

**XV REUNIÓN ASEPELT-ESPAÑA**  
**(A Coruña, 21 y 22 de Junio de 2001)**

Título: *“Crecimiento económico y dinámica de distribución de la renta en las regiones de la UE: un análisis no paramétrico”*

Autores: Julián Ramajo Hernández y María del Mar Salinas Jiménez

Universidad de Extremadura, Facultad de CC.EE. y EE., Badajoz

Correo electrónico: [ramajo@unex.es](mailto:ramajo@unex.es), [msalinas@unex.es](mailto:msalinas@unex.es) FAX: 924-289-547

Resumen: El objetivo del presente trabajo consiste en analizar el crecimiento económico en la regiones de la UE desde el punto de vista de la dinámica en la distribución de la renta (PIB *per capita*) disponible. A diferencia de la regularidad empírica encontrada en recientes trabajos empíricos sobre el tema a nivel de países, que sugiere una tendencia temporal hacia la bimodalidad (*twin-peaks*) en la distribución transversal, en el caso aquí estudiado a nivel de regiones NUTS-2 de la Unión Europea durante el período 1981-1996, se observa el fenómeno contrario, pasándose de una situación inicial bimodal a una situación final unimodal.

# CRECIMIENTO ECONÓMICO Y DINÁMICA DE DISTRIBUCIÓN DE LA RENTA EN LAS REGIONES DE LA UE: UN ANÁLISIS NO PARAMÉTRICO

## 1. Introducción

En los últimos años ha habido un auge importante de estudios teóricos y empíricos dedicados a analizar las causas del crecimiento económico, los procesos de convergencia entre áreas económicas (países o regiones) y la dinámica en la distribución de la renta entre dichas áreas.

La mayor parte de ese conjunto de investigaciones se basa en regresiones de crecimiento promedio que utilizan datos de corte transversal (o de panel, si se dispone de varias observaciones temporales para cada una de las unidades analizadas), las cuales por definición intentan explicar el comportamiento de la media condicional, pero aportan poca información sobre la distribución transversal de la riqueza y sobre la evolución de la misma, y no dicen casi nada acerca de los procesos de polarización y estratificación que pueden experimentar las diferentes economías.

Este trabajo intenta aportar algo más de información sobre estos últimos fenómenos, utilizando para ello un enfoque eminentemente no paramétrico, pero usando a la vez las regresiones paramétricas de convergencia “clásicas” como fuente para analizar las causas de la evolución observada en la distribución transversal de la renta *per capita*. Como base empírica se utilizarán datos de ámbito europeo; en concreto, se utilizarán datos de una muestra de 164 regiones europeas, correspondientes al nivel NUTS-2 de clasificación estadística territorial establecido por EUROSTAT. El período considerado abarca desde 1981 hasta 1996.

## 2. La metodología: el enfoque no paramétrico

Tal como se ha dicho en la introducción, la mayor parte de los análisis empíricos sobre convergencia se basan en regresiones de crecimiento transversales del tipo  $y_i = m(X_i) + \varepsilon_i$ , donde  $y$  es alguna medida del crecimiento experimentado por cada unidad económica durante cierto período de tiempo,  $X$  una matriz de variables explicativas de dicho crecimiento,  $\varepsilon$  un término de perturbación aleatorio (de media nula y varianza desconocida,  $\sigma^2(X)$ ) representativo de shocks específicos, y  $m$  una función prefijada que

depende de un conjunto de parámetros desconocidos que hay que estimar,  $m(X) = m(X, \beta)$ . Dado el carácter de estas regresiones, es obvio que la función  $m$  se entiende como el valor esperado (promedio) de crecimiento para unos determinados valores de las variables explicativas,  $E[y | X = X_i]$  y, en ausencia de información a priori que indique lo contrario, suele asumirse que la función  $m$  es lineal, es decir,  $m(X, \beta) = X\beta$ .

En este trabajo se utilizarán las técnicas de estimación no paramétricas desde una doble perspectiva. En primer lugar, a la hora de analizar la distribución del producto *per capita* de las regiones europeas, se han usado estimaciones *kernel* (núcleo) para aproximar la función de densidad de probabilidad ( $f(\bullet)$ ) de la variable aleatoria correspondiente a dicho producto regional. En segundo lugar, se han utilizado también dichas técnicas para estimar la relación funcional existente entre la variable  $y$  y los regresores  $X$ , es decir, la esperanza condicional  $E[y | X] = m(X)$ ; esta forma de determinación de la función de regresión es más flexible que el enfoque paramétrico ya que no exige la pertenencia de la función  $m$  a una determinada familia de funciones, exigiendo únicamente ciertas condiciones de regularidad que debe cumplir la función  $m(\bullet)$ .

Sin querer entrar en demasiados detalles de tipo técnico<sup>1</sup>, en las siguientes líneas se exponen los fundamentos básicos del enfoque no paramétrico.

Suponiendo que se dispone de  $n$  observaciones independientes,  $y_1, y_2, \dots, y_n$ , de la variable aleatoria  $Y$ , el estimador *kernel* del valor de la función de densidad de dicha variable en un punto  $y$ ,  $f(y)$ , viene dado por

$$\hat{f}_h(y) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{y_i - y}{h}\right)$$

donde  $K(\bullet)$  es la denominada función núcleo y  $h$  el llamado parámetro de suavizado (*bandwidth*) que controla la regularidad de la curva estimada.

Las propiedades estadísticas (sesgo, varianza, consistencia, etc.) del estimador propuesto dependen de la elección de la función núcleo y del parámetro de suavizado, aunque en la práctica la elección del *kernel* es menos

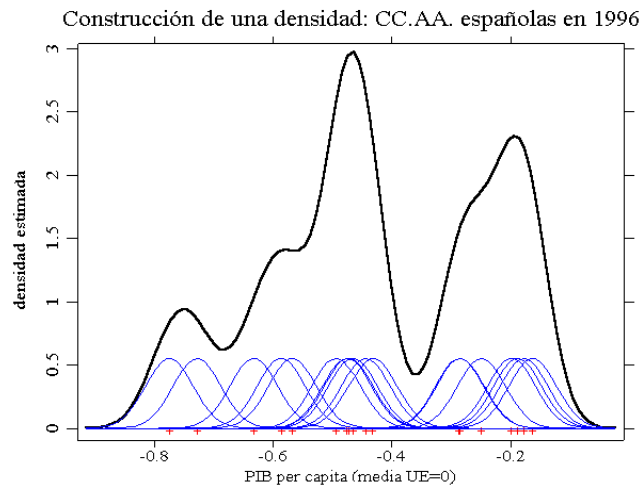
---

<sup>1</sup> Los cuales se pueden encontrar en diversos libros y artículos especializados, como por ejemplo Härdle y Linton (1994), Silverman (1986) o los surveys más recientes de Cao *et al.* (1997) y Delgado y Robinson (1995).

importante que la del parámetro  $h$  ya que se pueden reescalar las funciones núcleo para que la diferencia entre dos estimaciones basadas en *kernels* diferentes sea prácticamente despreciable (Marron y Nolan, 1988).

En nuestro caso se han empleado en todos los cálculos funciones núcleo de tipo Gaussiano ( $K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$ ) y se ha seleccionado el parámetro de suavización en base el criterio de validación cruzada mínimo-cuadrático (LSCV)<sup>2</sup>.

Para ilustrar cómo funciona la estimación *kernel* en un ejemplo sencillo, consideremos el caso de la distribución del PIB *per capita* de las regiones españolas en 1996<sup>3</sup>. En este caso, la función  $\hat{f}_h(\bullet)$  viene dada por la expresión  $\hat{f}_h(u) = \frac{1}{17\sqrt{2\pi}h} \sum_{i=1}^{17} \exp\left(-\frac{1}{2}(u_i - u/h)^2\right)$  y la Figura 1 ilustra la construcción de la estimación *kernel* basada en los 18 puntos, usando en la misma una función núcleo Gaussiana y eligiendo el  $h$  óptimo en base al criterio LSCV<sup>4</sup>.



**Figura 1**

Respecto a la utilización del enfoque no paramétrico para estimar la relación funcional entre una variable de respuesta ( $y$ ) y un conjunto de  $p$  variables explicativas ( $X$ ), dada por  $E[y | X] = m(X)$ , en este trabajo se ha usado el estimador de Nadaraya-Watson, el cual en el caso multivariante toma la forma

<sup>2</sup> Para más detalles sobre los distintos métodos de selección del *bandwidth* pueden consultarse los trabajos de Härdle (1991), Marron (1989), Park y Turlack (1992), Sheater y Jones (1991) o Turlack (1993).

<sup>3</sup> Los datos forman parte de las 164 observaciones disponibles para ese año a nivel europeo.

$$\hat{m}_h(X) = \frac{\sum_{i=1}^n \hat{K}\left(\frac{x_{i1} - x_1}{h_1}, \frac{x_{i2} - x_2}{h_2}, \dots, \frac{x_{ip} - x_p}{h_p}\right) y_i}{\sum_{i=1}^n \hat{K}\left(\frac{x_{i1} - x_1}{h_1}, \frac{x_{i2} - x_2}{h_2}, \dots, \frac{x_{ip} - x_p}{h_p}\right)}$$

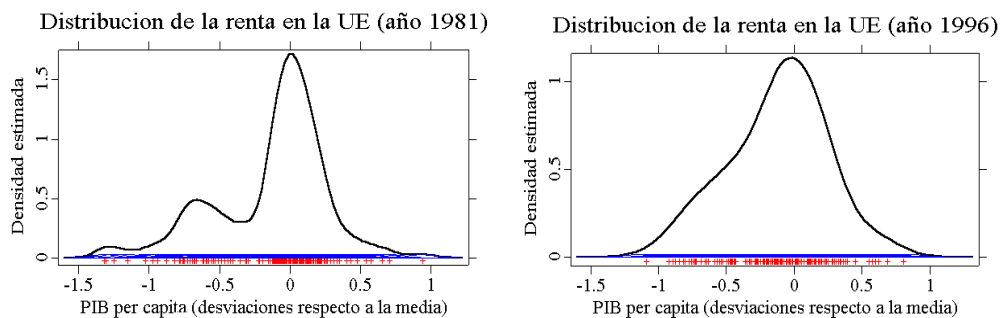
donde  $\hat{K}(u) = \hat{K}(u_1, \dots, u_p) = K(u_1) \dots K(u_p)$  es una función núcleo multi-dimensional de tipo multiplicativo, con  $K$  denotando la función *kernel* univariante, y el *bandwidth* es ahora un vector  $p$ -dimensional dado por  $h = (h_1, h_2, \dots, h_p)$ .

### 3. Análisis empírico

Una vez expuestos los fundamentos básicos de la metodología empleada, en este apartado se estudia la distribución del producto *per capita* de las regiones europeas y su evolución a lo largo del período 1981/1996. Este análisis nos proporcionará información no sólo acerca del comportamiento relativo de las diferentes regiones, sino también sobre la dinámica del proceso de convergencia europeo.

La variable básica de análisis, el producto relativo, viene expresado como el logaritmo natural del PIB *per capita* de cada región normalizado con relación al producto *per capita* medio europeo.

En la Figura 2 se presentan las densidades *kernel*-estimadas de la renta per capita relativa de las 164 regiones consideradas.

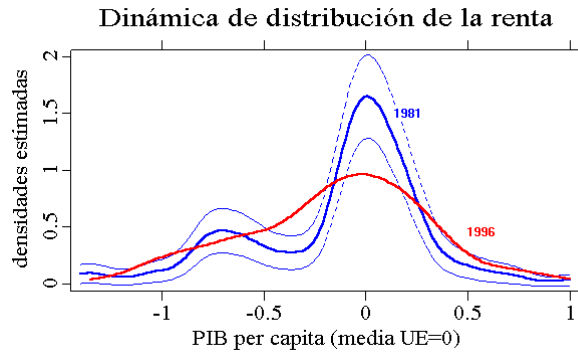


**Figura 2**

A fin de comprobar si existen diferencias estadísticamente significativas entre las densidades estimadas, en la Figura 3 se muestra ambas

<sup>4</sup> Todos los cálculos econométricos de este trabajo se han llevado a cabo con el programa XploRe (Härdle et al, 2000).

simultáneamente junto con la correspondiente banda de confianza<sup>5</sup> –para  $\alpha = 0.05$  - calculada para la densidad de 1981.



**Figura 3**

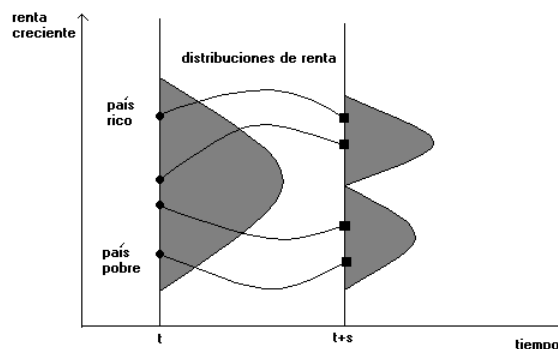
Tal como se observa en la gráfica, existen diferencias significativas entre las densidades estimadas para los años 1981 y 1996. Por una parte, la distribución de la renta en 1981 es claramente bi-modal, lo que pone de manifiesto que las regiones europeas hasta esa fecha siguieron diferentes trayectorias de crecimiento, tendiendo a agruparse en diferentes clases de renta. Este es el escenario de clubes de convergencia sugerido por Baumol (1986), e ilustrado, entre otros, por Quah (1996), en el que algunas regiones alcanzaron y se unieron al club de regiones “ricas” –con rentas similares a la media europea– mientras otras cayeron en la ‘trampa de la pobreza’.

Por otro lado, la distribución de renta quince años después, a mediados de la década de los noventa, parece acercarse a la unimodalidad, con una mayor cantidad de regiones agrupándose en torno a la renta media europea.

Las figuras anteriores sugieren algunas ideas interesantes. En primer lugar, a nivel comparado, los resultados obtenidos contrastan con los encontrados a nivel de países; por ejemplo, Quah (1997) encuentra como hecho estilizado lo que denomina ‘*emerging twin peaks*’, o bipolarización de la distribución transversal de la renta entre países. La Figura 4 muestra a grandes rasgos el fenómeno documentado en dicho trabajo: la población de economías

<sup>5</sup> Se han calculado los intervalos de confianza para la densidad estimada  $\hat{f}_h(\bullet)$  a través de la expresión  $\left[ \hat{f}_h(y) - z_{1-\alpha/2} \sqrt{\frac{\hat{f}_h(y) \|K\|_2^2}{nh}}, \hat{f}_h(y) + z_{1-\alpha/2} \sqrt{\frac{\hat{f}_h(y) \|K\|_2^2}{nh}} \right]$ , donde el valor  $z_{1-\alpha/2}$  representa el  $(1-\alpha/2)$  cuartil de la distribución normal estándar (ver Härdle, 1991).

subyacente ha tendido en el tiempo (s períodos) a polarizarse en torno a valores distintos.



**Figura 4**

Nuestros resultados para el caso de las regiones europeas parecen ir en la dirección contraria al caso de los países<sup>6</sup>, habiéndose suavizado notablemente el agrupamiento inicial de las regiones ricas y pobres. En cierto modo, esta evidencia indicando una reducción importante de las desigualdades regionales señala el efecto positivo que han tenido las distintas políticas regionales llevadas a cabo por la Unión Europea a través de los Fondos de Desarrollo Regional o de los Fondos de Cohesión, habiéndose cumplido –al menos en parte- con uno de los objetivos del Tratado de Maastricht<sup>7</sup>.

En segundo lugar, las Figuras 2 y 3 plantean diversas cuestiones acerca de cuál ha sido la dinámica de intra-distribución de la renta regional: ¿todas las regiones ricas en 1981 han continuado siendo ricas en 1996?, ¿todas las pobres se han mantenido pobres? (una respuesta afirmativa a estas preguntas implicaría una fuerte *persistencia* en la distribución); ¿o ha habido un proceso de acercamiento de las regiones ricas y pobres hacia valores promedio –empobrecimiento o enriquecimiento relativos-? (tal fenómeno de convergencia señalaría un fuerte *movilidad* en el posicionamiento relativo de la regiones).

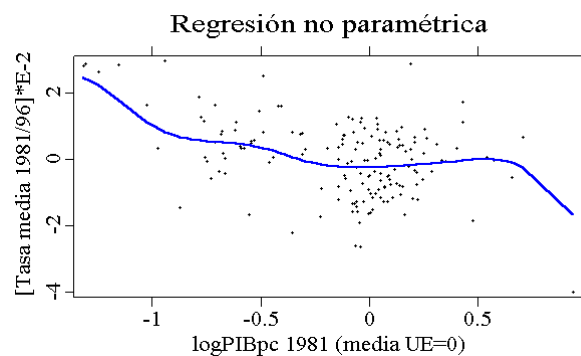
Las interrogantes anteriores plantean un análisis empírico complementario al ya realizado. ¿Ha existido convergencia en los niveles de

<sup>6</sup> No se trata de contraponer unos resultados empíricos con los otros puesto que el período temporal analizado por Quah (1961/1988) no coincide con el aquí utilizado (1981/1996), sino de comparar la regularidad empírica observada.

<sup>7</sup> En el artículo 130A de dicho Tratado se dice que “...la Comunidad se propondrá, en particular, reducir las diferencias entre los niveles de desarrollo de las diversas regiones y el retraso de las regiones menos favorecidas”.

producción?<sup>8</sup>. ¿Han tendido a desaparecer las diferencias existentes en los niveles del PIB *per capita* regional?. ¿Ha sido significativo el grado de movilidad regional intra-distribución?. ¿Cuáles han sido las causas de esta movilidad dentro del sistema y sobre el período analizado?. En los próximos párrafos se intentará responder a cada una de las cuestiones planteadas.

En primer lugar, para responder a la cuestión acerca de la existencia de convergencia regional en Europa –y su velocidad–, analizaremos la relación existente entre las tasas de crecimiento y los niveles iniciales de renta, tal como se deduce del modelo propuesto por Barro y Sala-i-Martin (1991,1992). La Figura 5 ilustra el resultado de la regresión no paramétrica entre dichas variables.



**Figura 5**

Como puede apreciarse en la gráfica, existe una relación inversa entre la tasa media de crecimiento del PIB *per capita* entre 1981 y 1996 y el valor inicial de dicha variable, lo que indica que aquellas regiones con menor nivel inicial de renta son las que, en promedio, han tenido un mayor crecimiento durante este período ( $\beta$ -convergencia)<sup>9</sup>. Realizando la regresión paramétrica correspondiente, dada por  $\log y_{iT} - \log y_{i,t_0} = a - b \log y_{i,t_0} + \varepsilon_{it}$ , se obtiene una velocidad de convergencia<sup>10</sup> absoluta del 1.3% anual, relativamente lenta si se compara con la obtenida a partir de las distintas estimaciones a nivel nacional e internacional para países o regiones.

<sup>8</sup> Esto implicaría que las regiones inicialmente más atrasadas han crecido en el período 1981/1996 a un ritmo mayor que las más avanzadas al comienzo de dicho período.

<sup>9</sup> Por otro lado, cuando se analizó la dispersión del PIB *per capita* a lo largo del período 1981/1996 ( $\sigma$ -convergencia), la gráfica correspondiente mostró un descenso de  $\sigma_t$  en el tiempo, aunque en los primeros años tendió a aumentar, y sólo a partir de la segunda mitad de la década de los ochenta el grado de dispersión comenzó a disminuir, aunque de forma reducida.



A pesar de la utilidad del análisis anterior, sin embargo, no se obtiene información acerca de si las regiones pobres se están acercando a las ricas, o sobre la posibilidad de que se esté observando un proceso de convergencia entre las regiones de un mismo grupo (club), pero no entre los distintos grupos de economías. A continuación abordamos estas cuestiones (hipótesis de clubes de convergencia) analizando la evolución relativa del producto *per capita* de las distintas regiones europeas.

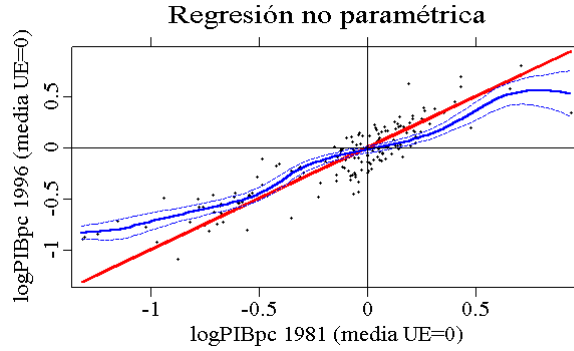
Siguiendo la formulación propuesta por Chatterji (1992,1993), si definimos la variable  $z_{it}$  como la diferencia entre el logaritmo de la renta *per capita* de cada región y el logaritmo la renta media en un cierto instante de tiempo, para analizar el proceso de convergencia de una economía es necesario analizar una relación del tipo  $z_{it} = \Phi(z_{i,t_0})$ , donde 0 representa el año base.

Representando la variable  $z_{iT}$  frente a  $z_{i,t_0}$ , es decir, la distancia de cada región con relación a la media en los momentos inicial y final, y una recta de 45°, correspondiente a una situación de estabilidad en la que la distancia respecto a la media no cambia en ambos momentos, pueden presentarse distintas situaciones de equilibrio dependiendo de la posición relativa inicial: puede convergerse hacia un único punto de equilibrio o aparecer múltiples estados estacionarios, localmente estables, a los que convergerán las distintas regiones, poniendo así de manifiesto la existencia de distintos clubes de convergencia.

En la Figura 6 se presentan los resultados obtenidos al estudiar la convergencia de las regiones europeas entre 1981 y 1996.

---

<sup>10</sup> Para calcular la velocidad de convergencia se ha utilizado la fórmula  $\hat{b} = (1 - e^{-\hat{b}T})/T$  donde  $T$  es la duración del intervalo de estimación (en nuestro caso 15 años).



**Figura 6**

Tal como parece desprenderse del ajuste no paramétrico realizado, no se observa evidencia a favor de la hipótesis de existencia de distintos clubes de convergencia entre las regiones europeas: entre 1981 y 1996 tanto las regiones más avanzadas como las más atrasadas (en términos relativos) han tendido a reducir la distancia que las separa de la media europea. Sin embargo, en el cuadrante inferior izquierdo –en torno al valor -0.5– el intervalo de confianza<sup>11</sup> (al 95%) alrededor de la curva estimada llega a tocar a la recta de 45° lo que evidencia, siquiera marginalmente, la existencia de un grupo de regiones cercana a un punto de equilibrio “inferior” al resto de economías regionales.

Una vez analizados el proceso de convergencia y la evolución relativa del producto *per capita* regional, a continuación abordamos el tema de la dinámica de movilidad regional.

Una de las formas de cuantificar la dinámica intra-distribución es a través de las tablas de movilidad, en las cuales se discretiza el espacio de la variable renta y se contabilizan las transiciones observadas entre las distintas celdas. Si representamos en la primera columna la distribución de frecuencias del producto *per capita* en 1981, en la primera fila la distribución correspondiente a 1996 y en cada celda  $(i, j)$  se recuenta el número de regiones que comenzaron

<sup>11</sup> Como en el caso de la estimación de densidades, puede obtenerse un intervalo de confianza aproximado para la función  $\hat{m}_h(x)$ ; la expresión es similar, substituyendo la expresión dentro de la raíz cuadrada por  $\hat{\sigma}_h^2(x) \|K\|_2^2 / nh \hat{f}_h(x)$ , donde  $\hat{f}_h(x)$  es la densidad estimada de la variable explicativa  $x$  y  $\hat{\sigma}_h^2(x)$  es la estimación de la varianza  $Var[y | x] = \sigma^2(x)$ , dada por la

$$\text{expresión } \hat{\sigma}_h^2(x) = \frac{\frac{1}{n} \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) \{y_i - \hat{m}_h(x)\}^2}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}.$$

en el grupo correspondiente a la fila  $i$ -ésima y terminaron en 1996 en la columna  $j$ -ésima, se obtiene la siguiente información acerca de la movilidad interna existente entre las 164 regiones europeas consideradas para el período 1981-1996<sup>12</sup>.

1996→ 1981 ↓		-1.5	-1.25	-1	-0.75	-0.5	-0.25	0	0.25	0.5	0.75	1
	frec	0	0	3	15	20	33	51	32	7	3	0
-1.5	0											
-1.25	4			1	3							
-1	3			1	1	1						
-0.75	16			1	9	6						
-0.5	16				1	10	4	1				
-0.25	15				1	1	10	3				
0	71					2	19	39	11			
0.25	30							8	19	2	1	
0.5	6								1	4	1	
0.75	2									1	1	
1	1								1			

**Cuadro 1**

A partir de la tabla anterior se puede construir una tipología de regiones en función de su posición relativa al comienzo y al final del período analizado.

Así, se pueden distinguir: regiones con una posición inicial por debajo de la media europea y que al final del período mantienen la distancia que la separa de esta (diagonal del segundo cuadrante); regiones con una posición inicial por encima o en torno a la media y que mantienen su ventaja comparativa al final de 1996 (diagonal del cuarto cuadrante); regiones que partieron en 1981 en una posición por debajo de la media y empeoraron su situación relativa (parte inferior del segundo cuadrante), o aquellas que redujeron la distancia al final del período pero se mantuvieron aún por debajo de la media (parte superior del segundo cuadrante); regiones que mejoraron su situación, moviéndose de una posición por debajo de la media a otra en torno a la media o por encima de ésta (primer cuadrante completo), o aquellas que empeoraron considerablemente su situación pasando de estar por encima o en torno a la media a situarse por debajo de ésta (tercer cuadrante completo).

<sup>12</sup> El valor de las primeras columna y fila representa el centro de cada uno de los intervalos considerados en la distribución (marca de clase).

Finalmente, pueden considerarse aquellas regiones con una posición inicial por encima o en torno a la media que redujeron la distancia que la separa de la misma –empeoraron–, aunque manteniéndose por encima o en torno a la media (parte inferior del cuarto cuadrante), y aquellas que partiendo de dicha posición aumentaron aún más la distancia respecto al promedio europeo (parte superior del cuarto cuadrante).

Clasificando las 164 regiones europeas en los ocho grupos considerados se obtuvieron los siguientes porcentajes (en el mismo orden en el que se han definido los distintos tipos de regiones): 18.3% en las regiones tipo I, 38.4% en las regiones tipo II, 2.4% en las regiones tipo III, 9.8% en las regiones tipo IV, 2.4% en las regiones tipo V, 12.8% en las regiones tipo VI, 6.7% en las regiones tipo VII y, finalmente, 9.2% en las regiones tipo VIII.

Como puede apreciarse, a partir de la tipología anterior se deduce que entre 1981 y 1996 la mayor parte –56.7%– de las regiones ha mantenido su posición relativa (tipos I y II). También se observa que las regiones avanzadas son las que más –19.5%– ven empeorar su posición (tipos VI y VII), mientras que son muchas menos –12.2%– las regiones más atrasadas que ven mejorar su posición (tipos IV y V). Finalmente, son pocas –11.6%– las regiones que aumentan la distancia que las separa de la media europea, tanto entre las más ricas (tipo VIII) como entre las más pobres (tipo III).

Para complementar los comentarios hechos a partir de la tabla de movilidad regional, a continuación se ha descompuesto la misma en dos sub-tablas: la tabla de inmovilidad, que comprende aquellas regiones cuya posición relativa no ha cambiado a lo largo del período considerado (diagonal de la tabla); y la tabla de cambio, en la que aparecen las regiones situadas inicialmente en el grupo  $i$  y que al final del período se sitúan en el grupo  $j$  ( $i \neq j$ ).

Sumando las frecuencias absolutas de las celdas de cada sub-tabla y dividiendo por el número total de regiones se obtienen, respectivamente, los índices de inmovilidad ( $I_I$ ) y de cambio ( $I_C$ )<sup>13</sup>. Además, al estar ordenados los grupos, en la sub-tabla de cambio las frecuencias de las celdas de la matriz triangular superior muestran cambios ascendentes (regiones que han mejorado de grupo a lo largo del período), mientras que las de la matriz inferior indican

---

<sup>13</sup> Por definición, ha de cumplirse que  $I_C = 1 - I_I$ .

cambios descendentes. Por tanto, puede descomponerse el índice de cambio de la forma  $I_C = I_{CA} + I_{CD}$ .

El resultado de aplicar la descomposición anterior a la matriz de movilidad original dio como resultado los valores (en %):  $I_I=56.7$ ,  $I_C=43.3$ ,  $I_{CA}=21.3$  y  $I_{CD}=22.0$ . Como puede apreciarse, el índice de inmovilidad es superior al de cambio (persistencia elevada) y, entre las regiones que modifican su posición relativa entre 1981 y 1996, el número de las que mejoran su posición es similar al de aquellas que la empeoran<sup>14</sup>.

A pesar de la utilidad del análisis realizado en los párrafos anteriores a la hora de cuantificar el grado de movilidad intra-distribución, la discretización *ad hoc* realizada puede distorsionar la dinámica real subyacente puesto que las observaciones que constituyen nuestros datos se supone que proceden de variables continuas. La solución a este problema consiste en no discretizar la variable renta y cuantificar la movilidad a través de la estimación no paramétrica de una densidad conjunta para las observaciones originales del producto *per capita* regional de los años 1981 y 1996<sup>15</sup>.

En el caso multivariante, la densidad estimada a partir de un conjunto de  $n$  observaciones independientes  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  (donde ahora  $\mathbf{y}_j = (y_{j1}, y_{j2}, \dots, y_{jp})'$ ) viene dada por la expresión

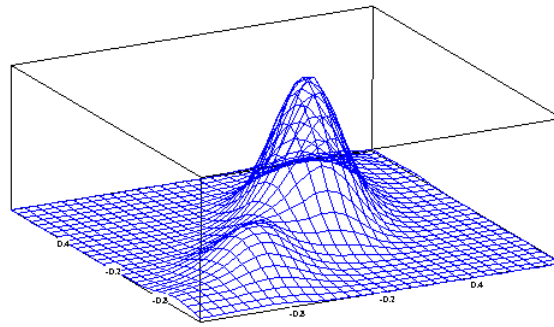
$$\hat{f}_h(\mathbf{y}) = \hat{f}_h(y_1, \dots, y_p) = \frac{1}{nh_1 \dots h_p} \sum_{i=1}^n \hat{K} \left( \frac{y_{i1} - y_1}{h_1}, \frac{y_{i2} - y_2}{h_2}, \dots, \frac{y_{ip} - y_p}{h_p} \right)$$

donde  $\hat{K}(u) = \hat{K}(u_1, \dots, u_p) = K(u_1) \dots K(u_p)$  es la función núcleo multivariante.

La Figura 7 muestra la densidad conjunta estimada para las observaciones  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{164}$ , con  $\mathbf{y}_j = (y_{j1981}, y_{j1996})'$ . Tal gráfica indica cómo ha evolucionado la distribución transversal de la renta regional desde el año 1981 hasta 1996.

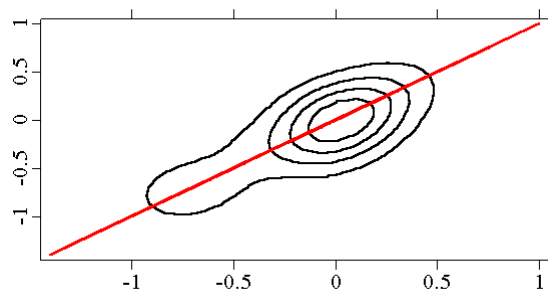
<sup>14</sup> Habría que matizar los índices de cambio ascendente y descendente para diferenciar los cambios ascendentes de las regiones más atrasadas y los descendentes de las más avanzadas, al ser éstos los que indicarían una tendencia a la convergencia.

<sup>15</sup> Esto es equivalente a hacer que el número de intervalos considerados con anterioridad en la distribución de frecuencias tienda hacia infinito.



**Figura 7**

Esta densidad bivalente puede visualizarse en detalle a través de los gráficos de contorno, las cuales representan las curvas  $\hat{f}_h(y_{1981}, y_{1996}) = c$  para diferentes valores de  $c$ . La Figura 8 siguiente muestra las líneas de contorno de la densidad estimada  $\hat{f}_h(y_{1981}, y_{1996})$  para los niveles 0.2, 0.4, 0.6 y 0.8.



**Figura 8**

Como puede apreciarse, la mayor parte del gráfico se concentra alrededor de la línea diagonal<sup>16</sup>, lo que corrobora el análisis realizado con la tabla de movilidad, en el sentido de que señala que existe una parte importante de regiones que se mantienen al final del período considerado en la misma posición que inicialmente ocupaban (por encima o por debajo de la media europea)<sup>17</sup>.

Finalmente, para concluir el análisis realizado cabe preguntarse, tal como se hacía algunos párrafos atrás, cuáles son las causas que pueden

<sup>16</sup> La diagonal de 45° grados mostrada aquí se corresponde con la diagonal de 135° marcada en la tabla de movilidad.

<sup>17</sup> Si, por el contrario, la mayor parte de la masa de la densidad estimada se concentrara en la diagonal de 45° rotada otros 90° en el sentido contrario a las agujas del reloj (en total 135°), señalaría que muchas regiones ricas han empeorado su posición relativa –regiones tipo VI– mientras que otras tantas de las regiones pobres se han movido de una situación por debajo de la media a otra en torno o por encima de ésta –regiones tipo V–.

explicar el crecimiento de las distintas economías regionales europeas en el período 1981/1996 o, más generalmente, cuales son los factores explicativos de la (in)movilidad observada en la distribución regional.

En este sentido, se utilizará una regresión de convergencia condicional del tipo  $\log y_{iT} - \log y_{i,t_0} = a - b \log y_{i,t_0} + X_{i,t_0} \theta + \varepsilon_{it}$  para intentar cuantificar los determinantes del crecimiento y la convergencia observados durante el período analizado; es decir, una regresión donde se intenta explicar la tasa de crecimiento en función del nivel inicial de renta y un conjunto de variables que determinan la posición de estado estacionario de cada región.

En este trabajo se han incluido como variables adicionales al efecto de convergencia neoclásico el peso del sector agrícola en cada región (expresado en términos de participación del empleo en el sector primario), la tasa de participación del empleo en la población total y un conjunto de variables ficticias de carácter geográfico que hacen referencia a la pertenencia de las regiones a alguno de los ejes de desarrollo identificados por la propia Comisión de las Comunidades Europeas<sup>18</sup>.

En el Cuadro 2 se recogen los resultados obtenidos al estimar la regresión anterior con los datos correspondientes a las 164 regiones europeas.

	Coef. estimado	Estadístico <i>t</i> (White)	P-valor
Constante	-0.226	-2.048	0.042
log(PIBpc) <sub>81</sub>	0.386	10.217	0.000
%AGRIC <sub>81</sub>	-0.004	-2.253	0.026
%OCUP <sub>81</sub>	0.009	3.942	0.001
Arco SP	-0.204	-5.313	0.000
Arco AT	-0.207	-5.573	0.000
Arco CC	-0.068	-2.273	0.024

<sup>18</sup> La Comisión de las CE considera la existencia de los siguientes arcos de desarrollo: arco Alpino (AL) formado por las regiones alpinas de Alemania, Francia e Italia; el arco Septentrional (SP), con las regiones costeras del mar del Norte; el arco Atlántico (AT); el arco de las Capitales Centrales (CC), formado por las regiones más avanzadas del Reino Unido, Alemania, Países Bajos, Bélgica, Luxemburgo y Francia; la Diagonal Continental (DC), o regiones continentales interiores de Francia y España; y dos zonas de desarrollo en el Mediterráneo: el arco Mediterráneo Central (MC), que comprende las regiones del Mezzogiorno italiano y Grecia, y el arco Mediterráneo Occidental (MO), que se extiende desde Andalucía hasta Nápoles.

Arco DC	-0.092	-2.883	0.004
Arco MC	-0.211	-4.384	0.000
Arco MO	-0.088	-2.381	0.019
<b>Ajuste del modelo →</b>	$R^2=0.46$	$\hat{\sigma}=0.129$	F=14.57 [P-val=0.000]

**Cuadro 2**

A partir de las cifras anteriores pueden extraerse varias conclusiones. En primer lugar, la velocidad de convergencia que se deriva de la regresión anterior es del 3.2% anual, sensiblemente superior a la tasa estimada a partir de la regresión absoluta. En segundo lugar, los efectos del peso del sector agrícola y de la participación en el empleo resultan estadísticamente significativos y presentan el signo esperado. Por último, las variables geográficas relativas a los distintos arcos regionales resultan conjunta e individualmente significativas, señalando cómo la convergencia en los niveles de producto *per capita* entre las regiones europeas se produce a ritmos diferentes según la pertenencia a determinadas áreas geográficas.

A continuación, se han utilizado los resultados de la regresión condicional para descomponer la tasa de crecimiento observada en sus distintas fuentes diferenciales. En concreto, se han estimado cuatro factores de crecimiento diferencial con respecto a una región representativa hipotética, básicamente la región promedio durante el período 1981/1996: el efecto de convergencia neoclásico, el impacto del sector agrícola, la contribución del empleo y el impacto de la ubicación geográfica de cada región.

Una vez realizada la descomposición diferencial, se ha planteado el siguiente ejercicio de simulación: ¿cuál hubiera sido la distribución de la renta regional en 1996 si la renta relativa de cada región hubiera cambiado sólo debido a alguno de los factores diferenciales estimados (*ceteris paribus*, es decir, manteniéndose constantes –en su valor promedio– el resto de factores). Para responder a esta cuestión se han calculado los valores de renta teórica que corresponderían a la regiones sumando al nivel inicial de producto *per capita* en 1981 el valor promedio del período completo y el factor diferencial correspondiente. Una vez realizados los cálculos anteriores, se han estimado las correspondientes densidades y se han comparado con la densidad



estimada en 1981. En la Figura 9 siguiente se presentan la densidades estimadas en cada caso (a efectos comparativos aparece en primer lugar la gráfica conjunta de las densidades estimadas con los datos reales de 1981 y 1996).

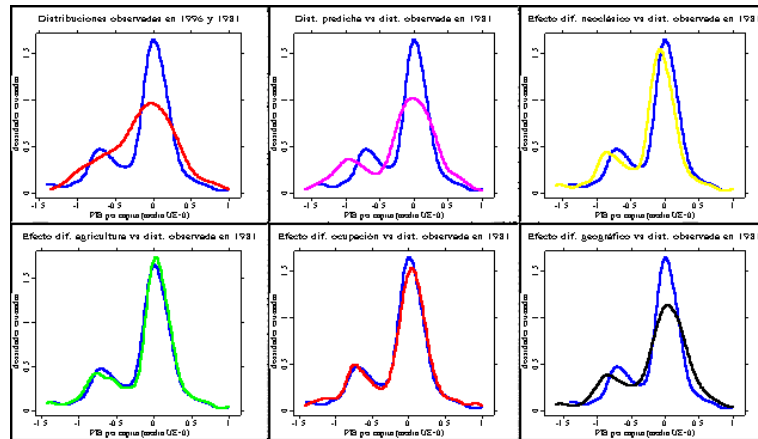


Figura 9

Como puede apreciarse a través de la segunda gráfica superior, la distribución que se obtendría en 1996 a partir del modelo –distribución predicha–, obtenida como suma de todos los factores diferenciales, no se corresponde con la realmente observada en 1996 (primera gráfica superior). La distribución sigue mostrando claramente dos modas, e incluso se aprecia como aumenta la polarización entre regiones puesto que la segunda moda se sitúa aún más a la izquierda de la que se observaba en 1981.

Por otra parte, de los cuatro factores explicativos introducidos, los únicos que contribuyen a modificar la distribución inicial de renta, aunque sin conseguir eliminar la bimodalidad, son los efectos neoclásico y geográfico.

Como conclusión al ejercicio de simulación realizado, podemos anotar que el modelo de regresión empírico utilizado *no es capaz* de explicar la desaparición –al menos parcial– de los dos modas que se observaban en 1981. Además, el análisis sugiere que los únicos factores que han contribuido a modificar la situación inicial han sido el efecto de convergencia neoclásico y las variables ficticias geográficas.

Es obvio que deben haber existido otras causas que han originado cambios importantes en la distribución transversal del producto *per capita* europeo a nivel regional. Entre las variables que explicarían dichos cambios se encuentran las diferencias en las dotaciones de capital público, tecnológico o

humano, o los movimientos migratorios entre las distintas regiones europeas. Sin embargo, estas variables no se han podido incluir en el modelo por las deficiencias existentes en la información estadística a nivel regional en la Unión Europea.

#### **4. Conclusiones**

En este trabajo se ha analizado la dinámica de crecimiento económico de las regiones europeas durante el período 1981/1996. Del análisis realizado se desprenden varias conclusiones.

En primer lugar, ha habido una reducción importante de las desigualdades regionales, pasándose de una situación inicial polarizada en torno a dos grupos de renta a otra caracterizada por una mayor homogeneidad en la distribución. Además, las ecuaciones de convergencia estimadas muestran que ha existido una  $\beta$ -convergencia significativa, indicando que las regiones inicialmente más atrasadas son las que han experimentado un mayor crecimiento durante los quince años considerados.

En segundo lugar, a pesar de los cambios que se han producido, ha existido una fuerte persistencia en el comportamiento regional, es decir, el grado de inmovilidad intra-distribución ha sido muy elevado.

Finalmente, a la hora de buscar factores explicativos del cambio producido, se ha encontrado que las variables condicionantes empleadas (efectos neoclásico, geográfico, agrícola y de empleo) no son capaces de explicar por sí solas la situación económica regional que se ha dado finalmente en 1996. Ello indica claramente que la investigación aquí realizada no está cerrada y que habrá que seguir buscando factores explicativos adicionales a los aquí utilizados.

#### **5. Bibliografía**

Barro R.J. y Sala-i-Martin X. (1991), "Convergence across states and regions", *Brookings Papers on Economic Activity*, n. 1, pp. 107-182.

Barro R.J. y Sala-i-Martin X. (1992), "Convergence", *Journal of Political Economy*, Vol. 100(2), pp. 223-251.

Baumol W.J. (1986). "Productivity growth, convergence, and welfare", *American Economic Review*, Vol. 76(5), pp.1072-1085.

Cao R., Delgado M.A. y González-Manteiga W. (1997), "Nonparametric curve estimation: an overview", *Investigaciones Económicas*, Vol. XXXI(2), pp. 209-252.

Chatterji M. (1992), "Difusión, endogenous growth and múltiple equilibria", Discussion Paper n. 38, University of Dundee.

Chatterji M. (1993), "Convergence clubs and endogenous growth", *Oxford Review of Economic Policy*, Vol. 8(4), pp. 57-69.

Delgado M.A. y Robinson P.M. (1995), "Nonparametric and semiparametric methods for economic research", en *Survey in Econometrics*, Oxley et al. (ed.), Basil Blackwell, Cambridge, pp. 350-396.

Härdle W. (1991), *Smoothing techniques, with implementations in S*, Springer, New York.

Härdle W. y Linton O. (1994), "Applied nonparametric methods", en *Handbook of Econometrics*, Vol. IV, Engle y McFadden (ed.), Elsevier Science, Amsterdam, pp. 2295-2339.

Härdle W., Klink S. y Müller M. (2000), *XploRe®-Learning Guide*, Springer-Verlag, Berlin.

Marron J.S. (1989), "Comments on a data based bandwidth selector", *Computational Statistics and Data Analysis*, Vol. 8, pp. 155-170.

Marron J.S. y Nolan D. (1988), "Canonical kernels for density estimation", *Statistics and Probability Letters*, Vol. 7(3), pp. 195-199.

Park B.U. y Turlach B.A. (1992), "Practical performance of several data driven bandwidth selectors", *Computational Statistics*, Vol. 7, pp. 251-270.

Quah D.T. (1996), "Twin peaks: growth and convergence in models of distribution dynamics", *The Economic Journal*, Vol. 106, pp. 1045-1055.

Quah D.T. (1997), "Empirics for growth and distribution: stratification, polarization, and convergence clubs", Discussion Paper n. 324, Centre for Economic Performance.

Sheater S.J. y Jones B.C. (1991), "A reliable data-based bandwidth selection method for kernel density estimation", *Journal of the Royal Statistical Society, Ser. B*, Vol. 53, pp. 683-690.

Silverman B.W. (1986), *Density Estimation for Statistics and Data Analysis*, Monographs on Statistics and Applied Probability, Vol. 26, Chapman and Hall, London.

Turlach B.A. (1993), "Bandwidth selection in kernel density estimation: a review", Discussion Paper n. 9307, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin,