

La convergencia real de Andalucía: Un análisis de cointegración del mercado de trabajo

Alicia Avilés
Consuelo Gámez
José L. Torres¹
Universidad de Málaga

BIBLID [0213-7585 (1997); 47: 15-36]

PALABRAS CLAVE: convergencia real, cointegración, tendencias comunes, análisis regional.

RESUMEN

En este trabajo analizamos la convergencia real y la existencia de tendencias comunes entre algunas variables del mercado de trabajo de Andalucía y España menos Andalucía. Para ello utilizamos un enfoque basado en el estudio de la existencia de una relación de cointegración y, por tanto, de la existencia de una tendencia común entre las variables representativas de dicho mercado y las del resto de regiones de España. Los resultados muestran que existe una tendencia común entre las tasas de paro, tasas de ocupación y población activa. Analizando en cada caso el vector de cointegración, obtenemos que existe un proceso divergente en términos de la tasa de paro y la tasa de ocupación, siendo convergente únicamente en términos de la población activa.

ABSTRACT

This work analyses convergence and the existence of common trends among Andalusian and Spanish labour markets. In doing that, we use an approach based on the existence of a cointegration relationship or alternatively, the existence of a common trend between labour market variables in both Andalusia and the rest of Spain's regions. We find the existence of a common trend in the unemployment rate, the employment rate, and labour force. Looking at the cointegration vectors we find the existence of divergence in both, unemployment rate and employment rate, whereas we find the existence of convergence in terms of the labour force.

1. Este trabajo ha recibido apoyo financiero de la Junta de Andalucía a través del Grupo de Investigación SEJ 0122. Todos los posibles errores son responsabilidad de los autores.

1. INTRODUCCIÓN

El objetivo de este trabajo es analizar la existencia de convergencia real (en otros términos, equidad o cohesión económica) entre las regiones españolas, en concreto, entre Andalucía y el resto de España. En términos generales, la convergencia expresa la aproximación de los niveles de renta per cápita de un conjunto de regiones o países durante un período determinado. Otro indicador de convergencia real es la similitud o diferencia entre las tasas de paro interregionales. Estas dos variables, renta per cápita y tasa de desempleo, son las que tradicionalmente se utilizan como indicadores del grado de convergencia real alcanzado. Aunque hay abundancia de trabajos sobre la convergencia de regiones o países en términos de la renta per cápita, en términos de la tasa de desempleo su estudio en el ámbito regional no se ha abordado con profusión. Por esta razón, y por la elevada tasa de desempleo de Andalucía con respecto a la nacional, analizamos el grado de convergencia real que se ha producido en los últimos 20 años centrándonos en la tasa de desempleo, junto con otras variables representativas de la evolución del mercado de trabajo: la tasa de ocupación y la población activa.

Con este análisis intentamos aportar respuestas a la siguiente pregunta: ¿Tiende a converger Andalucía hacia niveles de tasa de desempleo, tasa de ocupación y población activa, similares al resto de España o por el contrario, tiende a persistir o incluso a aumentar la disparidad existente? A partir de este estudio pretendemos obtener una información fundamental sobre la evolución del mercado de trabajo andaluz con respecto al resto de regiones, que va a ser determinante de la senda futura de la economía andaluza.

Para analizar el tipo de relación que existe entre estas variables del mercado de trabajo, en Andalucía y el resto de España, utilizaremos la metodología de la cointegración, siguiendo a Johansen (1988) y Johansen y Juselius (1990). La existencia de una relación de cointegración entre las citadas variables implicaría que hay una tendencia común, por lo que habría una relación estable de equilibrio a largo plazo. Es precisamente el análisis de esta relación de equilibrio a largo plazo la que nos permitirá deducir si el proceso que siguen ambas variables es convergente, divergente, o simplemente se mantienen las diferencias existentes a lo largo del tiempo.

Si observamos la evolución de las tasas de paro y las tasas de ocupación, de Andalucía y del resto de España, éstas aumentan o disminuyen de forma similar a lo largo del tiempo. Ahora bien, este movimiento

sincronizado de ambas series no arroja luz alguna sobre la existencia de convergencia o divergencia en las mismas. La pregunta que nos hacemos es si este proceso similar muestra convergencia o divergencia. En términos generales se entiende por proceso convergente el que tiende a disminuir las diferencias entre ambas variables a través del tiempo. Si las diferencias aumentan, estaríamos hablando de un proceso divergente. Por último en el caso de que las diferencias permanezcan constantes a lo largo del tiempo, no existiría convergencia ni divergencia, indicando que las tasas de crecimiento de ambas variables serían similares.

Las variables que definen la situación del mercado de trabajo se ven afectadas por dos tipos de fuerzas: una que afecta al agregado nacional y otra que tiene un marcado componente regional. Los mercados de trabajo regionales se ven afectados por las condiciones económicas e institucionales generales del país, como las rigideces a la movilidad, nivel del precio oficial del dinero y del tipo de cambio, políticas micro y macroeconómicas y en general, por la situación internacional. Estos condicionantes afectan tanto a nivel nacional como a nivel regional, si no en términos cuantitativos, sí en la misma dirección, lo que provoca que las variables del mercado de trabajo nacional y regional se muevan, en parte, de forma similar. Sin embargo, la disparidad de las características estructurales y particulares de los mercados de trabajo regionales, estructura productiva, nivel de formación profesional, dotación de infraestructuras, dotaciones de ciertos factores productivos,..., originan una dispersión de las tasas de paro entre regiones. Por tanto, si bien podemos esperar que ambas variables tengan una evolución temporal similar provocada por la marcha general de la economía nacional, en dicha evolución puede existir o no un proceso de convergencia, que va a estar determinado por las características propias de los mercados de trabajo regionales. Tradicionalmente, la situación laboral de Andalucía ha sido más negativa que la del resto de España: mantiene una tasa de paro superior, una tasa de ocupación inferior y una menor población activa en relación con la población total. Estas diferentes condiciones iniciales van a afectar al grado de convergencia real, en tanto que determinadas definiciones de la misma (véase Marcet 1994) dependen de dichas condiciones.

La estructura del trabajo es la siguiente. En la sección segunda se exponen las diferentes mediciones de convergencia que se han utilizado en la literatura sobre el tema, así como un breve análisis de los estudios empíricos que se han realizado. En la sección tercera se discute la metodología econométrica para el análisis del proceso de convergencia. En la cuarta se presenta la evidencia empírica: tests de integración, coin-

tegración y obtención de las tendencias comunes. Por último, en la sección quinta se señalan algunas conclusiones que se pueden derivar de los resultados obtenidos.

2. MEDICIÓN DE LA CONVERGENCIA

En este trabajo, como se ha indicado en la introducción, se analiza el proceso de convergencia real de Andalucía con respecto al resto de regiones de España en función de algunas de las variables que definen el mercado de trabajo. Para ello, utilizamos un enfoque basado en la teoría de la cointegración, siguiendo a Bernard y Durlauf (1995), como una alternativa a otros enfoques suficientemente conocidos y utilizados para el análisis de la convergencia en los últimos años, como las técnicas de Barro (1991) y Barro y Sala-i-Martin (1992), que consisten en analizar en una muestra transversal (países, regiones,...) la relación entre la situación inicial o de partida de la variable analizada (normalmente la renta per cápita) en cada área económica y las tasas de variación de la misma en un período concreto, o bien la técnica de Rogoff (1985) que utiliza el cálculo de las varianzas condicionadas como medidas de dispersión.

En general, cuando hablamos de convergencia hacemos referencia al hecho de que la diferencia entre dos o más series tiende a hacerse más pequeña o a desaparecer con el transcurso del tiempo².

2. El significado del término convergencia más ampliamente conocido es el propuesto por Barro (1991) y Sala-i-Martin (1992), en su doble conceptualización más tradicional, esto es, la convergencia β y la convergencia σ , así como la convergencia condicionada. Atendiendo a Sala-i-Martin (1994, 1996), se dice que en una muestra cruzada de economías, hay convergencia β si se encuentra una relación negativa entre la tasa de crecimiento de los ingresos per cápita y el nivel inicial de ingresos, en otras palabras si las economías más pobres tienen un crecimiento más rápido que las ricas. Este concepto se confunde con una definición alternativa, la denominada convergencia σ , según la cual la dispersión de los ingresos reales per cápita entre grupos de economías tiende a disminuir a lo largo del tiempo, existiendo convergencia si la dispersión disminuye con el tiempo. Sala-i-Martin argumenta que los dos conceptos, aunque no sean idénticos, están relacionados, para mayor detalle veáse Sala-i-Martin (1994). En palabras de Mas y otros (1993), la convergencia β es un concepto de largo plazo que analiza si en promedio y en un período suficientemente largo, las tasas de crecimiento del producto per cápita han sido superiores en las economías más pobres. La constatación de esta evidencia es, sin embargo, compatible con la presencia de perturbaciones en determinados subperíodos de la muestra, que aumentan transitoriamente la dispersión. Respecto a la convergencia condicional, ésta se da cuando existen diferencias interregionales en el output en el estado estacionario, la convergencia está condicionada por determinadas características particulares de cada región, y cada una de ellas converge hacia su propio estado estacionario.

Sala-i-Martin (1994) argumenta que, para distintas muestras, parece darse convergencia, siendo las velocidades a las que convergen las regiones de países diferentes, durante períodos distintos, sorprendentemente similares, alrededor de un dos por ciento, explorándose una gran variedad de posibles explicaciones a este hecho. Sin embargo, este importante resultado ha sido criticado recientemente por Quah (1996) que, examinando esta cuestión desde el punto de vista de la dinámica estocástica de raíces unitarias en grandes secciones cruzadas, concluye que esta tasa de crecimiento del 2 por ciento podría surgir de elementos ortogonales a los modelos de crecimiento.

Para contrastar la convergencia se han propuesto diversos enfoques (véase Camarero y otros (1995)):

1. Utilizar medidas de dispersión de las varianzas condicionadas: Tienen la desventaja de que se basan en la utilización de índices, siendo, por tanto, nula la dispersión en el año base, además de verse afectadas por acontecimientos exógenos, como pueden ser las perturbaciones de oferta o de cualquier otra índole.
2. A través de los parámetros de relaciones econométricas, de forma que si los coeficientes resultasen ser similares sería una prueba de convergencia. Una estrategia paralela sería analizar la existencia de cambios estructurales.
3. La moderna teoría del crecimiento contrasta la existencia de convergencia verificando el menor crecimiento de los países de renta per cápita más elevada respecto a los países con rentas menores.

Para España, la evidencia empírica sobre la convergencia es amplia. García y Raymond (1994) estudian la convergencia σ y β de las regiones españolas a través del PIB per cápita. Los resultados muestran un importante proceso de convergencia hasta finales de los 70, interrumpiéndose a partir de esta fecha. García, Raymond y Villaverde (1995) analizan la convergencia σ de las provincias españolas comparándolas con las correspondientes a las regiones. Los resultados indican que existió un proceso de convergencia real entre 1955 y 1991, más intenso en aquellas provincias que partían de un desarrollo medio-alto, aunque dicho proceso se quebró en los años ochenta. El análisis de la convergencia β condicionada muestra un agotamiento de la convergencia interprovincial. García y Raymond (1996) analizan la convergencia σ y la β de las regiones españolas, a través de una ecuación de convergencia en la que se incluyen los efectos temporales y la existencia de efectos individuales específicos, al tiempo que se estudia la influencia de los procesos migratorios sobre la convergencia. Los resultados muestran que existe un cierto paralelismo entre la convergencia σ y el fenómeno migratorio. Mas y otros (1994) analizan la convergencia condicionada y no condicionada de las regiones españolas en relación con la estructura productiva y el capital público. Artís y otros (1985) analizan la convergencia de las regiones españolas en términos de las variaciones en el empleo y la tasa de inflación.

Concretamente, se estima la regresión del producto per cápita de un país y su nivel anterior, y si existe convergencia el coeficiente debe ser negativo.

4. La existencia de cointegración entre las variables objeto de análisis se ha considerado, por muchos autores, como una condición necesaria aunque no suficiente de convergencia, dado que la convergencia exige que las diferencias de las series no pueden tener varianzas infinitas.
5. Contraste sobre los vectores de cointegración. Utilizando las mismas técnicas se han realizado trabajos que contrastan si, una vez que existe cointegración, esas relaciones satisfacen una serie de restricciones de acuerdo con la definición de convergencia que se adopte.

En definitiva, dentro de la literatura empírica sobre el tema existen, básicamente, dos formas diferentes de verificar la existencia de convergencia: mediante el análisis de datos de corte transversal y a través del análisis de series temporales. Este segundo enfoque es el seguido en el presente trabajo. Los tests de corte transversal estudian la correlación entre el nivel inicial de renta per cápita con su tasa de crecimiento, para un conjunto de regiones o países. La existencia de una correlación negativa entre ambas variables se interpreta como evidencia de convergencia, ya que supone que, en media, las regiones o países con un reducido nivel de renta per cápita están experimentando un crecimiento más elevado que aquellos que disponen de un mayor nivel de renta per cápita. Los tests de series temporales consisten en el estudio del comportamiento a largo plazo de las diferencias existentes entre los niveles de renta per cápita (o cualquier otro indicador de convergencia real) entre diferentes regiones o países. En este sentido, existe convergencia si estas diferencias tienen un carácter transitorio, de forma que la predicción a largo plazo de las diferencias converge a cero a medida que aumenta el horizonte de predicción. Por tanto, la convergencia requiere que las diferencias entre los niveles de renta de los distintos países no contengan una raíz unitaria o tendencia determinista a lo largo del tiempo, lo que implicaría a su vez, que los niveles de renta deberían estar cointegrados. Bernard y Durlauf (1995 y 1996) demuestran que los tests de corte transversal están asociados con una noción débil de convergencia en relación con los tests de series temporales. En este sentido, los tests de corte transversal no pueden rechazar la hipótesis de no convergencia cuando los datos son generados por economías con diferentes estados estacionarios en el largo plazo. Por contra, los tests de series temporales no

rechazan espúreamente la no convergencia cuando existen múltiples estados estacionarios³.

Como se ha especificado anteriormente, los innumerables estudios sobre el crecimiento económico, y más concretamente los que analizan las causas de las disparidades existentes entre regiones o países y el posible proceso de convergencia y divergencia, lo han realizado en términos de la renta per cápita. Por este motivo, en este trabajo se estudia la posible existencia de convergencia real en los niveles de empleo, paro y población activa, esto es, se analiza la existencia de una trayectoria común en la evolución de las magnitudes del mercado de trabajo de la Comunidad Autónoma de Andalucía frente al resto de regiones, es decir, frente al total nacional menos Andalucía.

3. METODOLOGÍA UTILIZADA PARA LA CONTRASTACIÓN DE CONVERGENCIA REAL

El método que vamos a utilizar está basado en el propuesto por Bernard y Durlauf (1995), a través del análisis de la existencia de cointegración y del estudio de los vectores de cointegración. Sin embargo, estos tests de convergencia suponen que las condiciones iniciales son similares para las distintas economías, hipótesis que no se cumple, como hemos indicado anteriormente en el análisis del mercado de trabajo de Andalucía y del resto de España, por lo que es necesario redefinir dichos tests con objeto de recoger la existencia de condiciones iniciales diferentes.

Cada una de las variables que vamos a analizar suponemos que se desenvuelven de acuerdo con el siguiente proceso:

$$a(L)X_{i,t} = \mu_i + \epsilon_{i,t} \quad (1)$$

donde i representa a una región o país, $a(L)$, tiene una raíz en el círculo unitario y $\epsilon_{i,t}$ es un proceso estacionario con media cero. Este tipo de formulación permite la existencia en las series tanto de tendencias lineales determinísticas como estocásticas. La interacción de ambos tipos de tendencias entre los ámbitos seleccionados puede reformularse en términos de convergencia y de tendencias comunes.

3. Sin embargo, el enfoque de series temporales requiere que la economía bajo estudio esté lo suficientemente cerca de su estado estacionario. En el caso de que existan importantes efectos transitorios, este análisis sería incorrecto (véase Bernard y Durlauf (1996)).

Bernard y Durlauf (1995) definen la convergencia del siguiente modo. La región A (por ejemplo, Andalucía) converge a E (en este caso resto de España) en términos de una determinada variable, si la predicción de ambas variables a largo plazo, en un tiempo fijo t , es la misma, es decir, si:

$$\lim_{k \rightarrow \infty} E(X_{E,t+k} - X_{A,t+k} | I_t) = 0 \quad (2)$$

Esta definición de convergencia implica que la diferencia entre ambas series disminuye a lo largo del tiempo hasta desaparecer completamente. Por tanto, según esta definición, existiría convergencia si ambas variables estuviesen cointegradas con un vector de cointegración $[1, -1]$. Esta definición implica que $X_{E,t} - X_{A,t}$ es una variable estacionaria con media cero. Pero, tal y como muestran Bernard y Durlauf (1996), si la serie $X_{E,t} - X_{A,t}$ contiene una raíz unitaria o tiene una media diferente de cero, la definición de convergencia dada por (2) no se cumple. En efecto, este método intenta verificar la compatibilidad de la diferencia de las series analizadas $X_{E,t} - X_{A,t}$ con la representación de Wold de la forma:

$$X_{E,t} - X_{A,t} = \mu_{E,A} + \sum_{r=0}^{\infty} \pi_{E,A,r} \epsilon_{E,A,t-r} \quad (3)$$

donde $\mu_{E,A} = 0$ y $\pi_{E,A,r}$ es un escalar. La presencia de una raíz unitaria o de un componente determinístico en dicha diferencia supondría, por tanto, una violación de dicha definición de convergencia, ya que la presencia de estos componentes provocaría que la predicción de las diferencias no convergieran a cero.

Baumol (1986) y Barro (1991) usan una definición de convergencia similar pero menos estricta, que implica que existe convergencia cuando las diferencias se reducen entre un tiempo t y $t+T$ ⁴, es decir si:

$$E(X_{E,t+T} - X_{A,t+T} | I_t) < (X_{E,t} - X_{A,t}) \quad (4)$$

La definición de convergencia en base a las tendencias comunes puede plantearse en términos de que las predicciones a largo plazo entre ambas variables sean proporcionales en un tiempo fijo t , es decir, cuando:

$$\lim_{k \rightarrow \infty} E(X_{E,t+k} - \alpha X_{A,t+k} | I_t) = 0 \quad (5)$$

En este caso, existiría una tendencia común bajo la existencia de cointegración con un vector de cointegración $[1, -\alpha]$. En el caso de que α

4. Bernard y Durlauf (1996) demuestran que la definiciones de convergencia dadas por (7) y (9) son equivalentes para un determinado T .

fuese 1 las definiciones de convergencia y de tendencia común serían equivalentes. Esta definición supone que las condiciones iniciales son iguales para ambas variables. Sin embargo, si las condiciones iniciales son diferentes, las definiciones anteriores no pueden ser aplicadas estrictamente. Por ejemplo, a lo largo de todo el período analizado (1976-1996), la tasa de paro de Andalucía se sitúa permanentemente a un nivel superior a la tasa de paro del resto de España. Por tanto, en nuestro caso, la existencia de una relación de cointegración supondría que la diferencia entre ambas variables es una variable estacionaria pero con respecto a una media no nula. En este caso, un α igual a 1 implicaría que las tasas de crecimiento a largo plazo de las variables son iguales, pero que se mantendrían las diferencias de partida a lo largo del tiempo. De este modo, habría convergencia en términos de tasas de crecimiento pero no existiría convergencia en los niveles por cuanto que existiría un componente determinista. En este caso, la definición de convergencia dada anteriormente implicaría que:

$$\lim_{k \rightarrow \infty} E(X_{E,t+k} - X_{A,t+k} | I_t) = \mu \quad (6)$$

por lo que la presencia de un vector de cointegración [1 -1], sólo implicaría que se mantienen las diferencias iniciales a lo largo del tiempo, siendo dicha diferencia μ . Dado que ambas variables se mueven de forma similar, la relación entre ellas sería positiva, por lo que α sólo puede tomar valores mayores que cero. Claramente, en el caso en que α fuese negativo, el proceso que siguen ambas variables sería divergente.

La definición de convergencia propuesta depende, por tanto, de las condiciones iniciales ya que viene definida en términos de las variables en niveles, y no en términos de sus tasas de crecimiento. Si $X_{E,t} - X_{A,t} > 0$, la media sería positiva y si $X_{E,t} - X_{A,t} < 0$, sería negativa. En base a estos hechos, las medidas de convergencia que se van a aplicar en nuestro caso son las siguientes:

- A. Si existe una tendencia común entre ambas variables y su evolución es creciente en el tiempo, es decir, si las primeras diferencias de las variables son estacionarias con una constante positiva, entonces si $X_{E,t} - X_{A,t} > 0$, la convergencia requeriría la existencia de una tendencia común y que α fuese menor que 1, es decir:

$$E(X_{E,t+T} - \alpha X_{A,t+T} | I_t) < (X_{E,t} - X_{A,t}); \text{ si } \alpha < 1 \quad (7)$$

Por el contrario, si $X_{E,t} - X_{A,t} < 0$, la convergencia requeriría la existencia de una tendencia común y que además α fuese mayor que 1:

$$|E(X_{E,t+T} - \alpha X_{A,t+T} | I_t)| < |(X_{E,t} - X_{A,t})|; \text{ si } \alpha > 1 \quad (8)$$

- B. Si existe una tendencia común entre ambas variables y fuese decreciente, es decir, si las primeras diferencias fuesen estacionarias con una constante negativa, la definición de convergencia sería la contraria a la postulada en A. Si $X_{E,t} - X_{A,t} > 0$, la convergencia requeriría un α mayor que 1 y si $X_{E,t} - X_{A,t} < 0$ la convergencia requeriría un α menor que 1.
- C. En el caso de que no existiese una relación de cointegración entre las variables analizadas, esta metodología no indicaría nada acerca de la convergencia, ya que sólo implicaría la existencia de una relación de equilibrio a largo plazo entre las mismas. En este caso, podría existir tanto convergencia como divergencia, pero estos procesos no serían estables a lo largo del tiempo.

4. EVIDENCIA EMPÍRICA: INTEGRACIÓN, COINTEGRACIÓN Y TENDENCIAS COMUNES

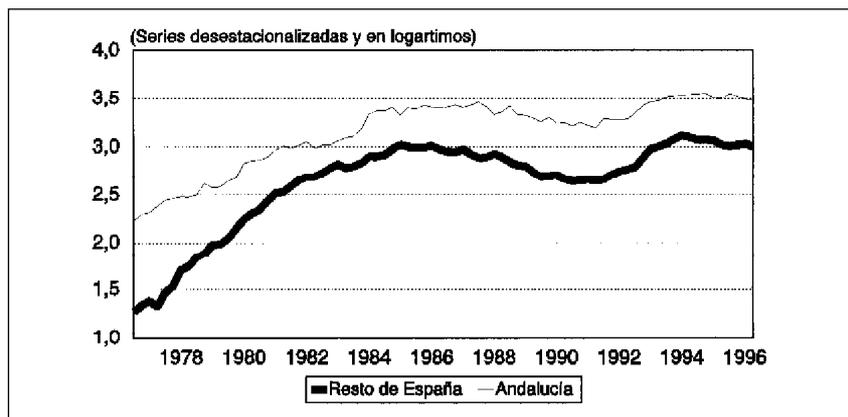
En este epígrafe se aporta evidencia respecto al proceso de convergencia real de Andalucía y el resto de regiones españolas. Los datos utilizados son la tasa de paro, la tasa de ocupación y la población activa y proceden de la Encuesta de Población Activa, del Instituto Nacional de Estadística. Los datos son trimestrales y el período muestral comprende desde el tercer trimestre de 1976 al segundo trimestre de 1996, lo que supone un total de 80 observaciones. Todas las series originales han sido desestacionalizadas con el método X-11 y convertidas a logaritmos, incluidas las tasas de paro y de ocupación, siguiendo a Nelson y Plosser (1982), Barro (1988) y Perron (1990). El período analizado se caracteriza por una tendencia creciente en las tasas de paro y en la población activa (Gráfico 1 y 3) y decreciente en las tasas de ocupación (Gráfico 2), observándose un comportamiento que podría ser cíclico.

4.1. Raíces unitarias

Con objeto de detectar la existencia de tendencias comunes entre las variables que analizamos, en primer lugar tenemos que conocer el grado de integración de las mismas⁵. Las variables (para la región y el

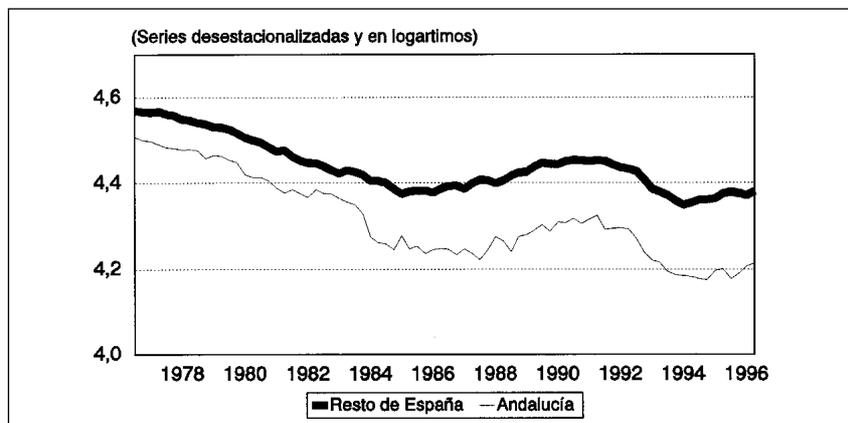
5. Siguiendo a Engle y Granger (1987), se dice que una variable xt es integrada de orden d , notándola $[x_t \sim I(d)]$, si se puede representar como un proceso ARMA (autorregresivo de media móvil) no determinístico, invertible y estacionario después de diferenciar la serie d veces. Por tanto, una serie es integrada de orden cero si es estacionaria en

GRÁFICO 1
TASAS DE PARO



Fuente: E.P.A.

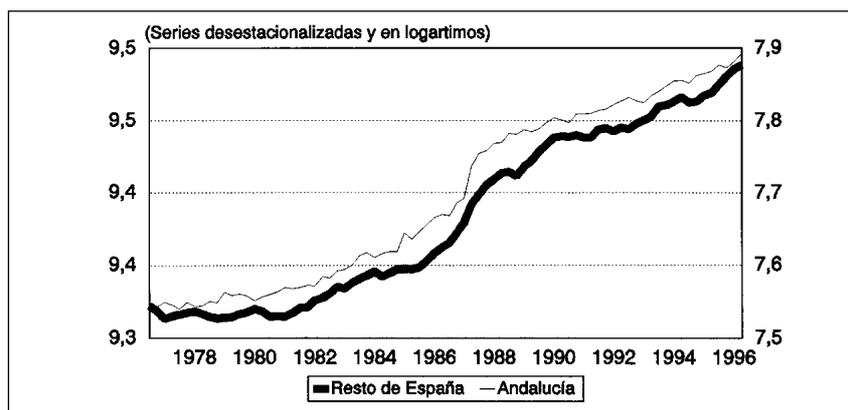
GRÁFICO 2
TASAS DE OCUPACIÓN



Fuente: E.P.A.

niveles, mientras que sería integrada de orden uno si su primera diferencia es estacionaria. La diferencia entre ambos conceptos estriba en que una serie estacionaria tiene una media constante que no varía con el tiempo y su varianza es constante y finita. Por el contrario, en una serie integrada, tanto la media como la varianza no son constantes y dependen del tiempo.

GRÁFICO 3
POBLACION ACTIVA



Fuente: E.P.A.

resto nacional) estarán cointegradas sólo si cada una de ellas es no estacionaria, ya que el análisis se realiza sólo sobre dos variables, las correspondientes a Andalucía y al resto de España. Dado que en el período analizado pueden existir diversas fases cíclicas, los tests se realizarán tanto con respecto a una media como con respecto a una tendencia. Para ello usamos los tests de integración tradicionales: el Dickey-Fuller ampliado y el test de Phillips-Perron, ambos tanto con respecto a una constante como con respecto a una tendencia⁶. La hipótesis nula es que las variables son no estacionarias mientras que la alternativa es que son estacionarias con respecto a una media o con respecto a una tendencia determinística. La Tabla 1 recoge los diferentes test basados en Dickey y Fuller (1979) y Phillips y Perron (1986). En la Tabla 1a se presentan los resultados para las variables de Andalucía y en la Tabla 1b para el total nacional menos Andalucía.

Para la tasa de paro de Andalucía, los tests con respecto a una tendencia indican la existencia de una raíz unitaria por lo que la serie sería $I(1)$, mientras que los tests con respecto a una media rechazan la existencia de una raíz unitaria. Sin embargo, dada la evolución observada de las series analizadas, es más probable la existencia de una tendencia, a pe-

6. Aunque los contrastes de integración son de sobra conocidos, para el lector no familiarizado con estos instrumentos se presenta en el apéndice I un breve resumen de la forma de dichos contrastes.

sar de que en el período analizado existen diversas fases cíclicas. En cuanto a la tasa de paro de España menos Andalucía, excepto el test de Phillips-Perron con respecto a una constante, todos los demás indican que se trata de una serie $I(1)$. Con respecto a las tasas de ocupación, los tests no ofrecen divergencias, mostrando en todos los casos la existencia de una raíz unitaria. El orden de integración de la población activa tampoco ofrece dudas, indicando todos los tests que se trata de una serie $I(1)$. Estos resultados son consistentes con los encontrados en numerosos estudios empíricos (Nelson y Plosser (1982), Perron (1990), entre otros), en los que las series del mercado de trabajo son procesos $I(1)$. Con respecto a las series en primeras diferencias, los anteriores tests no ofrecen duda alguna para las variables referentes a Andalucía, rechazando la hipótesis de integración en todos los casos, lo que confirma el comportamiento $I(1)$ de dichas series. En el caso de España menos Andalucía, los tests Dickey-Fuller ampliados plantean algunas dudas con respecto a la tasa de paro y la tasa de ocupación, pudiéndose deducir que ambas series podrían ser $I(2)$. Sin embargo, los tests de Phillips-Perron rechazan dicha hipótesis indicando la existencia de una raíz unitaria. En base a estos resultados podemos suponer que todas las series analizadas presentan una raíz unitaria.

TABLA 1A
TESTS DE INTEGRACIÓN. ANDALUCÍA

	k	ADFK	k	ADFT	PPK	PPT
Tasa de paro	0	-3,22	0	-1,83(*)	-3,18	-1,88(*)
Tasa de paro	2	-8,55	0	-9,16	-9,20	-9,43
Tasa ocupación	0	-1,75(*)	0	-1,47(*)	-1,65(*)	-1,75(*)
Tasa ocupación	0	-8,65	0	9,15	-9,15	-9,21
Población activa	0	0,71(*)	0	-2,27(*)	0,41(*)	-2,47(*)
Población activa	0	-9,16	0	-9,22	-9,59	-9,63

Nota: ADFK y ADFT son los tests de Dickey-Fuller ampliados con respecto a una constante y a una tendencia, respectivamente. PPK y PPT son los tests de Phillips-Perron con respecto a una constante y una tendencia, respectivamente. La hipótesis nula es que la variable es estacionaria en sus primeras diferencias. Los valores críticos al 5 por ciento son -2,89 para los casos de una constante y -3,46 para los casos de una tendencia. El número de retardos de los tests, k, de Dickey-Fuller ampliados, ha sido elegido siguiendo el criterio de Hannan y Quinn (1979). Los valores con () aceptan la hipótesis nula al 5 por ciento.*

TABLA 1B
TESTS DE INTEGRACIÓN. RESTO DE ESPAÑA

	k	ADFK	k	ADFT	PPK	PPT
Tasa de paro	6	-1,91(*)	9	-3,15(*)	-3,64	-2,28(*)
Tasa de paro	5	-2,75	5	-2,58	-6,86	-7,64
Tasa ocupación	5	-2,41(*)	5	-2,55(*)	-1,94(*)	-1,72(*)
Tasa ocupación	5	-1,39	4	-2,54	-6,06	-6,16
Población activa	1	0,92(*)	1	-2,40(*)	0,98(*)	-2,79(*)
Población activa	1	-4,44	0	-5,95	-6,10	-6,31

Nota: ADFK y ADFT son los tests de Dickey-Fuller ampliados con respecto a una constante y una tendencia, respectivamente. PPK y PPT son los tests de Phillips-Perron con respecto a una constante y a una tendencia, respectivamente. La hipótesis nula es que la variable es estacionaria en sus primeras diferencias. Los valores críticos al 5 por ciento son -2,89 para los casos de una constante y -3,46 para los casos de una tendencia. El número de retardos, k , de los tests de Dickey-Fuller ampliados, ha sido elegido siguiendo el criterio de Hannan y Quinn (1979). Los valores con (*) aceptan la hipótesis nula al 5 por ciento.

4.2. Tendencias comunes y convergencia

Dado que todas las series analizadas anteriormente presentan un comportamiento $I(1)$, procedemos a estudiar si existe alguna relación de equilibrio a largo plazo entre ellas. Con objeto de analizar la existencia de cointegración, usamos la técnica de cointegración multivariante desarrollada por Johansen (1988) y Johansen y Juselius (1990)⁷. Aunque esta técnica es suficientemente conocida expondremos brevemente sus fundamentos. Si una combinación lineal de variables $I(1)$ es estacionaria, entonces se dice que dichas variables están cointegradas y la combinación lineal estacionaria puede ser interpretada como un error de equilibrio. Supongamos que X_t es un vector $N \times 1$ de variables $I(1)$ y asumimos que este vector tiene una representación VAR (vector autorregresivo) de orden k con errores gaussianos ϵ_t :

$$X_t = \mu + \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \dots + \Pi_k X_{t-k} + \epsilon_t \quad (9)$$

7. El concepto de cointegración fue desarrollado inicialmente por Engle y Granger (1987). Se dice que las variables de un vector X_t están cointegradas de orden (d, b) :

$$X_t \sim CI(d, b)$$

si dichas variables son integradas de orden d , $I(d)$, y existe un vector α , denominado vector de cointegración, distinto de cero tal que:

$$Z_t = \alpha' X_t \sim I(d-b), \quad b > 0$$

El equilibrio estático a largo plazo que corresponde a esta ecuación es:

$$\Pi X = 0 \quad (10)$$

donde la matriz de coeficientes a largo plazo, Π , se define como:

$$I - \Pi_1 - \Pi_2 - \dots - \Pi_k = \Pi \quad (11)$$

Π es una matriz $N \times N$ cuyo rango determina el número de vectores de cointegración distintos que existen entre las variables que forman X . Definimos dos matrices $N \times r$, α y β , tal que:

$$\Pi = \alpha \beta' \quad (12)$$

Las filas de β' de los r distintos vectores de cointegración son:

$$\beta_i' X_t \sim I(0) \quad (13)$$

donde β_i' es la fila i de β' . El rango de la matriz de largo plazo determina el número de vectores de cointegración. Pueden existir tres posibilidades. En primer lugar, Π puede ser de rango pleno. En este caso, para asumir la estacionaridad del término de error se requiere que los niveles de X_t sean estacionarios, implicando la ausencia de cualquier tendencia estocástica en los datos, en contra de nuestra especificación inicial $I(1)$. En segundo lugar, Π puede ser de rango cero, por ejemplo, puede ser una matriz nula y por tanto no existir ninguna relación de cointegración entre los elementos de X_t . Finalmente, y en tercer lugar, Π puede tener un rango intermedio r , ($0 < r < n$). Si Π tienen rango r , entonces existen r relaciones de cointegración entre los elementos de X_t , o de forma equivalente, $n-r$ tendencias estocásticas comunes.

Johansen (1988) demuestra que la función de máxima verosimilitud de este problema es proporcional a:

$$L = (\prod_{i=1}^N (1 - \lambda_i))^{-\frac{t}{2}} \quad (14)$$

donde $\lambda_1, \lambda_2, \dots, \lambda_N$, muestran las N correlaciones canónicas al cuadrado entre las series X_{t-k} y X_t (véase Johansen (1988)). El número de vectores de cointegración se puede demostrar es igual al número de λ no nulas. Por tanto, el estadístico de máxima verosimilitud para la hipótesis nula de la existencia de, al menos, r vectores de cointegración, el *estadístico Traza*, es:

$$Traza = t \sum_{i=r+1}^N \ln(1 - \lambda_i) \quad (15)$$

Por otro lado, el estadístico de máxima verosimilitud de la existencia de al menos r vectores de cointegración, contra la alternativa de $r+1$ vectores de cointegración, el *estadístico del valor propio máximo*, $\lambda_{m\acute{a}x}$, es:

$$\lambda_{m\acute{a}x} = T \ln(1 - \lambda_{r+1}) \quad (16)$$

Con objeto de realizar el análisis de la existencia de tendencias comunes aplicamos los anteriores test a un sistema VAR, para cada una de las variables, de la siguiente forma:

$$X_t = [\text{España-Andalucía}, \text{Andalucía}] \quad (17)$$

Los resultados del análisis de cointegración se presentan en la tabla 2. Como podemos observar existe una relación de cointegración en cada variable regional con respecto a su correspondiente nacional, excluyendo dicha variable regional. Tanto el estadístico *Traza* como el $\lambda_{m\acute{a}x}$ rechazan la hipótesis nula de que no existe ninguna relación de cointegración. Con respecto al valor del vector de cointegración, observamos que en los tres casos es inferior a la unidad⁸. En el caso de la tasa de paro el parámetro α tiene un valor de 0,72, implicando que la relación de equilibrio a largo plazo entre la tasa de paro de Andalucía y la del resto de España presenta una evolución divergente, es decir, la tasa de paro andaluza, lejos de aproximarse a la del resto de España, ha ido aumentando en una mayor proporción. La convergencia, como se expuso anteriormente en la ecuación (9), requeriría que $\alpha > 1$. En el caso de la tasa de ocupación los resultados son similares, mostrando también un proceso divergente con un valor de parámetro α , aún más bajo, de 0,47, lo que verifica una mayor divergencia. Por último, para la población activa, el parámetro es de 0,50 indicando que la tasa de crecimiento de esta variable en Andalucía es superior a la del resto de España, por lo que esta relación de equilibrio a largo plazo indicaría un proceso de convergencia, que implicaría la reducción progresiva de las diferencias existentes. (Téngase en cuenta que dada la tendencia de esta serie, contraria a las anteriores, la convergencia en este supuesto requiere un valor de α inferior a la unidad).

8. Se ha realizado un test sobre la hipótesis nula de que α es la unidad. Si α es 1, no existiría convergencia ni divergencia, sino que las diferencias existentes se mantendrían, como hemos visto anteriormente, ya que sus tasas de crecimiento serían iguales. En los tres casos se rechaza la hipótesis nula, a través de una $\chi^2(1)$ al 5 por ciento, por lo que los parámetros obtenidos son significativamente inferiores a la unidad.

TABLA 2
ANÁLISIS DE COINTEGRACIÓN

		VAR(p)	Traza	$\lambda_{\text{máx}}$	α
Tasa de paro	r 1	6	6,91	6,91	0,72
	r 0		32,98(*)	26,06(*)	
Tasa de ocupación	r 1	5	3,20	3,20	0,47
	r 0		21,86(*)	18,65(*)	
Población activa	r 1	7	1,47	1,47	0,50
	r 0		23,47(*)	22,05(*)	

Nota: los valores con () representan la existencia de un vector de cointegración. Los valores críticos al 5 por ciento son 8,8 para r 1 y 17,84 para r 0 para el estadístico Traza y de 8,8 para r 1 y 17,59 para r 0 para el estadístico $\lambda_{\text{máx}}$. El número de retardos del VAR ha sido elegido siguiendo el criterio de Hannan y Quinn (1979).*

4. CONCLUSIONES

En este trabajo se ha analizado la convergencia real entre Andalucía y el resto de regiones españolas, centrándonos en la tasa de desempleo, tasa de ocupación y población activa. Intentamos aportar evidencia, utilizando el análisis de cointegración, sobre la existencia de un proceso convergente o divergente entre las citadas variables a nivel regional y nacional menos Andalucía. En concreto, se intenta detectar la existencia de tendencias comunes de equilibrio a largo plazo entre algunas variables de los mercados de trabajo presentando, al mismo tiempo, una definición de convergencia que viene determinada por dicha relación de equilibrio a largo plazo y su relación con la existencia de tendencias comunes.

Los resultados obtenidos son muy significativos. Mientras que se produce un proceso de convergencia entre la población activa, se produce una divergencia en términos de la tasa de paro y de la tasa de ocupación. Este proceso divergente en términos de la tasa de paro y tasa de ocupación puede explicarse, en parte, por el significativo proceso de convergencia que experimenta la población activa. Pero subyacen tras él, otras causas preocupantes y que podrían sugerir determinadas actuaciones públicas para intentar eliminarlas.

En términos de la población activa, se observa como en Andalucía, las tasas de crecimiento son muy superiores a las que se registran a nivel nacional. Esto es debido principalmente a la menor tasa de actividad que

registra Andalucía, y a la creciente incorporación de la mujer al mercado de trabajo. Por tanto, en función de esta variable, el comportamiento del mercado de trabajo en Andalucía tiende hacia los niveles que se registran en el total nacional. Sin embargo, este mayor crecimiento de la población activa, implica que las tasas de creación de empleo tienen que ser muy superiores en Andalucía a las del resto de España, si no se quiere aumentar la tasa de paro, lo cual implicaría que las tasa de crecimiento de la economía andaluza tendría que ser superior a la del resto de regiones españolas. Incluso, con un mayor incremento del empleo en Andalucía, puede suceder que aumente en mayor medida la tasa de paro, como consecuencia de los significativos incrementos que experimenta la población activa. Esta es una de las explicaciones posibles que subyacen al resultado obtenido de la existencia de divergencia en las tasas de paro.

No obstante, también encontramos un proceso divergente en cuanto a las tasas de ocupación. Las tasas de ocupación, tanto en Andalucía como en el resto de España han ido disminuyendo, excepto a finales de los ochenta y principios de los noventa, en los que aumentó. Sin embargo, la disminución en Andalucía ha sido muy superior a la que se ha registrado a nivel nacional, resultado que deriva en la existencia de un proceso divergente entre ambas variables. Este resultado supone que el mercado de trabajo andaluz no ha sido capaz de absorber la creciente población activa, lo que explica el proceso divergente entre las tasas de paro.

La evidencia que presentamos es consistente con la recientemente aportada por Villaverde (1996), que utiliza otra técnica de análisis (el método *shift-share*) para analizar la convergencia regional en la tasa de actividad y su incidencia sobre la dinámica del empleo y que concluye: "... en los próximos años no parece fácil que se produzca un acortamiento significativo en el gap existente entre las tasas de empleo en el ámbito regional y nacional".

Esta desigualdad interregional o divergencia entre Andalucía y el resto de España podría venir motivada porque el ciclo económico afecta a las regiones de forma distinta, de acuerdo con su especialización sectorial. La escasa capacidad de generación de empleo de Andalucía puede venir también motivada por insuficiencias de capital humano (menor nivel de educación y cualificación de los trabajadores) y tecnológico de las empresas (relación entre la proporción entre I+D y PIB, y la tasa de paro entre regiones). La divergencia en las tasas de paro mantenidas de forma sostenida a lo largo del período analizado, reflejan la existencia de desventajas estructurales en Andalucía, es decir, deficiencias en sus dota-

ciones de ciertos factores productivos. En este sentido, se requeriría un aumento de la inversión en infraestructuras, incentivos a la actividad privada y mejora en los niveles de formación ocupacional, como armas para reducir la elevada tasa de desempleo andaluza en relación con el resto de España. Nuestro análisis muestra que el camino recorrido en los últimos 20 años no ha sido precisamente el adecuado para alcanzar ni la convergencia real en la tasa de desempleo ni la reducción de la misma.

APÉNDICE I CONTRASTES DE RAÍCES UNITARIAS

Los contrastes de raíces unitarias utilizados han sido los de Dickey y Fuller (1979) y los de Phillips y Perron (1988).

En el primer caso se ha utilizado el contraste de Dickey-Fuller ampliado. Para el caso de la existencia de una constante, la regresión estimada ha sido la siguiente:

$$x_t = \mu + \gamma x_{t-1} + \sum_{i=1}^k \beta_i x_{t-i} + u_t \quad (\text{A.I.18})$$

donde x_t es la variable y μ es la constante.

Para el caso de una constante y una tendencia, la regresión estimada ha sido:

$$x_t = \mu + \delta t + \gamma x_{t-1} + \sum_{i=1}^k \beta_i x_{t-i} + u_t \quad (\text{A.I.19})$$

donde t es la tendencia. La hipótesis nula en ambos casos es que $\gamma=0$, a través del estadístico- t .

Los contrastes de Phillips y Perron parten de las siguientes ecuaciones:

$$x_t = \mu + \gamma x_{t-1} + u_t \quad (\text{A.I.20})$$

$$x_t = \mu + \delta t + \gamma x_{t-1} + u_t \quad (\text{A.I.21})$$

definiendo el estadístico como:

$$Z(T_\mu) = \left(\frac{\sum_{t=1}^T x_t}{T} \right) \tau_\mu^{-1/2} \left(\frac{\sum_{t=1}^T x_t^2}{T} - \left(\frac{\sum_{t=1}^T x_t}{T} \right)^2 \right)^{-1/2} T \left[\frac{\sum_{t=2}^T (x_t - x_{t-1})^2}{T} \right]^{-1/2} \quad (\text{A.I.22})$$

donde T es el período muestral, m es el número de autocorrelaciones estimadas, τ_{t-1} es la varianza muestral calculada como:

$$x\bar{\nabla}_{t-1} = (T-1)^{-1} \sum_{l=2}^T X_{t-1} \quad (\text{A.I.23})$$

τ_u es el estadístico- t asociado con la estimación de los parámetros de las ecuaciones (3) y (4), (con la hipótesis nula $\gamma=0$) y s_{Tm}^2 es la varianza a largo plazo estimada como:

$$s_{Tm}^2 = T^{-1} \sum_{t=1}^T u_t^2 + 2T^{-1} \sum_{s=1}^l w_{sm} \sum_{t=s+1}^T u_t u_{t-s} \quad (\text{A.I.24})$$

donde \hat{u}_t son los residuos de las regresiones (3) y (4) y donde:

$$w_{sm} = (1-s(m+1)), \quad s=1, \dots, m \quad (\text{A.I.25})$$

APÉNDICE II EL CRITERIO DE HANNAN-QUINN

El criterio de Hannan y Quinn (1979) para la obtención de los valores óptimos de los retardos consiste en minimizar la siguiente expresión:

$$\frac{1ns + 2(k+2)1nT}{T} \quad (\text{A.II.1})$$

donde k es el número de retardos, T es el período muestral y s es la

varianza.

BIBLIOGRAFÍA

- ARTIS, M., LÓPEZ, E., SANZO, A. y SURINACH, J. (1995): *Análisis económico regional. Nociones básicas de la teoría de la cointegración*. Barcelona: Antoni Bosh.
- BARRO, R. (1988): "The persistence of unemployment", *American Economic Review*, vol. 78, págs. 32-37.
- BARRO, R. (1991): "Economic growth in a cross section of countries", *Quarterly Journal of Economics*, vol. 106, págs. 407-443.
- BARRO, R. y SALA-I-MARTIN, X. (1992): "Convergence", *Journal of Political Economy*, vol. 100, págs. 223-251.
- BAUMOL, W. (1986): "Productivity growth, convergence, and welfare: what the long-run data show", *American Economic Review*, vol. 76, págs. 1072-1085.
- BERNARD, A. y DURLAUF, S. (1995): "Convergence in international output", *Journal of Applied Econometrics*, vol. 10, págs. 97-108.
- BERNARD, A. y DURLAUF, S. (1996): "Interpreting tests of the convergence hypothesis", *Journal of Econometrics*, vol. 71, págs. 161-173.
- CAMARERO, M., ESTEVE, V. y TAMARIT, C. (1995): "Medición de la convergencia en tasas de inflación: España versus Alemania y el SME", *Papeles de Economía Española*, nº 63, págs. 62-76.
- DICKEY, D. y FULLER, W. (1979): "Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, vol. 74, págs. 429-431.
- ENGLE, R. y GRANGER, C. (1987): "Cointegration and error correction: representation, estimation and testing", *Econometrica*, vol. 55, págs. 251-276.
- GARCÍA, B. y RAYMOND, J.L. (1994): "Las disparidades en el PIB per cápita entre comunidades autónomas y la hipótesis de convergencia", *Papeles de Economía Española*, nº 59, págs. 37-58.
- GARCÍA, B. y RAYMOND, J.L. (1996): "Distribución regional de la renta y movimientos migratorios", *Papeles de Economía Española*, nº 67, págs. 185-201.
- GARCÍA, B.; RAYMOND, J.L. y VILLAVERDE, J. (1995): "La convergencia de las provincias españolas", *Papeles de Economía Española*, nº 64, págs. 38-53.
- HANNAN, E. y QUINN, B. (1979): "The determination of the order of an

- autoregression", *Journal of the Royal Statistic Association, Serie B*, vol. 41, págs. 190-195.
- JOHANSEN, S. (1988): "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, vol. 12, págs. 231-254.
- JOHANSEN, S. y JUSELIUS, K. (1990): "Hypothesis testing for cointegration vectors for Gaussian vector autoregressive processes with constant and seasonal dummies", *Oxford Bulletin of Economics and Statistics*, vol. 52, págs. 169-210.
- MARCET, A. (1994): "Los pobres siguen siendo pobres: convergencia entre regiones y países, un análisis bayesiano de panel de datos", en J. Esteban y X. Vives, (eds.), *Crecimiento y convergencia regional en España*, Instituto de Análisis Económico. Barcelona.
- MAS, M., MAUDOS, J., PÉREZ, F. y URIEL, E. (1994): "Disparidades regionales y convergencia en las Comunidades Autónomas", *Revista de Economía Aplicada*, vol. 2, págs. 129-148.
- NELSON, C. y PLOSSER, C. (1982): "Trends and random walks in macroeconomic time series", *Journal of Monetary Economics*, vol. 10, págs. 139-162.
- PERRON, P. (1990): "The great crash, the oil price shock and the unit root hypothesis", *Econometrica*, vol. 57, págs. 1361-1401.
- PHILLIPS, P. y PERRON, P. (1986): "Testing for a unit root in a time series regression", *Biometrika*, vol. 75, págs. 335-346.
- QUAH, D. (1996): "Empirics for economic growth and convergence", *European Economic Review*, vol. 40, págs. 1353-1375.
- ROGOFF, K. (1985): "Can exchange rate predictability be achieved without monetary convergence? Evidence from the EMS", *European Economic Review*, vol. 28, págs. 93-115.
- SALA-I-MARTIN, X. (1994): "La riqueza de las regiones. Evidencia y teorías sobre el crecimiento regional y convergencia sobre el crecimiento regional y convergencia", *Moneda y Crédito*, nº 198, págs. 13-80.
- SALA-I-MARTIN, X. (1996): "Regional cohesion: evidence and theories of regional growth and convergence", *European Economic Review*, vol. 40, págs. 1325-1352.
- VILLAVARDE, J. (1996): "Impacto de la producción y de la productividad sobre el empleo", *Papeles de Economía Española*, nº. 67, págs. 202-221.