

Jordi Pons Novell*
Daniel A. Tirado Fabregat*

LOS DETERMINANTES DE LA DESIGUALDAD ECONÓMICA REGIONAL EN ESPAÑA

En este trabajo se efectúa una cuantificación de la importancia relativa de los distintos factores que tradicionalmente se han utilizado en la literatura para explicar la desigual distribución territorial de la actividad económica en España a lo largo del Siglo XX. En particular, se propone un estudio de los determinantes de la densidad relativa del PIB en las provincias españolas en cuatro cortes temporales: 1920, 1955, 1975 y 2003, siguiendo una metodología que permite identificar la importancia de los elementos de naturaleza puramente geográfica frente a los llamados de geografía de segunda naturaleza (es decir, la moldeada por la mano del hombre a través de la actividad económica) y de aquellos que se asocian a la existencia de economías de aglomeración y que amplifican las diferencias iniciales en uno u otro sentido.

Palabras clave: geografía económica, aglomeración, desigualdades, análisis regional.

Clasificación JEL: R12, O52.

1. Introducción

En los últimos años, los estudios dedicados al análisis de la distribución desigual de la actividad en el espacio, o a su elevada concentración en un número muy limitado de regiones, han vuelto a otorgar un peso explicativo elevado a la existencia de diferencias entre los territorios. No obstante, no hay acuerdo acerca de la impor-

tancia que debería atribuirse a las distintas características territoriales potencialmente relevantes.

En este sentido, Gallup, Sachs y Mellinger (1999), Sachs (2000) o Rapaport y Sachs (2003) han sostenido que son las características geográficas (e.g. clima o localización) las que determinan que ciertas regiones se encuentren ante muros infranqueables para el avance en su proceso de desarrollo. Por tanto, la geografía pura sería el elemento central a considerar a la hora de explicar las diferencias existentes en la distribución de la actividad en el espacio.

De forma no necesariamente enfrentada a esta percepción, la Nueva Geografía Económica (NGE) ha mostrado que la existencia de rendimientos crecientes en algunas actividades productivas podría ser clave en

* Universitat de Barcelona.

Los autores agradecen a Alfonso Herranz, Julio Martínez, Javier Silvestre y un evaluador anónimo sus útiles comentarios y sugerencias. Asimismo, reconocen el apoyo financiero prestado por el Ministerio de Educación y Ciencia a través de los proyectos SEJ2005-03196/ECON y SEJ2005-02498/ECON.

la explicación de la desigualdad económica territorial¹. La razón es que sólo con su presencia sería posible explicar la elevada concentración de la actividad en el espacio, a partir de la consideración de la existencia de procesos de causación acumulativa que tenderían a ampliar en el tiempo las diferencias territoriales preexistentes, con independencia de su origen o tamaño inicial. En este sentido, Krugman (1993) o Venables (1999) insisten en el hecho de que las condiciones diferenciales de partida, además de marginales, podrían tener un origen espurio, generando igualmente diferencias regionales relevantes a partir de la interacción entre presencia de rendimientos crecientes y reducción de los costes de transporte. Es decir, originadas a partir de la acción del hombre en la actividad económica.

En cualquier caso, la distinción entre una y otra concepción teórica se ha trasladado a la literatura económica en la definición de dos visiones de las diferencias geográficas existentes entre las regiones: la consideración geográfica *à la* Sachs, o *first nature geography*, y la realizada por la NGE, o *second nature geography*.

Este debate también se ha reproducido en la historia económica, y en particular en la española. De hecho, esta economía se caracteriza por la existencia de grandes diferencias en la distribución regional de la actividad, siendo, por tanto, un campo abonado para el estudio de los determinantes de las mismas. En este sentido, también han convivido dos líneas básicas de argumentación. Por una parte, algunos trabajos (Dobado, 2004) han apuntado que la explicación de las diferencias en densidad económica o de población entre los territorios españoles debe relacionarse con la existencia de caracteres geográficos regionales muy dispares. Otros, aun reconociendo la existencia de diferencias geográficas de primera naturaleza, han argumentado que los grandes desequilibrios existentes entre las regiones españolas hoy sólo pueden ser explicados consi-

derando la existencia de economías de aglomeración en las actividades productivas (Ayuda, Collantes y Piniella, 2005).

En estas condiciones, el trabajo que se desarrolla a continuación sigue la propuesta metodológica de Roos (2005) para abordar una cuantificación de la importancia relativa de unos y otros argumentos en la explicación de la desigual distribución territorial de la actividad económica en España a lo largo del Siglo XX. En particular, se propone un estudio de los determinantes de la densidad relativa del PIB en las provincias españolas en cuatro cortes temporales: 1920, 1955, 1975 y 2003, siguiendo una metodología que permite identificar la importancia de los elementos de naturaleza puramente geográfica frente a los llamados de geografía de segunda naturaleza (es decir, la moldeada por la mano del hombre a través de la actividad económica) y de aquéllos que se asocian a la existencia de economías de aglomeración y que amplían las diferencias iniciales en uno u otro sentido.

Del ejercicio se desprende que los elementos de geografía de primera naturaleza, por si solos, nunca han servido para explicar más allá del 20 por 100 de la varianza de la densidad relativa del PIB entre las provincias españolas, y que la capacidad explicativa de este conjunto de variables se ha ido reduciendo a lo largo del proceso de desarrollo económico español en el Siglo XX. Por el contrario, los elementos de segunda naturaleza, combinados con la existencia de economías de aglomeración, explican más del 70 por 100 de la varianza de la densidad relativa del PIB en todos los cortes temporales analizados, con porcentajes crecientes a lo largo del siglo. Finalmente, se muestra que los elementos de la llamada geografía de segunda naturaleza han ido progresando como determinantes de la desigual localización de la actividad, de forma que, aunque en los años veinte del siglo pasado, su contribución neta a la explicación de la desigualdad económica territorial era la mitad de la correspondiente a la geografía pura, en la actualidad (2003) multiplican por tres a la misma por lo que se refiere a su capacidad explicativa.

¹ FUJITA, KRUGMAN y VENABLES (1999) sintetiza los fundamentos teóricos de esta línea de análisis.

El resto del trabajo se organiza de la siguiente forma. En el segundo apartado se realiza un breve repaso de los estudios que, en los últimos años y desde diferentes ópticas, han ofrecido explicaciones sobre los determinantes de la elevada concentración de la actividad económica en el territorio en España. En el tercero se presenta la metodología seguida en el presente estudio y los datos empleados para la realización del mismo. El apartado siguiente ofrece los principales resultados obtenidos en el análisis empírico y se discuten en el marco de los obtenidos en otros trabajos relacionados con el tema tratado. Finalmente, el último epígrafe se dedica al resumen de las principales conclusiones.

2. Revisión de la literatura

La desigual localización de la actividad económica o de la población en el espacio ha vuelto a centrar la atención de los analistas económicos españoles de forma reciente. En este contexto de renovado interés, la mayor parte de trabajos han destacado la persistencia y la profundidad alcanzada por el fenómeno a lo largo de los últimos dos siglos. Sin embargo, han discrepado, en mayor o menor medida, a la hora de apuntar cuáles han sido sus determinantes fundamentales.

Por lo que respecta a la evidencia descriptiva, y en caso de tomar como referencia territorial para el análisis la provincia y el PIB por unidad de superficie como indicador de la densidad económica regional, la distancia existente entre las provincias españolas ha sido, y es, enorme. Dobado (2006), a partir del dato de PIBpm por km² en 1955 proveniente de Fundación BBV (1997) apunta que la densidad económica de la provincia más densa (Barcelona) era más de 50 veces superior a la de Teruel, la última clasificada al analizar esta variable para el conjunto de las provincias españolas. El mismo tipo de información para 2000 muestra un cambio de protagonistas (ahora la provincia más densa es Madrid y la menos densa es Soria) pero una distancia entre los mismos todavía más espectacular. Madrid presentaba en 2000 una densidad económica casi 100 veces superior a la de Soria.

La evidencia de tipo descriptivo elaborada para la realización del ejercicio que aquí se desarrolla (véase Anexo A) refuerza la impresión acerca de la existencia de una elevada desigualdad en la distribución de la actividad en el territorio, presente ya en 1920. En este sentido, destaca la elevada persistencia que caracteriza la desigualdad económica regional, de forma que el club de las cuatro provincias más densas de la economía española está compuesto por los mismos miembros (Madrid, Barcelona, Vizcaya y Guipúzcoa) en todos los cortes temporales analizados. Lo mismo ocurre con el grupo de provincias menos densas. En este caso son: Soria, Cuenca, Teruel, Guadalajara y Cáceres las provincias que ocupan las últimas posiciones del *ranking* en todos los años analizados. Sin embargo, en términos relativos, se observa que la brecha existente entre las provincias españolas más y menos densas habría iniciado una tímida reducción desde mediados de los años setenta y hasta la actualidad. Es decir, los valores alcanzados por los indicadores de desigualdad habrían seguido una trayectoria creciente a lo largo de los cortes temporales correspondientes a 1920, 1950 y 1975. Por el contrario, entre esta última fecha y 2003, la desigualdad en la distribución territorial de la actividad económica en España habría iniciado un camino descendente².

En el caso de que el indicador utilizado para medir la desigualdad sea la localización de la población en el espacio, se puede sostener el mismo tipo de conclusiones basadas, en este caso, en la evidencia correspondiente a un arco temporal más amplio. Ayuda, Collantes y Piniella (2005) apuntan que la evolución de la desigualdad interprovincial en términos de densidad de población, medida a través de un índice de Gini, ha sido creciente desde el inicio del Siglo XIX (es decir, con el dato de 1800 elaborado a partir del Censo de la Población Española de 1787) hasta 2000. Además, en ese mismo trabajo se ofrece evidencia contundente acerca de la persistencia

² Ver Anexo A.

de la distribución de la población en el territorio. En este sentido, se señala que el coeficiente de correlación por rangos de Spearman entre la distribución de las provincias más densas en 1800 y en 2000 es de 0,916. Cuando se calcula un coeficiente de correlación simple con los datos provinciales de densidad en 1787 y en 1995 (Dobado, 2004) el valor obtenido es de 0,665.

En resumen, se puede sostener la existencia de grandes diferencias en la distribución espacial de la actividad económica en España que, con la excepción de los últimos 25 años, lejos de desaparecer podrían haber ido ampliándose a lo largo de los dos últimos siglos. Además, se ha mostrado que los territorios que tenían una mayor densidad a inicios de este período han tendido a mantener esta posición. No obstante, se debe admitir la existencia de cierta movilidad en la ordenación de las provincias españolas en lo que se refiere a la concentración de la actividad económica. Así, frente a la presencia de territorios, como Asturias, que han perdido posiciones a lo largo de este lapso, otros, como Madrid, las han ganado. Sin embargo, en general, los Cuadros A.3 y A.4 del Anexo A muestran que la característica que define al *ranking* de densidad económica de las provincias españolas es la persistencia.

No obstante, el consenso sobre los determinantes de esta realidad dista de ser unánime entre los analistas. Así, algunos trabajos han ofrecido explicaciones diferentes, aunque no siempre contradictorias, de los elementos clave para su comprensión. En este sentido, se puede apuntar la existencia de dos corrientes de opinión que se sitúan, de forma más o menos clara, en cada una de las dos líneas de interpretación esgrimidas en la literatura internacional. Por una parte se encuentran los que abogan por que la geografía pura es el elemento fundamental en la comprensión de la existencia y persistencia de estas desigualdades. Es decir, que se alinean con las propuestas explicativas sostenidas por Sachs y colaboradores. Y aquellos que, atendiendo a que la explicación de las mismas debería también dar respuesta al crecimiento de la desigualdad durante fases importantes de la historia económica española,

fundamentan su explicación en la presencia de elementos favorecedores de la aglomeración en los procesos productivos. Es decir, ofreciendo explicaciones en consonancia con las propuestas de la llamada NGE.

Así, Dobado (2004 y 2006) ha recogido la tradición historiográfica española que identificaba a la geografía como uno de los problemas (males) que había de afrontar la economía española (algunas regiones) a la hora de avanzar por su senda de desarrollo (esta visión aparece en referencias de autores de finales del Siglo XIX)³ y la ha trasladado a un marco explicativo y analítico contemporáneo. En distintos trabajos, este autor ha defendido la validez de este marco conceptual para la comprensión del caso español y ha ofrecido evidencia al respecto. En particular, Dobado (2006) sigue la propuesta empírica de Gallup, Sachs y Mellinger (1999) y analiza la existencia de una relación estadística entre diferentes características geográficas puras (altitud media, pluviosidad, localización interior o costera de los territorios) o geoinstitucionales (capitalidad) y la densidad económica provincial (aproximada a través del PIB o la población por unidad de superficie). La conclusión que se deriva del ejercicio, en palabras del autor, es que «en España, las desigualdades demográficas y económicas reflejan diferencias geográficas (localización, clima). Se trataría, pues, de una desigualdad geográficamente determinada».

Otros trabajos dedicados al estudio del caso español a lo largo de su proceso de desarrollo, sin embargo, han apuntado que la explicación de los determinantes de la localización de la actividad industrial en el territorio y de su elevada concentración geográfica parece estar relacionada con la existencia de economías de escala en los procesos productivos. Economías de escala que, al actuar en conjunción con la reducción de los costes de transporte, habrían favorecido la aglomeración industrial en determinadas regiones. Éstas, por lo tanto, tal vez estuvieran dotadas de unas condiciones geográficas favorables, pero también por otras de segunda na-

³ MALLADA (1994).

turalidad. El tamaño de sus mercados, por ejemplo, podría ser un factor que explicara la aglomeración final de la producción en las mismas en una proporción superior a la de las ventajas de localización iniciales. Es decir, la geografía de la industrialización española ofrecería evidencia relevante en favor del potencial explicativo real de algunos de los modelos teóricos planteados por la literatura de la NGE.

La presencia de este tipo de elementos ha sido destacada por Rosés (2003) para la primera fase del proceso industrializador, entre inicios y mediados del Siglo XIX. Por Tirado, Paluzie y Pons (2002) para el período que transcurre desde mediados del Siglo XIX hasta la Primera Guerra Mundial. O por Betrán (1999) o Pons, Silvestre, Tirado y Paluzie (2007) para el período de entreguerras. Es decir, en todos estos trabajos se ha señalado que las economías de aglomeración son un factor explicativo de la localización de la industria en España a lo largo de distintas fases del proceso de industrialización español. Finalmente, períodos más acotados y recientes han sido objeto de estudio detallado por muchos otros trabajos (Alonso, Chamorro y González, 2003, o Viladecans, 2004). En estas circunstancias, se hace difícil comprender que los elementos de tipo aglomerativo estén presentes en la explicación de la existencia y avance de la desigualdad en la localización de la industria en el territorio, pero que no sean relevantes a la hora de explicar la desigual densidad económica general de las regiones españolas.

La evidencia existente sobre la distribución de la población en el espacio y su evolución en el tiempo arroja nuevas dudas acerca de la validez de las conclusiones derivadas de los trabajos planteados desde una perspectiva geográfica pura. Así, algunos trabajos dedicados a estudiar de forma específica la evolución de esta variable han apuntado la validez de explicaciones más complejas. En este sentido, el trabajo de Ayuda, Collantes y Pinilla (2005) muestra que las diferencias en densidad de población entre las provincias españolas han sido crecientes desde inicios del Siglo XIX.

A partir de varios análisis de tipo empírico, en dicho trabajo se defiende que aunque las variables de tipo geográfico pueden aportar una explicación de las posibilidades de crecimiento de una región en el contexto de una economía agraria, como la española del Siglo XIX, éstas han ido perdiendo capacidad explicativa a medida que la economía analizada se ha vuelto más compleja. De la misma forma, el cambio tecnológico en la esfera de la producción y la distribución de bienes ha restado importancia al clima, a la localización interior o costera de un territorio, o a la disposición de ciertos recursos naturales como determinantes de la especialización productiva regional en los sectores más dinámicos de la actividad. En estas condiciones, en el citado trabajo se sostiene que la explicación de la desigualdad territorial en la distribución de la población a lo largo de los Siglos XIX y XX debería responder a la conjunción entre elementos de geografía de *second nature* y la presencia de economías de aglomeración en determinados procesos productivos.

La validez de las afirmaciones vertidas por unos u otros autores depende, en última instancia, de la solidez de los análisis empíricos en que éstas descansan. En este sentido, por ejemplo, cabría apuntar que los trabajos de corte empírico en los que se sustenta la tesis geográfica son poco robustos. La existencia de relaciones estadísticamente significativas entre variables que captan la desigualdad en la distribución de la actividad y las características geográficas de las regiones no permite inferir la existencia de una relación de causalidad entre las mismas, y mucho menos descartar la relevancia de las hipótesis explicativas alternativas propuestas, por ejemplo, desde la NGE. Así, la obtención de coeficientes de determinación elevados en regresiones de este tipo puede deberse a la inclusión de un amplio número de variables explicativas o incluso a la presencia de relaciones meramente espurias entre las mismas. Además, cabría señalar que al tratarse de modelos de tipo heurístico es difícil afirmar que las variables presentadas en el modelo sirvan para recoger, de forma exclusiva, los elementos sostenidos por la perspectiva teórica defendida en el ejercicio. Es decir, este tipo de aproxi-

mación empírica no permite descartar la relevancia de hipótesis explicativas alternativas ya que, en ella, no se realiza un análisis conjunto de la potencia explicativa de las hipótesis alternativas que señale la importancia comparada de las mismas⁴.

Sin duda, algunos de los trabajos de corte empírico que sostienen que las diferencias relativas en densidad económica están relacionadas con la presencia de elementos de geografía de segunda naturaleza, o con la existencia de economías de aglomeración, ofrecen estudios más sólidos sobre la presencia de este tipo de elementos. La razón es que presentan una evidencia empírica más cercana a la teoría, incluso estimaciones de modelos de corte estructural que acercan mucho el análisis empírico realizado a la propuesta teórica objeto de análisis. En este sentido, Rosés (2003) o Pons, Silvestre, Tirado y Paluzie (2007) ofrecen estimaciones de corte casi estructural de modelos teóricos de la Nueva Geografía Económica y que, por tanto, permiten sostener con mayor solidez la presencia de estos elementos explicativos. En última instancia, sin embargo, tampoco estos trabajos sitúan en un mismo plano de análisis a las hipótesis defendidas por una y otra corriente de estudio.

Por lo tanto, sería especialmente interesante realizar un ejercicio que permitiera asignar a cada uno de ellos la carga explicativa correspondiente, avanzando así en la resolución de este debate. En este sentido, en el próximo apartado se plantea una metodología, ya utilizada por Roos (2005), para el estudio de la desigualdad existente entre las regiones alemanas en la actualidad, y la evidencia necesaria para la realización de un análisis que sirva para dirimir, desde una perspectiva empírica, el peso que debe imputarse a las distintas alternativas de explicación de la gran desigualdad existente en la densidad relativa de la actividad económica entre las regiones españolas en distintos momentos del tiempo.

⁴ ROOS (2005) ofrece una presentación pormenorizada de los límites a los que se enfrenta este tipo de análisis.

3. Datos y metodología

El análisis empírico de la importancia de los distintos elementos que compiten en la explicación de la desigual distribución de la actividad en el territorio es una tarea extremadamente compleja. La realidad observada, sin ningún tipo de duda, es producto de la acción conjunta, de la interacción, de elementos de geografía de primera y segunda naturaleza. Por ello, realizar una simple regresión de la densidad relativa de la actividad entre los territorios sobre las variables geográficas típicas, de *first* o de *second nature*, no resuelve el problema.

Con ese procedimiento se tiende, de forma inevitable, a magnificar la importancia de los elementos analizados. Un simple razonamiento permite entender la razón de este hecho. Supongamos, en primer lugar, la existencia de dos territorios iguales en cuanto a características geográficas y la ausencia de economías de aglomeración en los procesos productivos. En este caso, la distribución esperada de la actividad debería ser similar, homogénea, en ambos territorios. En segundo lugar, podemos pensar en un escenario igual en cuanto a características territoriales, pero en el que la actividad económica se desarrolla en un marco de economías de aglomeración. En este caso, la solución de distribución homogénea de la actividad en el territorio es posible, pero no estable. Cualquier diferencia territorial, por pequeña que sea y con independencia de su carácter (*first* o *second nature*), conduciría a esta hipotética economía a una situación de aglomeración de la producción en el territorio con una geografía más favorable. Sin embargo, sería injustificado asignar a la existencia de pequeñas diferencias entre los territorios la responsabilidad de la desigual distribución. Ésta sería fruto de la existencia de diferencias iniciales, en algunos casos muy poco relevantes, pero también de la presencia de economías de aglomeración en los procesos productivos.

Sin embargo, los análisis empíricos estándar no permiten realizar esta importante distinción y, por ello, tienden a asignar la totalidad de la varianza explicada de la densidad económica regional a los elementos geográficos.

cos diferenciales de partida, amplificando de forma injustificada su poder explicativo. Es por ello que en este trabajo se propone una alternativa a este tipo de aproximación: la realización de un análisis de los determinantes de la desigualdad económica regional en España, utilizando una metodología que identifica la importancia relativa que cabe imputar a cada uno de los elementos que compiten en la explicación.

En particular, el estudio se realiza sobre cuatro cortes temporales: 1920, 1955, 1975 y 2003. La consideración de estos cortes temporales permite cubrir un amplio período temporal correspondiente a la culminación del proceso de industrialización de la economía española en el Siglo XX, siendo la variable dependiente la densidad del PIB provincial relativizada por la superficie provincial (dr_k) expresada en términos logarítmicos:

$$dr_k = \frac{PIB_k / SUP_k}{\sum PIB_k / \sum SUP_k} = \frac{PIB_k / \sum PIB_k}{SUP_k / \sum SUP_k}$$

Para su cálculo se ha tomado el producto interior bruto de cada una de las provincias consideradas en miles de euros corrientes de cada año, por lo que la información de 1920, 1955 y 1975, originalmente en pesetas, se ha transformado utilizando el tipo de cambio oficial entre euro y peseta; y la superficie de cada una de estas unidades territoriales medida en km². En el Anexo A se ofrece detalle de los estadísticos descriptivos de esta variable en las cuatro cortes temporales.

Las variables geográficas de primera naturaleza consideradas han sido: la altitud media provincial (ALTI), la pluviosidad de cada provincia medida como la media registrada por su capital los dos años previos a cada corte temporal analizado (PREC), una variable ficticia que toma valor 1 en el caso que la provincia analizada se sitúe en la costa (COSTA), una ficticia para la capital del Estado (Madrid), y una ficticia que toma valor 1 en el caso que la provincia analizada albergue a la capital de la Comunidad Autónoma corres-

pondiente (CAPITAL). Esta última variable sólo se ha considerado en el estudio del último de los cortes temporales analizados, para el que se entiende que el desarrollo del Estado de las Autonomías en España podría hacer relevante esta posibilidad. Finalmente, se ha considerado la posición de cada una de las provincias en el territorio a través de una variable que refleja la distancia media existente entre cada capital de provincia y las 46 restantes (DIST).

Por lo que respecta a las variables de segunda naturaleza utilizadas se ha optado por, siguiendo a Roos (2005), aproximar los elementos destacados por la NGE a través de un indicador del tamaño de mercado de las regiones, medido a partir del tamaño de población provincial (POB) y, de forma adicional, mediante la consideración directa de un indicador de la existencia de economías de aglomeración: la productividad del factor trabajo en cada una de las provincias, medida a través del cociente entre el PIB provincial y la población ocupada. La variable así construida (PROD) se mide en euros corrientes de cada año por ocupado, con lo que, de nuevo, la información de 1920, 1955 y 1975 se ha transformado utilizando el tipo de cambio oficial entre euro y peseta.

Las fuentes a partir de las que se ha recogido y elaborado la información necesaria para la construcción de las variables se detallan en el Anexo C. En estas líneas sólo se quiere hacer constar que, para el estudio del corte temporal correspondiente a 1920, y atendiendo a la ausencia de datos o estimaciones previas referentes al PIB provincial español, para la realización de este trabajo se ha tomado la estimación realizada por Tirado, Pons y Paluzie (2006), que obtienen los datos referentes al PIB provincial empleando la metodología propuesta por Geary y Stark (2002) para la estimación de las series históricas de las cuatro grandes regiones de Gran Bretaña. Esta metodología exige la utilización de datos referentes a salarios y empleo por provincias y ramas de actividad que fueron recopilados a partir del Censo de la Población Española de 1920 y de los datos de salarios reales

medios diarios por provincias en agricultura e industria ofrecidos por Silvestre (2001)⁵.

En relación al método de análisis, en este trabajo se propone descomponer la varianza total de la variable empleada para reflejar la desigualdad en la distribución de la actividad en el espacio en cuatro conceptos. Una parte que permanece inexplicada, otra que es explicada de forma neta por la existencia de diferencias geográficas de *first nature* entre los territorios, otra que se vincula a la existencia de diferencias de segunda naturaleza y, finalmente, una última porción que estaría asociada a la presencia de economías de aglomeración que amplifican las diferencias territoriales iniciales. Es decir, que recoge la interacción entre elementos diferenciales en la geografía de primera y segunda naturaleza de las regiones.

Para el cálculo de cada una de estas varianzas se procede de la siguiente forma. En primer lugar, se estima el efecto total bruto de los elementos de primera y segunda naturaleza. Para ello se realiza una regresión de la variable endógena sobre unos y otros elementos. No obstante, atendiendo a que ambos pueden estar relacionados, los elementos de segunda naturaleza se incorporan en la regresión una vez se ha eliminado la parte de los mismos que puede estar correlacionada con las características geográficas puras. Con este objetivo, de forma previa, se regresan las variables de tipo geográfico de segunda naturaleza sobre las variables de primera naturaleza. Los residuos de esta regresión pueden ser considerados elementos de segunda naturaleza que no están relacionados con las características geográficas, y se utilizarán, *a posteriori*, como las variables encargadas de captar el poder explicativo neto de la

geografía de segunda naturaleza, sin interacción alguna con la geografía de primera naturaleza.

La regresión de la variable endógena sobre las variables de segunda naturaleza permite obtener la varianza explicada, de forma conjunta, por la suma del efecto neto de los elementos de segunda naturaleza y la interacción de los mismos con los de primera naturaleza. El efecto neto de la geografía de segunda naturaleza se deduce de una regresión entre la endógena y las variables de segunda naturaleza neta, calculadas a partir de la regresión descrita en el párrafo previo. La diferencia entre la varianza explicada de forma bruta y de forma neta por los elementos de segunda naturaleza nos permitirá aislar el efecto asociado a la interacción entre geografía de primera y segunda naturaleza. Finalmente, el efecto neto de la geografía de primera naturaleza se deriva por sustracción entre la varianza total, explicada por la acción conjunta de unos y otros elementos, y la explicada de forma bruta por los de segunda naturaleza.

De forma analítica, el ejercicio planteado es el siguiente. En primer lugar, se estiman las siguientes regresiones de los elementos de segunda naturaleza sobre las características geográficas territoriales:

$$\ln(pob_k) = \gamma_0 + \sum \gamma_i \cdot f_{ik} + \pi_k \quad [1]$$

$$f_{ik} = \{costa_k, \ln(alti_k), capital_k, madrid_k, \ln(prec_k), \ln(dist_k)\}$$

$$\ln(prod_k) = \delta_0 + \sum \delta_i \cdot f_{ik} + \rho_k \quad [2]$$

$$f_{ik} = \{costa_k, \ln(alti_k), capital_k, madrid_k, \ln(prec_k), \ln(dist_k)\}$$

Los residuos de estas regresiones se incluyen de forma conjunta a las variables geográficas para obtener el total de varianza explicada por la geografía según:

$$\ln(dr_k) = \alpha_0 + \sum \phi_i \cdot f_{ik} + \sum \theta_j \cdot \tilde{s}_{jk} + \varepsilon_k$$

$$f_{ik} = \{costa_k, \ln(alti_k), capital_k, madrid_k, \ln(prec_k), \ln(dist_k)\} \quad [3]$$

$$\tilde{s}_{jk} = \{\hat{\pi}_k, \hat{\rho}_k\}$$

El efecto bruto de la segunda geografía se obtiene a través de la estimación de la siguiente ecuación:

⁵ La consideración de este corte temporal descansa en la utilización de una reciente estimación del PIB provincial español. Sin duda, la calidad de esta estimación es inferior a la de las empleadas para el análisis de los años 1955, 1975 o 2003, hecho que arrojaría dudas sobre la conveniencia de su inclusión en el estudio. No obstante, su consideración permite ampliar el espectro temporal analizado y que puedan ser extraídas conclusiones acerca de la evolución del potencial explicativo de cada una de las variables propuestas a lo largo del proceso de desarrollo económico español en el largo plazo.

$$\begin{aligned} \ln(dr_k) &= \alpha_0 + \sum \theta_j \cdot s_{jk} + \varepsilon_k \\ s_{jk} &= \{pob_k, prod_k\} \end{aligned} \quad [II]$$

Y el efecto neto de la geografía de segunda naturaleza proviene de la varianza explicada mediante la siguiente ecuación:

$$\begin{aligned} \ln(dr_k) &= \alpha_0 + \sum \theta_j \cdot \tilde{s}_{jk} + \varepsilon_k \\ \tilde{s}_{jk} &= \{\hat{\pi}_k, \hat{\rho}_k\} \end{aligned} \quad [III]$$

A partir de la información derivada de la realización de estos ejercicios es posible descomponer la varianza de la desigualdad en la distribución territorial de la actividad en sus cuatro componentes o efectos:

Efecto total	=	R_i^2
Efecto primera naturaleza neto	=	$R_i^2 - R_{ii}^2$
Efecto segunda naturaleza neto	=	R_{iii}^2
Efecto primera naturaleza vía segunda naturaleza o interacción	=	$R_{ii}^2 - R_{iii}^2$
Variación no explicada [1]	=	$1 - R_i^2$

Por lo que se refiere al procedimiento de estimación de los distintos modelos analizados en todos los casos se ha partido de una especificación que incluye a todas las variables y, a partir de la misma, se ha retirado la variable que ofreciera menor significatividad, siempre que ésta fuera inferior al 10 por 100. A continuación, se ha vuelto a estimar el modelo y a retirar de nuevo la variable menos significativa hasta obtener estimaciones en las que todas las variables fueran significativas al 10 por 100. Cabe señalar que el cálculo de los números de condición y otros estadísticos permite concluir que la posible existencia de multicolinealidad no ha afectado a la significatividad de las variables. Por tanto, las variables que se han eliminado al no ser significativas no lo han sido debido a que un grado de multicolinealidad elevado hubiese podido sesgar los resultados obtenidos.

Finalmente, esta aproximación presenta un problema adicional, puesto que las variables de segunda naturaleza consideradas, POB y PROD, pueden ser endógenas y simultáneamente determinadas con el PIB. Esto puede provocar una relación de simultaneidad en la estimación de la ecuación [II]. La solución tradicional que ha sugerido la econometría para resolver este problema consiste en la utilización de un método de estimación basado en las variables instrumentales. No obstante, en nuestro caso los tests de endogeneidad de Hausman y Sargan sugieren que no existe este problema.

4. Evidencia empírica

El primer paso en el análisis es la estimación de las ecuaciones [1] y [2] con el objeto de construir las variables que captarán el efecto neto explicado por la geografía de segunda naturaleza en el análisis de descomposición de varianza (ANOVA). El Cuadro 1 resume los resultados de la realización de este ejercicio para los distintos cortes temporales.

El ejercicio confirma la existencia de una relación estadística entre las variables de segunda naturaleza y las de primera en todos los cortes temporales. Es por ello que el residuo calculado en cada una de estas regresiones puede ser tomado como el componente propio de las variables de segunda naturaleza, no vinculado a la existencia de diferencias geográficas de primera naturaleza entre las regiones. Estos residuos son los que se incluyen, junto a las variables de primera naturaleza, en el modelo I encargado de captar el total de varianza de la endógena explicada por los elemento de tipo geográfico.

Con el cálculo de esta variable se puede pasar a estimar los modelos basados en las ecuaciones I, I y III (en adelante, modelos I, II y III) para cada uno de los cortes temporales y obtener, a partir de los resultados alcanzados, el porcentaje de variación explicado por cada uno de los componentes. A continuación, en el Cuadro 2 se sintetizan los principales resultados y su lectura puede facilitar la realización de una interpretación dinámica del

CUADRO 1
RELACIÓN ENTRE LOS ELEMENTOS DE GEOGRAFÍA DE SEGUNDA NATURALEZA Y LOS DE PRIMERA

	1920		1955		1975		2003	
	ln (pob)	ln (prod)						
Constante	4,492 (0,022)	2,404 (0,000)	13,980 (0,000)	10,680 (0,000)	14,538 (0,000)	13,918 (0,000)	14,664 (0,000)	10,567 (0,000)
Costa								0,052 (0,067)
ln (alti)		-0,042 (0,002)	-0,187 (0,000)	-0,125 (0,000)	-0,293 (0,000)	-0,131 (0,000)	-0,329 (0,000)	
Capital							0,508 (0,002)	
Madrid	1,494 (0,000)	0,535 (0,000)	1,843 (0,000)	0,4711 (0,000)	2,618 (0,000)	0,271 (0,001)	2,527 (0,000)	0,190 (0,0501)
ln (prec)		0,124 (0,004)		0,235 (0,007)				
ln (dist)	1,330 (0,000)			-0,901 (0,007)		-0,769 (0,014)		
R ²	0,353	0,437	0,392	0,442	0,572	0,289	0,702	0,134
R ² corregido	0,324	0,398	0,365	0,389	0,553	0,239	0,682	0,094
F	12,005 (0,000)	11,127 (0,000)	14,196 (0,000)	8,332 (0,000)	29,438 (0,000)	5,827 (0,002)	33,813 (0,000)	
N*	47	47	47	47	47	47	47	47

NOTAS: Estimación de las ecuaciones [1] y [2].

Entre paréntesis valor-p del contraste de significación individual del parámetro estimado en el caso de los parámetros y de significación conjunta en el caso del estadístico F-Snedecor.

Las desviaciones estándar estimadas de los parámetros estimados se han obtenido mediante el método consistente de White.

Se han eliminado aquellas variables explicativas con un valor-p superior al 10 por 100.

* N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife, Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

FUENTES: Anexo C y elaboración propia.

papel desempeñado por cada uno de los componentes diferenciados de la geografía económica en la explicación de la desigualdad económica regional española a lo largo del Siglo XX.

Por otra parte, en el Anexo B se detallan los resultados de la estimación de los modelos I, II y III obtenidos para cada uno de los cortes temporales. Además, se in-

cluyen los resultados que se derivan de una estimación en la que se tratara de explicar la densidad relativa del PIB provincial a partir de variables geográficas de primera naturaleza.

En el Cuadro 2, el efecto A se corresponde con los resultados de la estimación del modelo que incluye las variables que captan la totalidad de la varianza explicada

CUADRO 2
DESCOMPOSICIÓN DE LA VARIACIÓN EXPLICADA. METODOLOGÍA ANOVA

Variación		1920	1955	1975	2003
Efecto A	Explicada	0,899	0,926	0,915	0,884
	No explicada [1]	0,101	0,074	0,085	0,116
	Total	1,000	1,000	1,000	1,000
Efecto B	Efecto secundario agregado.	0,698	0,788	0,807	0,825
	No explicada [2]	0,302	0,212	0,193	0,175
	Total	1,000	1,000	1,000	1,000
Efecto C	Efecto secundario neto	0,107	0,128	0,196	0,141
	Efecto primario vía secundario	0,591	0,660	0,611	0,684
	No explicada [2]	0,302	0,212	0,193	0,175
	Total	1,000	1,000	1,000	1,000
Efecto D	Efecto primario neto	0,201	0,138	0,108	0,059
	Efecto secundario neto	0,107	0,128	0,196	0,141
	Efecto primario vía secundario	0,591	0,660	0,611	0,684
	No explicada [1]	0,101	0,074	0,085	0,116
	Total	1,000	1,000	1,000	1,000

Efecto A

Variación explicada: coeficiente determinación ecuación I = R_I^2 Variación no explicada [1] = $1 - R_I^2$

Efecto B

Efecto secundario agregado: coeficiente determinación ecuación II = R_{II}^2 Variación no explicada [2] = $1 - R_{II}^2$

Efecto C

Efecto secundario neto: coeficiente determinación ecuación III = R_{III}^2 Efecto primario vía secundario = $R_{II}^2 - R_{III}^2$ Variación no explicada [2] = $1 - R_{II}^2$

Efecto D

Efecto primario neto = $R_I^2 - R_{II}^2$ Efecto secundario neto = coeficiente determinación ecuación III = R_{III}^2 Efecto primario vía secundario = $R_{II}^2 - R_{III}^2$ Variación no explicada [1] = $1 - R_I^2$ **FUENTES: Anexo C y elaboración propia.**

por las variables de primera naturaleza y el componente neto de las de segunda naturaleza (modelo I). Se observa que esta estimación permite explicar, en todos los cortes temporales analizados, alrededor del 90 por 100 de la varianza de la densidad relativa de la actividad económica en las regiones españolas.

En segundo lugar, en el Cuadro 2 el efecto B se relaciona con los resultados de la estimación en la que se han considerado variables explicativas de geografía de

segunda naturaleza (tamaño provincial y productividad) (modelo II). Los resultados indican que los elementos de *second nature*, de forma conjunta, explican por sí solos un porcentaje muy elevado de la varianza de la distribución relativa del PIB. Además, este porcentaje es creciente a lo largo de los cuatro lapsos temporales analizados. Así, si bien en 1920 este porcentaje era cercano al 70 por 100, en 2003 alcanzaba el 82,5 por 100 del total.

En este sentido, en el Cuadro A3 se ofrecen los resultados correspondientes a la realización de un ejercicio similar al realizado en Gallup, Sachs y Mellinger (1999) con el objeto de acreditar la importancia de los elementos de geografía de primera naturaleza en la explicación de la desigualdad económica. Se comprueba que este tipo de variables también explican un porcentaje muy elevado de la varianza de la distribución del PIB provincial español. Éste oscila entre un máximo del 80 por 100 y un mínimo del 76 por 100. Sin embargo, si comparamos estos resultados con los obtenidos en el ejercicio correspondiente al modelo II, se comprueba que, salvo para el corte temporal correspondiente a 1920, en el resto de casos el porcentaje de la varianza total explicado de forma bruta por la geografía de segunda naturaleza supera al explicado solamente por la geografía física.

En cualquier caso, cabe señalar que ni uno ni otro ejercicio permite obtener, de forma aislada, el potencial explicativo de la geografía de primera o segunda naturaleza como determinante de la desigualdad territorial de la actividad. Cabe recordar que ambos ejercicios magnifican la importancia de los factores geográficos analizados puesto que en las regresiones se incorpora una parte de varianza que está asociada a la interacción de ambos tipos de elementos. Con el objeto de aislar este efecto, el siguiente paso ha sido analizar la varianza explicada sólo por los elementos de geografía *à la* Krugman (efecto C en el Cuadro 2). Para ello, recordemos, se han utilizado como explicativas de la desigualdad las dos variables denominadas de efecto neto de segunda naturaleza, construidas a partir de los residuos de la estimación de las ecuaciones [1] y [2] (modelo III). Los resultados de este ejercicio son muy ilustrativos. Se observa que, en términos netos, los elementos de segunda naturaleza explican desde un mínimo cercano al 10 por 100 de la varianza total en 1920 a un máximo cercano al 20 por 100 en la estimación correspondiente a 1975. Es decir, la geografía de segunda naturaleza no explica, por sí sola, una proporción mayoritaria de la desigualdad

económica territorial en ninguno de los períodos analizados.

No obstante, el resultado no implica la preeminencia del efecto generado por la existencia de diferencias en la geografía de primera naturaleza. El efecto D en el Cuadro 2 computa el porcentaje de la explicación de la varianza de la endógena que debería asociarse al impacto de esta variable. Se observa que el papel desempeñado por la misma alcanza un porcentaje máximo del 20 por 100 en el momento inicial del estudio, 1920. Sin embargo, va perdiendo relevancia explicativa a medida que se avanza en el Siglo XX. Es decir, en las estimaciones correspondientes a los años 1955, 1975 y 2003 se reduce la importancia del efecto primario neto para explicar las desigualdades en la distribución de la actividad económica entre las provincias españolas.

Si se compara la acción neta de estos dos tipos de variables se comprueba que la geografía de primera naturaleza ha pasado de explicar el doble de la varianza explicada por la geografía de segunda naturaleza en 1920 a explicar, aproximadamente, un tercio de la misma en 2003. No obstante, la presentación que se ofrece en el Cuadro 2 del efecto D muestra que es la interacción entre ambos tipos de variables el principal argumento explicativo de la desigualdad relativa en la densidad de la actividad económica entre las regiones españolas. Este efecto explica una proporción creciente de la varianza, con un mínimo del 59,1 por 100 en 1920 y un máximo superior al 68,4 por 100 en 2003.

Es decir, se podría afirmar que las diferencias económicas interprovinciales se relacionan con la existencia de diferencias iniciales en elementos de corte geográfico de primera o segunda naturaleza amplificados por la existencia de economías de aglomeración en los procesos productivos. Además, la importancia neta de los aspectos relacionados con la geografía *à la* Krugman, como son el tamaño de mercado medido por el volumen de población o la productividad del trabajo en cada una de las regiones, habría sido creciente a lo largo del Siglo XX.

Los resultados ofrecen una perspectiva de los determinantes de la desigualdad económica provincial consistente con la evidencia descriptiva existente para el caso español. Ésta se resume en el crecimiento de largo plazo de las desigualdades económicas y, sobre todo, por la persistencia en las posiciones relativas ostentadas por los territorios situados en las primeras o las últimas posiciones en la clasificación de las provincias españolas en función de su densidad económica. Además, también son consistentes con la evidencia ofrecida por Ayuda, Collantes y Pinilla (2005), que han mostrado que la relación existente entre la posición presente y la pretérita de un territorio se ha ido debilitando a lo largo del proceso de desarrollo económico. El crecimiento observado en el potencial explicativo de los elementos de segunda naturaleza en detrimento de los de primera permitiría ofrecer una explicación razonable a este hecho.

5. Conclusiones

El trabajo aquí presentado ha aportado evidencia sobre la importancia relativa de los elementos de geografía de primera o segunda naturaleza como determinantes de la desigualdad en la distribución de la actividad económica en el territorio. En particular, se ha ofrecido y estudiado evidencia correspondiente a la desigualdad económica territorial en España en los años 1920, 1950, 1975 y 2003. Es decir, se ha analizado un caso nacional y desde una perspectiva de largo plazo. De esta forma, ha sido posible observar la existencia de alteraciones en la importancia relativa asignada a las distintas variables explicativas consideradas.

Del ejercicio realizado se concluye que la geografía, entendida en sentido amplio, permite explicar un porcentaje muy elevado de la variación en la densidad económica provincial en todos los cortes temporales analizados. Sin embargo, se ha mostrado cómo la explicación requiere la consideración de la presencia de economías de aglomeración que actúan ampliando las divergencias geográficas iniciales. De hecho, en ausen-

cia de este tipo de efectos aglomerativos y de los elementos de geografía de segunda naturaleza (como tamaño de mercado o productividad del trabajo) modelados por la interacción de las mismas con la mano del hombre, la desigualdad entre los territorios españoles hubiera sido en el pasado, y sería hoy, mucho menor.

Además, se ha mostrado cómo a lo largo del proceso de desarrollo económico español se ha consolidado la desvinculación del nivel de desarrollo relativo de las regiones de sus características geográficas iniciales. Es en este contexto en el que cabría comprender el avance relativo de Madrid o la pérdida de densidad relativa de ciertas provincias de la periferia Norte, como Asturias. Así pues, el estudio ha ofrecido nueva evidencia que sirve para reforzar las conclusiones derivadas de los trabajos que han defendido la importancia de la presencia de economías de aglomeración y de elementos de geografía de segunda naturaleza en la explicación de la localización del sector industrial.

Referencias bibliográficas

- [1] ALONSO, O.; CHAMORRO, J. M. y GONZÁLEZ, X. (2003): «Spillovers geográficos y sectoriales de la industria», *Revista de Economía Aplicada*, 32, 77-95.
- [2] AYUDA, M. I.; COLLANTES, F. y PINILLA, V. (2005): «From Locational Fundamentals to Increasing Returns: The Spatial Concentration of Population in Spain, 1787-2000», *Documento de Trabajo 2005-05*, Facultad de Ciencias Económicas y Empresariales, Universidad de Zaragoza.
- [3] BETRÁN, C. (1999): «Difusión y localización industrial en España durante el primer tercio del Siglo XX», *Revista de Historia Económica*, 3, 663-696.
- [4] DOBADO, R. (2004): «Un legado peculiar: la geografía», en LLOPIS, E. (ed.), *El legado económico del Antiguo Régimen en España*, Editorial Crítica, Barcelona.
- [5] DOBADO, R. (2006): «Geografía y desigualdad económica y demográfica de las provincias españolas (Siglos XIX y XX)», *Investigaciones de Historia Económica*, 5, 133-170.
- [6] FUNDACIÓN BBV (1997): *Renta Nacional de España y su distribución provincial. Serie homogénea. Años 1955 a 1993 y avances 1994 a 1997*, tomo II, Fundación BBV, Madrid.
- [7] GALLUP, J. L.; SACHS, J. y MELLINGER, J. A. (1999): «Geography and Economic Development», *International Regional Science Review*, 22, 179-232.

- [8] GEARY, F. y STARK, T. (2002): «Examining Ireland's Post Famine Economic Growth Performance», *Economic Journal*, 112, 919-935.
- [9] FUJITA, M.; KRUGMAN, P. y VENABLES, A. (1999): *The Spatial Economy. Cities, Regions, and International Trade*, MIT Press, Cambridge, MA.
- [10] KRUGMAN, P. (1993): «First Nature, Second Nature, and Metropolitan Location», *Journal of Regional Science*, 33, 129-144.
- [11] MALLADA, L. (1994): *Los males de la patria y la futura revolución española*, Alianza Editorial, Madrid.
- [12] PONS, J.; SILVESTRE, J.; TIRADO, D. A. y PALUZIE, E. (2007): «Testing the New Economic Geography: Migrations and Industrial Agglomerations in Spain», *Journal of Regional Science*, 47, 289-313.
- [13] RAPPAPORT, J. y SACHS, J. (2003): «The United States as a Coastal Nation», *Journal of Economic Growth*, 8, 5-46.
- [14] ROOS, M. (2005): «How Important is Geography for Agglomeration?», *Journal of Economic Geography*, 5, 605-620.
- [15] ROSÉS, J. R. (2003): «Why Isn't the Whole of Spain Industrialized? New Economic Geography and Early Industrialization, 1797-1910», *Journal of Economic History*, 64, 995-1022.
- [16] SACHS, J. (2000): «Tropical Underdevelopment», *CID Working Paper*, 57.
- [17] SILVESTRE, J. (2001): *Competitive and Non Competitive Determinants of Interindustry Wage Differentials During Industrialisation: The Case of Spain (1900-1936)*, mimeo, Universidad de Zaragoza.
- [18] TIRADO, D. A.; PALUZIE, E. y PONS, J. (2002): «Economic Integration and Industrial Location: The Case of Spain Before World War I», *Journal of Economic Geography*, 2, 343-363.
- [19] TIRADO, D. A.; PONS, J. y PALUZIE, E. (2006): «Los cambios en la localización de la actividad económica, 1850-1936», *Revista de Historia Industrial*, 31, 41-63.
- [20] VENABLES, A. (1999): «But why does Geography Matter, and which Geography Matters?», *International Regional Science Review*, 22, 238-241.
- [21] VILADECANS, E. (2004): «Agglomeration Economies and Industrial Location: City Level Evidence», *Journal of Economic Geography*, 4, 565-582.

ANEXO A

ANÁLISIS DESCRIPTIVO¹

La muestra utilizada en este trabajo incluye 47 provincias puesto que se ha descartado la utilización de la información correspondiente a las provincias insulares (Gran Canaria, Santa Cruz de Tenerife y Baleares) y a los territorios autónomos de Ceuta y Melilla. En particular, el Cuadro A1 ofrece información sobre la evolución de dos indicadores de dispersión de la densidad económica regional en los cuatro cortes temporales analizados en este trabajo: la desviación estándar y el coeficiente de variación. Asimismo, se ofrece información sobre el valor medio del indicador de densidad relativa del PIB provincial en cada uno de los años considerados en el análisis, así como sus valores máximos y mínimos. Por otra parte, en el Cuadro A2 se muestra la evolución del indicador de densidad relativa del PIB provincial cuando se compara con el conjunto de España. En el Cuadro A3 se presenta el *ranking* ocupado por cada provincia por lo que respecta al indicador de densidad relativa en los distintos cortes temporales. Finalmente, en el Cuadro A4 se muestran los coeficientes de correlación por rangos de Spearman de la información contemplada en el Cuadro A3.

CUADRO A1

DENSIDAD RELATIVA DEL PIB PROVINCIAL. EVIDENCIA DESCRIPTIVA

	Media	Desviación estándar	Coefficiente variación	Valor mínimo	Valor máximo	Máx./mín.	N*
1920	1,318	1,473	1,118	0,236	7,514	31,8	47
1955	1,503	2,485	1,654	0,186	11,478	61,7	47
1975	1,540	2,661	1,728	0,113	11,138	98,6	47
2003	1,443	2,364	1,639	0,109	11,716	107,5	47

NOTA: * N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife, Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

FUENTES: Anexo C y elaboración propia.

CUADRO A2

DENSIDAD RELATIVA DEL PIB PROVINCIAL. ESPAÑA = 1

	1920	1955	1975	2003
$0 < dr_k < 0,25$	1	5	10	11
$0,25 \leq dr_k < 0,5$	11	15	14	12
$0,5 \leq dr_k < 1$	17	10	7	6
$1 \leq dr_k < 2$	11	6	9	11
$2 \leq dr_k < 3$	2	2	3	3
$dr_k \geq 3$	5	4	4	4

FUENTES: Anexo C y elaboración propia.

¹ Ver apartados 1 y 2.

ANEXO A (continuación)

CUADRO A3
DENSIDAD RELATIVA DEL PIB PROVINCIAL. RANKING

	1920	1955	1975	2003
Álava	20	15	8	9
Albacete	41	42	41	39
Alicante	6	6	6	5
Almería	25	39	29	22
Asturias	8	9	9	17
Ávila	35	40	42	40
Badajoz	32	38	39	37
Barcelona	1	1	2	2
Burgos	33	33	30	31
Cáceres	44	43	43	43
Cádiz	10	12	14	12
Cantabria	14	8	11	14
Castellón	19	17	20	18
Ciudad Real	38	37	36	38
Córdoba	22	22	25	27
Coruña (La)	7	11	12	13
Cuenca	45	46	45	47
Girona	11	10	13	11
Granada	29	26	27	24
Guadalajara	46	44	44	44
Guipúzcoa	3	3	4	4
Huelva	26	27	23	26
Huesca	40	41	38	42
Jaén	24	25	28	29
León	31	29	33	33
Lleida	28	32	26	25
Lugo	16	30	32	32
Madrid	4	4	3	1
Málaga	12	16	15	8
Murcia	18	19	18	16
Navarra	21	20	19	19
Orense	17	24	24	28
Palencia	39	28	35	36
Pontevedra	5	7	7	7
Rioja (La)	23	18	21	21
Salamanca	36	31	34	34
Segovia	37	34	37	35
Sevilla	15	13	16	15
Soria	47	45	47	46
Tarragona	13	14	10	10
Teruel	43	47	46	45
Toledo	34	35	31	30
Valencia	9	5	5	6
Valladolid	30	21	17	20
Vizcaya	2	2	1	3
Zamora	42	36	40	41
Zaragoza	27	23	22	23

FUENTES: Anexo C y elaboración propia.

ANEXO A (continuación)

CUADRO A4

DENSIDAD RELATIVA DEL PIB PROVINCIAL. COEFICIENTE DE CORRELACIÓN POR RANGOS DE SPEARMAN

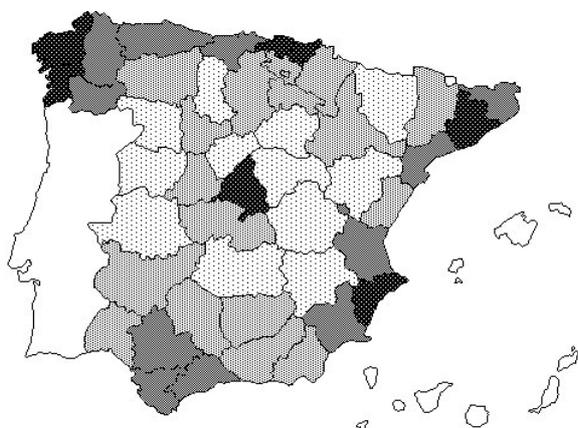
	1920	1955	1975
1920.	—	—	—
1955.	0,942	—	—
1975.	0,944	0,973	—
2003.	0,942	0,952	0,983

NOTA: En todos los casos se rechaza la hipótesis de que el coeficiente de correlación por rangos de Spearman es nulo con un nivel de significación del 1 por 100.

FUENTES: Anexo C y elaboración propia.

MAPA A1

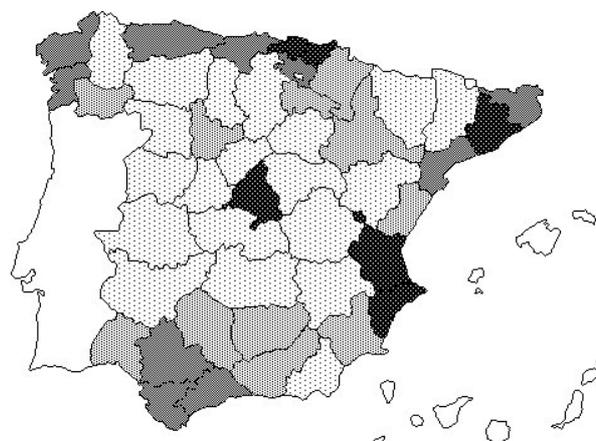
DENSIDAD RELATIVA, 1920. ESPAÑA = 1



FUENTE: Anexo C.

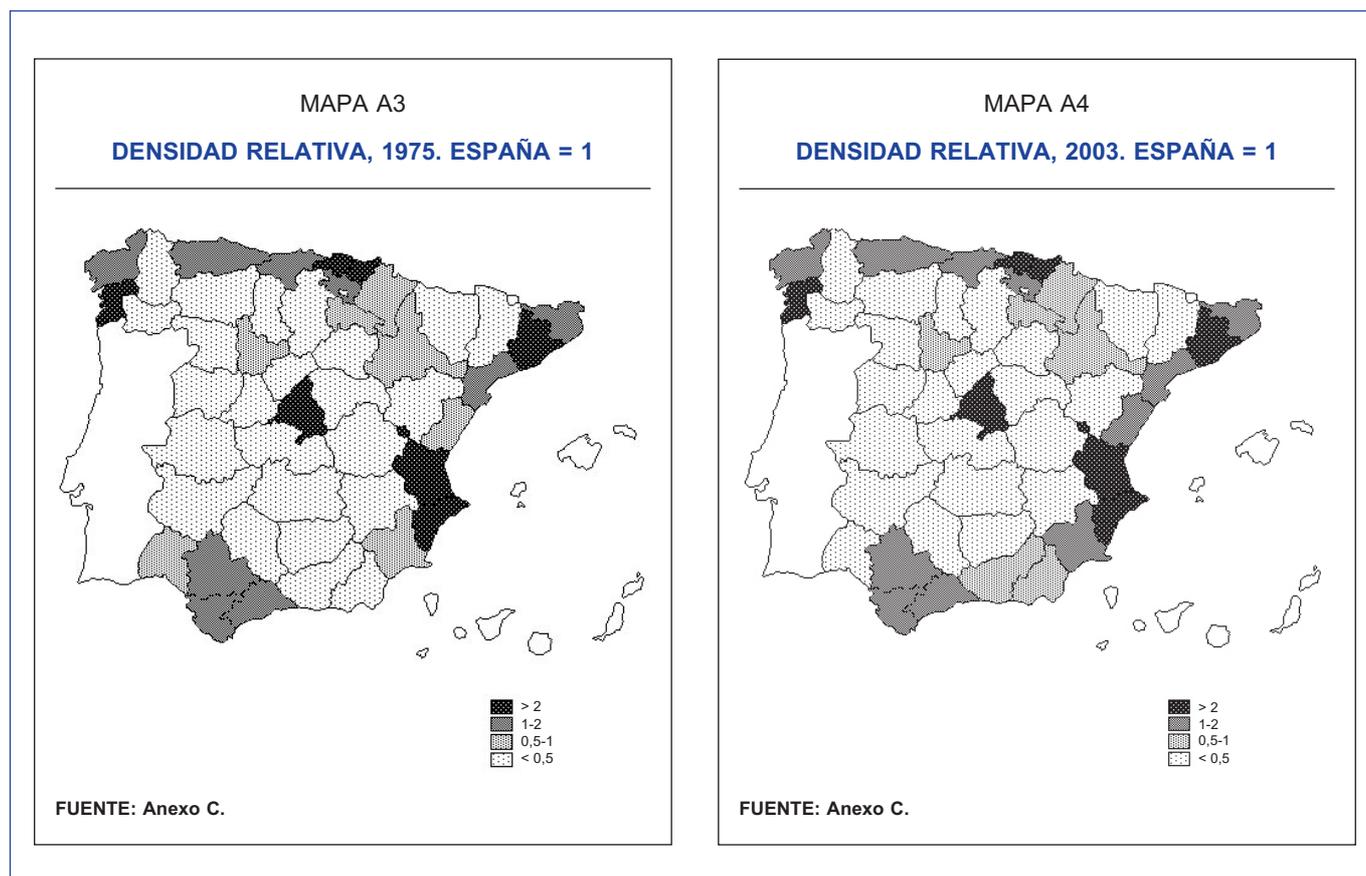
MAPA A2

DENSIDAD RELATIVA, 1955. ESPAÑA = 1



FUENTE: Anexo C.

ANEXO A (continuación)



ANEXO B

RESULTADOS DE LA ESTIMACIÓN DE LOS DIFERENTES MODELOS¹

CUADRO B1

AÑO 1920

	[I]	[II]	[III]
Constante	-2,799 (0,000)	-16,786 (0,000)	-0,101 (0,383)
Costa	0,444 (0,002)		
ln (alti)	-0,209 (0,000)		
Capital			
Madrid	2,162 (0,000)		
ln (prec)	0,560 (0,000)		
ln (dist)			
ln (pob)		0,678 (0,000)	
ln (prod)		2,668 (0,000)	
$\hat{\pi}$	0,398 (0,002)		0,491 (0,074)
$\hat{\rho}$	1,518 (0,000)		1,316 (0,158)
R ²	0,899	0,698	0,107
R ² corregido	0,884	0,685	0,067
F	59,229 (0,000)	50,914 (0,000)	2,649 (0,082)
N*	47	47	47

NOTAS: Entre paréntesis valor-p del contraste de significación individual del parámetro estimado en el caso de los parámetros y de significación conjunta en el caso del estadístico F-Snedecor.

Las desviaciones estándar de los parámetros estimados se han obtenido mediante el método consistente de White.

En la ecuación [I] sólo se han incluido las variables geográficas que han resultado significativas ($\alpha = 0,1$), con la excepción de la variable «capital» que únicamente se ha incluido en la estimación correspondiente a 2003.

* N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

FUENTE: Elaboración propia.

CUADRO B2

AÑO 1955

	[I]	[II]	[III]
Constante	4,351 (0,093)	-22,602 (0,000)	-0,255 (0,081)
Costa	0,309 (0,049)		
ln (alti)	-0,371 (0,000)		
Capital			
Madrid	2,742 (0,000)		
ln (prec)	0,902 (0,000)		
ln (dist)	-1,336 (0,002)		
ln (pob)		0,785 (0,000)	
ln (prod)		2,074 (0,000)	
$\hat{\pi}$	0,592 (0,000)		0,363 (0,199)
$\hat{\rho}$	1,312 (0,000)		1,316 (0,018)
R ²	0,926	0,788	0,128
R ² corregido	0,913	0,779	0,088
F	69,573 (0,000)	81,960 (0,000)	3,223 (0,049)
N*	47	47	47

NOTAS: Entre paréntesis valor-p del contraste de significación individual del parámetro estimado en el caso de los parámetros y de significación conjunta en el caso del estadístico F-Snedecor.

Las desviaciones estándar de los parámetros estimados se han obtenido mediante el método consistente de White.

En la ecuación [I] sólo se han incluido las variables geográficas que han resultado significativas ($\alpha = 0,1$), con la excepción de la variable «capital» que únicamente se ha incluido en la estimación correspondiente a 2003.

* N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

FUENTE: Elaboración propia.

¹ Ver apartados 2 y 4.

ANEXO B (continuación)

CUADRO B3

AÑO 1975

	[I]	[II]	[III]
Constante	6,957 (0,018)	-28,279 (0,000)	-0,397 (0,017)
Costa	0,379 (0,023)		
ln (alti)	-0,495 (0,000)		
Capital			
Madrid	3,319 (0,000)		
ln (prec)	0,552 (0,000)		
ln (dist)	-1,348 (0,003)		
ln (pob)		0,945 (0,000)	
ln (prod)		1,836 (0,000)	
π	0,616 (0,000)		0,558 (0,072)
$\hat{\rho}$	1,322 (0,000)		1,621 (0,001)
R ²	0,915	0,807	0,196
R ² corregido	0,899	0,798	0,160
F	59,622 (0,000)	91,974 (0,000)	5,376 (0,008)
N*	47	47	47

NOTAS: Entre paréntesis valor-p del contraste de significación individual del parámetro estimado en el caso de los parámetros y de significación conjunta en el caso del estadístico F-Snedecor.

Las desviaciones estándar de los parámetros estimados se han obtenido mediante el método consistente de White.

En la ecuación [I] sólo se han incluido las variables geográficas que han resultado significativas ($\alpha = 0,1$), con la excepción de la variable «capital» que únicamente se ha incluido en la estimación correspondiente a 2003.

* N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

FUENTE: Elaboración propia.

CUADRO B4

AÑO 2003

	[I]	[II]	[III]
Constante	-0,436 (0,593)	-62,926 (0,000)	-0,397 (0,018)
Costa	0,344 (0,079)		
ln (alti)	-0,399 (0,000)		
Capital	0,445 (0,007)		
Madrid	3,399 (0,000)		
ln (prec)	0,270 (0,042)		
ln (dist)			
ln (pob)		1,022 (0,000)	
ln (prod)		4,627 (0,000)	
$\hat{\pi}$	0,630 (0,000)		0,691 (0,061)
$\hat{\rho}$	3,679 (0,000)		4,267 (0,033)
R ²	0,884	0,825	0,141
R ² corregido	0,864	0,817	0,102
F	42,665 (0,000)	103,625 (0,000)	3,616 (0,035)
N*	47	47	47

NOTAS: Entre paréntesis valor-p del contraste de significación individual del parámetro estimado en el caso de los parámetros y de significación conjunta en el caso del estadístico F-Snedecor.

Las desviaciones estándar de los parámetros estimados se han obtenido mediante el método consistente de White.

En la ecuación [I] sólo se han incluido las variables geográficas que han resultado significativas ($\alpha = 0,1$).

* N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

FUENTE: Elaboración propia.

ANEXO B (continuación)

CUADRO B5
ESTIMACIÓN MODELO TIPO GALLUP,
SACHS Y MELLINGER

$$\ln(d_{ik}) = \beta_0 + \sum \beta_i \cdot f_{ik} + \varepsilon_k$$

	1920	1955	1975	2003		1920	1955	1975	2003
constante	-2,281 (0,002)	5,091 (0,159)	3,596 (0,339)	-0,771 (0,471)	ln (prec)	0,533 (0,000)	0,741 (0,000)	0,721 (0,002)	0,285 (0,098)
costa	0,268 (0,032)	0,348 (0,074)	0,441 (0,058)	0,488 (0,035)	ln (dist)		-1,298 (0,034)	-1,021 (0,098)	
ln (alti)	-0,264 (0,000)	-0,375 (0,000)	-0,445 (0,000)	-0,368 (0,000)	R ²	0,801	0,757	0,758	0,777
capital				0,534 (0,022)	R ² corregido	0,783	0,728	0,729	0,750
madrid	2,160 (0,000)	2,730 (0,000)	3,439 (0,000)	3,354 (0,000)	F	42,381 (0,000)	25,578 (0,000)	25,693 (0,000)	28,582 (0,000)
					N*	47	47	47	47

NOTAS: Entre paréntesis valor-p del contraste de significación individual del parámetro estimado en el caso de los parámetros y de significación conjunta en el caso del estadístico F-Snedecor.

Las desviaciones estándar de los parámetros estimados se han obtenido mediante el método consistente de White.

Se han eliminado aquellas variables explicativas con un valor-p superior al 10 por 100.

* N es el número de observaciones. En el análisis no se contemplan Baleares, Las Palmas de Gran Canaria, Santa Cruz de Tenerife, Ceuta y Melilla, por lo que el número de provincias utilizado es igual a 47.

ANEXO C

FUENTES ESTADÍSTICAS PARA LAS DISTINTAS VARIABLES UTILIZADAS¹

- PIB:
 - 1920: Tirado, Pons y Paluzie (2006)
 - 1955 y 1975: Fundación BBV
 - 2003: INE. Contabilidad Regional de España
- POBLACIÓN EMPLEADA:
 - 1920: Censo de la Población Española de 1920
 - 1955 y 1975: Fundación BBV
 - 2003: INE. Contabilidad Regional de España
- SUP, ALTI, POB y PREC:
 - INE. Anuario Estadístico de España
- DIST:
 - 1920: Estadísticas de Obras Públicas. Ministerio de Fomento
 - 1955, 1975 y 2003: Ministerio de Fomento. Atlas de Carreteras de España

¹ Ver apartados 2 y 3.