,04**
Jaén.....	4,50***	19,81***	-15,10***	-14,18***	-1,60	-4,19***
León.....	-0,70	5,54***	-10,99***	-9,73***	4,78***	-9,00***
Lérida.....	-0,97	4,28***	-6,80***	-1,43	3,32***	-0,82
La Rioja.....	0,80	3,11***	-4,87***	-2,66***	-1,50	-2,83***
Lugo.....	-2,66***	2,84***	-8,06***	-10,63***	19,39***	-7,68***
Málaga.....	2,15**	12,77***	-15,07***	-7,97***	3,80***	-2,01**
Murcia.....	5,91***	11,01***	-14,82***	0,44	-0,28	0,31
Navarra.....	1,62	-1,42	-1,40	-7,89***	2,36**	-3,21***
Orense.....	-0,48	10,14***	-12,35***	-8,74***	9,99***	-11,60***
Asturias.....	-2,32**	8,37***	-5,85***	-11,65***	7,09***	-9,92***
Palencia.....	0,18	10,05***	-5,73***	-7,93***	1,10	-6,49***
Las Palmas.....	2,23**	13,16***	-16,77***	-4,85***	6,54***	-4,27***
Pontevedra.....	0,37	6,46***	-13,32***	-15,34***	15,84***	-8,02***
Salamanca.....	-0,04	-1,90*	-3,68***	-7,07***	1,13	-5,18***
Sta. Cruz Tenerife.....	1,20	16,69***	-13,48***	-10,48***	5,55***	-7,53***
Cantabria.....	-1,21	-0,26	-6,02***	-10,65***	9,64***	-5,02***
Segovia.....	1,57	3,87***	-8,92***	-2,72***	-2,50**	-4,62***
Sevilla.....	3,17***	8,64***	-11,42***	-15,81***	0,85	3,00***
Soria.....	0,17	6,64***	-6,68***	-2,75***	1,67*	-5,27***
Tarragona.....	1,78*	8,05***	-9,98***	4,55***	-1,10	2,43**
Teruel.....	1,45	3,38***	-9,63***	-2,57**	-1,14	-4,85***
Toledo.....	2,93***	7,05***	-14,85***	-5,90***	0,50	0,37
Valencia.....	0,16	0,78	-7,74***	-3,71***	1,41	-3,44***
Valladolid.....	1,79*	-0,93	-1,31	-8,44***	0,92	-4,46***
Vizcaya.....	-2,17**	-4,27***	2,06**	-5,52***	4,36***	-5,43***
Zamora.....	1,67*	0,81	-8,43***	-7,11***	1,70*	-6,51***
Zaragoza.....	0,71	7,93***	-5,73***	-1,98**	0,61	-4,48***

Notas: Resultados del contraste de igualdad de proporciones para las variables discretas (Jóvenes, Educ. baja, Educ. sup, Extranj. y Mayores) y de igualdad de medias para la continua (Núm. Hijos). ***, **, * denotan rechazo de la hipótesis nula de igualdad a un nivel de significación del 10, 5 y 1% respectivamente. Jóvenes: edad entre 16 y 24 años; Educ. baja: nivel educativo inferior a secundario; Educ. sup: nivel educativo superior al secundario; Extranj.: individuos sin nacionalidad española; Mayores: individuo en un hogar con algún miembro de edad superior a 65 años; Núm. Hijos: número de hijos menores del individuo.

cial que guía este trabajo, dado que no podemos descartar que parte de la información contenida en las características individuales, que es imperfectamente capturada por los valores medios, contamine

la estimación de los efectos puramente espaciales que podrían estar operando en los mercados de trabajo regionales. La siguiente sección trata de realizar una valoración de esta posibilidad.

CUADRO N.º 6

ESTADÍSTICOS DE DEPENDENCIA ESPACIAL PARA LAS TASAS DE PARO REALES Y CONTRAFACUALES

2013	MORAN'S I			GEARY'S C		
	I	z	p-val	c	z	p-val
W contigüidad						
Real	0,747	8,139	0,0000	0,207	-7,705	0,0000
Contrafact. PSM	0,589	6,523	0,0000	0,374	-5,698	0,0000
Contrafact. NNM	0,622	6,844	0,0000	0,358	-6,03	0,0000
Contrafact. RW	0,561	6,384	0,0000	0,375	-4,909	0,0000
W inversa distancia²						
Real	0,614	11,324	0,0000	0,413	-7,334	0,0000
Contrafact. PSM	0,436	8,214	0,0000	0,516	-5,106	0,0000
Contrafact. NNM	0,481	8,98	0,0000	0,498	-5,735	0,0000
Contrafact. RW	0,461	8,876	0,0000	0,482	-4,046	0,0000
2006						
2006	MORAN'S I			GEARY'S C		
	I	z	p-val	c	z	p-val
W contigüidad						
Real	0,694	7,686	0,0000	0,29	-6,249	0,0000
Contrafact. PSM	0,499	5,518	0,0000	0,482	-5,009	0,0000
Contrafact. NNM	0,634	7,012	0,0000	0,381	-5,629	0,0000
Contrafact. RW	0,561	6,206	0,0000	0,421	-5,364	0,0000
W inversa distancia²						
Real	0,436	8,252	0,0000	0,442	-5,447	0,0000
Contrafact. PSM	0,278	5,339	0,0000	0,653	-4,283	0,0000
Contrafact. NNM	0,369	7,004	0,0000	0,554	-4,682	0,0000
Contrafact. RW	0,326	6,223	0,0000	0,566	-4,771	0,0000

Nota: ver cuadro n.º 2.

IV. TASAS DE PARO CONTRAFACUALES

En esta sección, en primer lugar, se obtienen tasas de paro contrafactuales para cada una de las provincias. Estas tasas de paro resultan de descontar el efecto de las diferencias provinciales en la distribución de las características observables de los individuos. En otras palabras, se calculan unas tasas de paro provinciales bajo el supuesto de que la población activa de todas las provincias tuviese características similares. En segundo lugar, se analiza la distribución espacial de esas tasas contrafactuales para valorar hasta qué punto las diferencias regionales observadas en las tasas de paro reales, y los rasgos distintivos de su distribución espacial, son ocasionadas por diferencias entre provincias en las características de los individuos.

La idea subyacente a las tasas de paro contrafactuales corresponde a lo que se ha venido en denominar como la descomposición de la brecha en un determinado estadístico para una variable de interés en dos grupos de la población. La más popular es la

descomposición propuesta inicialmente por Oaxaca (1973) y Blinder (1973) para la brecha en la media de una variable, por ejemplo el salario, entre dos grupos de trabajadores, por ejemplo hombres y mujeres. En el caso que nos ocupa en este trabajo, la diferencia entre las tasas de paro de dos provincias A y B incorpora el efecto de las diferencias entre ellas en las características observables de los trabajadores y una componente no explicada, atribuible a diferencias entre esos territorios en el efecto de las características sobre la probabilidad de estar desempleado, y a factores inobservables. El objetivo es, por tanto, determinar la contribución de las diferencias en las características observables a la diferencia en las tasas de paro entre las provincias A y B. Para ello, el objetivo es obtener la tasa de paro que hubiésemos observado en la provincia A si sus trabajadores hubiesen sido como los de la provincia B, es decir, si no difiriesen en sus características.

Para obtener las tasas de paro contrafactuales se puede recurrir a diversas propuestas sugeridas

en la literatura (véase, por ejemplo, Fortin, Lemieux y Firpo, 2011). Todas ellas requieren el cumplimiento de la condición de *ignorabilidad*, o de independencia condicionada, y de soporte común. La primera implica que los factores inobservables que determinan la probabilidad de estar desempleado no afectan a estar localizado en A o B. Por tanto, esta condición se incumple si: i) la distribución de características observables e inobservables es distinta en A y B; ii) la elección de trabajar en A o B depende de los inobservables; o iii) la distribución de inobservables depende de la región (por ejemplo, que dadas las características observables, el grado de habilidad de los individuos es diferente en A y B). En general, asumimos que si se dispone de un conjunto suficientemente amplio de controles para las características observables, la ubicación en una u otra región, condicionada a esos factores observables, es aleatoria. Por su parte, la condición de soporte común (para observables e inobservables) implica que no haya trabajadores con determinadas características solo en una de las regiones. O lo que es lo mismo, que cualquier trabajador en la población tenga opciones de estar localizado tanto en la provincia A como en la B. En consecuencia, la validez de los resultados del ejercicio empírico depende, en gran medida, de la consideración de un grupo suficientemente amplio de características observables, que determinan la probabilidad de desempleo, y de que esas características se encuentren suficientemente representadas en todas las provincias.

En este trabajo, las tasas de paro contrafactuales se construyen empleando tres métodos alternativos. El primero de ellos está basado en la reponderación de la muestra de trabajadores en A para hacerla lo más parecida posible a la de la provincia B en cuanto a sus características observadas, siguiendo la propuesta de DiNardo, Fortin y Lemieux (1996). Para ello se definen los siguientes pesos:

$$\psi(X) = \frac{Pr(X|Prov=B)}{Pr(X|Prov=A)} = \frac{Pr(Prov=B|X) / Pr(Prov=B)}{Pr(Prov=A|X) / Pr(Prov=A)}$$

donde X denota el conjunto de características observables y Pr la probabilidad. $Pr(Prov=r|X)$, $r=A,B$, es la probabilidad de que un individuo se localice en la provincia correspondiente condicionada a las características observables, mientras que $Pr(Prov=r)$, $r=A,B$, es la probabilidad incondicionada de localización en esa provincia. La estimación de los pesos, $\psi(X)$, se puede obtener fácilmente combinando los datos de los individuos en las regiones A y B y estimando $\hat{Pr}(Prov=B|X)$ mediante un modelo *logit* o

probit para la probabilidad de estar localizado en la provincia B. De esta manera, $\hat{Pr}(Prov=A|X) = 1 - \hat{Pr}(Prov=B|X)$. Por su parte, $\hat{Pr}(Prov=B)$ y $\hat{Pr}(Prov=A)$ se obtienen mediante las proporciones muestrales de las observaciones de las provincias B y A respectivamente.

Los pesos estimados se emplean para obtener la siguiente media ponderada:

$$U_A^{C1} = \frac{1}{N_A} \sum_{i \in A} \psi(X_i) u_i$$

donde u es una variable dicotómica que toma valor 1 si un individuo está desempleado y 0 en caso contrario, y N_A denota el número de individuos en la muestra de la provincia A. Definimos a dicha media ponderada como la tasa de paro contrafactual para la provincia A con una composición de la población activa similar a la de B. Nótese que si todos los pesos toman valor 1 obtenemos la tasa de paro observada en la provincia A; es decir, con la composición de esa provincia. Por tanto, la diferencia entre la tasa de paro efectivamente observada en la provincia A (U_A) y su correspondiente tasa de paro contrafactual (U_A^{C1}) es atribuible al efecto composición. Es decir, que la tasa de paro contrafactual de la provincia A descuenta el efecto de las diferencias en las características entre los trabajadores de A y B.

Los otros dos métodos utilizados para obtener las tasas de paro contrafactuales para cada provincia se basan en procedimientos habituales en las estimaciones del efecto de tratamiento mediante técnicas de emparejamiento, *matching* (véase, por ejemplo, Imbens, 2015). De manera muy breve, y aplicado al caso que nos ocupa, se trata de determinar un conjunto de individuos de A con características (observadas) lo más parecidas posible a las de B para, a partir de ellos, calcular la tasa de paro contrafactual de la provincia A. Dado que se selecciona el conjunto de individuos de A con características similares a las de B, la correspondiente tasa de paro contrafactual descuenta el efecto de las diferencias en características entre A y B (1). La selección de la muestra de individuos de A que minimiza las diferencias en características con la provincia B está basada en una medida de similitud, es decir, en un estadístico que indica lo parecidas que son dos observaciones. En este trabajo utilizamos las dos más habituales. Por una parte, el emparejamiento basado en el vecino más próximo (*nearest-neighbor matching*, *NNM*) y por otra el que utiliza el emparejamiento basado en el valor de la propensión (*propensity-score matching*, *PSM*).

En el caso del *NNM*, para cada trabajador de B se determinan aquellos de A con características menos distantes (utilizando, por ejemplo, la distancia de Mahalanobis). De esa forma se obtiene la muestra emparejada de A de características similares a la muestra de trabajadores de B, con la que se calcula la tasa de paro contrafactual como:

$$U_A^{C2} = \frac{1}{N_{A^*}} \sum_{i \in A^*} u_i$$

donde A^* denota la muestra de la provincia A resultante del proceso de emparejamiento basado en el *NNM*.

Por su parte, el método de *PSM* realiza el emparejamiento de los trabajadores en A con los de B no directamente a través de las características observadas, sino mediante la estimación de las probabilidades predichas de pertenencia a la provincia B (valor de la propensión). Ello implica la estimación de un modelo probabilístico (*logit* o *probit*) de, en nuestro caso, estar localizado en la provincia B *versus* en la A, utilizando el conjunto de características observadas en X. La estimación de dicho modelo se utiliza para obtener el *propensity score* de cada individuo en A y B. A partir de ello se determina la muestra emparejada de A, con la que se calcula la tasa de paro contrafactual U_A^{C3} , de manera equivalente a como se hace con U_A^{C2} .

Para obtener la tasa de paro contrafactual en cada una de las provincias con cualquiera de los tres métodos es necesario identificar una provincia de referencia. En el ejercicio empírico se ha seleccionado como referencia a Madrid, dado que es una de las provincias con menor tasa de desempleo, su mercado de trabajo es amplio y suficientemente heterogéneo y se encuentra en el centro del territorio analizado. En todo caso, conviene indicar que las conclusiones derivadas tomando a esa provincia como referencia resultaron robustas a la consideración de otras provincias alternativas. Por otra parte, cabe indicar que en el cálculo de las tasas contrafactuales basadas en los métodos de *matching* se ha empleado la opción de reemplazo de las observaciones, lo que implica que la observación de un individuo puede utilizarse más de una vez como contraparte de la correspondiente a la de otro trabajador de la provincia de referencia.

Las tasas de paro contrafactuales obtenidas para cada una de las provincias con los tres métodos comentados anteriormente se reproducen en el último bloque de columnas del cuadro n.º 2. Se ha

empleado un modelo *probit* para implementar los métodos basados en la reponderación y en el *propensity-score*, utilizándose en los tres casos las características del individuo y del hogar definidas en el cuadro n.º 1. Se pueden apreciar dos rasgos generales. El primero referido al efecto composición, ya que las tasas netas del efecto de las características individuales resultan menores que las tasas observadas en las provincias de desempleo elevado (por ejemplo, las tasas contrafactuales en Cádiz resultan alrededor de siete puntos porcentuales inferiores), mientras que aumentan en las de bajo desempleo (en Navarra son entre dos y tres puntos superiores). El segundo, relativo a las discrepancias de las tasas contrafactuales obtenidas con cada uno de los métodos. En general, los tres métodos proporcionan valores no demasiado discrepantes, aunque hay algunas discordancias que conviene tener en consideración. Por ello, en lo que sigue, se presentan los resultados para las tres tasas contrafactuales.

Una imagen más precisa del efecto de las características de la población activa sobre la tasa de paro la obtenemos a través de la estimación de la función de densidad de las tasas contrafactuales, y su comparación con la del paro observado. Las mismas se incluyen en el gráfico 1. Lo más llamativo es la disminución en la masa de probabilidad para los valores más elevados de la tasa de paro (25 por 100 en adelante) y un aumento en las zonas centrales de la distribución (entre el 15 y el 25 por 100). Así y todo, las distribuciones contrafactuales mantienen indicios bastante claros de bimodalidad, aunque con valores de los dos grupos de provincias mucho más próximos que en el caso de las tasas de paro observadas. Finalmente, indicar que se aprecia un moderado desplazamiento hacia la derecha de las distribuciones del paro contrafactual en la zona de bajo desempleo. En su conjunto, el análisis de las funciones de densidad del paro observado y contrafactual confirma que una parte de las diferencias provinciales en la incidencia del desempleo se debe a diferencias de composición de la población activa, siendo su contribución más intensa para explicar los valores en las provincias de mayor desempleo. Sin embargo, no puede ni debe obviarse que la distribución de las tasas de paro netas del efecto composición continúan presentando un elevado grado de desigualdad, lo que justifica el estudio de las mismas desde una perspectiva espacial.

Para ahondar en el análisis de las características espaciales de la distribución de las tasas de paro netas del efecto composición, en el gráfico 2 se muestran los mapas correspondientes, mientras que

los estadísticos de dependencia espacial asociados se reproducen en el cuadro n.º 3. En el caso de la representación a través de los mapas, cabe indicar, en primer lugar, que se ha mantenido la definición de las categorías de provincias basada en los cuartiles de la distribución de las tasas de paro observadas para, de esa manera, facilitar su comparación con la de las tasas contrafactuales. La imagen que se deriva de los mapas es pues consistente con la disminución, comentada anteriormente, de las diferencias tras condicionar por el efecto composición. Sin embargo, y a pesar de ello, los mapas siguen mostrando agrupaciones espaciales de provincias con valores más elevados, básicamente en la mitad sur, y de valores más moderados en el noreste, quizá con la excepción de algunas de las provincias de la costa mediterránea. Por su parte, los estadísticos de los contrastes I de Moran y c de Geary, con las dos matrices de pesos espaciales consideradas, presentan valores para las tasas de paro contrafactuales algo menores a los de las tasas reales, aunque en ningún caso conducen a no rechazar claramente la hipótesis nula de ausencia de correlación espacial. Por tanto, los resultados sugieren que la mayor parte de la dependencia espacial observada en las tasas de paro provinciales no es debida al efecto de la distribución espacial de las características de la población activa, sino que muy posiblemente se deba a factores esencialmente territoriales, más allá de los que pueden estar relacionados con las características observables de los trabajadores.

V. DIFERENCIAS ENTRE FASES EXPANSIVAS Y RECESIVAS

Hasta este punto los resultados se han obtenido y presentado para la información referida al segundo trimestre de 2015. En ese año, el desempleo en el conjunto de España había descendido ligeramente hasta el 22.4 por 100, desde su punto más álgido en 2013 donde se sobrepasó ligeramente el 26 por 100, y muy por encima del 8.4 por 100 de 2006 que marca el punto de referencia de la última fase expansiva de la economía española antes del impacto de la Gran Recesión. Dadas esas importantes diferencias en las tasas de paro entre el período expansivo y el de recesión, en este apartado se analiza si la contribución de las características individuales a las diferencias regionales en la incidencia del desempleo es estable en el tiempo o bien si varía entre períodos de expansión y de crisis. Para ello se utiliza la información muestral contenida en la *EPA* para los segundos trimestres de 2006 y 2013.

El procedimiento es el mismo que el seguido en el caso del apartado anterior para los últimos datos disponibles: en primer lugar, se obtienen las tasas de paro contrafactuales para cada provincia, netas del efecto composición de las características individuales y, en segundo lugar, se compara la distribución de esas tasas contrafactuales con las realmente observadas.

La estimación de las correspondientes funciones de densidad se muestra en los gráficos 4 y 5 para 2013 y 2006, respectivamente. Las conclusiones que se derivan para el año de mayores tasas de paro son muy parecidas a las comentadas anteriormente

GRÁFICO 4
FUNCIONES DE DENSIDAD DE LAS TASAS DE PARO REALES Y CONTRAFCTUALES, 2013

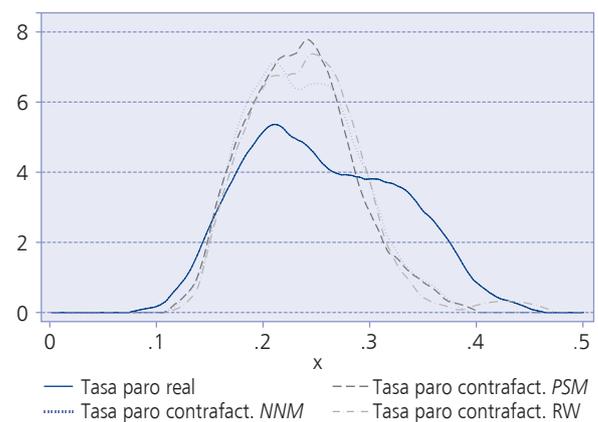
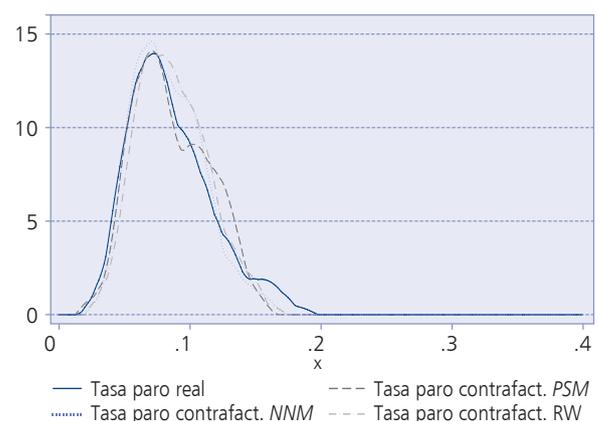


GRÁFICO 5
FUNCIONES DE DENSIDAD DE LAS TASAS DE PARO REALES Y CONTRAFCTUALES, 2006



para 2015: el efecto composición permite explicar buena parte del diferencial para las provincias con tasas de paro elevadas. Así, tras controlar por las diferencias en características de los trabajadores se obtiene una distribución más concentrada en torno a la tasa media de paro. Por el contrario, las distribuciones contrafactuales para 2006 difieren solo marginalmente de la del paro realmente observado. Únicamente se aprecia un descenso en la masa de probabilidad en los valores superiores más extremos que se compensa con un aumento en los valores intermedio-altos. Por tanto, de la comparación de los resultados para ambos años se deduce que las diferencias provinciales en las características de la población activa tienen una mayor repercusión en los períodos de crisis que en los expansivos.

Los mapas de los gráficos 6 y 7 permiten valorar las peculiaridades de la distribución geográfica de

las tasas de paro (reales y contrafactuales) en los dos años considerados. En 2013, la distribución de las tasas de paro observadas no difiere excesivamente de lo comentado anteriormente para 2015; es decir, valores elevados concentrados en el suroeste y bajos (en términos relativos) en el norte, aunque con mayor heterogeneidad en esta zona que entre las provincias meridionales. Lo mismo cabe decir respecto a la distribución de las tasas contrafactuales, dado que la relativa concentración de la distribución detectada anteriormente mediante las funciones de densidad no conduce a la desaparición de la distinción norte-sur o, en cualquier caso, a que se observen grupos de provincias próximas geográficamente con tasas similares. Los resultados de los contrastes de dependencia espacial, recogidos en el bloque superior del cuadro n.º 5, confirman esta circunstancia, dado que en todos los casos se rechaza la hipótesis nula de ausencia de

GRÁFICO 6
TASAS DE PARO REALES Y CONTRAFCTUALES (MISMA COMPOSICIÓN QUE MADRID), 2013

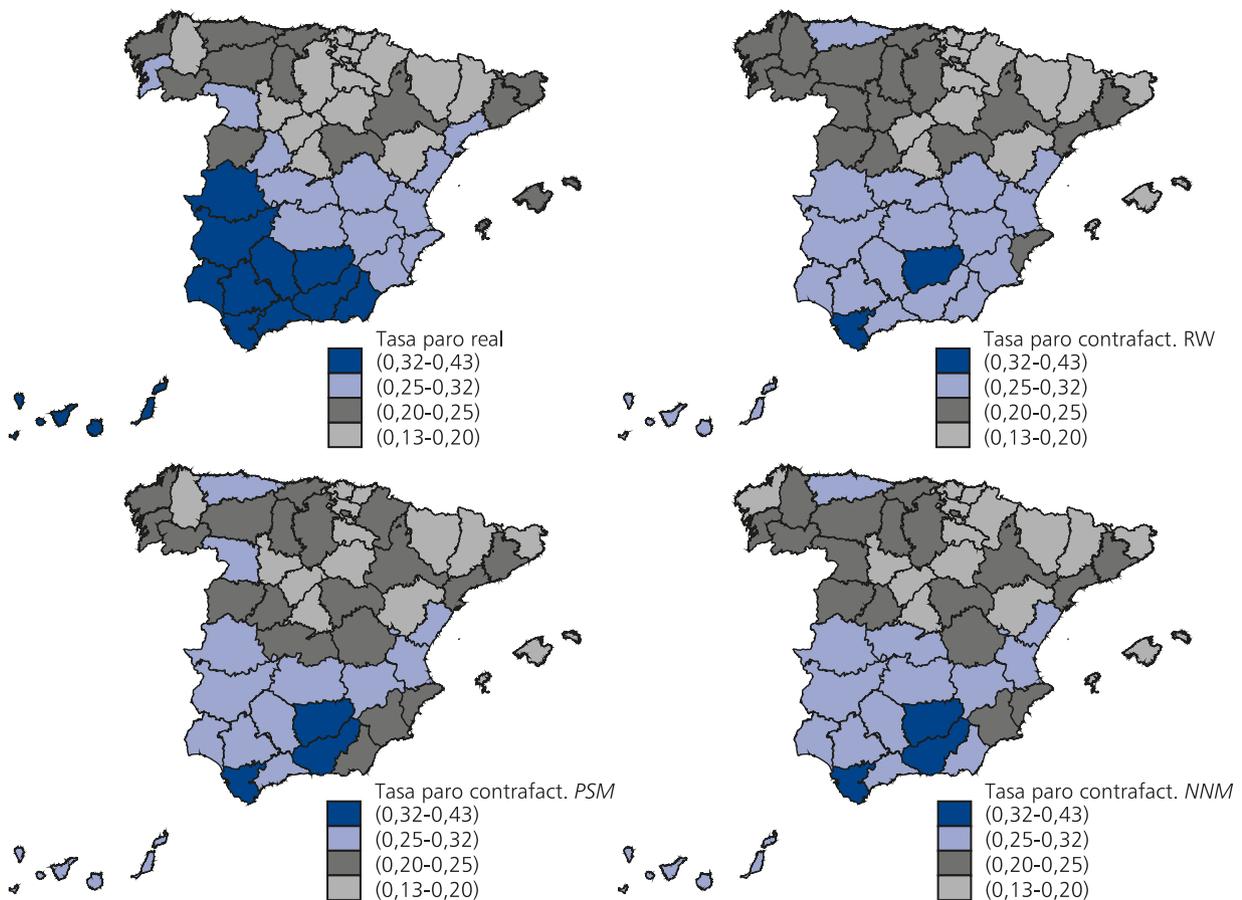
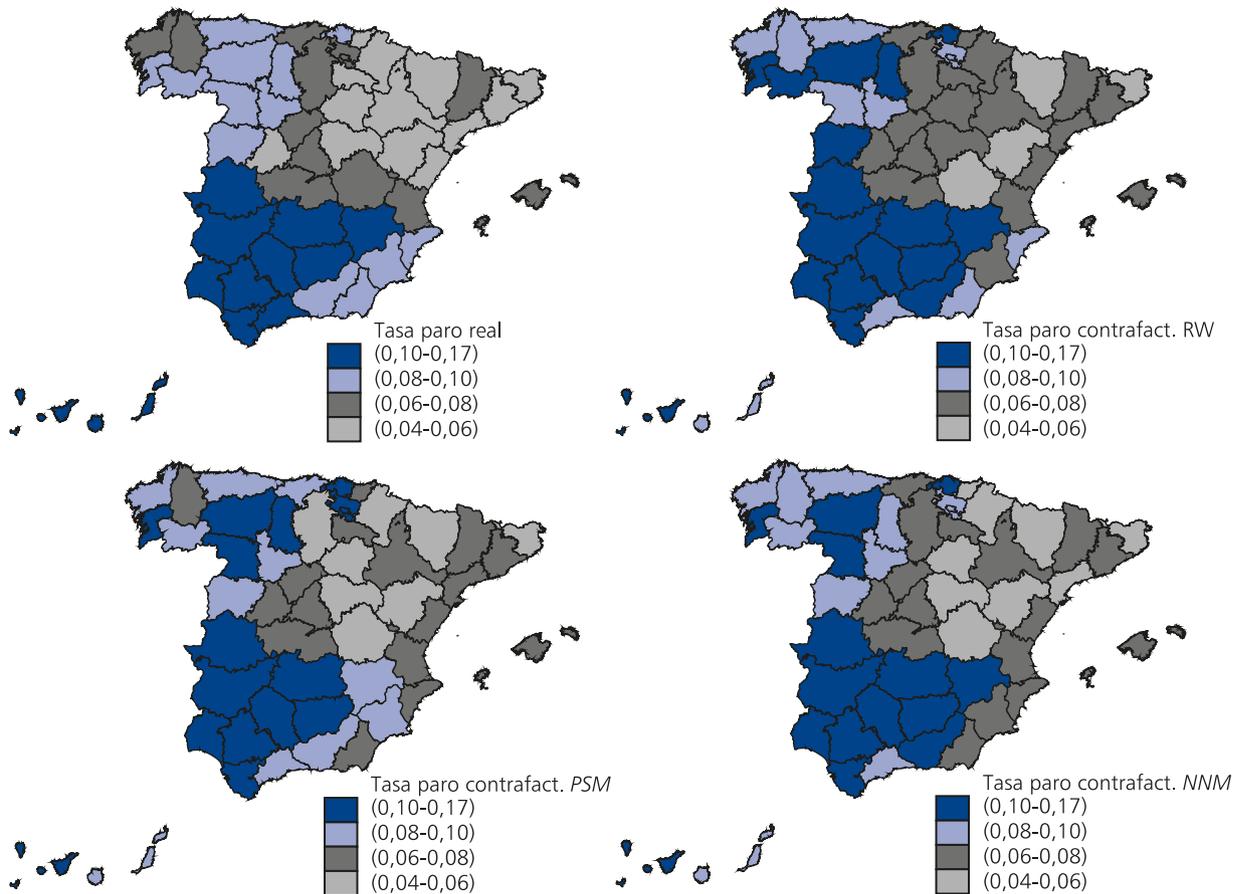


GRÁFICO 7

TASAS DE PARO REALES Y CONTRAFACUALES (MISMA COMPOSICIÓN QUE MADRID), 2006



correlación espacial. Por lo que respecta al año de menor desempleo, 2006, el mapa correspondiente a las tasas de paro observadas muestra la división suroeste-noreste habitualmente comentada para el desempleo en España. De hecho, ese patrón geográfico es mucho más claro en ese año del período expansivo que para los años posteriores a la crisis. En cualquier caso, lo más relevante para el objetivo específico de este trabajo es que tras condicionar por las características de los trabajadores la imagen general de la distribución geográfica del desempleo permanece en gran medida inalterada. Obviamente, se producen modificaciones en la clasificación en los grupos definidos por los cuartiles de las tasas de paro observadas en algunas provincias, pero estas no son lo suficientemente numerosas ni intensas para alterar la coincidencia con la imagen general de la distribución de las tasas realmente observadas en 2006. Al igual que en los casos anteriores, los

resultados de los contrastes I de Moran y c de Geary para ese año confirman la existencia de dependencia espacial tanto en las tasas observadas como en las contrafactuales.

En resumen, los resultados de este apartado sugieren que el efecto de las características de los trabajadores sobre las diferencias provinciales en las tasas de paro es más intenso en los períodos de elevado desempleo que cuando el paro es más moderado, pero también que ni en un caso ni en el otro el efecto composición permite explicar la mayor parte de la distribución espacial del desempleo.

VI. CONCLUSIONES

El objetivo prioritario de este trabajo ha sido el de valorar el impacto que tiene el efecto compo-

sición de la población activa sobre las diferencias territoriales en la tasa de paro; es decir, el que las características que determinan la propensión al desempleo de los trabajadores difieran entre unas provincias y otras. La evidencia obtenida a partir de los microdatos de la EPA se puede sintetizar en los siguientes puntos:

- Se confirma la intensa desigualdad espacial en la incidencia del desempleo en España. Como sucedió en períodos de crisis anteriores, la brecha entre territorios en las tasas de paro se ha ampliado respecto a la situación observada en el período de bajo desempleo previo a la Gran Recesión.
- Se observan significativas diferencias provinciales en la composición de la fuerza de trabajo. Las provincias difieren en cuanto a ciertas características de los trabajadores, fuertemente relacionadas con la propensión al desempleo (como por ejemplo la edad y el nivel educativo).
- La comparación de las tasas de paro reales y contrafactuales revela que solo parte de las diferencias entre provincias se explican por diferencias en composición. El factor composición explica, en gran medida, las diferencias en provincias con tasas de paro muy elevadas, pero tiene poca capacidad para explicar la situación de las de tasas de paro bajas.
- El control por el efecto composición disminuye la intensidad de la dependencia espacial en las tasas de paro. Pero aun así, ésta permanece como claramente significativa. Por tanto, se puede concluir que gran parte de la dependencia espacial en las tasas de paro no tiene su origen en la distribución espacial de las características observadas de los trabajadores. Trabajadores similares localizados en distintas provincias muestran distinta propensión al desempleo, y esa propensión en una provincia está relacionada con la de las provincias cercanas geográficamente.
- Ello deja margen para considerar como causantes de la misma a mecanismos generadores de externalidades intrínsecos a los mercados de trabajo regionales.
- El análisis para 2006 y 2013 revela que el factor composición tiene mayor impacto en períodos de recesión; es decir, de elevado

desempleo. Ello es consistente con el aumento en la probabilidad de desempleo para los trabajadores dotados de peores características. En períodos de expansión, con bajo desempleo, no hay diferencias relevantes antes y después de controlar por diferencias entre provincias en las características observables de los trabajadores. En cualquier caso, la mayor parte de la dependencia espacial no es atribuible al factor composición en ninguna de esas circunstancias.

Asimismo, de la evidencia obtenida se pueden derivar las siguientes implicaciones para los trabajos empíricos relativos al desempleo regional:

- La importancia de filtrar por las diferencias en la distribución geográfica de las características de los trabajadores a la hora de comparar magnitudes agregadas del mercado de trabajo. Esto es especialmente importante en caso de que no se pueda controlar perfectamente por el factor composición con los datos agregados.
- La importancia de considerar explícitamente el efecto de las externalidades entre regiones en los modelos empíricos, dado que la mayor parte de la dependencia espacial observada en las tasas de desempleo no parece deberse a diferencias geográficas en la distribución de las características observables de los trabajadores.

Finalmente, recordar que la validez de la evidencia obtenida está sujeta al cumplimiento de las condiciones de ignorabilidad y de soporte común, indicadas en la cuarta sección de este artículo. Mientras que diversas pruebas realizadas nos llevan a validar el cumplimiento de la segunda, resulta más complicado empíricamente comprobar la primera. De hecho, a partir de Imbens (2015), podemos asumir que el cumplimiento del supuesto de ignorabilidad es más razonable de lo que *a priori* podríamos pensar dado que la información contenida en el amplio conjunto de características observables de los trabajadores de cada provincia estaría incorporando buena parte de la correspondiente a características inobservables. Aun así, el análisis detallado de la existencia de factores inobservables no relacionados con las características observadas y, de forma más genérica, la aplicación de este tipo de estrategia empírica al análisis de las disparidades territoriales en las magnitudes del mercado laboral deberá ser abordada en estudios futuros.

NOTAS

(1) El interés habitual de los estimadores de *matching* de los efectos del tratamiento residiría, trasladado a nuestra aplicación, en la diferencia entre la tasa de paro contrafactual en A y la tasa de paro observada en B. Dicha diferencia mediría el efecto del tratamiento, en nuestro caso estar localizado en la provincia B, condicionado a las características de los individuos.

BIBLIOGRAFÍA

- BANDE, R., y M. KARANASSOU (2009), «Labour market flexibility and regional unemployment rate dynamics: Spain 1980-1995», *Papers in Regional Science*, 88: 181-207.
- BLANCHARD, O. J., y L. F. KATZ (1992), «Regional evolutions», *Brookings Papers on Economic Activity. Economic Studies Program. The Brookings Institution*, 23: 1-76.
- BLINDER, A. S. (1973), «Wage discrimination: Reduced form and structural estimates», *Journal of Human Resources*, 8: 436-455.
- CRACOLICI, M. F.; CUFFARO, M., y P. NIJKAMP (2007), «Geographical distribution of unemployment. An analysis of provincial differences in Italy», *Growth and Change*, 38: 649-670.
- DECRESSIN, J., y A. FATÁS (1995), «Regional labor market dynamics in Europe», *European Economic Review*, 39: 1627-1655.
- DINARDO, J.; FORTIN, N. M., y T. LEMIEUX (1996), «Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach», *Econometrica*, 64: 1001-1044.
- ELHORST, J. P. (2003), «The mystery of regional unemployment differentials: Theoretical and empirical explanations», *Journal of Economic Surveys*, 17: 709-748.
- EVANS, P., y B. McCORMICK (1994), «The new pattern of regional unemployment: causes and policy significance», *Economic Journal*, 104: 633-647.
- FILIZTEKIN, A. (2009), «Regional unemployment in Turkey», *Papers in Regional Science*, 88: 863-878.
- FORTIN, N. M.; LEMIEUX, T., y S. FIRPO (2011), «Decomposition methods in economics», en ASHENFELTER, O., y CARD, D. (eds.): *Handbook of Labor Economics*, Vol. IV, Part A, Elsevier, Amsterdam, 1-102.
- HALL, R. E. (1972), «Turnover in the labor force», *Brookings Papers on Economic Activity*, 3: 709-764.
- IMBENS, G. W. (2015), «Matching methods in practice: Three examples», *Journal of Human Resources*, 50: 373-419.
- LAZAR, F. (1997), «Regional unemployment rate disparities in Canada: some possible explanations», *Canadian Journal of Economics*, 10: 112-129.
- LÓPEZ-BAZO, E.; DEL BARRIO, T., y M. ARTÍS (2002), «The regional distribution of Spanish unemployment: A spatial analysis», *Papers in Regional Science*, 81: 365-389.
- (2005), «Geographical distribution of unemployment in Spain», *Regional Studies*, 39: 305-318.
- LÓPEZ-BAZO, E., y E. MOTELLÓN (2013a), «The regional distribution of unemployment. What do micro-data tell us?», *Papers in Regional Science*, 92: 383-405.
- (2013b), «Disparidades en los mercados de trabajo regionales. El papel de la educación», *Papeles de Economía Española*, 138: 46-61.
- MARSTON, S. T. (1985), «Two views of the geographic distribution of unemployment», *Quarterly Journal of Economics*, 100: 57-79.
- MARTIN, R. (1997), «Regional unemployment disparities and their dynamics», *Regional Studies*, 31: 237-252.
- MARTIN, R., y P. SUNLEY (1999), «Unemployment flow regimes and regional unemployment disparities», *Environmental and Planning A*, 31: 523-550.
- MORENO, R., y E. VAYÁ (2000), *Técnicas econométricas para el tratamiento de datos espaciales: La econometría especial*, UB 44 manuals, Edicions Universitat de Barcelona, Barcelona.
- OAXACA, R. L. (1973), «Male-female wage differentials in urban labor markets», *International Economic Review*, 14: 693-709.
- OCDE (2011), *Regions at a glance 2011*, OECD Publishing, París.
- OVERMAN, H. G., y D. PUGA (2002), «Unemployment clusters across Europe's regions and countries», *Economic Policy*, 34: 115-147.
- PARTRIDGE, M. D., y D. S. RICKMAN (1997), «The dispersion in US unemployment rates: The role of market and nonmarket equilibrium factors», *Regional Studies*, 31: 593-606.
- RIOS, V. (2016), «What drives unemployment disparities in European regions? A dynamic spatial panel approach», *Regional Studies*, en prensa.
- ROSEN, S. (1974), «Hedonic prices and implicit prices», *Journal of Political Economy*, 86: 34-53.
- TAYLOR, J., y S. BRADLEY (1997), «Unemployment in Europe: A comparative analysis of regional disparities in Germany, Italy and the UK», *Kyklos*, 50: 221-245.