

## Estimación y análisis de la productividad aparente del trabajo en las ciudades españolas

Fernando Rubiera-Morollón \*, Esteban Fernández-Vázquez \*  
y Elizabeth Aponte-Jaramillo \*\*

**RESUMEN:** Existe una amplia literatura que estudia la productividad de las ciudades y las causas de las diferencias existentes. Tales diferencias pueden deberse a factores económicos, geográficos o características de la propia ciudad tales como su tamaño, estructura urbana o gestión entre otros. En algunos países las estadísticas oficiales permiten disponer de información sobre la producción desagregada a nivel local que posibilitan el hacer estos análisis. Sin embargo en otros muchos casos, como ocurre con el español, esta información no está disponible. El objetivo de este trabajo es desarrollar un procedimiento de inferencia ecológica basado en una estimación mediante entropía cruzada que nos permita obtener datos de productividad por municipios agregados según tamaño poblacional. Al analizar los datos estimados encontramos evidencias para España en la línea de las obtenidas en otros países aunque con algunas particularidades que se estudian en profundidad.

**Clasificación JEL:** C15, C21, R11 y R12.

**Palabras clave:** productividad, ciudades, economía urbana, inferencia ecológica, entropía y España.

### Estimation and analysis of labor productivity in Spanish cities

**ABSTRACT:** The relationship between city size and territorial productivity has attracted much attention in the urban economic literature. Some theories on the field claim for a strong positive correlation between the size of the municipalities and their income, mainly motivated by economical reasons, geographical characteristics or other factor of the urban environment. Unfortunately, in many countries

\* REGIOLab, Universidad de Oviedo (España).

\*\* Universidad Autónoma de Occidente (Colombia).

Dirección postal de contacto: Departamento de Economía Aplicada, Facultad de Economía y Empresa. Avda. del Cristo S/N, 33006 Oviedo (Asturias). Dirección electrónica de contacto: [frubiera@uniovi.es](mailto:frubiera@uniovi.es).

Este trabajo ha sido financiado por el Plan Nacional de I+D+i del Ministerio de Ciencia e Innovación (proyecto MICINN-08-ECO2008-01617). Los autores desean expresar su agradecimiento a los evaluadores de Investigaciones Regionales por sus valiosos comentarios que han mejorado la versión final de la investigación.

*Recibido: 4 de mayo de 2011 / Aceptado: 19 de octubre de 2011.*

the empirical research on this topic is not possible given the lack of data of income at a local level. This paper proposes the use of entropy econometrics to estimate urban income and urban productivity according to city size from aggregate information, which can be defined as an exercise of ecological inference. With the estimated data a regional classification based on the relevance of the cities size allows us to measure the relevance of agglomeration economics on the cities productivity in Spain.

**JEL Classification:** C15, C21, R11 y R12.

**Keywords:** productivity, cities, urban economics, ecological inference, entropy econometrics and Spain.

## 1. Introducción

Existe una amplia literatura que conecta el nivel de urbanización de los territorios con sus niveles de renta per cápita. Véase, a modo de ejemplo, los trabajos de Polèse (2005), Fay y Opal (2000), Jones y Koné (1996), Lemelin y Polèse (1995) o Tolley y Thomas (1987). Otros, como las aportaciones de Ciccola (1999) o Prud'homme (1997) prestan su atención al hecho de que sea en las grandes áreas urbanas donde se genera la mayor parte del producto nacional. Del mismo modo hay un amplio desarrollo de la literatura que conecta la productividad con el fenómeno de la urbanización y que enlaza con la amplia y relevante literatura que conforma la *Nueva Geografía Económica* [véase Fujita, Krugman y Venables (1999) o Fujita y Thisse (2002), como síntesis]. Así, algunas de las referencias más habitualmente citadas son los trabajos de Glaeser (1998), Quigley (1998), Ciccone y Hall (1996), Krugman (1991) o Henderson (2003 y 1988). En todos estos estudios se evidencia la importancia de las *economías de aglomeración* en la explicación de por qué las grandes ciudades tienen una productividad media mayor a las de menor tamaño.

Las *economías de aglomeración* generan una serie de efectos que pueden ser descompuestos en *economías de localización*, las derivadas de la concentración en un espacio reducido de un amplio número de empresas similares, y *economías de urbanización*, producidas ante las ventajas de accesibilidad a personal cualificado y tecnologías diversas en una gran concentración de población. Por estas dos vías una vez que un territorio acumula una cantidad determinada de población o una concentración sectorial relevante desata efectos externos positivos que generan ganancias de productividad y que atraen nuevas empresas provocando procesos acumulativos en la línea de lo descrito en los trabajos fundamentales de Lucas (1988, 1990 y 2001) o Romer (1986 y 1994).

Sin embargo, la abundancia de trabajos teóricos y estudios empíricos para algunos países contrasta con la ausencia de aplicaciones similares para otros. La economía española es un buen ejemplo de ello dado que las referencias de estudios sobre la productividad de las ciudades son mínimas.

Cuadrado *et al.* (1997) presentan uno de los primeros análisis sobre la evolución y comportamiento de la productividad en España con desagregación espacial por re-

giones. En el mismo se apunta ya a la existencia de fuertes divergencias regionales en productividad claramente relacionadas con los niveles de desarrollo y urbanización de los territorios que se mantiene en las conclusiones de investigaciones posteriores como las de Sánchez de la Vega *et al.* (2001) o Martínez (2001), entre otras. Hay análisis que combinan esta desagregación regional con una desagregación sectorial como Segarra (1997) que alcanza similares conclusiones para las actividades industriales. Ampliado el análisis a las regiones de la Unión Europea destaca el estudio de Maza y Villaverde (2006). Asimismo, Serrano (2001) presenta resultados sobre las diferencias sectoriales y regionales en productividad en España. Existen diversos intentos de estudio de la productividad por regiones en el sector concreto de los servicios, pero las conclusiones agregadas son más confusas y su interpretación regional más compleja, véase, a modo de revisión, Martínez y Rubiera (2000).

Otras investigaciones más recientes han puesto el acento en las relaciones entre la productividad y sus fuentes a nivel regional. Varios autores, como Lago y Caramés (1998), Gil *et al.* (2001), Alonso y Freire (2002), Pedraja *et al.* (2002) han estudiado la relación entre la productividad y el *stock* de capital público por regiones. Fernández *et al.* (2003) y Rodríguez-Vález (2006) analizan con modelos de estimación por máxima entropía la productividad de las infraestructuras y la rentabilidad de la misma en términos de su incidencia sobre la producción regional. La influencia de las economías externas ha sido también tratada por otros autores. Destaca el trabajo de Serrano (2000) analizando los efectos de las economías externas tecnológicas sobre la productividad aparente del trabajo para las regiones españolas.

Todos estos trabajos, así como otros muchos no citados, nunca pueden descender a un nivel de desagregación superior al provincial. De este modo, en la literatura aplicada a la economía española apenas hay referencias al efecto de las ciudades y las economías externas de aglomeración sobre la productividad. Existen algunos intentos de aproximación indirecta como Rubiera (2006) donde se estudia el efecto de las grandes aglomeraciones urbanas sobre el crecimiento usando datos de empleo y población.

La principal causa por la que la evidencia empírica sobre la relación entre productividad y aglomeración urbana sea tan escasa o nula en algunos casos como el español reside en la importante limitación de datos normalmente existente. Salvo algunas importantes excepciones en torno a las que suelen estar todas las aplicaciones, como EEUU o Francia, la mayor parte de los países, aunque con amplio desarrollo estadístico, carece de información oficial de variables como la Renta o el PIB con una desagregación espacial que nos permita identificar con claridad la productividad de las ciudades y calcular cómo ésta se ve afectada por su tamaño poblacional u otras variables.

El objetivo de esta investigación es salvar este problema de información en el caso de la economía española proponiendo una metodología de inferencia ecológica basada en la entropía cruzada. Esta técnica nos permitiría tener datos de renta y productividad por municipios agregados según tamaño para nuestro país, siendo una manera muy adecuada de obtener información con la que poder contrastar el vínculo entre tamaño poblacional y productividad en el caso concreto de la economía española.

Con este objetivo principal el artículo se estructura del siguiente modo. En un primer apartado, el siguiente, se presenta con detalle los fundamentos de la técnica mediante la cual se estiman datos desagregados a partir de información agregada. En el siguiente apartado se aplica esta técnica al caso de la economía española, para lo que se describen las bases de datos que se utilizan y que condicionan el análisis y se presentan los resultados obtenidos. Estos resultados son brevemente evaluados mediante una simulación Monte Carlo. Con la información obtenida de este modo podemos estudiar cómo el tamaño y la posición de los municipios afecta a su productividad mediante un procedimiento que esencialmente consiste en la agregación y representación de los datos de productividad obtenidos. Los resultados, aunque acordes en líneas generales con la literatura internacional, presentan interesantes singularidades en España que son comentadas y discutidas en profundidad en un penúltimo apartado. El último apartado se dedica a un resumen y extensiones futuras del trabajo.

## **2. Estimación de datos desagregados a partir de información agregada: la inferencia ecológica y los procedimientos basados en la entropía**

En muchos campos de investigación se requiere de información particular, individual o desagregada y, dadas las limitaciones de disponer de bases de datos con altos niveles de desagregación, mediante enfoques propios del análisis estadístico, cada vez es más frecuente recurrir a estimaciones. La Inferencia Ecológica (IE, en adelante) es el proceso a través del cual se extraen características individuales desde la información que se encuentra contenida en un conjunto de datos agregados. Aunque las técnicas básicas de IE surgen desde 1919 con la investigación de Ogburn y Goltra (1919), aplicadas a un problema de estimación del voto femenino, ya que éste era el momento en que se ejercía por primera vez ese derecho en los Estados Unidos, existe cierto consenso en asumir que los enfoques de formalización metodológica y el desarrollo consecuente de modelos surgen en los años cincuenta del siglo XX con los trabajos de Duncan y Davis (1953) y, especialmente, Goodman (1953). Las técnicas de estimación han evolucionado desde entonces y desde mediados de la década de los noventa ha adquirido especial auge, principalmente por las contribuciones realizadas por King (1997). La IE es un caso especial dentro de la inferencia estadística básicamente porque en los procesos de elaboración de datos agregados se pierde información valiosa referida a variables con niveles de mayor desagregación (véase King, Rosen y Tanner, 2004).

En la última década la IE se ha sofisticado mediante el uso de modelos de entropía<sup>1</sup>. Judge, Miller y Cho (2003) proponen una aplicación específica que nos servirá de punto de partida para el planteamiento metodológico que se adopta en la presente investigación y, por consiguiente, procederemos a detallar las bases de la técnica de estimación desarrollada en el mismo.

<sup>1</sup> Véase Golan, Judge y Miller (1996) para un estudio detallado de este tipo de técnicas de estimación.

Así pues, se puede partir de una matriz de flujos con el objetivo de obtener los valores desconocidos de las celdas (*matrix balancing problem*), utilizando la información disponible agregada por filas y columnas (véase Golan, 2006:105). La estructura básica de la matriz se presenta en el cuadro 1.

**Cuadro 1.** Datos conocidos y desconocidos en una matriz de flujos

	$z_{\cdot 1}$	...	$z_{\cdot j}$	...	$z_{\cdot T}$
$z_{1\cdot}$	$z_{11}$	...	$z_{1j}$	...	$z_{1T}$
...	...		...		...
$z_{i\cdot}$	$z_{i1}$	...	$z_{ij}$	...	$z_{iT}$
...	...		...		...
$z_{K\cdot}$	$z_{K1}$	...	$z_{Kj}$	...	$z_{KT}$

Los  $z_{ij}$  elementos de la matriz son las cantidades desconocidas a estimar, donde  $\sum_{j=1}^T z_{ij} = z_{i\cdot}$ ,  $\sum_{i=1}^K z_{ij} = z_{\cdot j}$  y  $\sum_{i=1}^K \sum_{j=1}^T z_{ij} = z$ . Estos elementos pueden ser expresados como un conjunto de distribuciones de probabilidad (columnas) al dividir las cantidades de la matriz por el total de la columna correspondiente  $z_{\cdot j}$ . Por consiguiente, la matriz previa puede ser expresada en una nueva matriz  $P$  conformada por un conjunto de  $T$  distribuciones de probabilidad, cuadro 2.

**Cuadro 2.** Matriz de flujos en términos de probabilidades

	$y_1$	...	$y_j$	...	$y_T$
$x_1$	$p_{11}$	...	$p_{1j}$	...	$p_{1T}$
...	...		...		...
$x_i$	$p_{i1}$	...	$p_{ij}$	...	$p_{iT}$
...	...		...		...
$x_K$	$p_{K1}$	...	$p_{Kj}$	...	$p_{KT}$

Donde las  $p_{ij}$  se definen como la proporción  $\frac{z_{ij}}{z_{\cdot j}}$  y las nuevas columnas y filas contienen las respectivas proporciones  $y_j = \frac{z_{\cdot j}}{z}$  y  $x_i = \frac{z_{i\cdot}}{z}$ , respectivamente. En consecuencia, los elementos  $p_{ij}$  satisfacen las siguientes igualdades<sup>2</sup>:

<sup>2</sup> En este caso, las  $p_{ij}$  pueden interpretarse como probabilidades condicionales de cada columna.

$$\sum_{j=1}^T p_{ij} y_j = x_i; \forall i = 1, \dots, K \quad (1)$$

$$\sum_{j=1}^K p_{ij} = 1; \forall j = 1, \dots, T \quad (2)$$

Estas ecuaciones sintetizan las características generales de los elementos de la matriz  $P$ . La primera ecuación muestra la relación cruzada entre cada elemento  $p_{ij}$  (desconocido) de la matriz y los totales por filas y columnas (conocidos). La segunda ecuación, por su parte, indica que las probabilidades  $p_{ij}$  se comportan como distribuciones de probabilidad por columnas. Nótese que en esta estructura sólo se dispone de  $K + T$  observaciones como información para estimar los  $K \times T$  elementos de la matriz  $P$ . Esto implica que el problema es indeterminado, situación usualmente denominada *pure linear inverseproblem*. No obstante, es posible obtener una solución a esta clase de problema mediante la minimización de una medida de divergencia respecto a una matriz  $Q$  de probabilidades *a priori*, sujeta a un conjunto de restricciones. El denominado problema de Entropía Cruzada (EC, en adelante) se expresa de la siguiente forma:

$$\text{Min}_P D(P \parallel Q) \sum_{i=1}^K \sum_{j=1}^T p_{ij} \ln \left( \frac{p_{ij}}{q_{ij}} \right) \quad (3)$$

Sujeto a las restricciones recogidas en las ecuaciones (1) y (2).

La medida de divergencia  $D(P \parallel Q)$  es la denominada divergencia de Kullback-Liebler<sup>3</sup>, que se establece entre las distribuciones posterior y previa. Es evidente que si la información previa  $Q$  es consistente con las observaciones, entonces,  $P = Q$ , y, por tanto, la medida de divergencia tendría un valor 0, lo que significaría que la base de datos no contiene información adicional a esta información de partida.

### 3. Aplicación de la entropía cruzada a la estimación de datos locales de productividad aparente del trabajo

#### 3.1. Adaptación de los modelos de entropía cruzada y entropía cruzada generalizada a la estimación de datos locales de PIB

Para explicar cómo el enfoque EC puede aplicarse para la obtención de datos locales de PIB, supóngase un área geográfica, un país por ejemplo, que puede dividirse en  $T$  unidades espaciales más pequeñas, regiones. Además de esta primera distribución geográfica, supongamos que es posible realizar otra división de acuer-

<sup>3</sup> En teoría de la probabilidad y teoría de la información, esta medida de divergencia, también conocida como de entropía relativa, establece la pseudo-distancia entre dos distribuciones de probabilidad, desde una supuesta distribución «verdadera»  $P$  a una distribución de probabilidad arbitraria  $Q$ .

do a alguna característica adicional. El segundo criterio de clasificación pueden ser las clases de unidades locales, municipios, que conforman el país atendiendo a un determinado criterio de clasificación obteniendo así  $K$  diferentes tipos. En este contexto el objetivo es estimar cómo una variable se distribuye entre las regiones de acuerdo con la clasificación de localidades, a partir de información agregada (cuadro 3), lo que representa una estructura igual al esquema presentado en el cuadro 2.

**Cuadro 3.** División espacial por regiones y tipo de municipio

		Regiones				
		$y_1$	...	$y_j$	...	$y_T$
Tipo de municipio	$x_1$	$p_{11}$	...	$p_{1j}$	...	$p_{1T}$
	...	...		...		...
	$x_i$	$p_{i1}$	...	$p_{ij}$	...	$p_{iT}$
	...	...		...		...
	$x_K$	$p_{K1}$	...	$p_{Kj}$	...	$p_{KT}$

Cada uno de los  $p_{ij}$  es ahora definido como la proporción (desconocida) de la variable estudiada dentro de los municipios de tipo  $i$  que están localizados dentro de la región  $j$ , formando una matriz  $P$  de tamaño  $(K \times T)$  con  $T$  distribuciones de probabilidad desconocidas. El vector fila  $y'$  de tamaño  $(1 \times T)$  representa las proporciones regionales de la variable, mientras que el vector columna  $x$  de tamaño  $(K \times 1)$  muestra la distribución nacional de la variable de acuerdo con el tipo de municipio. Obsérvese que estos dos vectores contienen la información agregada directamente observable por el investigador. Si un conjunto inicial de distribución de probabilidades  $Q$  también resulta observable, puede aplicarse un modelo de EC y el problema se resuelve mediante un proceso de minimización como:

$$\text{Min}_P D(P \| Q) \tag{4}$$

Sujeto a:

$$x = Py' \tag{5}$$

$$e'_K P = e'_K \tag{6}$$

Donde  $e_K$  representa un vector columna de unos con la dimensión apropiada. Nótese que, a diferencia de algunas de las técnicas de estimación empleadas habitualmente en la resolución de problemas de Inferencia Ecológica (véase Freedman, 2001), no es necesario imponer una forma funcional para estimar un modelo de regresión, ni asumir homogeneidad geográfica en los parámetros.

### 3.2. Aplicación al caso concreto español: bases de datos existentes y procedimiento de estimación del PIB local para 2001

Para poder disponer de datos agregados con una estructura similar a la presentada en el apartado anterior podemos recurrir a tres bases de datos oficiales del sistema de estadísticas de España: la Encuesta Continua de Presupuestos Familiares (ECPF, en adelante), la Contabilidad Regional de España (CRE, en adelante) y el Censo de Población y Viviendas de 2001.

El cuadro 4 presenta cómo se organizan los datos territorialmente en la ECPF. Como se ve, aunque dicha clasificación no corresponde exactamente con una ordenación de acuerdo con el tamaño dado que la categoría «capital de provincia» no refleja el volumen de población, ésta puede asumirse como un indicador del tamaño municipal porque sólo en unas pocas provincias existen ciudades con poblaciones mayores que los municipios que son la capital de provincia (Cádiz, Asturias, Toledo y Pontevedra son las excepciones).

**Cuadro 4.** Clasificación de los municipios de España en la ECPF

<i>Tipo de municipio</i>	<i>Descripción</i>
m1	Capital de provincia (independiente del tamaño de población)
m2	Municipios con más de 100.000 habitantes
m3	Municipios con población entre 50.000 y 100.000 habitantes
m4	Municipios con población entre 20.000 y 50.000 habitantes
m5	Municipios con población entre 10.000 y 20.000 habitantes
m6	Municipios con menos de 10.000 habitantes

A partir de la ECPF y la CRE puede disponerse de los datos correspondientes a los vectores  $x$  e  $y$  del cuadro 3. Ahora el vector  $x$  de dimensión  $(6 \times 1)$  contiene la distribución del Ingreso por tipos de municipio, según la ECPF, y el vector  $y$  de dimensión  $(1 \times 50)$  está formado por la participación de las provincias en el PIB, de acuerdo con la CRE. A partir de estos datos agregados se aplica una estimación de entropía para estimar la distribución del PIB provincial dependiendo del tipo de municipio. Además, se especifica una matriz inicial de distribución  $Q$ , como se requiere en estimaciones por EC, definida como la distribución de ocupados por tipo de municipio y obtenida a partir de la información del Censo de Población y Viviendas del año 2001.

De acuerdo con estos criterios, y para el desarrollo del ejercicio, se plantea el siguiente problema de EC:

$$\underset{P}{\text{Min}} D(P \parallel Q) \sum_{i=1}^6 \sum_{j=1}^{50} p_{ij} \ln \left( \frac{p_{ij}}{q_{ij}} \right) \quad (7)$$

Sujeto a:

$$x_i = \sum_{j=1}^T p_{ij} y_j; \forall i = 1, \dots, 6. \tag{8}$$

$$\sum_{i=1}^K p_{ij} = 1; \forall j = 1, \dots, 50. \tag{9}$$

En la solución de EC para el caso propuesto, el valor del PIB fue dividido por la población para calcular el PIB per cápita en miles de euros.

### 3.3. Resultados y evaluación

Los resultados obtenidos para el año 2001 se presentan en el cuadro 5. Las celdas en blanco significan que no existe esa tipología de municipio en esa provincia.

**Cuadro 5.** Estimación del PIB de España según tipos de municipio (2001). Miles de euros

Provincia	m1	m2	m3	m4	m5	m6
Almería	3.875.375		1.354.475	308.413	877.944	1.918.577
Cádiz	2.416.496	3.221.939	4.176.952	1.805.128	1.506.657	922.151
Córdoba	4.524.466			1.262.518	840.459	1.643.018
Granada	3.954.834		481.639	609.574	1.293.000	2.750.427
Huelva	2.453.599			126.514	1.495.565	1.920.191
Jaén	1.695.053		913.720	1.264.557	1.178.770	1.852.607
Málaga	8.745.076	817.609	1.115.958	2.982.577	877.457	1.036.882
Sevilla	9.564.492	1.574.079	837.711	4.064.018	2.949.890	2.659.736
Huesca	1.090.383				1.048.324	1.457.128
Teruel	651.815				202.610	1.488.927
Zaragoza	11.464.413				1.071.150	2.537.793
Asturias	2.763.521	4.384.843	1.228.314	2.823.665	2.164.160	1.636.263
Baleares	9.285.276			3.896.914	1.587.015	2.536.939
Las Palmas	6.877.198		1.846.916	3.784.210	1.529.314	1.013.536
Santa Cruz de Tenerife	3.184.773	2.066.385	585.896	2.711.234	1.500.556	2.647.496
Cantabria	3.377.022		911.171	769.643	1.136.117	2.262.559

**Cuadro 5.** (Continuación)

Provincia	m1	m2	m3	m4	m5	m6
Ávila	733.260					1.384.919
Burgos	3.532.909			1.256.996		1.554.684
León	2.166.603		1.104.857	321.103	758.280	2.430.348
Palencia	1.411.126					1.234.619
Salamanca	2.476.414				593.601	1.086.875
Segovia	1.086.965					1.305.935
Soria	619.506					836.318
Valladolid	5.891.808			387.038	312.983	1.794.594
Zamora	870.824				235.687	1.249.970
Albacete	1.713.433			773.318	118.716	1.938.293
Ciudad Real	1.109.636		850.227	837.476	1.072.887	2.473.067
Cuenca	636.893				178.511	1.715.156
Guadalajara	1.105.938			290.222		1.170.532
Toledo	833.052		1.642.155		444.698	4.162.049
Barcelona	38.251.040	15.301.673	7.778.199	8.955.466	9.026.187	17.103.913
Girona	1.907.619			2.503.482	1.883.655	5.454.748
Lleida	3.116.894				1.010.132	3.331.939
Tarragona	2.951.294		1.298.485	2.228.794	1.876.503	4.607.994
Alicante	5.985.140	3.210.052	3.974.599	3.717.699	2.180.110	2.793.190
Castellón de la Plana	3.258.034			3.088.270	661.177	2.178.918
Valencia	13.879.147		2.428.537	6.937.374	5.185.719	6.982.023
Badajoz	1.542.110		673.415	928.136	692.678	3.147.763
Cáceres	1.139.316			504.479	334.896	2.317.439
Coruña	3.565.160		2.325.280	2.082.955	2.669.922	3.891.553
Lugo	1.391.990				899.493	2.209.362
Ourense	1.581.986				433.618	2.058.187
Pontevedra	965.183	4.463.904		1.672.344	2.765.989	1.994.392
Madrid	76.838.909	18.392.299	5.154.271	4.878.483	2.668.061	12.980.273
Murcia	5.240.671	2.084.010	791.665	3.919.660	2.480.531	2.076.322

**Cuadro 5.** (Continuación)

Provincia	m1	m2	m3	m4	m5	m6
Navarra	4.901.123			1.104.184	1.032.285	4.605.374
Álava	5.053.206				633.902	897.495
Guipúzcoa	4.492.116		1.594.975	1.960.694	3.964.798	2.385.947
Vizcaya	7.707.989		4.378.931	4.418.365	1.865.180	2.974.257
La Rioja	2.839.859			454.293	275.706	1.506.414

Fuente: Estimaciones EC a partir de ECPF, CRE y Censo de Población y Viviendas (2001).

La evaluación de la capacidad de este modo de inferir datos desagregados a partir de la información agregada ha sido validada mediante un ejercicio de Monte Carlo. El punto de inicio del mismo es el vector y de proporciones observadas del PIB para las provincias españolas en 2001, el cual es mantenido fijo a través del proceso de simulación. En cada repetición del proceso de simulación se genera una matriz aleatoria  $P$ , conformada por los elementos  $p$  que se extraen desde una distribución uniforme como  $p_{ij} \sim U[0,0.2]$ ;  $i = 1, \dots, 5$ ; y  $p_{6j} = 1 - \sum_{i=1}^5 p_{ij}$  con el fin de asegurar que éstos se comportan como un conjunto de distribuciones de probabilidad. A partir de la relación lineal  $x = Py$ , el vector  $x$  se obtiene en cada repetición, y junto con las observaciones del vector  $y$ , representa la información agregada para obtener las estimaciones de la matriz desconocida  $P$ . Otra parte importante dentro del proceso de estimación es la matriz  $Q$ . Para reflejar la idea de que la especificación de la matriz  $a$  priori es relativamente similar a la matriz  $P$ , en este experimento las celdas de  $Q$  se generan desde  $P$ , un término de error aleatorio  $u$ , es establecido.

Las siguientes ecuaciones expresan de manera general el proceso <sup>4</sup> donde  $u \sim N(1, \sigma)$  y  $\sigma$  es un escalar. Obsérvese que si  $\sigma = 0$ , se cumple que  $p = q$  para todas las celdas de ambas matrices, obteniéndose entonces el menor valor posible (0) para la divergencia entre las matrices  $P$  y  $Q$ . Esta consecuencia es muy lógica, dado que una buena especificación de la matriz  $Q$  (cerca a la matriz  $P$  real) es una ayuda útil en el proceso de estimación. Al contrario, si la matriz  $Q$  especificada difiere significativamente de la matriz  $P$  actual, los datos observados (los vectores  $x$  e  $y$ ) tendrán dificultades para dar una solución cercana a los valores reales.

$$\left. \begin{aligned} q_{ij} &= (p_{ij}) \cdot (u_{ij}); \forall i = 1, \dots, 5; \forall j = 1, \dots, 50. \\ q_{6j} &= 1 - \sum_{i=1}^5 p_{ij}; \forall j = 1, \dots, 50. \end{aligned} \right\} \quad (10)$$

Dentro del experimento realizado, se han simulado seis escenarios diferentes para varios valores del escalar  $\sigma$ : 0,1; 0,2; 0,25; 0,35; 0,4 y 0,5. En cada uno de

<sup>4</sup> Basado en el experimento efectuado en Golan *et al.* (1996:63 y 64). Para evitar valores negativos no convenientes en  $q$  cuando el número generado resulta negativo éste se reemplaza por  $q = 10$ .

estos seis escenarios se realizaron 1.000 repeticiones y se calculó el promedio de las medidas de error obtenidas, a saber: la raíz cuadrada del error cuadrático medio  $\left[ RECM = \sqrt{\frac{1}{50 \times 6} \sum_{i=1}^6 \sum_{j=1}^{50} (\tilde{p}_{ij} - p_{ij})^2} \right]$  y el error absoluto medio  $\left[ EAM = \frac{1}{50 \times 6} \sum_{i=1}^6 \sum_{j=1}^{50} |\tilde{p}_{ij} - p_{ij}| \right]$ , donde los  $\tilde{p}_{ij}$  denotan las estimaciones.

En el Cuadro 6 se presentan los resultados de estas medidas de error.

**Cuadro 6.** Medidas de error en el proceso de simulación de Monte-Carlo

Valor de sigma	$\sigma = 0,5$	$\sigma = 0,4$	$\sigma = 0,35$	$\sigma = 0,25$	$\sigma = 0,2$	$\sigma = 0,1$
Estimación EC						
RECM	0,005	0,003	0,003	0,001	0,001	0,000
EAM	0,049	0,040	0,035	0,025	0,020	0,010

Estos resultados dan una idea general sobre la cuantía del error que puede tener la estimación del ejercicio a desarrollar. Si se compara para el año 2001 la distribución del PIB por provincia (según la CRE) con la distribución provincial del empleo (según el Censo de Población y Viviendas) por medio de un cociente, el cual es similar al error  $u$  considerado en el experimento de Monte Carlo, se obtiene un vector de tamaño  $(50 \times 1)$  que se comporta aproximadamente como una distribución normal con desviación típica de 0,19. Este resultado sugiere, entonces, que la estimación realizada de los elementos  $p$  para el caso de España en 2001 puede presentar errores en torno al 2% (para el caso del EAM) lo que representa un indicio que reafirma su fiabilidad.

### 3.4. Del PIB a la productividad aparente del trabajo

A partir del PIB total estimado para 2001 para poder obtener el indicador de la productividad aparente del trabajo (PIB por empleado) necesitamos disponer de información sobre la población ocupada para lo que se han utilizado los datos del Censo de Población y Viviendas de dicho año. Los resultados de dicha operación se presentan en el cuadro 7 construido a partir de los resultados del cuadro 5 y los datos del Censo.

Una vez obtenida esta información podemos proceder a su análisis mediante una reagregación de las unidades espaciales en regiones funcionales con significado económico desde la perspectiva de las *economías de aglomeración*. Éste es el objetivo de la penúltima sección del trabajo.

**Cuadro 7.** Estimación de la productividad aparente del trabajo de España según tipos de municipio (2001). Euros

<i>Provincia</i>	<i>m1</i>	<i>m2</i>	<i>m3</i>	<i>m4</i>	<i>m5</i>	<i>m6</i>
Almería	46.360		45.295	45.258	45.597	45.841
Cádiz	42.732	41.929	41.399	41.261	41.589	41.651
Córdoba	34.465			34.028	33.959	33.626
Granada	34.902		34.578	34.401	34.337	33.987
Huelva	40.332			39.777	39.875	39.840
Jaén	31.314		31.265	31.098	30.947	30.538
Málaga	36.980	36.634	36.145	35.876	35.897	35.469
Sevilla	37.837	37.296	36.624	36.280	36.353	35.847
Huesca	40.980				40.698	40.694
Teruel	40.429				40.194	49.310
Zaragoza	42.863				41.626	41.764
Asturias	42.078	41.283	40.731	40.572	40.882	40.894
Baleares	50.723			48.001	48.968	49.868
Las Palmas	44.123		42.519	42.401	42.844	43.046
Santa Cruz de Tenerife	40.746	40.171	39.723	39.570	39.773	39.698
Cantabria	42.017		41.231	41.146	41.337	41.366
Ávila	37.569					37.376
Burgos	45.215			44.419		44.800
León	42.067		41.432	41.364	41.520	41.547
Palencia	43.145					42.955
Salamanca	39.581				39.227	39.181
Segovia	34.067					33.821
Soria	37.767					37.636
Valladolid	40.956			40.144	40.304	40.291
Zamora	41.136				40.950	40.950
Albacete	37.872			37.526	37.554	37.462
Ciudad Real	38.635		38.217	38.125	38.179	38.073
Cuenca	36.562				36.393	36.323
Guadalajara	43.829			43.527		43.650

**Cuadro 7.** (Continuación)

<i>Provincia</i>	<i>m1</i>	<i>m2</i>	<i>m3</i>	<i>m4</i>	<i>m5</i>	<i>m6</i>
Toledo	30.562		36.218		36.091	35.882
Barcelona	52.764	43.901	40.022	39.729	43.516	46.641
Girona	44.301			42.937	43.294	43.465
Lleida	47.844				47.133	47.422
Tarragona	50.207		48.203	48.201	48.903	49.538
Alicante	42.527	41.354	40.549	40.321	40.777	40.802
Castellón de la Plana	43.354			42.332	42.585	42.683
Valencia	38.859		36.805	36.250	36.393	35.603
Badajoz	33.865		33.676	33.530	33.449	33.134
Cáceres	30.928			30.804	30.704	30.438
Coruña	37.521		36.738	36.498	36.529	36.159
Lugo	30.881				30.646	30.367
Ourense	29.074				28.881	28.585
Pontevedra	35.088	34.986		34.427	34.348	33.899
Madrid	56.735	42.801	37.924	38.138	44.279	51.105
Murcia	38.400	37.915	37.385	37.133	37.233	36.902
Navarra	47.912			46.265	46.805	47.244
Álava	50.752				50.073	50.449
Guipúzcoa	47.683		45.753	45.703	46.330	46.816
Vizcaya	49.990		46.843	46.814	47.882	48.811
La Rioja	44.325			43.716	43.885	43.979

Fuente: Estimaciones EC a partir de ECPF, CRE y Censo de Población y Viviendas (2001).

#### **4. Un primer análisis de los resultados. Relación entre productividad aparente del trabajo, *tamaño* poblacional y *posición* de las ciudades españolas**

Una vez que disponemos de datos de productividad a escala local a modo de primer análisis de los resultados obtenidos, proponemos ahondar en la comprensión de cómo el *tamaño* y la *posición* de cada ciudad inciden sobre su productividad.

Para ello proponemos reagregar los datos en función del *tamaño* poblacional lo que permitiría evaluar de modo aproximado la importancia de las *economías de aglomeración*. Sin embargo, no sólo importa el *tamaño* de la propia unidad espacial sino también la distancia a otras grandes metrópolis. En este sentido la *distancia* al *tamaño*, es decir, la posición respecto a las principales metrópolis del país, resulta igualmente crucial.

La propuesta que vamos a seguir en esta investigación es, esencialmente, la elaborada por Coffey y Polèse (1988) para Estados Unidos, aplicado por Polèse y Champagne (1999) para México, Polèse y Shearmur (2004) para Canadá y, Rubiera (2006), Polèse, Rubiera y Shearmur (2007) y Viñuela, Rubiera y Cueto (2010) para España. Una explicación detallada de esta clasificación y su conexión con las teorías más importantes en Economía Urbana y Regional está contenida en el texto Polèse y Rubiera (2009). La clave de la organización del espacio que se propone con esta clasificación consiste en tener en cuenta los dos elementos que consideramos clave: el *tamaño* poblacional y la *posición* o *distancia* respecto a las principales concentraciones de población (principales metrópolis).

En primer lugar, se clasificarán los territorios por *tamaño* poblacional. Una primera división puede consistir en separar las áreas urbanas (AU), entendiendo como tales aquellas en las que existe un núcleo poblacional que supere un cierto tamaño mínimo determinado, de las áreas rurales (AR). Posteriormente, se puede tener una mayor precisión dentro de las AU, áreas urbanas, distinguiendo varios niveles atendiendo a su tamaño poblacional: AU1 (metrópolis de máximo tamaño), AU2 (metrópolis menores pero de gran tamaño), AU3 (ciudades grandes), AU4 (ciudades medianas), etc. Se pueden aplicar cuantos niveles se consideren oportunos, dada la realidad empírica que se estudie y las limitaciones estadísticas que existan. Supóngase que se acepta trabajar con seis grados de tamaño urbano, dos (1 y 2) para las grandes metrópolis, y cuatro (3, 4, 5 y 6) para ciudades y núcleos urbanos de menor tamaño. Lógicamente, el nivel 6 es el tamaño mínimo considerado para calificar como urbano a un territorio, de modo que ante tamaños menores se considera el lugar como área rural (AR).

En segundo lugar, se puede definir cada una de estas áreas como central (AC) o periférica (AP) en relación a su *distancia* con una gran metrópoli (una AU1 o AU2). Como señalan Wood y Parr (2005), Phelps (2004) o Parr (2002), entre otros, según nos alejamos de una gran concentración urbana los efectos positivos de las economías de aglomeración, así como el beneficio de atraer actividades expulsadas de la gran ciudad, por su mayor sensibilidad a las *deseconomías de aglomeración*, van reduciéndose. Obviamente es difícil calcular el valor exacto de esta relación inversa, así como delimitar una frontera concreta hasta dónde el efecto de una gran metrópoli llega a notarse. A pesar de ello, Desmet y Fafchamps (2005) encuentran que esa frontera para la economía norteamericana puede fijarse en aproximadamente 50 kilómetros. Coffey y Polèse (1988), Polèse y Champagne (1999) y Polèse y Shearmur (2004), tras analizar distintas realidades nacionales (Estados Unidos, Canadá y México), llegan a la conclusión de que la distancia máxima que se puede tomar como límite para considerar a un territorio como central, por disfrutar de los efectos positivos de estar próximo a una gran metrópoli, es aproximadamente una hora de transporte por ca-

rrertera o ferrocarril. Tomando estas experiencias podemos considerar como centrales (AC) a todos los territorios localizados a no más de una hora de transporte por tierra de una AU1 o AU2. El resto serán consideradas áreas periféricas (AP).

Lo expuesto se puede resumir en cuadro sintético como el cuadro 8. En la figura 1 se ilustra cómo se aplicarían los criterios de clasificación territorial en un caso imaginario. En el mapa 1 se presenta la aplicación real al caso concreto de la economía española para cuya concreción se han utilizado las bases de datos de información geográfica del Instituto Geográfico Nacional que permiten ubicar cada municipio con precisión y calcular así su carácter central o periférico.

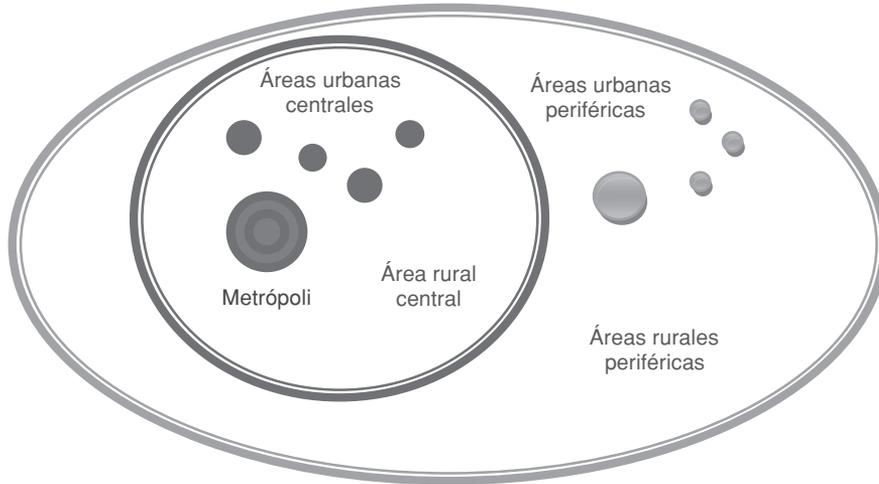
**Cuadro 8.** Síntesis de la clasificación de los territorios conforme a su *tamaño* poblacional y *distancia* respecto a una gran metrópoli

	<b>AU</b> <i>Áreas Urbanas</i> (poseen un núcleo urbano principal del área por encima de una población mínima, tamaño AU6)		<b>AR</b> <i>Áreas Rurales</i> (no poseen núcleo urbano o éste no alcanza una población mínima, tamaño AU6)
<b>AC</b> <i>Áreas Centrales</i> (a una distancia menor a una hora de transporte respecto a la gran metrópoli más cercana)	<b>ACU</b> <i>Áreas Centrales Urbanas</i>	ACU1 ACU2	<b>ACR</b> <i>Áreas Centrales Rurales</i>
		ACU3 ACU4 ACU5 ACU6	
<b>AP</b> <i>Áreas Periféricas</i> (a una distancia mayor a una hora de transporte respecto a la gran metrópoli más cercana)	<b>APU</b> <i>Áreas Periféricas Urbanas</i>	APU3 APU4 APU5 APU6	<b>APR</b> <i>Áreas Periféricas Rurales</i>

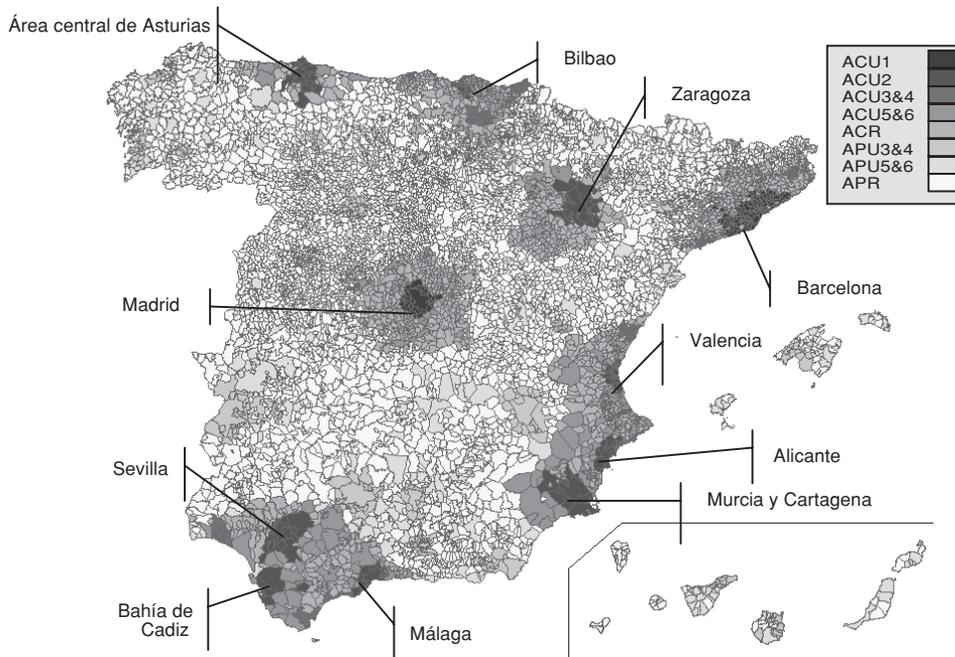
*Nota:* El número que acompaña a las *áreas urbanas* (AU), ya sean *centrales* (ACU) o *periféricas* (APU), indica el *tamaño* de la misma en una graduación que va de 1 a 6 en este ejemplo, y que puede adaptarse a la realidad empírica que se desee estudiar. Las áreas 1 y 2 son grandes urbes. Las áreas 4, 5 y 6 son ciudades de diferentes *tamaños* de mayor a menor. Las áreas con núcleos poblacionales de menor *tamaño* del tomado para el nivel 6 serán *áreas rurales* (AR). Obviamente, por su propia definición, las zonas 1 y 2 sólo son posibles en el caso central.

Una vez realizada esta reagregación en el gráfico 1 presentamos el comportamiento de la productividad aparente del trabajo agregado según tamaños poblacionales conforme a los seis niveles propuestos en el cuadro 8 y para el caso de cinco provincias que presenten todos los tamaños de municipios (ver cuadro 7 para resultados). Incluyen las dos principales áreas metropolitanas del país, Madrid y Barcelona, pero casos concretos con diferentes tamaños o estructuras urbanas: Asturias, con una conurbación en el área central compuesta por tres ciudades de tamaño medio que suman en un radio de 30 km una población cercana al millón de habitantes, Málaga, con una gran ciudad principal de más de medio millón de habitantes

**Figura 1.** Representación esquemática de la clasificación de los territorios conforme a los criterios de *tamaño* poblacional y *distancia* respecto a una gran metrópoli

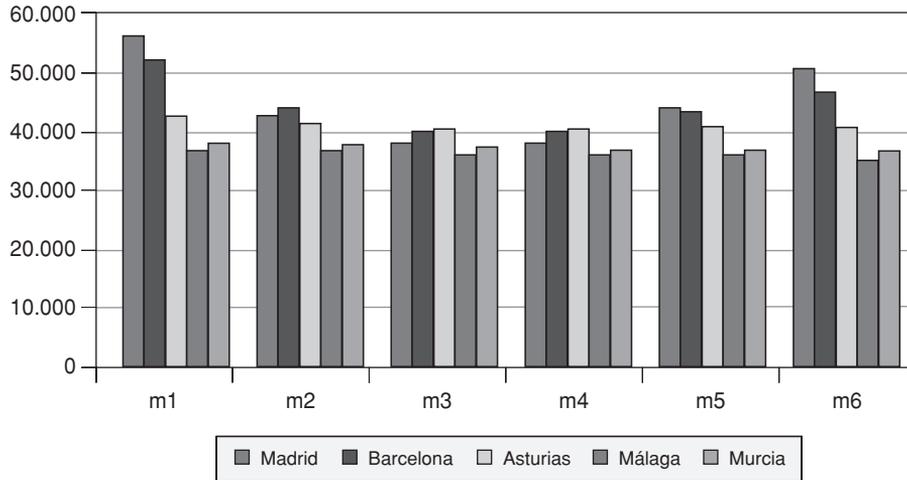


**Mapa 1.** Representación de la distribución de los territorios en España conforme a la clasificación propuesta



Fuente: Tomado de Viñuela, Rubiera y Cueto (2010).

**Gráfico 1.** Productividad aparente del trabajo estimada de las ciudades según tamaño de municipios de las provincias de Madrid, Barcelona, Asturias, Málaga y Murcia (2001). Euros



Fuente: Elaboración propia a partir de las estimaciones ECG.

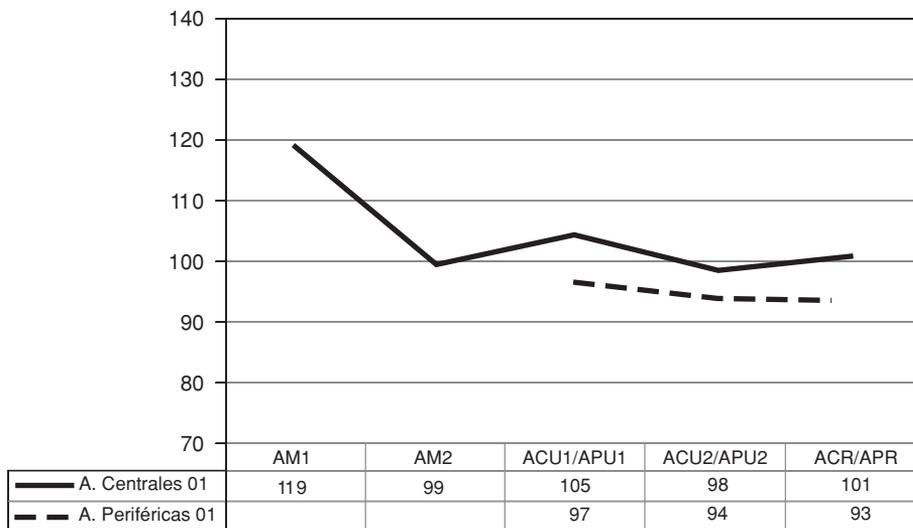
y Murcia, que junto con Cartagena, constituye otra conurbación de más de medio millón de habitantes.

Como era de esperar, las dos grandes metrópolis del país sobresalen por tener los niveles más altos de productividad que la media nacional, sin embargo, estos niveles son claramente superiores en las principales metrópolis mientras que en el resto, salvo en las zonas rurales, la productividad se asemeja más a la de las otras zonas urbanas recogidas en el gráfico.

En el gráfico 2 se sofisticaba la representación gráfica de la productividad aparente del trabajo diferenciando entre las áreas centrales (línea continua) y las periféricas (discontinua). En ambas, según se avanza por el eje de abscisas desciende el tamaño poblacional de modo que el último dato corresponde a las áreas rurales según la propuesta contenida en el cuadro 8.

La clara pendiente negativa de las líneas que representan las áreas centrales y periféricas refleja que el tamaño tiene una importante incidencia en el comportamiento de la productividad aparente del trabajo de las ciudades españolas confirmando lo apuntado en el gráfico 1, pero con mayor claridad gracias a la distinción que aporta la división entre central y periférico. Sólo se escapa de este comportamiento general la presencia de un ligero ascenso de la productividad estimada en las ciudades que conforman las áreas AU1. Este hecho puede atribuirse a que en España buena parte de las industrias están ubicadas en áreas relativamente cercanas y conectadas a las principales áreas metropolitanas, toda vez que se trata de actividades económicas con alto valor añadido pero que son sensibles a los costes de materia primas y/o

**Gráfico 2.** Niveles de productividad aparente del trabajo estimada según tamaño y distancia (2001). Índice España = 100



Fuente: Elaboración propia a partir de las estimaciones ECG.

de trabajo (estructura de *desconcentración contenida*). Una segunda salvedad observable es que las áreas AM2 se sitúan por debajo de las ciudades tipo AU1 en el comportamiento de la productividad aparente del trabajo. Este resultado obedece seguramente a que algunas de las ciudades que se localizan en esta clase de región han asumido procesos de reconversión industrial o tienen un bajo nivel de desarrollo general.

En términos generales, el comportamiento de la productividad aparente del trabajo para el resto de *tamaños* una vez diferenciados por *distancias a tamaño*, concuerda con lo que es previsible según la literatura en el campo de la Economía Regional y Urbana. De este modo, a menor *tamaño* el nivel de la productividad aparente del trabajo es menor y lo mismo ocurre cuando una ciudad se aleja de las grandes metrópolis. El gráfico muestra estas situaciones y permite concluir que la *distancia* respecto a las grandes metrópolis es un factor importante para entender la productividad agregada de los territorios.

## 5. Resumen y extensiones futuras

La conexión entre ciudades y productividad está ampliamente explorada en la literatura internacional. Según la mayor parte de los trabajos hay una relación positiva entre el *tamaño* de las ciudades, en términos de población, y su productividad. Esta relación positiva se puede explicar por la creciente importancia de las *economías de aglomeración* y, dentro de ellas, especialmente al tipo de *urbanización*.

La abundancia de estudios empíricos para algunas economías, como EEUU, contrasta con la ausencia en casos, como el español. Detrás del escaso número de análisis de cómo las ciudades afectan a la productividad o la Renta, está un problema de información al no disponerse en nuestro país, como ocurre en otros muchos casos, de estadísticas oficiales con un grado suficiente de desagregación espacial como para obtener datos por ciudades de productividad.

En este trabajo se trata de salvar este problema mediante el desarrollo de un procedimiento de inferencia ecológica basado en la técnica de entropía cruzada. Utilizando los datos de la Contabilidad Nacional de España por provincias y la Encuesta Continua de Presupuestos Familiares por municipios agregados según tamaños, se estima de este modo la Renta de los municipios españoles por tamaños y provincias. Sobre estos datos de la Renta es posible una extensión del trabajo para obtener datos de productividad aparente del trabajo con el mismo nivel de desagregación.

Esta información así estructurada resulta valiosa para realizar un trabajo sobre la relación entre el *tamaño* y *posición* relativa de las ciudades y su productividad estimada. Reagregando la información en una clasificación que tenga en cuenta el tamaño poblacional de cada municipio y su posición respecto a las grandes metrópolis del país, podemos llegar a una representación en la que se conecta la productividad aparente del trabajo con el tamaño de los municipios y su posición. Observamos una relación clara y directa entre tamaño y productividad. Sin embargo en el caso de España es muy evidente el gran salto de productividad que experimentan las dos principales ciudades: Madrid y Barcelona. Estas ciudades presentan niveles muy superiores al resto. Otra conclusión que podemos identificar con claridad es que las ciudades de tamaño medio situadas cerca, a menos de una hora de distancia por carretera, de estas dos grandes metrópolis presentan niveles de productividad superiores a los de las ciudades de igual tamaño alejadas más de una hora de distancia por carretera de los principales núcleos urbanos del país.

La mejora de esta metodología y la consolidación de la validez de los datos obtenidos se puede lograr mediante el desarrollo de posibles extensiones futuras de esta investigación previstas por los autores. Por una parte es posible aplicar una metodología similar a la utilizada en este trabajo para estimar datos a escala local de salarios. El análisis conjunto de salarios, producto y productividad robustecerá las conclusiones esbozadas en este primer artículo y permitirá extraer valiosas consideraciones económicas. Otra vía para la revisión y mejora de la metodología que los autores pretenden explorar es la aplicación de la misma a otros países. El caso de EEUU es especialmente interesante ya que podemos reproducir un punto de partida similar al que tenemos en España para estimar datos existentes en las estadísticas oficiales del país lo que permitirá evaluar la capacidad de la técnica. México ofrece un escenario incluso más interesante, ya que para algunos años concretos se dispone de información local mientras que para otros ésta no ha sido elaborada, de modo que la aplicación de una metodología similar a la desarrollada en éste permitiría reconstruir los datos ausentes disponiendo de un controlador de la fiabilidad de las estimaciones en los años en los que hubiera información desagregada.

## Referencias

- Alonso, J., y Freire, M. J. (2002): «Infraestructuras sociales: sus efectos sobre el crecimiento de la productividad de las CCAA españolas», *Revista de Estudios Regionales*, 64, 167-186.
- Ciccola, P. (1999): «Globalización y dualización en la región metropolitana de Buenos Aires. Grandes inversiones y reestructuración socio-territorial en los años noventa», *Revista Latinoamericana de Estudios Urbano-Regionales*, XXV (76), 5-28.
- Cicccone, A., y Hall, R. (1996): «Productivity and the density of economic activity», *American Economic Review*, 86, 54-70.
- Coffey, W., y Polèse, M. (1988): «Locational shifts in Canada employment, 1971-1981, decentralisation versus decongestion», *Canadian Geographer*, 32, 248-255.
- Cuadrado, J. R.; Garrido, R., y Mancha, T. (1997): «Tendencias de la productividad regional española, 1964-1993», *Revista de Economía*, 762, 87-110.
- Desmet, K., y Fafchamps, M. (2005): «Changes in the spatial concentration of employment across US Counties: A sectoral analysis 1972-2000», *Journal of Economic Geography*, 5, 261-284.
- Duncan, O. D., y Davis, B. (1953): «An alternative to ecological correlation», *American Sociological Review*, 18, pp. 665-666.
- Fay, M., y Opal, C. (2000): *Urbanization without growth: a non common phenomenon*, Working Paper 2.412, The World Bank, Washington.
- Fernández, J.; Orea, L., y Álvarez, A. (2003): «La productividad de las infraestructuras en España», *Papeles de Economía Española*, 95, 47-65.
- Freedman, D. (2001): «Ecological Inference and the Ecological Fallacy», *International Encyclopedia for the Social and Behavioral Sciences*, 6, 4027-4030.
- Fujita, M.; Krugman, P., y Venables, J. A. (1999): *The spatial economy- Cities regions and international trade*, The MIT Press.
- Fujita, M., y Thisse, J. F. (2002): *Economics of agglomeration*, Cambridge University Press.
- Gil, C.; Rapún, M., y Pascual, P. (2001): «Productividad, capital público y convergencia de las regiones españolas», *Hacienda Pública Española*, 156, 135-154.
- Glaeser, E. L. (1998): «Are cities dying?», *Journal of Economic Perspectives*, 12, 139-160.
- Golan, A.; Judge, G., y Miller, D. (1996): *Maximum entropy econometrics: Robust estimation with limited data*, John Wiley & Sons Ltd.
- Golan, A. (2006): «Information and entropy econometrics. A review and synthesis», *Foundations and Trends in Econometrics*, 2.
- Goodman, L. (1953): «Ecological regressions and the behavior of individuals», *American Sociological Review*, 18, pp. 663-664.
- Henderson, J. V. (2003): «Marshall's scale economies», *Journal of Urban Economics*, 53, 1-28.
- Henderson, V. (1988): *Urban development: theory, fact and illusion*, Oxford University Press, New York (USA).
- Instituto Nacional de Estadística (1997): *Encuesta Continua de Presupuestos Familiares Base 1997. Ficheros longitudinales de usuarios, características anuales de los hogares, diseño de registros*, [www.ine.es](http://www.ine.es).
- (2001): *Censo de población y viviendas de 2001*, [www.ine.es](http://www.ine.es).
- (2006): *Principales características de la Encuesta de Presupuestos Familiares 2006*, [www.ine.es](http://www.ine.es).
- (2007): *Contabilidad regional de España, base 2000 — Serie homogénea 1995-2006. Nota metodológica*, [www.ine.es](http://www.ine.es).
- Jones, B., y Koné, S. (1996): «An exploration of the relationships between urbanization and per capita income: United States and countries of the word», *Papers in Regional Science*, 75 (2), 135-153.

- Judge, G.; Miller, D., y Cho, W. (2003): *An information theoretic approach to ecological estimation and inference*, Documento de Trabajo, 946, University of California, Berkeley.
- King, G. (1997): *A solution to the Ecological Inference Problem: Reconstructing individual behavior from aggregate data*, Princeton, Princeton University Press.
- King, G.; Rosen, O., y Tanner, M. A. (2004): *Ecological Inference: New Methodological Strategies*, Cambridge University Press, UK.
- Krugman, P. (1991): «Increasing returns and economic geography», *Journal of Political Economy*, 99 (3), 483-499.
- Lago, S., y Caramés, L. A. (1998): «Capital público y productividad de las regiones españolas», *Cuadernos de Información Económica*, 136-137, 79-86.
- Lemelin, A., y Polèse, M. (1995): «What about the Bell-shaped relationship between primacy and development?», *International Regional Science Review*, 18, 313-330.
- Lucas, R. (1988): «On the mechanics of economic development», *Journal of Monetary Economics*, 22, 3-42.
- (1990): «Why doesn't capital flow from rich to poor countries», *American Economic Review*, 80 (2), 92-96.
- (2001): «Externalities and cities», *Review of Economics Dynamics*, 4, 245-274.
- Martínez, E. (2001): «Convergencia, productividad y empleo en las regiones españolas: 1985-1999», *Revista Asturiana de Economía*, 20, 27-53.
- Martínez, S. R., y Rubiera, F. (2000): «Algunas reflexiones acerca de la productividad de los servicios en España», *Estudios de Economía Aplicada*, 16, 157-170.
- Maza, A., y Villaverde, J. (2006): «La productividad industrial en las regiones de la Unión Europea: 1980-2003», *Papeles de Economía Española*, 107, 66-79.
- Ogburn, W. F., y Goltra, I. (1919): «How women vote: A study of an election in Portland, Oregon», *Political Science Quarterly*, 3 (34), 413-433.
- Parr, J. (2002): «Agglomeration economies: Ambiguities and confusions», *Environment and Planning A*, 34, 717-731.
- Pedraja, F.; Salinas, J., y Salinas, M. (2002): «Efectos del capital público y el capital humano sobre la productividad de las regiones españolas», *Papeles de Economía Española*, 93, 135-147.
- Phelps, N. (2004): «Clusters, dispersion and the spaces in between: For an Economic Geography of the Banal», *Urban Studies*, 46 (5/6), 971-989.
- Polèse, M. (2005): «Cities and national economic growth: A reappraisal», *Urban Studies*, 42 (8), 1429-1451.
- Polèse, M., y Champagne, E. (1999): «Location matters: Comparing the distribution of economic activity in the Mexican and Canadian urban systems», *International Journal Science Review*, 22, 102-132.
- Polèse, M., y Shearmur, R. (2004): «Is distance really dead?: Comparing industrial location patterns over time in Canada», *International Regional Science Review*, 27 (4), 1-27.
- Polèse, M.; Rubiera, F., y Shearmur, R. (2007): «Observing regularities in location patterns. An analysis of the spatial distribution of economic activity in Spain», *European Urban and Regional Studies*, 14(2), 157-180.
- Polèse, M., y Rubiera, F. (2009): *Economía urbana y regional. Introducción a la geografía económica*, Thompson-Civitas, Madrid.
- Prud'homme, R. (1997): «Urban transportation and economic development», *Régionnet Développement*, 5, 40-53.
- Quigley, J. M. (1998): «Urban diversity and economic growth», *Journal of Economics Perspectives*, 12 (2), 127-138.
- Rodríguez-Vález, J. (2006): «Productividad y rentabilidad de las infraestructuras regionales a partir de estimaciones por máxima entropía», *Investigaciones Regionales*, 8, 123-139.
- Romer, P. (1986): «Increasing returns and long-run growth», *Journal of Political Economy*, 94, 1002-1037.

- (1994): «Origins of endogenous growth», *Journal of Economic Perspective*, 8 (1), 3-22.
- Rubiera, F. (2006): *Ciudades, crecimiento y especialización territorial: Dinámicas espaciales de concentración del empleo y la población en España*, Consejo Económico y Social del Principado de Asturias.
- Sánchez de la Vega, J. C.; Buendía, J., y Llago, M. (2001): «La productividad del trabajo en las regiones españolas: un análisis desagregado», *Análisis Regional*, 181-191.
- Segarra, G. (1997): «Las disparidades regionales de la productividad industrial: 1978-1992», *Economía Industrial*, 317, 21-34.
- Serrano, G. (2000): «Economías externas y productividad del trabajo», *Revista de Economía Aplicada*, 8 (24), 105-138.
- (2001): «Productividad total y sectorial en las regiones españolas», *Análisis Regional*, 165-179.
- Tolley, G. S., y Thomas, V. (1987): *The economics of urbanization and urban policies in development nations*, The World Bank, Washington.
- Viñuela, A.; Rubiera, F., y Cueto, B. (2010): «An analysis of Urban Size and Territorial Location Effects on Employment Probabilities: the Spanish Case», *Growth and Change*, 41 (4), 495-519.
- Wood, G. A., y Parr, J. (2005): «Transaction costs, agglomeration economies and industrial location», *Growth and Change*, 36 (1), 1-15.

