



Universidad de Valladolid

Un análisis de la convergencia de las tasas de paro regionales en España: 1977-2000

Alfonso Moral de Blas

Tesis de Doctorado

Facultad de Ciencias Económicas y Empresariales

Director: Dr. D. José Miguel Sánchez Molinero

2003

UNIVERSIDAD DE VALLADOLID

FACULTAD DE CIENCIAS ECONÓMICAS Y
EMPRESARIALES



UN ANÁLISIS DE LA CONVERGENCIA DE
LAS TASAS DE PARO REGIONALES EN
ESPAÑA

Alfonso Moral de Blas

Proyecto de Tesis Doctoral dirigido por el Profesor Doctor:

D. José Miguel Sánchez Molinero

VALLADOLID, JUNIO DE 2003

A Susana



No quisiera comenzar esta tesis sin hacer una mención especial a todos aquellos sin cuya inestimable ayuda hubiera sido imposible su finalización.

Sin duda la magistral dirección del profesor Sánchez Molinero ha sido la base fundamental de este trabajo. En su total dedicación y apoyo he encontrado los alicientes necesarios para su elaboración y solo mi impericia en la aplicación de sus enseñanzas es la causa de las deficiencias que pudieran aparecer en ella.

Esta tesis también es deudora de las aportaciones hechas por los profesores Fernández-Jardón y Pérez Domínguez. Su inestimable ayuda en los aspectos econométricos, y sus adecuadas correcciones, no han hecho más que enriquecer el contenido de la misma.

No quisiera dejar pasar esta ocasión sin mostrar un especial agradecimiento a Jesús María Alonso y Angel Luis Martín. Su colaboración, más allá del ámbito académico e intelectual, ha resultado determinante y ha sido un soporte adecuado para superar todos los obstáculos presentados.

Finalmente extendiendo mi gratitud al Departamento de Fundamentos del Análisis Económico por su ayuda desinteresada y por darme la oportunidad de elaborar esta tesis.

ÍNDICE

CAPÍTULO 1

INTRODUCCIÓN.....	1
1.1.- ANTECEDENTES.....	2
1.2.- OBJETIVO PRINCIPAL Y ESQUEMA GENERAL DEL TRABAJO.....	3
1.3.- ORGANIZACIÓN DEL TRABAJO.....	4

CAPÍTULO 2

DIFERENCIAS DEL DESEMPLEO ENTRE REGIONES: UN ANÁLISIS DESCRIPTIVO..... 7

2.1.- INTRODUCCIÓN	8
2.2.- ANÁLISIS ESTÁTICO.....	9
2.3.- ESTRUCTURA DE LA POBLACIÓN Y PARO.....	14
2.4.- OTROS FACTORES EXPLICATIVOS.....	31
2.4.1.- <i>Composición sectorial y paro</i>	31
2.4.2.- <i>Presión sindical y paro</i>	38
2.5.- ANÁLISIS DE LOS COMPONENTES DEL DESEMPLEO REGIONAL A TRAVÉS DE LA TÉCNICA “SHIFT-SHARE”.....	43
2.6.- ANÁLISIS DINÁMICO.....	51

CAPÍTULO 3

BASES TEÓRICAS PARA EL ANÁLISIS DE CONVERGENCIA..... 56

3.1.- INTRODUCCIÓN.....	57
-------------------------	----

3.2.- ESCENARIOS PARA EL ANÁLISIS DE CONVERGENCIA.....	58
3.2.1.- <i>Criterios básicos</i>	58
3.2.2.- <i>Los cuatro escenarios</i>	62
3.2.2.a <i>Primer escenario</i>	63
3.2.2.b <i>Segundo escenario</i>	65
3.2.2.c <i>Tercer escenario</i>	68
3.2.2.d <i>Cuarto escenario</i>	69
3.2.2.e <i>Conclusiones</i>	70
3.3.- ¿CONVERGENCIA DE LAS TASAS DE PARO REGIONALES?.....	71

CAPÍTULO 4

MÉTODOS PARA EL ANÁLISIS DE CONVERGENCIA EN TÉRMINOS

EMPÍRICOS..... 74

4.1.- INTRODUCCIÓN.....	75
4.2.- LA HIPÓTESIS DE β -CONVERGENCIA ABSOLUTA.....	75
4.3.- LA HIPÓTESIS DE β -CONVERGENCIA CONDICIONADA.....	75
4.4.- LA HIPÓTESIS DE σ -CONVERGENCIA.....	86
4.5.- MÉTODOS DE COINTEGRACIÓN.....	91
4.6.- MÉTODO DE DESCOMPOSICIÓN DE LA VARIANZA.....	96
4.7.- RECAPITULACIONES Y OBSERVACIONES FINALES.....	98

CAPÍTULO 5

ANÁLISIS EMPÍRICO..... 100

5.1.- INTRODUCCIÓN.....	101
5.2.- FUENTES ESTADÍSTICAS.....	102

5.3.- ANÁLISIS DE COINTEGRACIÓN.....	105
5.4.- ANÁLISIS DE DESCOMPOSICIÓN DE LA VARIANZA.....	115

CAPÍTULO 6

RESUMEN Y CONCLUSIONES	137
-------------------------------------	------------

APÉNDICE 1

TEORÍA DE LA COINTEGRACIÓN.....	143
--	------------

APÉNDICE 2

TEST DE RAÍCES UNITARIAS.....	147
--------------------------------------	------------

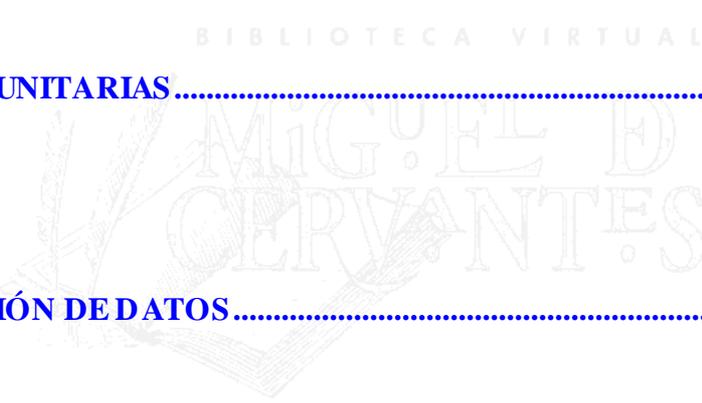
APÉNDICE 3

HOMOGENEIZACIÓN DE DATOS.....	161
--------------------------------------	------------

APÉNDICE 4

COEFICIENTE DE CORRELACIÓN ENTRE VARIABLES.....	168
--	------------

BIBLIOGRAFÍA.....	174
--------------------------	------------



BIBLIOTECA VIRTUAL

**CAPÍTULO 1:
INTRODUCCIÓN**



1.1.- Antecedentes

Durante las últimas décadas del siglo XX ha habido una auténtica proliferación de los estudios sobre convergencia económica; y la razón de esto habría que buscarla en los recientes procesos de globalización de la economía y en la formación de entidades supranacionales. Esta tesis podría encuadrarse dentro del panorama general de estos estudios sobre convergencia.

Centrándonos en el caso español, este proceso de globalización se ha visto concretado en dos hechos clave, como son nuestra integración primero en la Comunidad Económica Europea y después en la Unión Monetaria Europea. Este proceso de integración vino precedido del cumplimiento de una serie de criterios referentes a la convergencia “nominal” entre los miembros de la Unión Europea, criterios que habría que considerar como requisitos previos para la consecución de una eventual convergencia “real” entre las diversas economías nacionales.

Pero la cuestión de la convergencia real no tiene por qué circunscribirse al ámbito internacional. También tiene sentido preguntarse si este progresivo acercamiento entre las economías de los distintos países se ha visto correspondido por un mayor grado de integración a escala nacional. Y esto es lo que hacemos en esta tesis: nos preguntamos si los procesos recientes de globalización de la economía han dado lugar a un mayor grado de homogeneidad (si ha habido o no ha habido convergencia real) entre las regiones españolas.

Una cuestión básica que hay que resolver antes de enfrentarse a un estudio de este tipo es la selección de algún índice o medida de convergencia. Naturalmente, la convergencia es un concepto multidimensional y cualquier índice que utilicemos tendrá que ser parcial. Por lo general, el grueso de la literatura dedicada al análisis de convergencia real se ha centrado en el estudio de las rentas per capita, olvidándose de otras variables como podrían ser los salarios o las tasas de desempleo. Este hecho, junto con la peculiar configuración institucional del mercado de trabajo español, nos ha llevado a centrarnos en un aspecto particular de la convergencia entre regiones: la convergencia de las tasas de paro regionales.

Debemos tener en cuenta que esas dos últimas décadas del siglo XX han supuesto también, dentro del mercado de trabajo español, un incremento espectacular de las tasas de desempleo, muy superior al sufrido por el resto de socios comunitarios. Este incremento ha ido acompañado a su vez de un aumento en las diferencias interregionales en lo referente a tasas de desempleo, como tendremos ocasión de precisar en el segundo capítulo de esta tesis. Por eso, parece razonable preguntarse si esas diferencias crecientes pueden reconciliarse con la idea generalizada de que las economías tienden a integrarse cada vez más.

1.2.- Objetivo principal y esquema general del trabajo

Como hemos apuntado anteriormente, esta tesis pretende aportar algunos elementos nuevos al análisis de la convergencia real entre las regiones españolas. Dado que se han hecho muchos estudios sobre la convergencia en términos de renta per capita, en esta tesis hemos optado por analizar la convergencia de las tasas de paro, que

es una cuestión muy importante en el caso español, y que ha recibido mucha menos atención dentro de la literatura económica. En concreto, en esta tesis trataremos de ver si tiene algún sentido hablar de convergencia entre regiones en términos de tasas de paro, a pesar de que, a simple vista, la evidencia parece apuntar en sentido contrario.

Nuestro estudio comenzará con un análisis básicamente descriptivo del problema de la convergencia de las tasas de paro regionales; a continuación, introduciremos las consideraciones teóricas necesarias para el desarrollo del análisis econométrico posterior; y finalmente presentaremos lo que constituye la aportación principal de esta tesis: un análisis econométrico de la convergencia de las tasas de paro de las regiones españolas a través de dos técnicas alternativas, a saber, las técnicas de cointegración y el método de descomposición de la varianza.

1.3.- Organización del trabajo

El capítulo segundo de esta tesis está dedicado al análisis descriptivo de las tasas de desempleo regionales, la evolución de las diferencias entre unas regiones y otras, y los posibles factores explicativos de todo esto. Este análisis descriptivo se complementa con un primer intento de explicación a través de la técnica *shift-share* cuyo contenido específico será descrito en su momento.

El capítulo tercero desarrolla los criterios teóricos básicos a partir de los cuales sería de esperar la existencia de algún tipo de proceso de convergencia entre las tasas de paro regionales. A partir de unos supuestos básicos acerca de la homogeneidad o heterogeneidad de la mano de obra y el grado de competitividad de los mercados

laborales, se construyen cuatro escenarios alternativos y se llega a ciertas conclusiones en relación con las diferencias regionales de desempleo y la evolución que cabe esperar que sigan dichas diferencias.

En el capítulo cuarto hacemos una revisión detallada de los principales métodos utilizados en el análisis de convergencia, tanto los que emplean series temporales como los basados en datos de corte transversal. Se describen los contextos en que dichos métodos han sido aplicados y se analizan las ventajas y los inconvenientes de cada uno de ellos. Terminamos este capítulo intentando justificar el porqué de los métodos empleados en esta tesis.

El capítulo quinto se centra en el análisis econométrico que, como ya dijimos antes, constituye la aportación principal de esta tesis. Se analiza en primer lugar el proceso de convergencia de las tasas de paro regionales utilizando un enfoque de cointegración. Dicho enfoque trata, básicamente, de encontrar relaciones de largo plazo entre las tasas de paro de cada región y el agregado nacional. El segundo enfoque utilizado consiste en la descomposición de la varianza de las tasas de paro regionales, con el fin de calibrar la importancia de los factores específicos de cada región en la explicación de las diferencias regionales.

El objetivo último de ambos enfoques consiste en mostrar que la convergencia entre regiones se hace tanto más verosímil (las diferencias regionales de desempleo se acortan) a medida que vamos acotando más y más el grupo de población y nos vamos centrando en grupos más y más homogéneos.

Por último, el capítulo sexto presenta un resumen de todo el trabajo donde se incluyen las principales conclusiones de esta tesis.



BIBLIOTECA VIRTUAL

CAPÍTULO 2:

**DIFERENCIAS DE DESEMPLEO ENTRE REGIONES: UN
ANÁLISIS DESCRIPTIVO**

2.1.- Introducción

Antes de adentrarnos en el estudio pormenorizado de la convergencia entre las tasas de paro regionales, parece lógico realizar una primera exploración, desde un punto de vista meramente descriptivo, sobre la situación que presenta el mercado de trabajo español a nivel regional. Este primer acercamiento nos permitirá, por un lado, tener una visión preliminar sobre la magnitud de las diferencias entre las tasas de paro regionales, así como sobre su evolución, y por otro, nos proporcionará las primeras explicaciones intuitivas sobre las causas de dichas diferencias.

Este análisis descriptivo lo realizaremos desde una doble perspectiva, estática y dinámica. Desde el punto de vista estático, veremos cuál es la situación de las tasas de paro regionales en el año 2000, haciendo una comparación simple con la situación de 1977. Desde el punto de vista dinámico, trataremos de dar una visión puramente intuitiva de cuál ha sido la evolución temporal de las diferencias regionales de desempleo a lo largo de los años del periodo 1977-2000.

El resto del capítulo se organiza de la siguiente manera. En el epígrafe 2 abordamos la situación del desempleo en el año 2000 y analizamos su distribución por regiones comparándola con lo que ocurría en el año 1977. En los epígrafes 3 y 4 hacemos un análisis similar, también de carácter estático, para las distintas variables que pueden influir sobre el desempleo, centrándonos en la estructura de la población, la composición sectorial y algunos aspectos institucionales potencialmente relevantes. El epígrafe 5 complementa los anteriores a través de un análisis *shift-share*, que nos proporciona una primera impresión acerca de la importancia relativa de los factores

regionales en el fenómeno del desempleo. Finalmente, en el epígrafe 6, se analiza la evolución del coeficiente de variación de las tasas regionales de desempleo con el fin de obtener una primera aproximación al problema de la convergencia desde el punto de vista dinámico.

2.2.- Análisis estático

Antes de comenzar con el análisis de los datos parece adecuado hacer mención a la procedencia de los mismos. El conjunto de ellos procede de la Encuesta de Población Activa (EPA) que proporciona el Instituto Nacional de Estadística (INE). También es preciso aclarar que, por tratarse de datos de periodicidad trimestral, se han utilizado las observaciones correspondientes al segundo trimestre de cada año con el fin de evitar el componente estacional en las comparaciones.

En cuanto a los años seleccionados, se ha optado por utilizar 1977 y 2000 simplemente, porque son los extremos de nuestro periodo de análisis y la comparación entre estos años puede proporcionarnos una primera impresión de los cambios más relevantes que han tenido lugar en la situación del mercado de trabajo español.

Las TABLAS 2.1 y 2.2 nos proporcionan una primera visión representativa de la situación del desempleo en España y de su distribución entre las distintas Comunidades Autónomas.

TABLA 2.1
Tasas de paro regionales

	TASA DE PARO	
	1977	2000
España	4.78	13.83
Andalucía	10.86	24.38
Aragón	2.37	6.98
Asturias	3.87	17.38
Baleares	4.11	4.53
Canarias	8.49	13.90
Cantabria	3.55	13.58
Castilla L	3.75	13.62
Castilla M	4.19	12.28
Cataluña	3.42	8.61
Valencia	3.12	11.40
Extremadura	8.04	24.31
Galicia	1.68	14.75
Madrid	4.26	11.47
Murcia	5.13	11.45
Navarra	2.46	4.46
País Vasco	3.53	11.94
Rioja	0.80	7.89

Fuente: Encuesta de Población Activa 2º trimestre de 1977 y 2000

TABLA 2.2
Posición relativa de las distintas regiones según su tasa de desempleo

	TASA DE PARO	
	1977	2000
España	5	6
Andalucía	1	1
Aragón	16	16
Asturias	9	3
Baleares	8	17
Canarias	2	5
Cantabria	11	8
Castilla L	10	7
Castilla M	7	9
Cataluña	13	14
Valencia	14	13
Extremadura	3	2
Galicia	17	4
Madrid	6	11
Murcia	4	12
Navarra	15	18
País Vasco	12	10
Rioja	18	15

Fuente: Elaboración propia a partir de TABLA 2.1

Lo primero que podemos destacar es el extraordinario aumento que ha sufrido la tasa de paro desde el año 1977 hasta el 2000. Simplemente con observar la media nacional ya podemos apreciar como ésta casi se triplica, pasando del 4,78% al 13.83%. Este aumento medio en la magnitud de la tasa de paro se produce, de forma generalizada, en todas las comunidades autónomas españolas. Únicamente Baleares mantiene un comportamiento relativamente estable, subiendo su tasa desde el 4.11% al 4.53%. En el extremo opuesto aparecen otras regiones donde el aumento en los niveles de desempleo ha sido espectacular. Aquí podríamos incluir a Galicia o a La Rioja, donde la tasa de paro se ha multiplicado por más de ocho.

El segundo aspecto importante tiene que ver con la distribución del desempleo y con cuáles son las regiones con mayores y menores tasas de paro. En cuanto a las regiones con tasas de desempleo superiores a la media nacional podemos observar que, tanto en 1977 como en el año 2000, se mantiene el núcleo de tres comunidades (Andalucía, Extremadura y Canarias) que son las que tienen tasas de paro más altas. En este grupo también habría que incluir a Murcia en el año 1977 y Galicia y Asturias en el 2000.

Por lo que se refiere a las regiones donde el problema del desempleo es menor también podemos observar ciertas generalidades. En primer lugar, son bastantes más las regiones que están por debajo de la media en términos de desempleo que las que están por encima. Esto se debe a que Andalucía y Extremadura tienen tasas de paro muy altas y cuentan negativamente a la hora de calcular la tasa de paro nacional. En segundo lugar, y refiriéndonos a las regiones con menos desempleo, podemos apreciar también un núcleo que se mantiene en los dos años de nuestro estudio y que es el formado por La Rioja, Navarra, Aragón, Cataluña y la Comunidad Valenciana. En el año 1977

El GRÁFICO 2.1 nos permite corroborar desde un punto de vista visual las conclusiones obtenidas hasta el momento. En primer lugar podemos apreciar que la distribución de puntos muestra claramente una tendencia positiva. Este resultado pone de manifiesto que hay una cierta estabilidad en las regiones que mayores y menores tasas de desempleo manifiestan. Es decir, las regiones que en 1977 presentaban unas mayores tasas de desempleo, por lo general, siguen teniendo altas tasas de desempleo en el año 2000, y las regiones que menos desempleo tenían en 1977, también siguen siendo las que menos problemas de paro presentan en el año 2000.

El segundo resultado importante que nos proporciona el GRÁFICO 2.1 tiene que ver con la distribución de los puntos. Podemos apreciar cómo todos ellos se encuentran por encima de la bisectriz del mismo. Este resultado pone de manifiesto que se ha producido un aumento generalizado de las tasas de desempleo entre el año 1977 y el 2000.

Después de esta primera exploración, vamos a intentar profundizar un poco en los principales factores que han podido afectar a las tasas de paro regionales y a las diferencias existentes entre unas regiones y otras.

2.3.- Estructura de la población y paro

Quizá sean los factores demográficos lo primero que haya que tener en cuenta para comprender el pasado y el futuro del paro en España. Características poblacionales como el sexo, la edad o el nivel de estudios afectan a variables como la actividad o el

empleo, y por lo tanto, también provocan diferencias en la tasa de paro. En este sentido, es de esperar que la distinta composición de la población regional en términos de sexo, edad o educación pueda explicar, en mayor o menor medida, la evolución de las tasas de paro regionales.

Las TABLAS 2.3 y 2.4 nos muestran una panorámica de las tasas de actividad y desempleo de los hombres y de las mujeres para los dos años extremos de nuestro estudio.

La TABLA 2.3 se refiere a las tasas de actividad regionales y nos permite observar una serie de generalidades. En primer lugar, se observa que la evolución es distinta para hombres y mujeres. Mientras que la tasa de actividad de los hombres ha sufrido una disminución a lo largo del período 1977-2000 (entre 10 y 15 puntos porcentuales), la tasa de las mujeres ha sufrido un aumento. En segundo lugar, y a pesar de la evolución mencionada, sigue siendo mayor la tasa de actividad masculina que la femenina (en torno al 65 ó 70% para los hombres y entre el 35 ó 40 % para las mujeres). Las regularidades previamente descritas se mantienen, en líneas generales, en las distintas comunidades autónomas. Por su parte, la TABLA 2.4 se refiere a las diferencias por sexo de las tasas de paro regionales. Podemos observar en primer lugar que las tasas de paro son mucho mayores para las mujeres que para los hombres. En esta tabla también podemos apreciar un aumento generalizado de las tasas de paro entre 1977 y el año 2000, pero ese aumento es mucho más acusado en el caso femenino que en el de los varones. Si ahora nos referimos a las diferencias regionales, también podemos observar que son las regiones con más problemas de desempleo las que también presentan mayores niveles en las dos tasas – principalmente Andalucía y Extremadura -.

TABLA 2.3
Tasas de actividad de 1977 y 2000 para las distintas Comunidades Autónomas y para España diferenciadas por sexo

		Tasa de actividad	
		1977	2000
ESPAÑA	Varones	76.94	66.32
	Mujeres	27.78	41.19
Andalucía	Varones	75.12	65.69
	Mujeres	20.18	38.86
Aragón	Varones	75.66	64.16
	Mujeres	25.55	37.65
Asturias	Varones	74.9	58.3
	Mujeres	30.63	34.85
Baleares	Varones	76.75	69.52
	Mujeres	33.96	46.92
Canarias	Varones	78.68	68.55
	Mujeres	25.48	43.52
Cantabria	Varones	74.91	63.3
	Mujeres	29.81	35.14
Castilla L	Varones	72.45	62.02
	Mujeres	27.42	36.88
Castilla M	Varones	73.98	65.68
	Mujeres	21.93	34.14
Cataluña	Varones	79.56	68.44
	Mujeres	28.7	45.92
Valencia	Varones	77.64	68.1
	Mujeres	28.48	42.21
Extremadura	Varones	75.44	65.42
	Mujeres	20.67	38.77
Galicia	Varones	77.87	62.56
	Mujeres	44.23	41.86
Madrid	Varones	78.82	68.36
	Mujeres	27.67	43.62
Murcia	Varones	76.05	69.12
	Mujeres	30.18	39.52
Navarra	Varones	76.75	68.03
	Mujeres	27.76	41.11
País Vasco	Varones	79.24	66.09
	Mujeres	28.33	42.33
Rioja	Varones	76.42	65.51
	Mujeres	30.98	38.09

Fuente: Elaboración propia a partir de datos procedentes del INE

TABLA 2.4
Tasas de paro de 1977 y 2000 para las distintas Comunidades Autónomas y para España diferenciadas por sexo

		Tasa de actividad	
		1977	2000
ESPAÑA	Varones	4.7	9.53
	Mujeres	4.99	20.38
Andalucía	Varones	11.01	17.96
	Mujeres	10.33	34.66
Aragón	Varones	1.7	3.52
	Mujeres	4.28	12.64
Asturias	Varones	3.2	11.61
	Mujeres	5.35	26.14
Baleares	Varones	4.89	2.63
	Mujeres	2.44	7.24
Canarias	Varones	8.24	10
	Mujeres	9.23	19.91
Cantabria	Varones	3.97	8.55
	Mujeres	2.58	22.07
Castilla L	Varones	2.82	7.83
	Mujeres	6.14	23.02
Castilla M	Varones	3.66	7.35
	Mujeres	5.86	21.55
Cataluña	Varones	3.94	6.06
	Mujeres	2.1	12.18
Valencia	Varones	2.95	7.28
	Mujeres	3.56	17.68
Extremadura	Varones	7.9	16.99
	Mujeres	8.49	36.23
Galicia	Varones	1.88	10.5
	Mujeres	1.37	20.56
Madrid	Varones	3.6	7.96
	Mujeres	5.89	16.46
Murcia	Varones	3.84	7.76
	Mujeres	8.14	17.64
Navarra	Varones	1.88	2.13
	Mujeres	4.03	8.25
País Vasco	Varones	2.36	7.59
	Mujeres	6.66	18.36
Rioja	Varones	0.98	3.1
	Mujeres	0.38	15.89

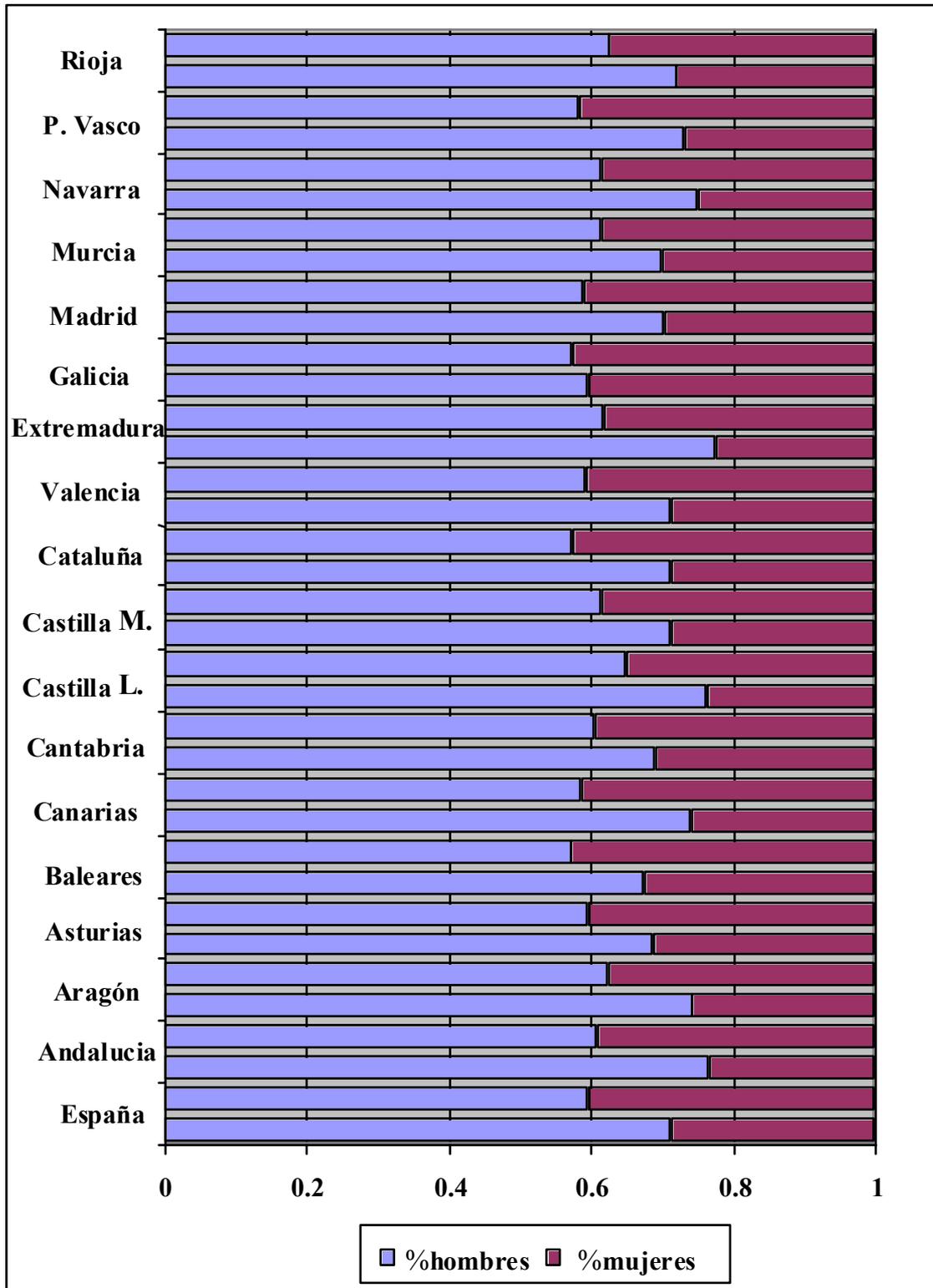
Fuente: Elaboración propia a partir de datos procedentes del INE

La principal conclusión que podemos extraer de las TABLAS 2.3 y 2.4 es que el sexo es un factor importante como variable explicativa del desempleo. Por tanto, la existencia de diferencias en la composición por sexo de la población activa de cada región puede dar lugar a diferencias regionales en cuanto a tasas de paro. También se puede concluir que una de las causas del gran aumento sufrido por el desempleo en el intervalo de tiempo estudiado ha podido ser el incremento registrado en la participación femenina en el mercado de trabajo.

El GRÁFICO 2.2 se refiere a la composición por sexo de la población activa de cada región y de España en su conjunto. Se ha analizado la composición de la población activa en lugar del conjunto de la población por dos razones: en primer lugar, porque la composición por sexo de la población general es muy similar en todas las regiones; y en segundo lugar, porque, si queremos saber los efectos de la composición por sexo sobre el desempleo, parece natural concentrarse en los que participan en el mercado de trabajo, ya que únicamente los participantes pueden contar como desempleados.

Para cada región aparecen dos filas, la superior referida a la composición en el año 2000 y la inferior con la composición de 1977. La primera conclusión que se puede extraer de este gráfico es el aumento considerable que experimenta el porcentaje de población activa formado por mujeres –para el conjunto de España pasa de ser inferior al 30 % en 1977, a alcanzar el 40 % en el año 2000-. En cuanto a las posibles divergencias a nivel regional, también podemos observar en ese mismo gráfico cómo las regiones donde mayor es el aumento en el porcentaje de población activa femenina son las que tienen mayores tasas de paro (Extremadura y Andalucía), aunque no sean las que acaban teniendo un mayor nivel de participación femenina.

GRÁFICO 2.2
Composición por sexo de la población activa de las Comunidades Autónomas y de España en 1977 y 2000



Fuente: Elaboración propia a partir de datos procedentes del INE

Una vez comprobada la importancia del sexo como variable explicativa, tanto de la evolución como de las diferencias interregionales en términos de tasas de paro, vamos a analizar lo que ocurre con la siguiente variable poblacional: la composición por edades de la población.

Las TABLAS 2.5 y 2.6 ofrecen las tasas de actividad y paro para los distintos grupos de edad, tanto nacionales como regionales, en el año 1977 y en el 2000.

La TABLA 2.5, se refiere a las tasas de actividad. En esta tabla podemos observar ya una serie de generalidades. En primer lugar, se produce una fuerte disminución de la tasa de actividad del grupo de población más joven – los comprendidos entre 16 y 19 años – para todas las regiones y para España, disminución que puede ser debida a un mayor índice de escolarización de la población, con lo cual se retrasa la participación en el mercado de trabajo. En segundo lugar, también se produce una disminución en la tasa de actividad del grupo de mayor edad, lo cual puede ser debido a la reducción en la edad de jubilación y a las jubilaciones anticipadas.

Por lo tanto, podemos decir que son los grupos de edad intermedia los que pasan a soportar el peso de la actividad en el mercado de trabajo, sobre todo el grupo comprendido entre los 25 y 54 años. No obstante, también se pueden observar ciertas diferencias regionales, sobre todo la baja actividad juvenil de País Vasco y Asturias, que se ve incluso triplicada por otras regiones como Murcia o Castilla la Mancha, o la baja participación de los jóvenes entre 20 y 24 años en Castilla y León, más de 15 puntos porcentuales por debajo de los que presentan Baleares o Cataluña.

TABLA 2.5
Tasas de actividad de 1977 y 2000 para las distintas Comunidades Autónomas y para España, diferenciando por grupos de edad.

	Tasa de actividad							
	1977				2000			
	16-19	20-24	25-54	55 y +	16-19	20-24	25-54	55 y +
España	54.53	58.42	62.41	28.64	25.67	60.1	77.83	16.42
Andalucía	51.99	50.04	57.06	22.93	27.86	60.49	73.63	14.33
Aragón	55.59	58.03	62.69	26.61	20.33	60.04	78.2	15.61
Asturias	42.04	51.84	63.08	34.6	14.74	54.51	71.77	11.88
Baleares	44.94	63.08	67.57	32.87	30.56	67.68	80.75	18.85
Canarias	45.27	51.94	61.54	30.61	20.84	61.57	75.44	19.44
Cantabria	43.77	51.46	64.77	30.5	13.68	52.35	73.96	13.78
Castilla L	44.17	53.28	63.36	28.51	18.63	51.53	78.36	13.98
Castilla M	55.3	55.15	59.05	24.7	34.3	61.11	74.06	13.81
Cataluña	61.09	64.34	62.93	27.63	31.05	67.73	82.04	18.11
Valencia	62.87	64.41	62.69	25.01	33.23	64.09	77.77	17.04
Extremadura	48.03	50.68	59.49	26.62	30.98	62.58	77.47	14.04
Galicia	50.81	58.73	72.47	42.48	21.1	54.68	79.13	17.77
Madrid	49.33	60.56	60.52	28.39	20.16	54.63	79.87	19.24
Murcia	59.33	56.78	63.17	26.57	36.85	61.39	75.05	16.45
Navarra	43.72	60.83	64.23	29.24	18.28	57.44	80.29	17.51
País Vasco	53.48	58.68	62.64	28.82	11.08	56.6	81.04	15.37
Rioja	61.04	67.02	67.14	23.74	17.97	61.86	77.45	15.54

Fuente: Elaboración propia a partir de datos del INE

TABLA 2.6
Tasas de paro de 1977 y 2000 para las distintas Comunidades Autónomas y para España, diferenciando por grupos de edad.

	Tasa de paro							
	1977				2000			
	16-19	20-24	25-54	55 y +	16-19	20-24	25-54	55 y +
España	13.6	9.03	3.53	2.11	34.04	22.8	12.34	9.31
Andalucía	25.34	17.62	7.97	5.62	43.69	34.11	22.16	20.37
Aragón	7.86	6.62	1.3	0.79	23.74	15.58	6.12	2.17
Asturias	17.89	11.85	2.19	0.59	30.86	35	16	6.9
Baleares	10.22	6.62	3.51	2.67	23.42	5.62	3.52	4.9
Canarias	20.42	13.54	6.8	3.34	34.69	22.28	12.76	6.7
Cantabria	9.11	7.44	2.62	2.14	31.15	31.62	11.65	7.15
Castilla L	11.88	12.81	1.86	0.95	41.22	26.45	12.38	6.18
Castilla M	10.6	7.28	2.95	2.19	22.18	18.45	11.37	7.5
Cataluña	7.43	4.42	2.86	2.41	31.43	14.15	7.08	7.03
Valencia	8.19	4.19	2.16	2.58	27.73	18.28	9.75	9.34
Extremadura	20.17	12.18	6.7	2.65	43.86	29.52	23.11	18.82
Galicia	6.68	6.37	0.83	0.45	31.98	27.03	13.6	8.43
Madrid	14.11	8.65	2.36	3.04	36.39	19.57	10.24	8.08
Murcia	13.58	10.83	3.23	2.06	23.76	21.57	9.3	8.58
Navarra	8.65	5.32	1.46	1.36	25.3	10.05	3.77	0.94
País Vasco	15.02	7.88	1.74	0.33	34.87	23.39	10.86	6.77
Rioja	0	3.35	0.6	0	19.46	18.27	6.52	5.07

Fuente: Elaboración propia a partir de datos del INE

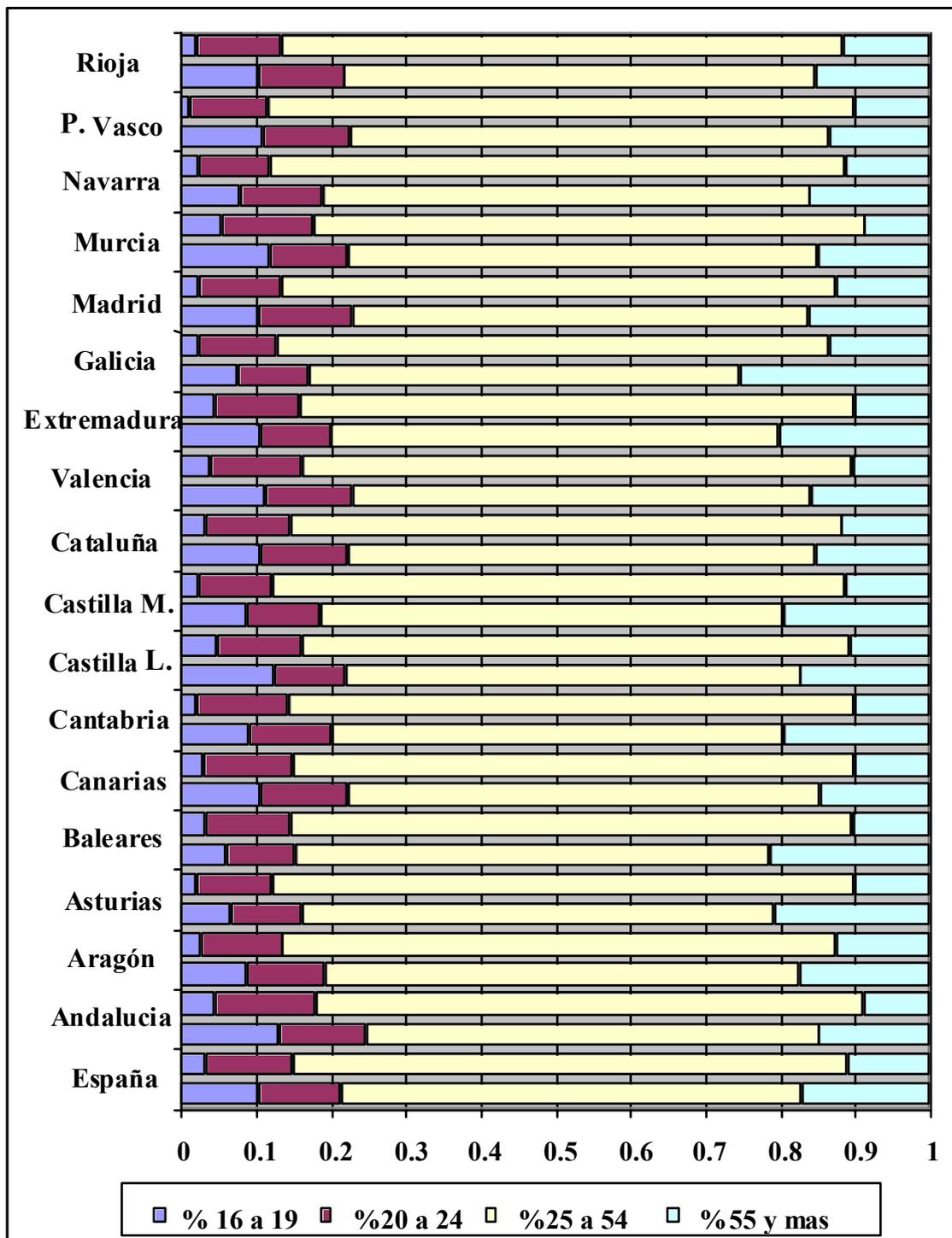
La TABLA 2.6 se refiere a las tasas de paro, y también podemos extraer una serie de conclusiones de su análisis. La principal conclusión es el gran peso que sobre las tasas de paro globales tienen los grupos de población más jóvenes, llegando incluso a cuadruplicar las tasas de paro de los grupos con edad superior a los 25 años.

También pueden observarse las bajas tasas de desempleo que presenta el estrato de población con más edad en todos los casos. Además, se comprueba que siguen existiendo diferencias entre unas regiones y otras, y también que son nuevamente las regiones con mayor problema de desempleo las que presentan mayores tasas de paro para todas las cohortes de edad. Especialmente relevantes son las tasas de paro superiores al 40% de los jóvenes comprendidos entre 16 y 19 en Andalucía, Extremadura y Castilla León.

Nuevamente observamos la importancia que tiene la edad de la población activa a la hora de analizar la tasa de paro de cada región, y cómo la composición por edades de cada Comunidad Autónoma puede ser un factor determinante a la hora de explicar los niveles de desempleo de la misma.

El GRÁFICO 2.3 muestra la composición por edades de la población activa de cada región y de España en general para los años 2000 (barra superior) y 1977 (barra inferior). Nuevamente nos centraremos en la composición de la población activa, en lugar de la población global por las razones anteriormente apuntadas.

GRÁFICO 2.3
Composición por grupos de edad de la población activa de las Comunidades Autónomas y de España en 1977 y 2000



Fuente: Elaboración propia a partir de datos procedentes del INE

Como conclusión más relevante, se observa para el año 2000 una disminución en el porcentaje de la población activa de los grupos de edad extremos, es decir los comprendidos entre los 16 y los 19 años y los que superan los 55, en relación con los valores correspondientes a 1997. También se observa un gran aumento en el porcentaje que representan los activos con edades comprendidas entre 25 y 54 años.

Como elemento diferenciador a nivel regional, podemos observar que Andalucía presenta en los dos años un porcentaje de población joven (entre 16 y 24 años) superior al del resto de regiones, lo cual podría explicar en parte las altas tasas de desempleo andaluzas. En el año 2000 Extremadura presenta una situación muy similar a la andaluza.

Así pues, en líneas generales, podemos apreciar dos fenómenos: Primeramente, el gran cambio que ha sufrido la estructura de la población activa entre los años 1977 y 2000. Y, en segundo lugar, las marcadas diferencias existentes en la composición por edades de las poblaciones activas de las diferentes regiones. Todo esto sugiere que la composición por edades ha de ser un factor importante que habrá que tener en cuenta en nuestro análisis de convergencia.

La última variable demográfica a la que vamos a hacer referencia tiene que ver con el nivel de cualificación de la mano de obra. Aquí nos vamos a centrar en el nivel de estudios (educación formal) ignorando, por tanto, la formación en el puesto de trabajo y otros aspectos de la cualificación igualmente difíciles de medir. Trataremos de ver qué influencia ejerce la educación formal sobre la participación laboral y sobre el desempleo.

Las TABLAS 2.7 y 2.8 muestran los efectos de la educación sobre el mercado de trabajo, en primer lugar sobre la tasa de actividad, y en segundo lugar sobre la tasa de desempleo. En ellas también pueden apreciarse las diferencias existentes, tanto entre regiones, como entre los dos años extremos de nuestro estudio.

La TABLA 2.7 es la referida a la tasa de actividad, y de ella podemos extraer una serie de observaciones. En lo referente a la comparación temporal, se observa, para el año 2000, una disminución de la tasa de actividad de la población con menos estudios, y un aumento en la población con estudios medios, todo esto en comparación con lo que ocurría en 1977. También se puede comprobar cómo la tasa de actividad de los más educados se ha mantenido bastante estable entre ambos años, presentando la población con mayor nivel de estudios las mayores tasas de actividad. En cuanto a la reducción de la tasa de actividad de los menos educados, ésta se asocia al aumento de los años de educación obligatoria. En cuanto a las divergencias registradas entre las distintas regiones se aprecia lo siguiente. Es en los grupos poblacionales de menor cualificación en los que las divergencias regionales en tasas de participación son mayores. En los cohortes más formados se aprecia, en cambio, más homogeneidad en las diferentes comunidades.

Por su parte, la TABLA 2.8 ofrece las tasas de paro de cada grupo de estudios, para cada región y para cada uno de los años considerados. Nuevamente podemos observar el aumento generalizado de las tasas de desempleo de cada grupo entre 1977 y el año 2000. Únicamente los trabajadores con estudios superiores de Aragón y Baleares sufren un mínimo repunte en sus tasas de desempleo. También podemos observar que el grupo formado por los individuos con estudios medios es el que presenta, por lo general, mayores tasas de paro, y el que más contribuye al problema del paro en España.

TABLA 2.7
Tasas de actividad de 1977 y 2000 para las distintas Comunidades Autónomas y para España, diferenciando por nivel de estudios.

	Tasa de actividad					
	1977			2000		
	Bajo	Medio	Superior	Bajo	Medio	Superior
España	49.14	49.29	78.05	31.12	62.93	79.28
Andalucía	44.65	41.35	74.74	34.47	59.47	78.16
Aragón	48.53	49.05	82.37	29.01	63.69	79.58
Asturias	50.80	44.77	74.25	25.27	55.71	75.12
Baleares	52.58	52.77	78.98	34.79	68.46	81.73
Canarias	50.28	45.45	82.71	40.01	61.23	82.05
Cantabria	51.83	41.17	67.80	24.46	56.23	70.41
Castilla L	45.96	39.43	81.23	29.61	64.61	83.25
Castilla M	48.80	39.87	71.35	28.27	59.42	78.40
Cataluña	49.45	58.35	83.82	28.46	68.86	82.98
Valencia	49.56	52.02	80.33	28.85	66.13	78.95
Extremadura	46.19	35.64	74.32	34.09	64.40	79.30
Galicia	59.95	44.89	70.80	34.10	63.61	78.53
Madrid	46.39	52.68	78.21	30.61	60.75	76.46
Murcia	49.05	49.78	82.69	33.66	63.49	81.76
Navarra	50.54	46.78	79.31	27.59	61.08	78.87
País Vasco	51.40	49.96	78.38	27.86	60.21	81.06
Rioja	51.43	47.37	79.27	29.69	63.88	80.61

Fuente: Elaboración propia a partir de datos del INE

TABLA 2.8
Tasas de paro de 1977 y 2000 para las distintas Comunidades Autónomas y para España, diferenciando por nivel de estudios.

	Tasa de paro					
	1977			2000		
	Bajo	Medio	Superior	Bajo	Medio	Superior
España	4.47	6.68	4.85	14.84	15.10	11.08
Andalucía	10.98	11.43	7.58	27.17	25.98	17.69
Aragón	1.41	5.04	8.77	4.66	8.43	7.49
Asturias	2.21	13.41	6.93	13.58	19.50	17.51
Baleares	3.93	3.85	4.03	3.53	5.48	3.70
Canarias	9.42	6.68	3.64	14.09	15.39	11.33
Cantabria	3.07	6.60	4.17	10.13	15.57	13.58
Castilla L	3.77	8.19	4.69	12.60	12.54	11.24
Castilla M	2.35	11.76	7.65	10.55	15.12	14.80
Cataluña	3.56	3.22	2.34	9.55	10.14	5.52
Valencia	2.75	4.95	5.95	11.49	12.29	9.65
Extremadura	7.98	9.26	4.85	28.53	24.71	13.85
Galicia	1.27	6.42	3.39	11.24	16.29	16.81
Madrid	3.88	6.29	3.67	12.29	13.66	8.41
Murcia	3.85	9.20	10.68	10.03	11.55	13.36
Navarra	1.97	4.70	2.17	3.19	5.02	5.38
País Vasco	2.50	8.97	4.77	10.68	13.45	11.39
Rioja	0.87	2.47	4.62	5.69	8.43	9.62

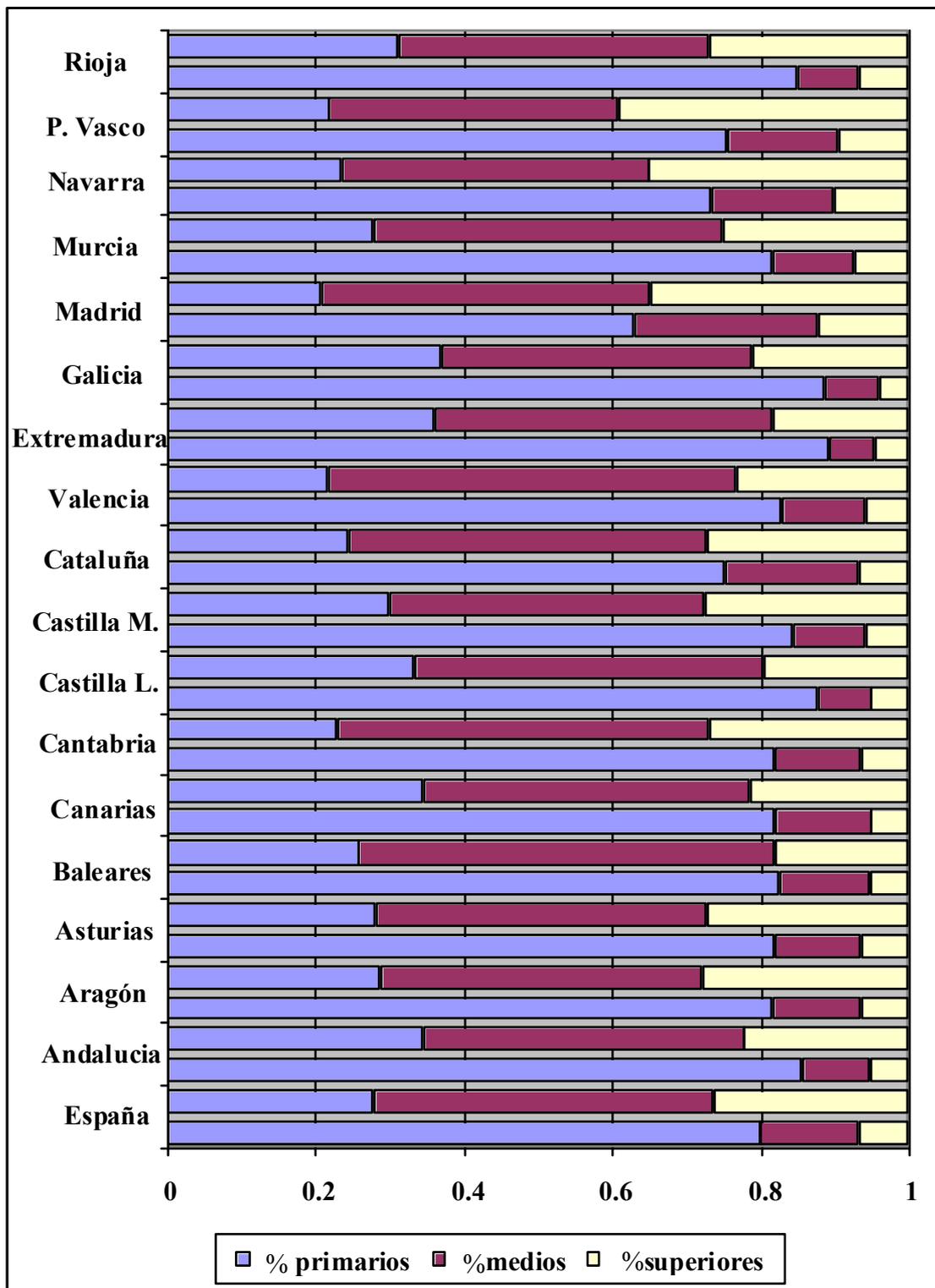
Fuente: Elaboración propia a partir de datos del INE

No existe ninguna razón a priori, suficientemente intuitiva, que explique por qué la incidencia del paro es mayor entre los trabajadores con estudios medios que entre los de estudios bajos. Una razón posible podría ser que los trabajadores con estudios medios tienen mejores expectativas de empleo que los de estudios bajos y no se conforman tan fácilmente con trabajos marginales y poco remunerados, prefiriendo la situación de paro.

Desde el punto de vista de las comparaciones interregionales hemos de observar que, en el año 2000, hay solamente tres regiones –Andalucía, Extremadura y Castilla León- que presentan mayores tasas de desempleo para el grupo de menos estudios que para el de estudios medios. Vemos pues que es en los grupos de estudios medios e inferiores donde estas regiones se separan de las demás. Los individuos con estudios superiores de estas tres regiones, aunque también tienen tasas de desempleo elevadas, están más cerca de los individuos del mismo grupo de las demás regiones.

El GRÁFICO 2.4 nos muestra la distribución de la población activa regional en función de la cualificación de los individuos. El primer aspecto importante es la gran transformación sufrida por la población activa española a lo largo del período considerado. Se puede observar que la población menos educada ha pasado de representar cerca del 80% de la población activa en 1977, a situarse en niveles inferiores al 30% en el año 2000. Por su parte, el grupo de individuos con estudios medios ha pasado a ser el más numeroso, pasando de estar en torno al 15% de la población activa en 1977 a alcanzar cotas superiores al 40% en el último año de nuestro estudio. Finalmente, el grupo de población formado por los individuos con estudios superiores también ha experimentado un gran aumento, y ha pasado de representar a menos del 10% de la población activa en 1977 a superar claramente el 30% en el año 2000.

GRÁFICO 2.4
Composición por niveles de estudios de la población activa de las Comunidades Autónomas y de España en 1977 y 2000



Fuente: Elaboración propia a partir de datos procedentes del INE

Por otro lado, también podemos extraer algunas conclusiones acerca de las diferencias regionales en lo referente a la composición de la población activa por niveles educación. Podemos observar cómo las regiones con mayores tasas de desempleo, como son Andalucía y Extremadura, presentan mayores porcentajes de población con estudios bajos que la media nacional, y también menores porcentajes de población activa con estudios superiores.

En resumen, vemos que la composición por sexo, edad y nivel de estudios son factores a tener muy en cuenta, si queremos entender las diferencias regionales en tasas de paro y la eventual evolución a largo plazo de estas diferencias.

2.4.- Otros factores explicativos

Aparte de la estructura de la población de cada zona, puede haber otros factores que nos permitan explicar en mayor o menor medida el comportamiento específico de cada Comunidad Autónoma. Entre estos factores quizá los más importantes sean la composición sectorial y los aspectos institucionales de cada región. A continuación vamos a apuntar algunos hechos relevantes sobre dichos factores.

2.4.1.- Composición sectorial y paro

La composición por sectores de la economía regional puede ser claramente otro de los factores explicativos de las diferencias en las tasas de paro. Hay que tener en cuenta que no todos los sectores económicos son igualmente dinámicos y que no todos

responden de igual manera ante las perturbaciones de la economía. Por lo tanto, la composición sectorial puede explicar en parte, las diferencias de desempleo entre unas Comunidades Autónomas y otras.

El análisis que vamos a hacer en este apartado también tiene un carácter esencialmente estático, y consiste en la comparación de la situación existente en los años 1977 y 2000. En lo referente a los datos, éstos siguen procediendo de la EPA.

La TABLA 2.9 recoge las tasas de paro de las distintas regiones y de España, pero diferenciando entre los cuatro grandes sectores productivos – agricultura, industria, construcción y servicios – para los mismos años que antes.

La primera conclusión relevante de esta tabla es el aumento generalizado de las tasas de paro entre 1997 y el año 2000 en tres de los cuatro sectores productivos, mientras que se produce una ligera disminución en el caso de la construcción.

También habría que destacar que los sectores más castigados desde el punto de vista del desempleo han sido la agricultura y el sector servicios. En la agricultura, en concreto, la tasa media nacional de paro se ha multiplicado por seis. Finalmente, también se aprecia un cambio en la ordenación los sectores en cuanto a su tasa de desempleo: la construcción ha pasado de ser el sector que más desempleo sufría en 1977 a ocupar el segundo lugar en el año 2000, claramente por detrás de la agricultura y muy cerca del sector servicios.

TABLA 2.9
Tasas de paro de 1977 y 2000 para las distintas Comunidades Autónomas y para España, diferenciando por sectores económicos.

	Tasa de paro							
	1977				2000			
	Agri.	Ind.	Cons	Ser	Agri.	Ind.	Cons	ser
España	2.72	2.71	10.64	1.98	17.87	6.84	10.14	8.32
Andalucía	9.73	7.04	23.66	4.32	39.21	12.69	18.39	13.45
Aragón	0.10	1.24	3.07	1.63	3.43	3.26	2.68	5.51
Asturias	0.00	1.88	9.12	1.45	2.15	6.65	15.57	10.22
Baleares	0.26	2.74	12.10	2.64	0.00	1.78	3.31	3.04
Canarias	3.67	5.48	23.84	3.48	8.73	6.11	10.94	9.56
Cantabria	0.00	2.73	7.14	1.06	4.08	5.73	5.43	8.98
Castilla L	0.40	1.76	4.56	1.96	4.16	5.30	6.88	8.42
Castilla M	2.22	1.59	9.37	1.49	10.02	10.07	9.01	10.85
Cataluña	1.02	2.58	10.63	0.93	5.83	5.85	5.49	4.90
Valencia	0.15	2.15	6.11	1.69	9.06	7.53	6.75	7.99
Extremadura	5.15	7.96	19.63	3.95	29.37	14.74	27.81	13.91
Galicia	0.13	1.02	3.10	0.71	3.53	9.38	10.48	9.65
Madrid	5.88	2.68	6.87	1.26	16.75	4.70	6.96	6.80
Murcia	0.71	2.87	5.80	2.16	4.10	7.44	6.49	6.33
Navarra	0.65	1.28	5.14	0.74	0.00	2.54	2.22	4.03
País Vasco	0.75	1.65	3.57	2.17	2.17	4.92	8.30	7.50
Rioja	0.00	1.06	1.16	0.69	2.11	5.51	0.00	5.11

Fuente: Elaboración propia a partir de datos del INE

Si nos fijamos ahora en las diferencias entre las tasas de paro sectoriales de las distintas regiones, podemos concluir lo siguiente. En primer lugar, hay que destacar las enormes tasas de paro agrícola de Andalucía y Extremadura en el año 2000, que contrastan con el desempleo, estadísticamente nulo que podemos observar en otras regiones, como Navarra o Baleares. En segundo lugar, también sorprende que, aun existiendo una leve tendencia a la disminución del desempleo en el sector de la construcción, aparecen regiones como Asturias, Castilla León, Extremadura, Galicia y el País Vasco donde las tasas de paro de la construcción sufren un notable aumento. Finalmente, también conviene destacar que regiones como Cataluña, Navarra o La Rioja, que son de las que menos desempleo presentan, tienen las mayores tasas en el sector industrial, que sin embargo es el sector que mejor se ha comportado a nivel nacional en términos de desempleo. Esto puede deberse a que algunos de los procesos más duros de reconversión industrial de los últimos años han estado localizados en estas regiones.

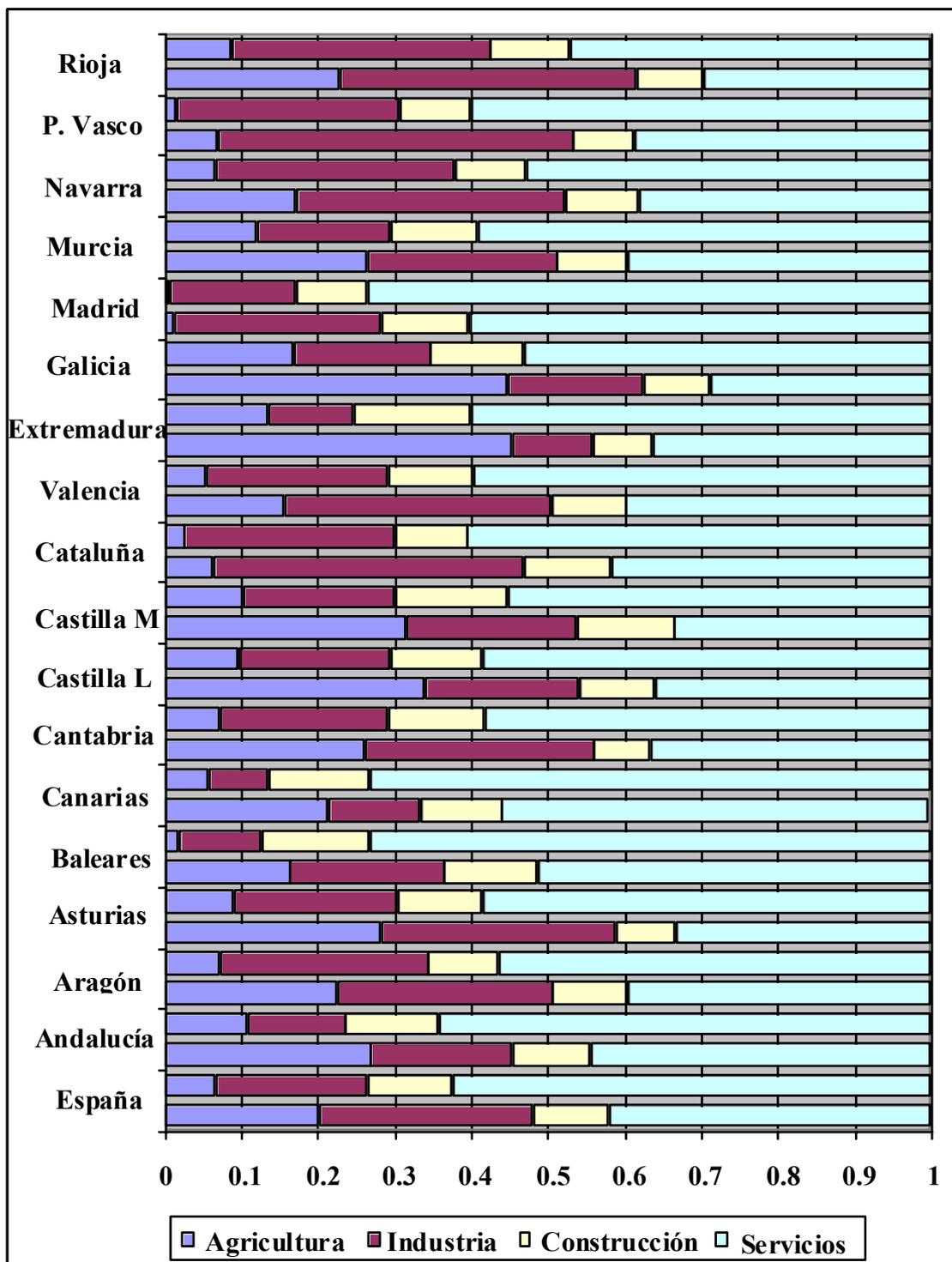
A continuación, trataremos de ver cómo se distribuye la población ocupada de cada región por sectores económicos. Con este fin se ha construido el GRÁFICO 2.5, en el que se recoge, para los años 1977 y 2000 (la barra superior de cada región se refiere al año 2000 y la inferior a 1977), la composición de la población ocupada según el sector al que pertenece. Esta distribución puede ser un buen indicador de la importancia que tiene cada sector productivo en las distintas regiones. También y como complemento a lo anterior, se ha elaborado el GRÁFICO 2.6, referente a la composición del Valor Añadido Bruto (VAB) regional. Este gráfico solo contiene datos del año 2000 ya que la fuente de información es la contabilidad regional y ésta no ofrece datos para 1977.

Lo primero que hay que observar en el GRÁFICO 2.5 es el aumento del porcentaje de población ocupada en el sector servicios entre 1977 y el año 2000. Este aumento se ha producido, básicamente, a costa de la disminución del porcentaje dedicado a la agricultura, que, para el conjunto de España, se ha reducido a menos de la mitad. Por su parte, la industria y la construcción han seguido una evolución más estable, con una ligera disminución en el porcentaje del sector industrial y un pequeño aumento en el caso de la construcción.

En cuanto a las diferencias existentes a nivel regional, el GRÁFICO 2.5 muestra lo siguiente. Se observa cómo Madrid, Baleares y Canarias son las Comunidades Autónomas que muestran mayores tasas de empleo en los servicios, debido en buena parte a que Madrid es el principal centro administrativo del estado, y a la importancia turística de los dos archipiélagos. También podemos apreciar cómo las regiones con mayor importancia del sector agrícola son Andalucía, Extremadura, Murcia, Asturias, Galicia y las dos Castillas, lo que puede explicar en parte la magnitud del desempleo de las mismas.

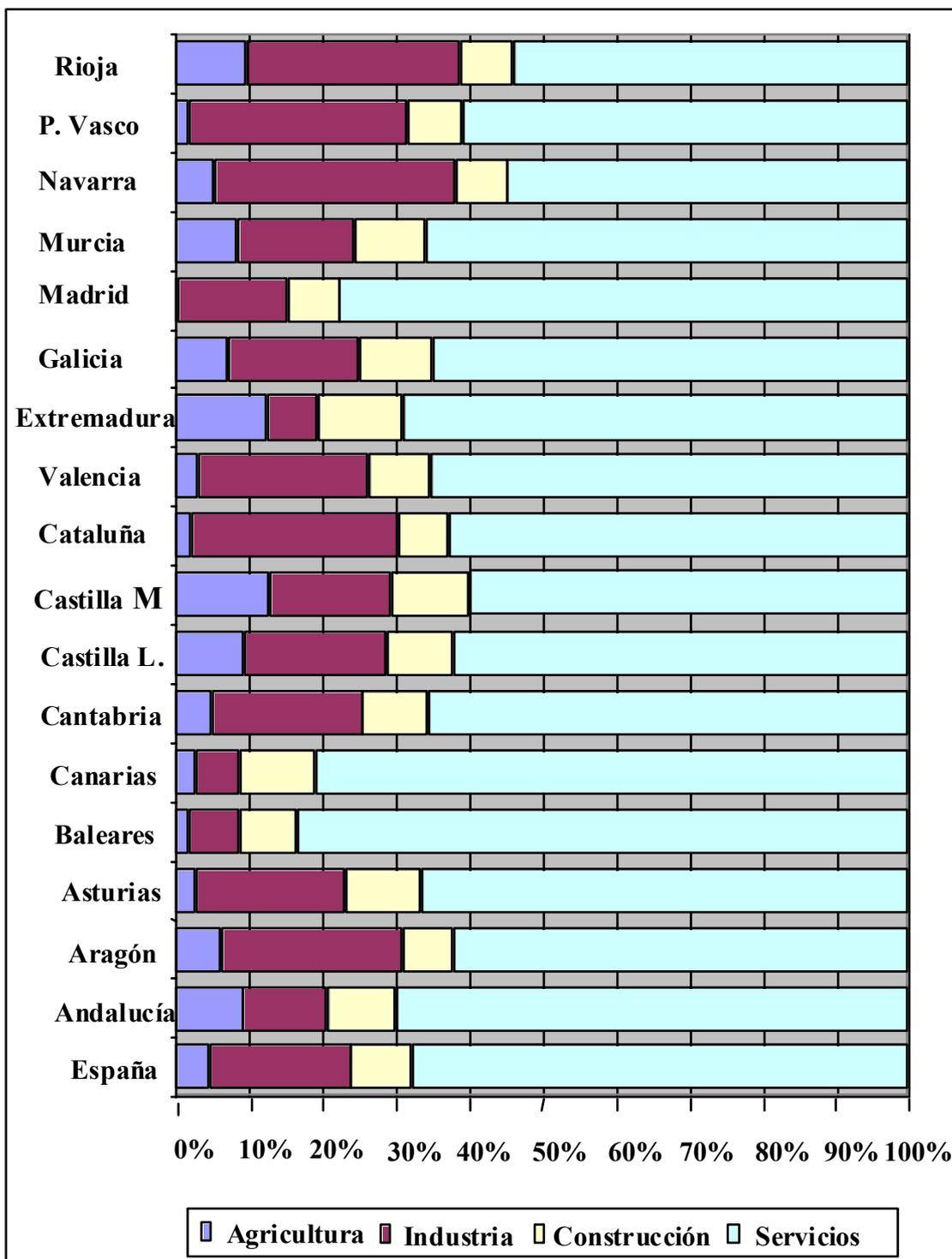
El GRÁFICO 2.6 nos permite confirmar las apreciaciones previas desde un punto de vista diferente. Podemos ver que el mayor porcentaje de VAB corresponde al sector servicios en todas las regiones, y por lo tanto también en España, donde el VAB del sector servicios llega a alcanzar un nivel cercano al 70% del total.

GRÁFICO 2.5
Composición por sectores económicos de la población ocupada. Datos para las distintas Comunidades y para España en los años 1977 y 2000



Fuente: Elaboración propia a partir de datos procedentes del INE

GRÁFICO 2.6
Composición del Valor Añadido Bruto por sectores. Datos para las distintas Comunidades y para España en el año 2000 (avance)



Fuente: Elaboración propia a partir de datos procedentes del INE

Las peculiaridades regionales en cuanto al VAB corroboran lo apreciado previamente en términos de paro; más en concreto, en aquellas regiones con mayor peso relativo del sector industrial (La Rioja, Navarra, Aragón, Cataluña y Comunidad Valenciana) se registran, a su vez, las tasas de desempleo más bajas.

Así pues queda demostrada la importancia que tiene la composición productiva de cada Comunidad Autónoma como factor explicativo del desempleo regional, así como su potencial efecto sobre la existencia de divergencias entre las tasas de paro regionales.

2.4.2- Presión sindical y paro

Otro de los factores que pueden afectar al comportamiento diferenciado del desempleo regional son los aspectos institucionales, y entre ellos quizá el más influyente sea el poder relativo de las organizaciones sindicales en cada Comunidad Autónoma. Por lo que respecta a otros aspectos institucionales potencialmente relevantes, como podrían ser las regulaciones en materia de contratos de trabajo, duración de la jornada laboral, seguridad e higiene en el trabajo, así como las leyes de salario mínimo, no serán tenidos en cuenta; y la razón es que todos estos aspectos institucionales se definen más bien a escala nacional y no permiten hacer demasiadas distinciones entre regiones.

De acuerdo con Jimeno (1992), dos son las principales razones que explican la influencia de los sindicatos españoles: en primer lugar, que todos los trabajadores, incluso los temporales, eligen (o pueden elegir) a sus representantes sindicales; y en

segundo lugar, que los convenios colectivos son legalmente vinculantes y se aplican a todos los trabajadores, independientemente de que estén afiliados o no.

Por otra parte, como señalan Jimeno y Toharia (1993), la tasa de afiliación sindical actualmente es relativamente baja (aunque no existen cifras fiables que lo demuestren), sobre todo si la comparamos con los primeros años de la democracia, en los que se llegaron a alcanzar niveles superiores al 30 % para el total de trabajadores y cerca del 50% en el caso de la industria (Milner y Nombela, 1994). A pesar de esto último, podría decirse que la “tasa efectiva de sindicación”, que intenta medir la proporción de los trabajadores cuyos intereses son efectivamente defendidos por los sindicatos en la negociación colectiva, tiene que ser bastante alta, sobre todo si se tiene en cuenta algo que ya se ha apuntado anteriormente: en España los convenios colectivos tienen fuerza de ley y éste parece ser un rasgo bastante insólito en la escena internacional.

En España existen varios niveles de negociación: nacional, donde intervienen gobiernos, sindicatos y patronal; sectorial, entre la patronal y los sindicatos de cada sector; y por empresas, dando lugar a los convenios de empresa. De todos ellos, son los convenios sectoriales los más generalizados. Por lo que respecta a los convenios nacionales hay que decir que tienen muy poca importancia relativa (Toharia *et al.*, 1998). Y en relación con los convenios de empresa se puede afirmar que, a pesar de que el número de éstos es superior al de convenios sectoriales, el porcentaje de trabajadores afectados es relativamente bajo.

TABLA 2.10
Número de convenios y trabajadores afectados durante el año 2000 según el tipo de convenio.

	Convenios de empresa		Convenios de otro ámbito	
	Número	Trabajadores	Número	Trabajadores
Andalucía	540	68235	185	1137344
Aragón	150	30929	61	147293
Asturias	100	31557	33	92217
Baleares	66	7319	21	174109
Canarias	178	18516	41	253523
Cantabria	80	12384	33	49189
Castilla L.	263	30140	167	240807
Castilla M.	134	17068	77	166233
Cataluña	545	107070	150	1376652
Valencia	279	33930	136	627753
Extremadura	49	9429	30	157126
Galicia	265	36892	104	249312
Madrid	307	90137	54	625558
Murcia	55	3975	42	166240
Navarra	89	19013	31	85211
País Vasco	342	52698	104	206537
Rioja	37	5786	19	25138

Fuente: Datos procedentes del Ministerio de Trabajo y Asuntos Sociales (MTAS)

La TABLA 2.10 nos proporciona información sobre los distintos tipos de convenios y los trabajadores afectados por cada tipo en el año 2000. En primer lugar, esta tabla nos permite observar que son más los convenios de empresa que los de otro ámbito (la gran mayoría de estos últimos, de carácter sectorial), pero también que son muchos menos los trabajadores afectados por los primeros. Este dato, que podría parecer sorprendente, no lo es tanto si tenemos en cuenta que el ámbito de aplicación de un convenio sectorial es mucho más amplio que el de un convenio de empresa.

Este predominio de los convenios sectoriales tiene consecuencias serias desde el punto de vista del funcionamiento del mercado de trabajo español, según veremos más adelante en el capítulo segundo de esta tesis. Esto nos llevará a concentrar nuestra atención en este tipo de convenios.

La TABLA 2.10 no nos permite obtener conclusiones claras desde el punto de vista de las comparaciones regionales. La razón de esto estriba en que dicha tabla está elaborada en valores absolutos y por lo tanto los datos que en ella aparecen están condicionados por el tamaño de cada Comunidad autónoma.

A fin de corregir esta deficiencia se ha elaborado la TABLA 2.11, que muestra el porcentaje de trabajadores afectados por cualquier clase de convenios, tanto en el caso de los ocupados como en el de los asalariados. Téngase en cuenta que el grupo de los ocupados incluye, además de a los asalariados, a los trabajadores autónomos y a algunos empresarios que además son empleadores.

TABLA 2.11
Número de convenios y trabajadores afectados en el año 2000, y porcentajes de afectados sobre total de ocupados y asalariados.

	Total convenios			
	Número	Trabajadores	% sobre Ocupados	% sobre Asalariados
España	4767	15261300	41.64	48.13
Andalucía	725	1205579	52.74	60.96
Aragón	211	178222	37.97	44.39
Asturias	133	123774	35.30	42.86
Baleares	87	181428	50.85	63.93
Canarias	219	272039	41.27	48.41
Cantabria	113	61573	32.39	39.19
Castilla L.	430	270947	30.20	38.32
Castilla M.	211	183301	29.94	38.30
Cataluña	695	1483722	54.89	62.84
Valencia	415	661683	40.50	46.25
Extremadura	79	166555	48.44	62.40
Galicia	369	286204	27.75	37.60
Madrid	361	715695	33.89	34.58
Murcia	97	170215	39.41	47.07
Navarra	120	104224	44.46	52.16
País Vasco	446	259235	30.75	34.02
Rioja	56	30924	29.04	34.36

Fuente: Elaboración propia a partir de datos procedentes del INE y del MTAS

Los principales resultados de la TABLA 2.11 son los siguientes: en primer lugar, se corrobora la importancia de la negociación colectiva, que afecta como media al 41% de los ocupados y al 48% de los trabajadores asalariados; y en segundo lugar, se comprueba que los trabajadores más afectados por la negociación colectiva son los que forman el grupo de los asalariados, los cuales superan en siete puntos porcentuales al grupo de ocupados.

En cuanto a las comparaciones regionales, habría que destacar algunos aspectos relevantes. En primer lugar, se observa que las regiones más afectadas por la negociación colectiva son Andalucía, Cataluña, Extremadura y Navarra. En segundo lugar, se pueden apreciar grandes diferencias entre regiones; así por ejemplo Cataluña duplica los niveles de Galicia en el caso de los ocupados, y los de Madrid, País Vasco o La Rioja en el caso de los asalariados.

En general podemos afirmar que el poder sindical es un factor importante el mercado de trabajo español. Además, existen suficientes discrepancias regionales como para pensar que dicho factor puede tener que ver mucho con las diferencias de desempleo entre las distintas Comunidades Autónomas.

2.5.- Análisis de los componentes del desempleo regional a través de la técnica “shift-share”

Con el fin de sintetizar de alguna manera la información provista en los apartados previos vamos a elaborar una aplicación de la técnica denominada *shift-share*, la cual nos permitirá distinguir dos grupos de factores explicativos del cambio de las

tasas de paro regionales, los referidos a características poblacionales de cada Comunidad Autónoma, y los que podemos considerar propios de cada región.

Para el cálculo de la influencia de cada componente vamos a recurrir al procedimiento utilizado por Jimeno (1997), donde la tasa de paro de cada región se desglosa como una media ponderada de las tasas de paro de los distintos grupos poblacionales que la integran. Por tanto, es posible descomponer el diferencial de la tasa de paro entre cada región y el resto de España de la manera siguiente:

$$\sum_{i=1}^N \alpha_i^j u_i^j - \sum_{i=1}^N \alpha_i^{RE} u_i^{RE} = \sum_{i=1}^N (\alpha_i^j - \alpha_i^{RE}) u_i^{RE} + \sum_{i=1}^N \alpha_i^j (u_i^j - u_i^{RE})$$

donde u_i^j y u_i^{RE} son, respectivamente, las tasas de paro del grupo i en cada región j y en el resto de España, y α_i^j y α_i^{RE} son, respectivamente la participación del grupo i en la población activa de cada región j y en el resto de España. El primer sumando del lado derecho de la expresión anterior representa la parte del diferencial de tasas de paro (diferencia entre la tasa de paro regional y la del resto de España) debida al hecho de que la composición de la población activa de la región no coincide con la del resto de España. El segundo sumando representa aquella parte del diferencial de tasas de paro que se debe a factores específicamente regionales, los cuales hacen que una región pueda tener tasas de paro más altas o más bajas (dentro de cualquier grupo de población) que el resto de la nación.

TABLA 2.12
Descomposición de la diferencia entre el desempleo regional y el del resto de España para 1977

	Efecto Población	Efecto Regional	Efecto Total
Andalucía	0,141	6,928	7,069
Aragón	-0,059	-2,493	-2,552
Asturias	-0,506	-0,483	-0,989
Baleares	-0,636	-0,247	-0,883
Canarias	0,042	4,077	4,119
Cantabria	-0,189	-1,067	-1,256
Castilla L	0,071	-0,746	-0,675
Castilla M	-0,214	-1,078	-1,292
Cataluña	0,217	-1,862	-1,645
Valencia	0,092	-1,863	-1,771
Extremadura	-0,098	3,308	3,210
Galicia	-0,672	-2,699	-3,371
Madrid	0,230	-0,623	-0,393
Murcia	0,042	0,126	0,168
Navarra	-0,040	-2,341	-2,381
País Vasco	0,241	-1,422	-1,181
La Rioja	0,007	-3,776	-3,769

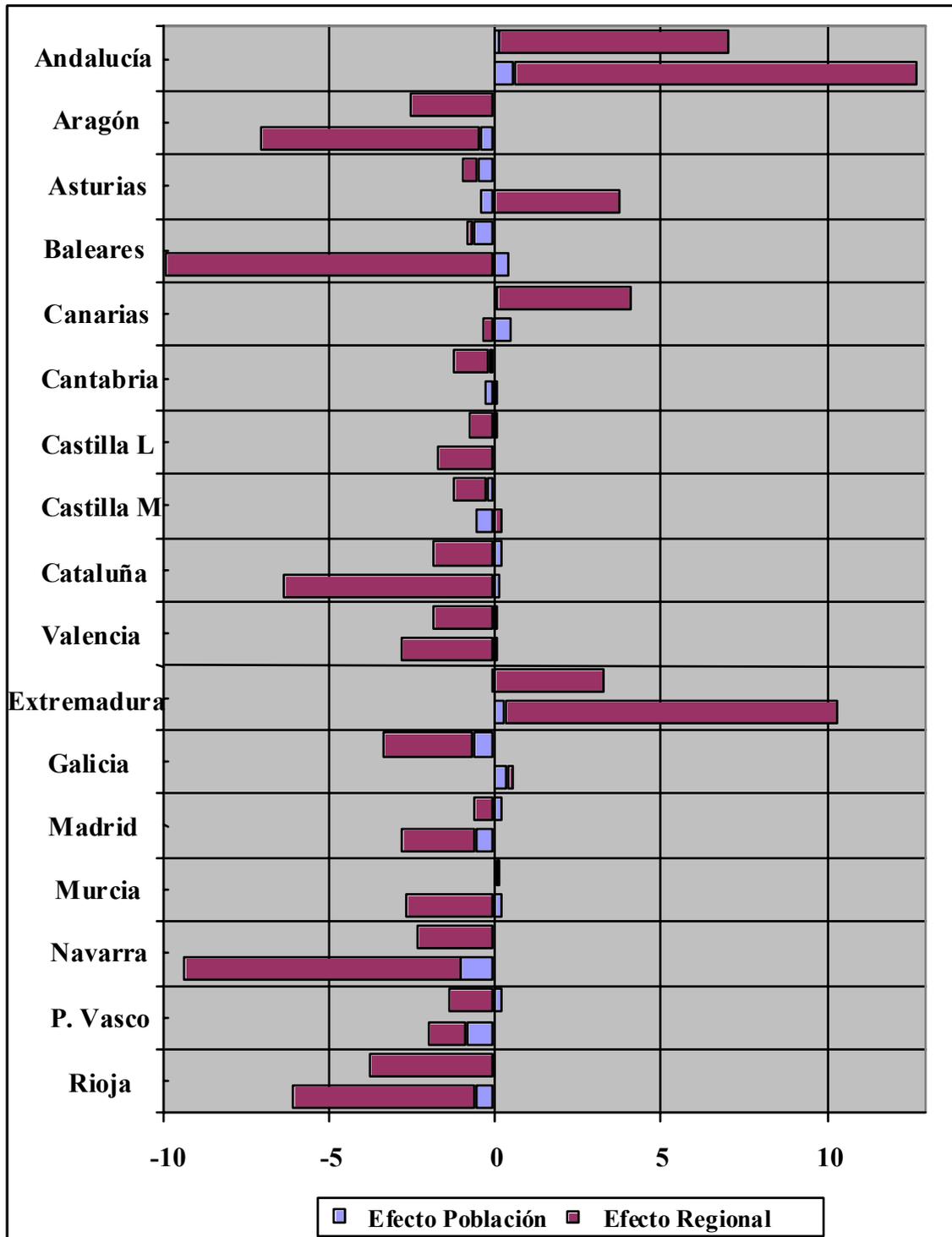
Fuente: Elaboración propia a partir de datos procedentes del INE.

TABLA 2.13
Descomposición de la diferencia entre el desempleo regional y el del resto de España para 2000

	Efecto Población	Efecto Regional	Efecto Total
Andalucía	0,566	12,200	12,766
Aragón	-0,440	-6,649	-7,089
Asturias	-0,411	3,811	3,400
Baleares	0,461	-9,967	-9,506
Canarias	0,496	-0,397	0,099
Cantabria	-0,284	0,101	-0,183
Castilla L	-0,065	-1,664	-1,729
Castilla M	-0,587	0,268	-0,319
Cataluña	0,167	-6,397	-6,230
Valencia	0,106	-2,861	-2,755
Extremadura	0,310	10,070	10,380
Galicia	0,407	0,201	0,608
Madrid	-0,586	-2,216	-2,802
Murcia	0,207	-2,662	-2,455
Navarra	-1,023	-8,355	-9,378
País Vasco	-0,878	-1,159	-2,037
La Rioja	-0,564	-5,541	-6,105

Fuente: Elaboración propia a partir de datos procedentes del INE.

GRÁFICO 2.7
Descomposición de las diferencias entre las tasas de paro regionales y las del resto de España para los años 1977 y 2000.



Fuente: Elaboración propia a partir de datos procedentes del INE

La desagregación por grupos se ha realizado de acuerdo con el sexo, la edad y el nivel de estudios del desempleado. En total se han construido doce grupos diferenciando entre hombres y mujeres, mayores y menores de veinticinco años, y tres niveles de estudios (superiores, medios e inferiores).

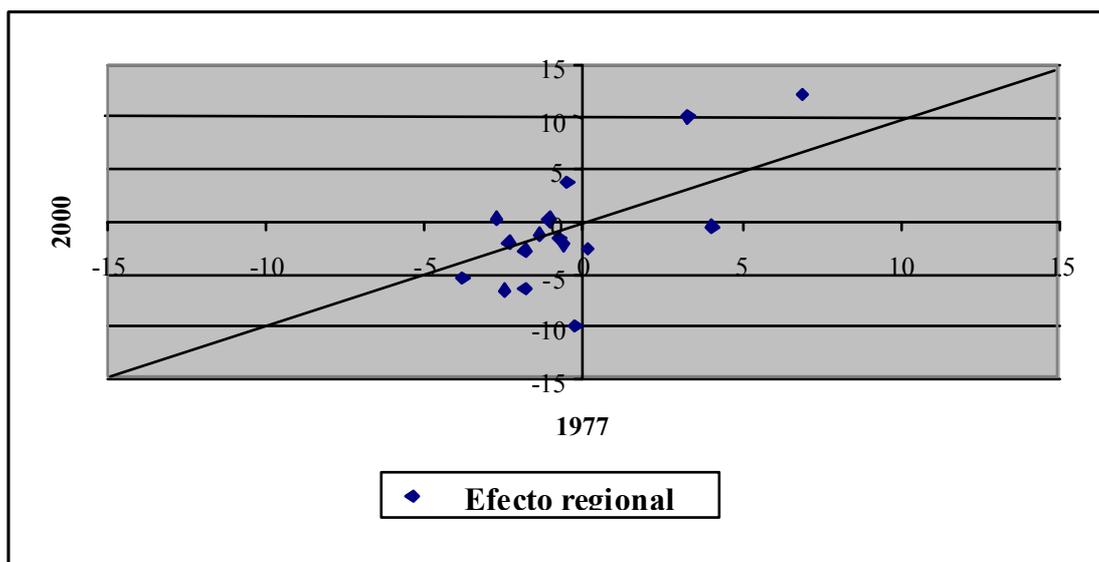
Los resultados de esta descomposición aparecen en las TABLAS 2.12 y 2.13 y el GRÁFICO 2.7. La TABLA 2.12 descompone el diferencial de desempleo para el año 1977 y la TABLA 2.13 hace lo mismo para el año 2000. En cada una de ellas encontramos tres columnas. La tercera columna nos indica la diferencia (positiva o negativa) entre la tasa de paro de la región y la del resto de España. Y las dos primeras nos indican las partes de esa diferencia que se deben a la estructura de la población (primera columna) y a los factores específicamente regionales (segunda columna). A través del GRÁFICO 2.7 puede verse que el efecto regional predomina claramente sobre el efecto de la composición de la población. Además, se ve también que los dos efectos no siempre actúan en el mismo sentido, ya que se observan tanto casos en los que los dos efectos se refuerzan mutuamente, como casos en los que ambos van en sentido opuesto. No obstante, debido a la distinta magnitud de los efectos, siempre suele ser el regional el que marca el sentido de las diferencias globales.

De lo dicho anteriormente se desprende que la composición de la población activa de una región no es la causa principal de que esa región tenga más o menos desempleo que el resto del país. Esto puede resultar un poco sorprendente, ya que, en principio, uno esperaría que cada grupo de población, en la medida en que es un grupo homogéneo, goce de una situación económica similar en todas las regiones: si no fuera así, se pondrían en marcha movimientos migratorios compensadores entre las regiones hasta que las diferencias desapareciesen; a menos que exista algún tipo de barreras a la

movilidad o alguna otra razón, que en todo caso habría que concretar. Por eso, será preciso afinar un poco más el análisis y aclarar en qué condiciones y hasta qué punto tiene sentido esperar que las tasas de desempleo varíen de unas regiones a otras. En el capítulo siguiente intentaremos dar respuesta a estas preguntas.

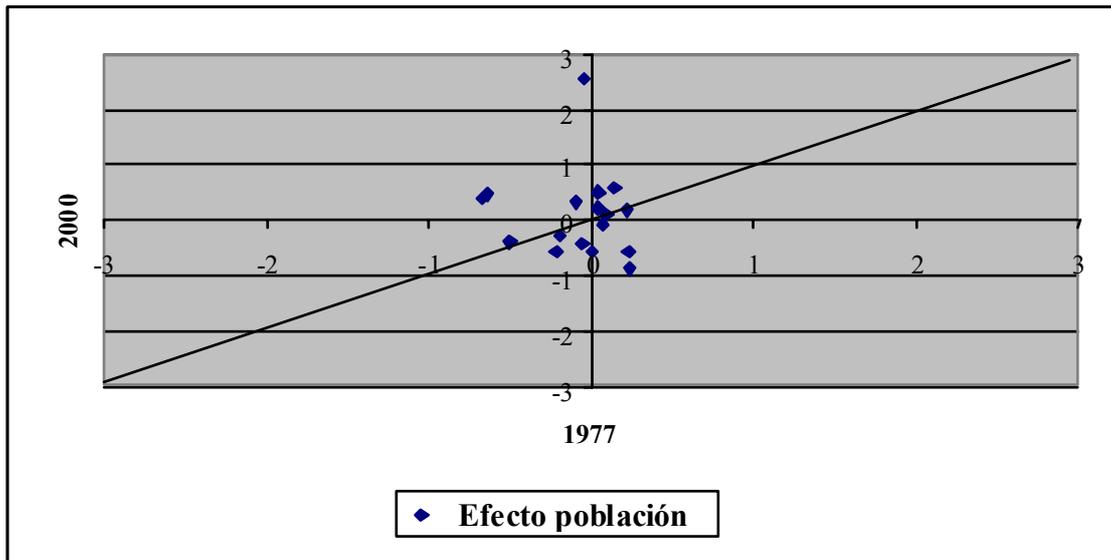
Para finalizar este epígrafe podemos observar los GRÁFICOS 2.8, 2.9 y 2.10, donde se muestra la dispersión de los tres efectos entre los dos años extremos de nuestro estudio. El primer resultado importante es que el efecto total parece estar determinado por factores regionales. Como se puede ver en el GRÁFICO 2.8 los efectos poblacionales son, por lo general, de magnitud muy reducida. Por otra parte, el hecho de que los GRÁFICOS 2.8 y 2.10 muestren unas distribuciones tan similares sugiere que los efectos regionales son que más cuentan como factores explicativos de las diferencias de desempleo. El GRÁFICO 2.10 también indica que, independientemente del año estudiado, las regiones con tasas de paro superiores o inferiores a las del agregado nacional siguen siendo las mismas.

GRÁFICO 2.8
Relación entre los efectos regionales de 1977 y 2000 para las distintas Comunidades Autónomas



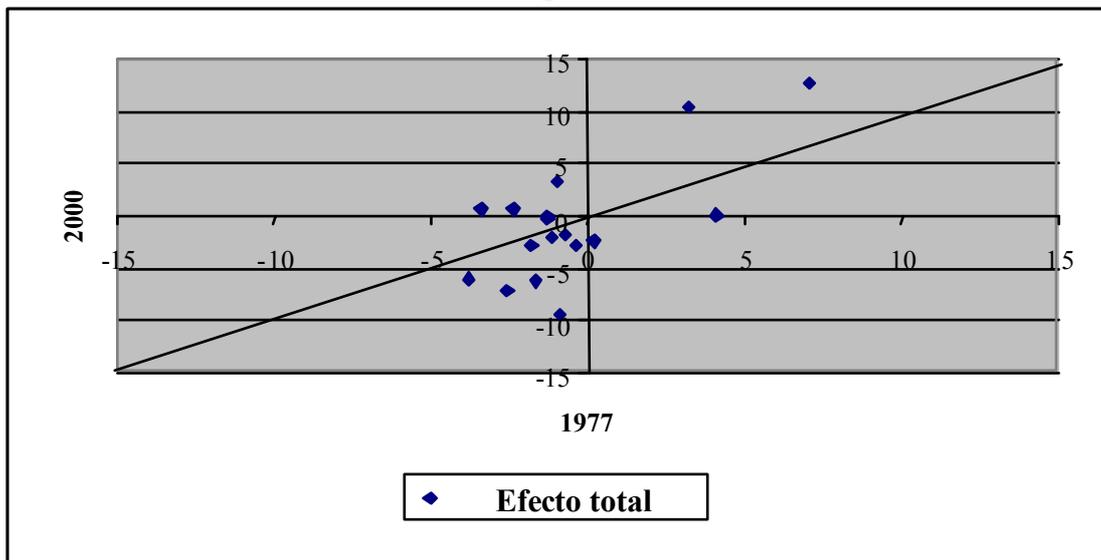
Fuente: Elaboración propia a partir de TABLA 2.13

GRÁFICO 2.9
Relación entre los efectos poblacionales de 1977 y 2000 para las distintas Comunidades Autónomas



Fuente: Elaboración propia a partir de TABLA 2.13

GRÁFICO 2.10
Relación entre los efectos totales de 1977 y 2000 para las distintas Comunidades Autónomas



Fuente: Elaboración propia a partir de TABLA 2.13

Por su parte, el GRÁFICO 2.9 referido al efecto población no permite obtener ninguna conclusión relevante. Únicamente podemos resaltar que tiene un valor muy bajo ya que los puntos se encuentran muy cerca del eje de coordenadas. Este gráfico parece indicar una cierta homogeneidad de la población activa de las distintas regiones.

2.6.- Análisis dinámico

En este último epígrafe intentaremos dar una visión de cómo ha evolucionado a lo largo del tiempo la dispersión de las tasas de desempleo regionales. Este análisis podrá darnos una primera impresión acerca de la posibilidad de convergencia entre dichas tasas, algo que más adelante abordaremos desde perspectivas distintas con el fin de confirmar la verosimilitud de esta primera impresión.

Para estudiar cómo ha evolucionado la dispersión de las tasas de paro regionales tendremos que definir en primer lugar una medida de dispersión. Con este fin vamos a recurrir al cuadrado¹ del coeficiente de variación de las tasas de desempleo², el cual se define como sigue:

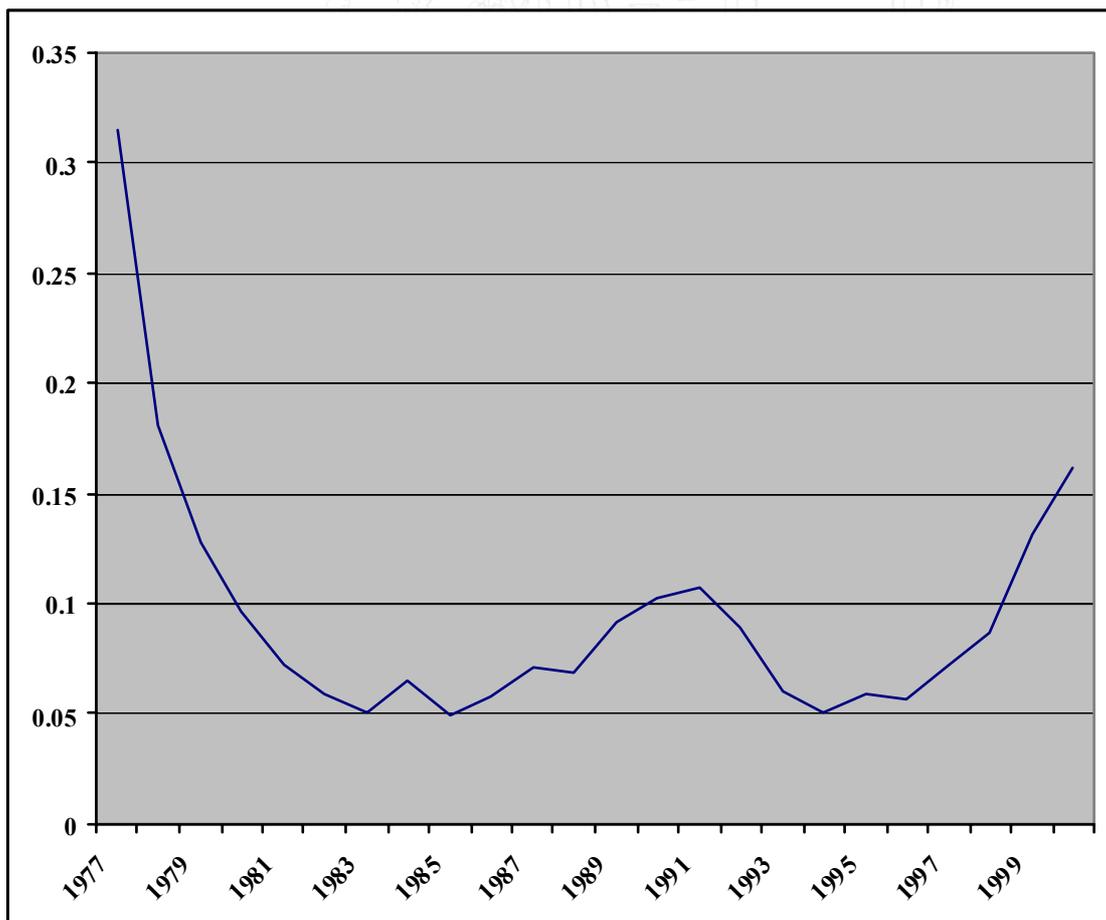
¹ Existen varias formas de calcular la dispersión. En primer lugar, podríamos calcular la *distancia media* que separa a los distintos valores de la variable de la media muestral. Estas distancias se pueden medir en valor absoluto, o bien se pueden elevar al cuadrado, a fin de que unas distancias no se compensen con otras, como ocurriría si utilizásemos diferencias simples. De esta forma lo que obtendríamos sería una medida de dispersión *absoluta*. Para tener una medida de dispersión *relativa* sería preciso ponderar esas distancias por el valor medio de la variable o por su cuadrado, dependiendo de si estamos utilizando distancias en valor absoluto o los cuadrados de esas distancias.

² Podemos encontrar un análisis similar para el periodo 1976-1985 en Plaza Acero (1994).

$$CV^2 = \frac{\sum_i (u_{i,t} - u_t)^2 (n_{i,t} / n_t)}{u_t^2}$$

donde $u_{i,t}$ es la tasa de paro de la Comunidad Autónoma i en el año t , u_t es la tasa de paro de la nación en el año t , $n_{i,t}$ es la población activa de la región i en el año t y n_t es la población activa de España en el año t .

BIBLIOTECA VIRTUAL
GRÁFICO 2.11
Dispersión interregional de las tasas de desempleo.



Fuente: Elaboración propia a partir de TABLA 2.14

TABLA 2.14
CV: Dispersión de las tasas de desempleo

	CV
1977	0,315
1978	0,180
1979	0,127
1980	0,096
1981	0,072
1982	0,059
1983	0,050
1984	0,064
1985	0,049
1986	0,057
1987	0,070
1988	0,068
1989	0,092
1990	0,102
1991	0,107
1992	0,089
1993	0,060
1994	0,050
1995	0,059
1996	0,057
1997	0,072
1998	0,087
1999	0,131
2000	0,162

Fuente: Elaboración propia a partir de datos procedentes del INE

En la TABLA 2.14 y en el GRÁFICO 2.11 encontramos los resultados obtenidos al calcular ese coeficiente de variación. Estos resultados nos van a permitir realizar un análisis de la dispersión del desempleo regional para el periodo comprendido entre 1977 y el año 2000. Se puede decir que la dispersión del desempleo ha disminuido a lo largo del periodo considerado, aunque la senda recorrida ha tenido fuertes altibajos. Tras un periodo de brusca disminución que va desde 1977 hasta 1984, pasamos a un segundo periodo donde se observa un estancamiento en la dispersión de las tasas de paro, con un cierto repunte entre 1989 y 1992. Finalmente, se observa un tercer período de crecimiento de la dispersión que comienza en 1996 y abarca los últimos cinco años de nuestro estudio.

Ya hemos dicho que el GRÁFICO 2.11 nos muestra la evolución que ha seguido la dispersión de las tasas de paro regionales a lo largo del período 1977-2000. Si se compara éste con el GRÁFICO 2.12, donde además del coeficiente de variación al cuadrado aparece la tasa de paro nacional, podemos ver el carácter claramente procíclico de la evolución de las diferencias regionales (la tasa de desempleo, que ahí aparece como indicador del ciclo económico, aumenta en las recesiones y disminuye en los auges).

El análisis dinámico que hemos llevado a cabo en este epígrafe no nos permite obtener conclusiones inequívocas sobre si existe o no un proceso de convergencia entre las tasas de desempleo. Únicamente nos permite admitir un comportamiento muy ligado al ciclo económico, con aumento de las disparidades en las fases expansivas del ciclo, y disminución en las recesivas. Esto puede ser una consecuencia del hecho de que las regiones con menores tasas de desempleo suelen tener un comportamiento cíclico más acusado que las otras (las oscilaciones de sus tasas de paro suelen ser más acusadas en

las primeras que en las segundas). Por lo tanto, en las fases expansivas del ciclo es de esperar un alejamiento entre los niveles desempleo de las distintas regiones, porque las que tienen menos desempleo son las que más lo reducen, en cambio, en las fases recesivas es de esperar el comportamiento contrario.

GRÁFICO 2.12
Comparación entre el coeficiente de variación y el ciclo económico.



Fuente: Elaboración propia a partir de datos procedentes del INE

En los capítulos posteriores se intentará profundizar en estas cuestiones, y se buscará, a través del análisis econométrico, una respuesta más concreta acerca de si ha existido, o no, un proceso de acercamiento entre las tasas de paro regionales.

BIBLIOTECA VIRTUAL

CAPÍTULO 3:

**BASES TEÓRICAS PARA EL ANÁLISIS DE
CONVERGENCIA**

3.1.- Introducción

Cualquier planteamiento teórico que pretenda analizar la convergencia de las tasas de paro regionales requiere un estudio previo de los mercados de trabajo de las distintas regiones, de sus características estructurales y del grado de integración entre los mismos. Dependiendo de cómo se caractericen dichos mercados, vamos a poder diferenciar varios tipos de escenarios que nos pueden llevar a conclusiones muy diferentes.

A continuación vamos a fijar una serie de escenarios y vamos a ver qué posibilidades ofrece cada uno de ellos con respecto al desempleo. Más concretamente, intentaremos ver si es de esperar que las tasas de desempleo de las distintas regiones tiendan a igualarse o si por el contrario es posible la existencia de diferencias permanentes entre las tasas de desempleo regionales³.

En esencia, lo que vamos a ver es que no tiene demasiado sentido esperar que las tasas de desempleo regionales converjan en torno a un mismo valor, aunque sí cabe esperar una cierta tendencia a la homogeneización de las tasas de paro en la medida en que los colectivos de trabajadores comparados sean también homogéneos: a mayor homogeneidad de las fuerzas de trabajo regionales, mayor tendencia a la convergencia de las tasas de paro.

³ En este sentido habría que destacar aquí la similitud entre la línea de razonamiento seguida en este capítulo y la que inspira toda una serie de trabajos explicativos de las diferencias regionales de desempleo. Entre estos trabajos se pueden destacar, entre otros, los de Brechling (1967), Metcalf (1975), Hyclack y Lynch (1980) y Ho fler y Murphy (1989).

3.2.- Escenarios para el análisis de convergencia

3.2.1.- Criterios básicos

Para construir los distintos escenarios en los que se va a desarrollar nuestro análisis de convergencia, nos vamos a fijar en dos aspectos fundamentales: el grado de homogeneidad de los trabajadores y el mecanismo de fijación de los salarios. Cada uno de estos aspectos va a presentar dos posibilidades alternativas, y dependiendo de cómo se combinen estas posibilidades, vamos a tener cuatro escenarios posibles.

Con respecto a los trabajadores, vamos a considerar dos posibilidades: que sean homogéneos, si todos los oferentes de trabajo poseen las mismas cualidades y cualquier trabajador es sustituto perfecto de otro, o que sean heterogéneos.

También vamos a distinguir dos mecanismos alternativos de fijación de salarios: el mecanismo competitivo y el sistema de negociación colectiva. A continuación pasamos a explicar esta distinción.

Decimos que los salarios se fijan de forma competitiva cuando éstos se determinan a través de negociaciones atomizadas. Cada empresario negocia individualmente con cada trabajador hasta que se alcanza un acuerdo. Si hay muchos negociadores (muchos empresarios que demandan trabajo y muchos trabajadores que lo ofrecen), todo el mundo se está observando mutuamente (todos los contratos son públicos), y existe la posibilidad de revisar los contratos (aunque sea después de un período de prueba), es de esperar que todas las negociaciones converjan en torno a un

salario que es el que llamamos “salario de equilibrio competitivo”. El mecanismo de las negociaciones atomizadas garantiza la consecución de un equilibrio entre oferentes y demandantes en el mercado de trabajo: todo el que quiere trabajar (al salario de equilibrio competitivo) encuentra trabajo, y todo empresario que desee contratar a alguien (a ese salario) encuentra empleados disponibles.

El otro mecanismo de fijación de salarios que vamos a considerar aquí es el de la negociación colectiva. En este caso el salario se fija a través de un único proceso negociador en el cual participan dos grupos separados de personas: los trabajadores (o sus representantes) por un lado y los empresarios (o sus representantes) por otro lado. En realidad, la negociación colectiva no tiene por qué centrarse únicamente en los salarios; puede referirse a muchas otras cuestiones. Por ejemplo, se puede negociar la duración de las vacaciones, la seguridad en el trabajo (sobre todo si se trata de trabajos arriesgados), etc. Y desde luego también se puede negociar el nivel de empleo de toda una industria o el número de empleados que tiene que contratar cada empresario en particular. A pesar de todo esto, el análisis de esta tesis va a estar basado en el supuesto de que el único aspecto verdaderamente relevante de la negociación colectiva es la cuestión salarial: los convenios colectivos fijan unos salarios que sirven de referencia para cada empresa individual y para cada trabajador individual. Al margen de este dato, podemos decir que el mercado de trabajo funciona como un mercado competitivo. Cada trabajador decide si trabaja o no trabaja al salario negociado y cada empresa decide cuántos trabajadores va a contratar a la vista del salario negociado. Es decir, siguen existiendo las funciones de oferta y demanda de trabajo típicas de cualquier mercado competitivo.

Para justificar esta visión simplificadora del sistema de negociación colectiva recurriremos a las siguientes consideraciones⁴:

En primer lugar podemos argumentar que cuando hablamos de salarios no solamente nos referimos a las remuneraciones monetarias. También existen “salarios psicológicos” o remuneraciones no monetarias. Las vacaciones, los horarios, las condiciones de seguridad, etc. son formas de remuneración no monetaria. Podemos pues concebir el salario como un vector con varios elementos. Con el fin de simplificar nuestro análisis nosotros optamos por reducir todos esos elementos a una única dimensión monetaria que llamamos “salario negociado”.

No obstante lo dicho en el párrafo anterior, es preciso reconocer que no todos los contenidos de la negociación colectiva admiten una reducción tan simplista. En particular, cuando la negociación se hace extensiva al empleo –es decir, cuando el convenio colectivo no solamente establece cuánto hay que pagar a cada trabajador sino cuánta gente tiene que contratar cada empresario– el argumento “reduccionista” del párrafo anterior no resulta en absoluto convincente.

El problema que acabamos de exponer constituye sin duda un problema grave desde el punto de vista teórico. Por eso, para resolver dicho problema, han surgido los modelos de “negociación eficiente”⁵. Tales modelos implican una visión del mercado laboral muy cercana a la idea del monopolio bilateral y no nos permiten analizar el

⁴ La discusión de esta cuestión sigue la línea argumental de Sánchez Molinero (2000)

⁵ La literatura sobre esta cuestión es muy amplia. Entre las referencias más antiguas puede citarse Leontief (1946). Otras referencias relativamente más recientes serían McDonald y Solow (1981), Chatterjee (1982), Grossman y Hart (1983), Hart (1983), Oswald (1984), Brown y Ashenfelter (1986), y Macurdy y Pencavel (1986).

funcionamiento de dicho mercado por medio de curvas convencionales de oferta y demanda. No obstante lo dicho, desde el punto de vista de esta tesis, el problema en cuestión no parece tener demasiada importancia y puede ser ignorado como vamos a ver a continuación.

Para entender esto último hay que tener en cuenta ciertas peculiaridades del sistema español de negociación colectiva que ya fueron expuestas en el capítulo anterior. En particular, conviene recordar la distinción entre convenios de empresa y convenios sectoriales. Estos últimos son por lo general de ámbito provincial y son los más importantes en términos agregados. Los asalariados españoles, en su inmensa mayoría, están cubiertos por convenios de esta clase. Los convenios de empresa son ciertamente más numerosos que los sectoriales, pero, como ya vimos en el capítulo anterior, cubren a un porcentaje muy reducido de asalariados. Concretamente, de acuerdo con los datos que proporciona el Ministerio de Trabajo y Asuntos Sociales para el año 2001, únicamente el 10.9% de los trabajadores está sujeto a convenios de empresa, mientras que el 88.4% lo está a convenios sectoriales. Estos datos ponen de manifiesto la importancia de los convenios sectoriales en España, y justifican que, al hablar de negociación colectiva en términos generales, nos centremos exclusivamente a dichos convenios.

La distinción entre convenios de sectoriales y de empresa es relevante porque las posibilidades de “negociación eficiente” se dan solo en los primeros. Los convenios sectoriales son ciertamente los más comunes en los mercados laborales “atomizados”, es decir, cuando hay muchas empresas que demandan trabajo, y muchos trabajadores que lo ofrecen. En este contexto, es de esperar que la negociación se limite exclusivamente a los salarios y a las condiciones de trabajo ya que los acuerdos sobre cuotas de empleo

resultan inviables. Piénsese que la imposición de cuotas de empleo a las empresas particulares exigiría un control muy exhaustivo de la actividad empresarial por parte de los sindicatos y sobre todo un sistema de sanciones que difícilmente podrían hacerse efectivas.

Podemos afirmar por lo tanto que los convenios sectoriales sirven fundamentalmente para negociar salarios. Y en la medida en que exista una presión salarial excesiva en este tipo de convenios⁶ (es decir, en la medida en que los salarios negociados sean superiores a los competitivos) hay que pensar que dicha presión va a dar lugar a un cierto nivel de desempleo⁷.

3.2.2.- Los cuatro escenarios

Teniendo en cuenta lo dicho hasta aquí podemos proceder a la construcción de los cuatro escenarios que van a servir de base para nuestro análisis de la convergencia de las tasas de paro regionales. A continuación pasamos a describir dichos escenarios.

⁶ Hay razones para creer que el sistema de convenios tiende a generar una presión salarial excesiva. Para entender dichas razones hay que entrar un poco en el campo de la elección social y analizar cuáles son los intereses individuales de los representantes de los trabajadores. Es de esperar que estos representantes actúen de acuerdo con los intereses de sus votantes, que son fundamentalmente los trabajadores empleados, ya que los desempleados no votan en las elecciones sindicales. A los empleados les interesa mucho más asegurarse una elevación salarial – un beneficio inmediato – que evitar el posible riesgo de caer en el desempleo, ya que el votante no percibe este riesgo como una amenaza directa. Por esto, en los convenios colectivos es probable que predomine más el interés por conseguir aumentos salariales que la preocupación por el desempleo.

⁷ Habría que aclarar que para que la negociación colectiva dé lugar a desempleo, es necesario que esté bastante generalizada. Si unos sectores están sujetos a negociación colectiva y otros no, los excesos de oferta generados en los primeros serán automáticamente absorbidos por los segundos. En este caso, la única consecuencia de la negociación colectiva sería un aumento de la dispersión salarial entre los sectores con salarios negociados y los no sujetos a convenios. En España la negociación colectiva afecta prácticamente a la totalidad de los asalariados, con la excepción de los funcionarios y los trabajadores sometidos al salario mínimo legal.

3.2.2.a. Primer Escenario

El primer escenario, y también el más restrictivo en cuanto a los supuestos, está caracterizado por la homogeneidad de la mano de obra y por la fijación competitiva de los salarios. La homogeneidad de los trabajadores nos indica que todos son perfectamente sustitutivos entre sí y que tienen el mismo nivel de cualificación. Por lo tanto, existe un único mercado de trabajo dentro de cada región. Por otra parte, la fijación competitiva de salarios nos permite esperar que siempre se va a alcanzar el equilibrio entre la oferta y la demanda de trabajo⁸.

En este caso es evidente que el único desempleo que puede surgir es el denominado friccional. Este tipo de desempleo tiene su origen en los desajustes o fricciones que se producen dentro del mercado de trabajo, y que están asociados a los procesos de búsqueda de empleo⁹.

Naturalmente, esto nos lleva a preguntarnos qué son y de qué dependen esas fricciones. En este sentido, hay que señalar que siempre existe un lapso de tiempo entre el momento de comenzar la búsqueda de un empleo y el momento en que éste se encuentra. Aunque en este primer escenario estamos suponiendo la homogeneidad de la mano de obra, y la perfecta sustituibilidad de los trabajadores, las ofertas y las

⁸ Desde un punto de vista teórico, puede darse el caso de que la fijación competitiva de los salarios no vacíe el mercado. Este caso extremo se produciría cuando para un salario igual a cero siguiera existiendo un exceso de oferta de trabajo. Gráficamente esto querría decir que la oferta y la demanda de trabajo no se cortan para ningún salario positivo. Esta situación parece muy poco probable desde el punto de vista empírico, ya que para que se diera una situación de este tipo tendríamos que negar las posibilidades de sustitución entre factores y también la posibilidad de sustituir unas producciones por otras.

⁹ Un resumen muy completo de los aspectos más importantes de la teoría de la búsqueda y de sus implicaciones con respecto a la cuestión del desempleo se puede encontrar en el capítulo dedicado a Mortensen (1986), que aparece como capítulo 15 del libro colectivo editado por Ashenfelter y Layard (1986).

demandas de trabajo no suelen estar perfectamente sincronizadas en el tiempo; y esto hace que los trabajadores tengan que permanecer ciertos períodos de tiempo desempleados, lo cual implica una cierta tasa de paro, que hemos denominado friccional. Detrás de esta terminología subyace el supuesto de que, en un mercado de trabajo competitivo, siempre va a haber un empleo disponible de manera más o menos inmediata para cualquiera que esté dispuesto a desempeñarlo.

En suma, es de esperar que haya fricciones, aunque los trabajadores sean perfectamente homogéneos. Pero estas fricciones pueden ser de mayor o menor magnitud, es decir, la duración de los procesos de búsqueda puede variar, y el número de buscadores (desempleados) puede ser más alto o más bajo. ¿De qué depende todo esto? Aquí nos vamos a fijar en dos factores principales: la estructura sectorial de cada región y los denominados factores “hedónicos”.

El grado de sincronización entre las ofertas y las demandas de trabajo no es el mismo en todos los sectores. Por eso la composición sectorial de la economía puede influir sobre la duración de los procesos de búsqueda y sobre la cuantía de paro friccional. Es perfectamente posible que la tasa de paro (el porcentaje de trabajadores desocupados en proceso de búsqueda) varíe de unas regiones a otras cuando la composición sectorial no es igual en todas las regiones.

Por otro lado están los factores hedónicos, que hacen referencia a los atractivos especiales de cada región. Aspectos tales como la calidad de las infraestructuras, la aplicación de una adecuada política medioambiental, o la existencia de un servicio adecuado de asistencia sanitaria, entre otros, hacen que una determinada zona o región sea más o menos deseable a la hora de desarrollar una actividad laboral.

Este tipo de factores suelen verse reflejados en las diferencias salariales. Los trabajadores en general pueden estar dispuestos a permanecer en un escalón salarial inferior, si se les ofrecen a cambio, compensaciones que mejoren su calidad de vida. Estas compensaciones pueden ser de tipo medioambiental, asistencial, o de cualquier otra clase. Pero los factores hedónicos también pueden influir sobre el nivel de paro friccional en la medida en que pueden aumentar o disminuir el tiempo de búsqueda: pueden aumentarlo, pueden aumentarlo si tenemos en cuenta que una mayor calidad de vida nos permite esperar durante más tiempo la llegada de un trabajo idóneo; o pueden disminuirlo, si consideramos que ciertas mejoras, por ejemplo, en el sistema de información o en las infraestructuras, nos facilitan el acceso a un mayor número de empleos alternativos.

En suma, podemos afirmar que, en este primer escenario, solamente va a existir un paro de carácter friccional, que es el único compatible con la fijación competitiva de los salarios. Por otro lado, es posible que existan divergencias entre las tasas de paro regionales, causadas éstas por la distinta composición sectorial de cada región, y por las diferencias que tienen que ver con la dotación de infraestructuras y con la calidad de vida en general.

3.2.2.b Segundo Escenario

Para la caracterización de este segundo escenario, se va a seguir manteniendo el supuesto de homogeneidad de la mano de obra. Sin embargo, ahora vamos a considerar que los salarios se fijan mediante negociación colectiva. En el escenario anterior la homogeneidad de la mano de obra implicaba que iba a haber un único mercado de

trabajo para todas las regiones. Sin embargo, la existencia de negociación colectiva hace que esto deje de ser así. El sistema de convenios colectivos reduce la flexibilidad salarial y genera salarios superiores a los competitivos; además, como estos salarios no tienen por qué ser iguales en todas las regiones, puede ocurrir que el mercado de trabajo quede fraccionado, tanto a escala regional como sectorial.

En este segundo caso también va a existir desempleo, pero ahora el paro será de dos clases. Además del desempleo friccional que ya aparecía en el primer escenario, surge un nuevo tipo de desempleo que, en contraste con el anterior, es básicamente involuntario¹⁰. Este nuevo tipo de paro está asociado a los excesos de oferta y es una consecuencia directa del sistema de fijación de salarios.

A continuación, podemos intentar resumir cuáles son los principales determinantes de las tasas de paro regionales en este segundo escenario.

Por un lado, el hecho de que una parte del desempleo sea friccional, nos permite seguir manteniendo como factores explicativos los dos que ya aparecían en el primer escenario: la composición sectorial que tenga cada región y los factores hedónicos relevantes en cada zona.

¹⁰ La existencia de excesos de oferta con carácter permanente implica que no todos los buscadores de trabajo van a poder acceder a un empleo. Por eso se trata de un desempleo involuntario. Sin embargo, cuando consideramos el paro desde un punto de vista friccional y permitimos que transcurra el tiempo suficiente para que tanto los oferentes como los demandantes puedan reajustar sus planes, es de esperar que, todo el que busca un empleo acabe encontrándolo en algún momento. Por eso el paro friccional parece más voluntario que involuntario.

Por otro lado, el desempleo tiene ahora un componente asociado a la magnitud del salario negociado en cada región. Este hecho nos permite incluir un tercer factor explicativo en nuestro análisis: las diferencias regionales en el grado de presión sindical.

Teniendo en cuenta lo que antes se dijo acerca de las características de la negociación colectiva, parece bastante evidente que la presión salarial de los sindicatos incide directamente sobre la tasa de desempleo. Cuanto más poder tengan las organizaciones sindicales en la negociación, tanto mayor será la diferencia entre el salario negociado y el de equilibrio competitivo (el que se generaría de manera espontánea, en ausencia de negociaciones). Por lo tanto, un mayor poder de los sindicatos se va a traducir en un mayor exceso de oferta de trabajo y en una mayor cantidad de desempleo involuntario.

Vamos a terminar el análisis de este segundo escenario extrayendo dos conclusiones principales.

En primer lugar, es de esperar que exista un cierto nivel de paro, que presumiblemente será superior al del primer escenario: además del desempleo provocado por la falta de sincronización entre ofertas y demandas de trabajo, existirá otra clase de desempleo, básicamente involuntario, que es consecuencia de unos salarios excesivamente altos: para el salario del convenio, son más los que quieren trabajar que los empleos disponibles.

En segundo lugar, en este escenario seguirá habiendo diferencias entre las tasas de paro regionales. Y las causas de estas diferencias son los tres factores antes

apuntados: la composición sectorial de cada región, los llamados factores hedónicos y las diferencias en la presión sindical sufrida por cada zona.

3.2.2 c Tercer Escenario

Este tercer escenario va a ser totalmente distinto al anterior. En primer lugar, vamos a suponer que los trabajadores son heterogéneos, y en segundo lugar, vamos a volver al supuesto inicial de fijación competitiva de los salarios.

El supuesto de heterogeneidad de los trabajadores refleja la existencia de distintos niveles de formación, de cualificación, o de capacidad dentro de la mano de obra. La principal implicación de este supuesto es que los trabajadores dejan de ser perfectamente sustitutivos, y por lo tanto, existen distintos mercados de trabajo en función de cuáles sean las características de los trabajadores que concurren en cada uno de ellos. Por otra parte, la fijación competitiva de los salarios, como ya vimos en el primer escenario, nos garantiza la consecución del equilibrio entre la oferta y la demanda de trabajo¹¹.

En este caso volvemos a tener únicamente un paro de carácter friccional. La flexibilidad de los salarios garantiza el vaciado del mercado de trabajo, y la única causa que puede generar desempleo son las fricciones relacionadas con los procesos de búsqueda. Pero estas fricciones no solamente tienen que ver con la falta de sincronización entre las ofertas y las demandas de trabajo como ocurría en los escenarios anteriores; también tienen que ver con la heterogeneidad del trabajo.

¹¹ Ver nota explicativa 6.

El hecho de que el trabajo sea heterogéneo introduce ahora un elemento de incertidumbre que magnifica las fricciones: cuando surge una vacante y hay que elegir entre varios aspirantes, antes de seleccionar a uno hay que investigar cuál es el mejor. Del mismo modo, cuando un trabajador busca un empleo entre varias opciones disponibles, tiene que investigar antes de decidir cuál es el más adecuado a sus cualificaciones y el que está más de acuerdo con sus aspiraciones.

Por todas estas razones podemos afirmar que la composición de la fuerza de trabajo en cuanto a niveles de cualificación, experiencia, etc., influye de alguna manera sobre el nivel de paro friccional. Las diferencias interregionales en lo que respecta a tasas de desempleo dependerán pues, no solamente de la composición sectorial y de los factores hedónicos de cada región; también dependerán de las dotaciones de capital humano –la estructura de cualificaciones de la mano de obra– de cada zona.

3.2.2.d. Cuarto Escenario

El cuarto escenario que vamos a describir, se caracteriza de la siguiente manera: heterogeneidad de la mano de obra, y fijación de salarios mediante negociación colectiva. Estos supuestos tienen dos implicaciones fundamentales: se produce una fragmentación del mercado de trabajo en función de las características que presenten los trabajadores, y es de esperar que los salarios negociados en cada uno de los mercados estén por encima de los niveles de equilibrio correspondientes.

La caracterización de este último escenario hace que nos encontremos nuevamente con dos tipos de desempleo: el friccional que ha aparecido con carácter

general en todos los casos estudiados, y el paro involuntario, al que ya nos hemos referido en el segundo escenario, y que es consecuencia de la falta de flexibilidad en el precio de la mano de obra.

Lo dicho anteriormente implica que a la lista de causas del desempleo identificadas en el escenario anterior (la composición sectorial, la estructura de cualificaciones de la mano de obra y los factores hedónicos de cada región), tenemos que agregar una más: el grado de presión salarial que ejercen los sindicatos de cada área geográfica en los procesos de negociación colectiva. En adelante nos referiremos a esa variable como “presión salarial”.

3.2.2.e. Conclusiones

A lo largo de todo este epígrafe hemos pretendido caracterizar los distintos escenarios que se nos pueden presentar a la hora de estudiar el desempleo y sus diferencias regionales, y en cada uno de ellos hemos intentado extraer las principales conclusiones al respecto.

El primer resultado que hemos observado es la existencia de un cierto nivel de desempleo, independiente del escenario seleccionado, que surge como consecuencia de los desajustes inherentes al mero funcionamiento del mercado de trabajo. Este desempleo, que hemos denominado friccional, puede tener una mayor o menor magnitud dependiendo de las particularidades que presente cada zona (la estructura sectorial, la dotación de infraestructuras y demás factores hedónicos, y la mezcla de cualificaciones de los trabajadores), pero nunca va a desaparecer del todo.

En segundo lugar, esa tasa de paro friccional puede verse incrementada por otra, de carácter involuntario, que es la propia de los escenarios donde los salarios se fijan a través de negociación colectiva. Este nuevo tipo de paro también puede ser de mayor o menor magnitud, dependiendo de las características de cada región, pero además, va a estar fuertemente condicionado por el poder negociador que tengan los sindicatos en cada zona.

3.3.- ¿Convergencia de las tasas de paro regionales?

Una vez descritos los distintos escenarios que nos podemos encontrar y conocido el comportamiento del empleo en cada uno de ellos, el siguiente paso es dar respuesta a la pregunta central de este trabajo: ¿Podemos esperar que exista un proceso de convergencia entre las tasas de paro regionales?.

La respuesta inmediata a esa pregunta, teniendo en cuenta el análisis desarrollado en el apartado anterior, es que no. En todos los escenarios analizados se ha visto que hay una serie de factores, cuya desigual incidencia según regiones, provoca diferencias permanentes en las tasas de paro regionales. Hemos visto que la estructura sectorial de una región afecta directamente a la magnitud de las fricciones que surgen en ese mercado de trabajo; y también que la mayor calidad de vida de una determinada zona desincentiva el flujo de parados hacia otras regiones con menos desempleo. También hemos visto que la existencia de salarios negociados a través de convenios altera los incentivos para la movilidad del trabajo, y esto repercute sobre las tasas de paro regionales.

No obstante, aún asumiendo la ausencia de un proceso de convergencia generalizada, sí que es posible que se produzca un cierto acercamiento entre las tasas de paro regionales cuando nos centramos en grupos poblacionales muy concretos: Cuanto más homogéneo sea el grupo considerado tanto menores serán las fuentes de distorsión. Por ejemplo, si nos centramos en la población con estudios superiores, estamos eliminando el nivel de cualificación como factor causante de disparidades entre las tasas de desempleo regionales; y será de esperar un mayor grado de convergencia que cuando se comparan poblaciones heterogéneas desde el punto de vista de la formación académica.

En este sentido, según vamos avanzando en la homogeneización de la población estudiada, parece más viable la existencia de un proceso de convergencia entre las regiones. Por ejemplo, si suponemos que nos encontramos en el escenario tres, y analizamos la población con estudios superiores, parece lógico suponer que se producirá un acercamiento entre las tasas de paro de las regiones que tengan una composición sectorial parecida y un nivel de vida similar. Si nos encontrásemos en el escenario cuatro, y nos centramos en el análisis de la población más cualificada de un sector de producción concreto, también sería de esperar un proceso de convergencia entre las tasas de paro de las regiones con presión sindical y nivel de vida parecido.

Naturalmente, si hubiera otras fuentes de distorsión que no estuvieran reflejadas en los escenarios anteriormente descritos, el grado de convergencia sería menor.

Estas son las cuestiones que intentaremos corroborar a través del análisis empírico que vamos a desarrollar en el capítulo cinco. En ese capítulo, a partir de una homogeneización progresiva del grupo estudiado, intentaremos demostrar el aumento

progresivo del número de regiones que convergen en tasas de paro con el correspondiente agregado nacional.



BIBLIOTECA VIRTUAL

CAPÍTULO 4:

**MÉTODOS PARA EL ANÁLISIS DE CONVERGENCIA EN
TÉRMINOS EMPÍRICOS**

4.1.- Introducción

En este capítulo intentaremos dar una visión de los métodos más significativos o más frecuentemente utilizados en el marco del análisis de convergencia. Nos referiremos a cinco métodos principales: los dos inspirados en la noción de β -convergencia (β -convergencia absoluta y β -convergencia condicionada); el método conocido como σ -convergencia; el método (o métodos) de cointegración; y finalmente el análisis de descomposición de la varianza.

Trataremos de explicar las ideas centrales en las que se basa cada uno de los métodos mencionados. En los apartados siguientes iremos viendo en qué tipo de contextos se ha empleado cada uno de ellos, y eso servirá para situar esta tesis dentro del panorama de la literatura económica relevante. Al tiempo que exponemos todo esto puntualizaremos los inconvenientes de cada método y trataremos de justificar por qué hemos seleccionado unos y hemos rechazado otros.

4.2.- La hipótesis de β -convergencia absoluta

En términos puramente formales la hipótesis de β -convergencia absoluta se contrasta mediante la estimación de la siguiente ecuación:

$$\ln\left(\frac{y_{i,t}}{y_{i,t-1}}\right) = \alpha + \beta \cdot \ln(y_{i,t-1}) + \varepsilon_t$$

donde $y_{i,t}$ e $y_{i,t-1}$ representan los valores de la variable que, según se presupone, debería converger a través de un conjunto de regiones, países o zonas geográficas designadas por el subíndice i . Esta variable podría ser la renta per capita, la tasa de desempleo o cualquier otra; y lo que la expresión anterior dice es que la tasa de crecimiento de esa variable en un momento determinado, t , depende linealmente del valor de dicha variable en el momento inmediatamente anterior. Se espera que cuanto más cerca se encuentre la variable de su valor de equilibrio estacionario, tanto más lento será su ritmo de crecimiento, de tal modo que ese ritmo irá disminuyendo hasta que la variable en cuestión se estabilice.

Dicho de otra forma, cuando la variable y deja de crecer se verifica que

$$\ln\left(\frac{y_{i,t}}{y_{i,t-1}}\right) = 0$$

lo cual implica que

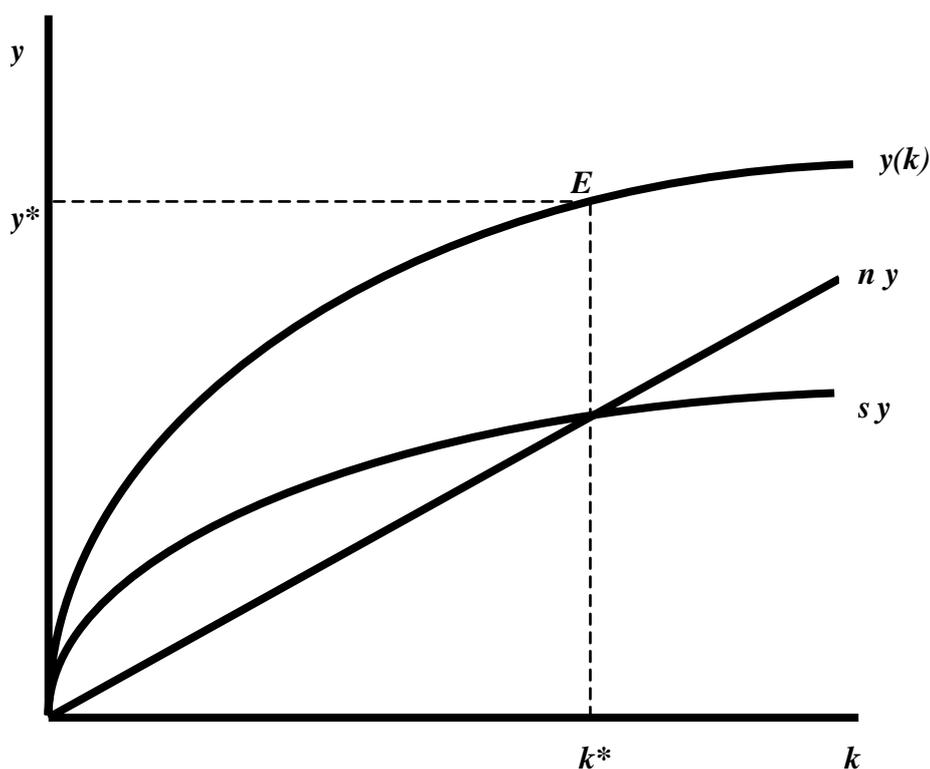
$$\ln(y_{i,t-1}) = \frac{\alpha}{(-\beta)}$$

que es el valor de equilibrio estacionario de $\ln y$. Se espera pues que β sea negativo y que exista una relación inversa entre el valor que tome la variable y su tasa de crecimiento. Aunque en un momento dado $\ln y$ varíe de unas zonas a otras, a lo largo del tiempo se espera que dicho valor se iguale a $(-\alpha/\beta)$ en todas partes. En esto consiste la β -convergencia absoluta.

La hipótesis de β -convergencia absoluta se podría defender en referencia a las rentas per capita de una serie de países, siempre que dichos países presenten unas características básicas muy similares; y para justificar este punto de vista podríamos acudir al modelo neoclásico tradicional de crecimiento económico.

Para aclarar esta cuestión, supongamos que tenemos un conjunto de economías, y que todas ellas comparten la misma función de producción, la cual depende exclusivamente de dos factores productivos (trabajo y capital), y presenta rendimientos a escala constantes. Supongamos también que no existe progreso tecnológico. De acuerdo con el modelo neoclásico de crecimiento, el comportamiento de este tipo de economías se podría describir por medio de la FIGURA 4.1, donde y es el nivel de renta per capita, k es la relación capital/trabajo, s es la propensión al ahorro y n la tasa de crecimiento de la población.

FIGURA 4.1



En estas condiciones, si suponemos que todas las economías tienen la misma tasa de crecimiento de la población, n , y la misma propensión al ahorro, s , sería previsible que todas ellas acabaran en el mismo estado estacionario, E . Por lo tanto, el resultado que se produciría sería la convergencia en términos de renta per capita en el nivel y^* . Este nivel de equilibrio de rentas se conseguiría porque todas las economías ajustarían su cantidad de capital por trabajador al valor de equilibrio a largo plazo, k^* .

De acuerdo con este modelo, las economías únicamente pueden diferir en su situación inicial –el valor inicial de k –, así como en los valores de la tasa de crecimiento de la población y de la propensión al ahorro. Aunque la situación inicial fuera diferente, el horizonte final –el estado estacionario– sería el mismo, siempre que las tasas de crecimiento de la población y las propensiones al ahorro fuesen iguales en todas las economías.

Cuando se acude al modelo neoclásico de crecimiento para justificar la hipótesis de convergencia, hay que ser conscientes de que se trata de un modelo de economía cerrada. Esto quiere decir que la convergencia se produce como consecuencia de que todos los países comparten ciertos rasgos dinámicos básicos, no a través del comercio internacional.

Ha habido algunos intentos de verificación de la hipótesis de β -convergencia absoluta aplicada a las rentas per capita de algunos países. Como ya hemos indicado antes, esta hipótesis puede tener sentido en el caso de las rentas per capita, si suponemos que todos los países tienen acceso a las mismas tecnologías (misma función de producción), si sus poblaciones respectivas crecen más o menos a la misma velocidad y si además las tasas de ahorro son parecidas en todos ellos. Desde una perspectiva muy

agregada –haciendo abstracción de toda clase de consideraciones referentes a la estructura sectorial o a la estructura de la población de las distintas economías– la hipótesis de β -convergencia absoluta podría ser una conjetura razonable. Si lo es o no lo es solo se puede saber a la vista de los resultados del análisis empírico.

Los trabajos de Barro y Sala i Martín (1990, 1991, 1992) son tal vez los puntos de referencia principales de la literatura empírica sobre crecimiento y convergencia. En el CUADRO 4.1 (basado en el que aparece en Barro y Sala i Martín (1992)) se recoge un resumen de los resultados obtenidos por estos autores para tres muestras distintas de datos de transversales: una muestra muy amplia con 98 países, otra con los 20 miembros originales de la OCDE, y una tercera con los 48 estados continentales de Estados Unidos.

CUADRO 4.1
Estimación de la β -convergencia absoluta

	B	R²
98 Países	0.0037	0.04
OCDE	-0.0095	0.45
Estados USA	-0.0218	0.38

El CUADRO 4.1 nos muestra evidencia acerca de la hipótesis de β -convergencia absoluta aplicada a un conjunto de zonas geográficas presumiblemente homogéneas. Se puede apreciar que el grado de convergencia es mayor cuando tenemos una muestra de regiones pertenecientes a un mismo país que en el caso de países diferentes.

Presumiblemente, el grado de homogeneidad también es mayor en el primer caso que en el segundo.

Centrándonos en el caso español, también hay que señalar la existencia de trabajos que contrastan la hipótesis β -convergencia absoluta tanto a nivel provincial como en el caso de las Comunidades Autónomas. En el CUADRO 4.2 se recogen los resultados de tres trabajos significativos: Dolado *et al.* (1994) que se centra en la evidencia por provincias, y Mas *et al.* (1994) y Villaverde (1999), que analizan la convergencia entre Comunidades Autónomas.

BIBLIOTECA VIRTUAL
CUADRO 4.2
Estimación de la β -convergencia absoluta

	B	Error std.
Dolado <i>et al.</i>	-0.02	0.0044
Mas <i>et al.</i>	-0.0194	0.0052
Villaverde	-0.02	0.020

Nuevamente volvemos a comprobar que sigue produciéndose un cierto grado de convergencia, y además, que se repite el valor (cercano al 2%) que ya obtenían Barro y Sala en el caso de los estados continentales de USA. Este resultado del 2%, como ya veremos más adelante, es consecuencia de imponer la condición de que todas las regiones tienden hacia un mismo estado estacionario, y a que estamos suprimiendo variables presumiblemente relevantes que reflejen las peculiaridades de las distintas regiones. En la medida en que esto último sea verdad, cabe esperar una subestimación del coeficiente β .

Todos estos problemas se podrían corregir si se sustituyera la hipótesis de β -convergencia absoluta por la de β -convergencia condicionada, la cual reconoce la posibilidad de que existan varios estados estacionarios, y permite tener en cuenta las peculiaridades de cada región.

Si en lugar de fijarnos en las rentas per capita nos fijamos en las tasas de paro y nos preguntamos por la convergencia de las tasas de paro regionales de un país como España, la hipótesis de la β -convergencia absoluta no parece tener demasiado sentido.

La razón de esto resulta bastante clara a la luz de los argumentos esgrimidos en el capítulo anterior: las regiones de un país están conectadas a través de una serie de mercados comunes, y esto implica cierta tendencia a la homogeneidad; no obstante, en la medida en que cada región tiene una estructura sectorial distinta, y en la medida que las poblaciones de las distintas regiones no tienen exactamente la misma composición por edades, cualificaciones etc., es de esperar que las tasas de desempleo “de equilibrio estacionario” varíen de unas zonas a otras. En el capítulo primero hemos visto también que no hay ninguna indicación de que las tasas de paro tiendan a igualarse entre las regiones españolas.

Por estas razones, en esta tesis no nos plantearemos siquiera la hipótesis de β -convergencia absoluta en relación con las tasas de paro regionales.

4.3.- La hipótesis de β -convergencia condicionada

Esta nueva hipótesis de convergencia se contrasta mediante la estimación de una ecuación del tipo siguiente:

$$\ln\left(\frac{y_{i,t}}{y_{i,t-1}}\right) = \alpha + \beta \cdot \ln(y_{i,t-1}) + \gamma \cdot X_i + \mu_t$$

Este tipo de convergencia es menos estricto que la β -convergencia absoluta, y además permite que se introduzcan variables explicativas adicionales que pueden resultar relevantes a la hora de identificar el proceso de convergencia. La diferencia entre esta hipótesis y la anterior es que ahora estamos admitiendo la posibilidad de que el estado estacionario de la variable y varíe de unas zonas a otras. Esto es así porque en el estado estacionario se tiene que cumplir que

$$\ln\left(\frac{y_{i,t}}{y_{i,t-1}}\right) = 0$$

lo cual implica que

$$\ln(y_{i,t-1}) = \frac{\alpha + \gamma \cdot X_i}{(-\beta)}$$

Es evidente que el estado estacionario de la variable y en cada zona geográfica i dependerá de una serie de variables, que serán las incluidas el vector X_i , y que pueden

variar de unas zonas a otras. Esta hipótesis nos permite también la identificación de “clubes de convergencia” (Villaverde, 1999), que se podrían definir como grupos de países o regiones de características parecidas (mismos valores de X_i), dentro de las cuales la variable y tendería a un mismo equilibrio estacionario.

La hipótesis de β -convergencia condicionada también puede defenderse en relación con las rentas per capita de un conjunto de países, aunque ahora ya no exigimos a priori que todas las rentas tiendan a alcanzar exactamente el mismo valor. Admitimos que ciertos países pueden ser “más parecidos” que otros; y por tanto esperamos que en cada grupo de países haya una renta de equilibrio estacionario distinta.

Para justificar teóricamente esta hipótesis podemos acudir a los modelos de crecimiento endógeno. Estos modelos introducen dos novedades respecto al modelo neoclásico tradicional: en primer lugar se admite la existencia de rendimientos a escala crecientes; y en segundo lugar se utiliza un concepto mucho más amplio de capital, que incluye el acervo de conocimientos –capital tecnológico o capital humano– como componente fundamental. En estos modelos los incrementos de la productividad del trabajo pueden deberse no solamente a un aumento del capital por trabajador, sino también a otros factores tales como la educación o las infraestructuras.

Los modelos de crecimiento endógeno permiten explicar la existencia de ciertas divergencias persistentes –o incluso crecientes– entre países o regiones. Esto estaría relacionado con la existencia del llamado “capital tecnológico”, el cual puede dar lugar a rendimientos crecientes: una mayor dotación de capital indica un nivel tecnológico más alto y cuanto más tecnología se posee más fácil es innovar; por eso, al aumentar el stock de capital no tiene por qué disminuir la productividad del trabajo sino que ésta

puede aumentar. En este contexto, podría ocurrir que el crecimiento económico agrandara la “brecha tecnológica” entre unas economías y otras, y eso haría que aumentaran las diferencias de renta per capita entre las mismas.

A pesar de todo, también en estos modelos se pueden encontrar razones a favor de un cierto grado de convergencia. Los avances tecnológicos tienen propiedades de bien público, y eso hace que los países menos desarrollados se beneficien de los avances obtenidos en las zonas más industrializadas; es decir, los países más pobres pueden acceder de forma relativamente rápida y fácil a las innovaciones sin tener que hacer frente a los mismos gastos de inversión inicial que tuvieron que realizar los países avanzados. Así pues, a medida que las innovaciones se van difundiendo, sería de esperar un cierto grado de convergencia.

La hipótesis de β -convergencia condicionada también se podría defender en relación con las tasas de desempleo de las distintas regiones de un país. Para justificar esto no es preciso recurrir a ningún modelo de crecimiento endógeno. Nos basta con admitir que se trata de economías diversificadas (muchos sectores) y además muy interconectadas unas con otras. En estas condiciones, cabe esperar que los procesos naturales de integración de los mercados regionales impongan una cierta tendencia a la homogeneización de las tasas de paro dentro de ciertos límites. Desde luego, no tendría demasiado sentido esperar la convergencia absoluta a la luz de los argumentos desarrollados en el capítulo anterior. En realidad, dichos argumentos podrían interpretarse como argumentos a favor de la hipótesis de la convergencia condicionada de las tasas de desempleo.

Son muchos los trabajos empíricos que han recurrido a la hipótesis de β -convergencia condicionada, sobre en el caso de las rentas per capita, tanto de países como de regiones. En la mayoría de estos trabajos se observa que la introducción de variables adicionales implica un aumento en el valor de β y por lo tanto un aumento en la velocidad de convergencia de cada zona hacia su correspondiente estado estacionario. Barro y Sala i Martín (1992) y Mankiw *et al.* (1992) demuestran, para distintas muestras de países, que, cuantas más variables explicativas se introducen en el análisis con el fin de poder distinguir mejor entre los estados estacionarios de los distintos países, (a) tanto mayor es la velocidad de aproximación a dicho estado; y (b) esta velocidad se va acercando al 2%. De la Fuente (1993) obtiene un resultado en la misma línea usando muestras distintas a las anteriores.

En cuanto a la β -convergencia condicionada entre regiones, el trabajo de Sala i Martín (1994) proporciona un resumen de la evidencia acumulada sobre el tema, tanto para Estados Unidos como para varios países europeos, además de Canadá y Japón. En ese trabajo también se observa que la velocidad de convergencia sigue situándose en torno al 2%.

Finalmente, y refiriéndonos al caso español, los principales trabajos siguen siendo los de Dolado *et al.* (1993), referido a las provincias, y Mas *et al.* (1993) y Villaverde (1999), que se refieren a las Comunidades Autónomas. En estos trabajos, además de la hipótesis de β -convergencia absoluta, se investiga también la de β -convergencia condicionada, demostrándose que en este segundo caso la velocidad de convergencia es mayor que en el primero.

Para terminar con esta revisión de la literatura empírica sobre β -convergencia condicionada sería conveniente referirnos a las principales variables que se han introducido para controlar la posible existencia de distintos estados estacionarios. En este sentido podemos decir que las más comunes han sido las ficticias regionales (o nacionales dependiendo del ámbito del trabajo), aunque en muchos trabajos también se introducen *proxies* referentes a la productividad, la inversión o el nivel de cualificación de la mano de obra.

A pesar de sus ventajas frente a la hipótesis de β -convergencia absoluta, la hipótesis β -convergencia condicionada tiene algunos inconvenientes que nos han llevado a no tenerla en cuenta en esta tesis. El principal inconveniente tiene que ver con el hecho de que este tipo de análisis se basa en datos de corte transversal y tiende a esconder la variabilidad de las tasas de desempleo regionales a lo largo del tiempo. Si las tasas de desempleo regionales siguieran una evolución suave y constante a lo largo del tiempo, una visión instantánea del proceso podría ser suficiente para indicarnos si hay o no hay convergencia. Ahora bien, la evolución de las tasas de desempleo regionales no es uniformemente creciente ni decreciente. Más bien está sujeta a multitud de influencias que generan una gran variabilidad. Por tanto, cualquier instantánea podría darnos una visión distorsionada del proceso de convergencia.

Este problema –la variabilidad de las tasas de crecimiento de la variable en la cual se centra el análisis de convergencia– podría evitarse en algunos casos recurriendo a la hipótesis de σ -convergencia, que vamos a discutir en el apartado siguiente.

4.4.- La hipótesis de σ -convergencia

El análisis de σ -convergencia se basa simplemente en el cálculo de la varianza de la variable entre las distintas áreas geográficas de acuerdo a la siguiente expresión

$$\sigma_t^2 = \left(\frac{1}{N}\right) \sum_{i=1}^N \left[\ln(y_{it}) - \eta_t \right]^2$$

donde η_t es la media de los logaritmos de y_i correspondiente al año t . En este caso la convergencia se produciría si σ_t^2 presentara una clara tendencia decreciente. Esto significaría la existencia de un acercamiento progresivo de las variables hacia un mismo equilibrio estacionario, que obviamente tendría que ser igual para todos los países o regiones. Es evidente que para saber si hay o no hay convergencia necesitamos calcular toda una serie de varianzas. Estamos pues ante un enfoque mixto, en el sentido de que requiere datos de corte transversal y datos de series temporales.

Ya hemos dicho que la existencia de σ -convergencia implica que todas las y_i tiendan a un mismo equilibrio estacionario. Podemos decir por lo tanto que esta nueva hipótesis constituye una forma más de convergencia absoluta. En este sentido la σ -convergencia es similar a la β -convergencia absoluta; aunque las dos hipótesis no son exactamente equivalentes, pudiéndose demostrar (Sala-i-Martin, 1994) que la β -convergencia absoluta es una condición necesaria, pero no suficiente para que se produzca σ -convergencia. En otras palabras, la β tiene que ser negativa para que se produzca una reducción progresiva de las disparidades entre las distintas áreas geográficas, pero una β negativa no garantiza *per se* que las disparidades geográficas se

van a reducir con el tiempo. Bernard y Durlauf (1996) muestran que es posible la existencia de β -convergencia, cuando las distintas economías tienden hacia dos equilibrios diferentes –uno para los países ricos y otro para los países pobres– sin que exista σ -convergencia, ya que la dispersión que supone este doble equilibrio es mayor que la que existía inicialmente.¹²

La hipótesis de σ -convergencia ha sido ampliamente utilizada en relación con las rentas per capita. En el CUADRO 4.3 encontramos un resumen de los trabajos más significativos en esta línea. La primera columna indica el ámbito geográfico del estudio; la segunda el autor o autores; la tercera el período de estudio; y la cuarta los resultados (si hay o no hay convergencia).

Sala i Martín (1994, 1996 a, 1996 b) solo encuentra evidencia de convergencia entre los países cuando existe cierta homogeneidad entre ellos, aunque esa convergencia se frena desde 1975. En estos trabajos también se realiza un análisis para regiones de distintos países, y los resultados son muy similares: también se observa un cierto grado de convergencia y se ve que ésta se va reduciendo a partir de 1975. Esteban (1994) no encuentra evidencia de convergencia para los países de la Unión Europea, pero el periodo analizado es demasiado corto.

En el caso español, Raymond y García (1994), García, Raymond y Villaverde (1995) y Villaverde (1999) encuentran resultados muy similares tanto para las

¹² Este resultado es muy similar al recogido en la literatura con el nombre de *falacia de Galton* (Galton, 1889 y 1892). Es posible que exista una relación negativa entre el nivel y la tasa de variación de una variable, aunque puede que no se produzca una disminución de la dispersión de esa variable a lo largo del tiempo. Esta conclusión también puede encontrarse en Quah (1993a y b).

Comunidades Autónomas como para las provincias, y concluyen que existe convergencia, aunque ésta se ve frenada a partir de 1980.

CUADRO 4.3¹³
 σ -convergencia entre países y regiones

Naciones	Autor	Periodo	¿Convergencia?
110 Países	Sala i Martín	1960-1990	Divergencia
OCDE (22)	Sala i Martín	1950-1990	Convergencia
		1975-1990	Mantenimiento
Europa (12)	Esteban	1980-1989	Divergencia
Regiones	Autor	Periodo	¿Convergencia?
USA (48)	Sala i Martín	1880-1992	Convergencia
		1975-1992	Mantenimiento
Japón (47)	Sala i Martín	1930-1990	Convergencia
		1975-1990	Mantenimiento
Alemania (11)	Sala i Martín	1950-1990	Convergencia
		1975-1990	Mantenimiento
España (17)	Sala i Martín	1950-1990	Convergencia
Francia (21)	Sala i Martín	1950-1990	Convergencia
UK (11)	Sala i Martín	1950-1990	Convergencia
		1975-1990	Divergencia
España (17)	Raymond y García	1955-1991	Convergencia
		1980-1993	Mantenimiento
España (17)	Villaverde	1955-1995	Convergencia
		1981-1995	Mantenimiento
España (51)	Raymond, García y Villaverde	1955-1991	Convergencia
		1980-1991	Mantenimiento

¹³ Fuente: Gil Canaleta (2001)

También conviene señalar que la varianza (o la desviación típica) no ha sido el único indicador utilizado a la hora de estudiar la convergencia de rentas. En el trabajo de Esteban (1994) se aplican otras medidas de dispersión, también referidas a la renta per capita, tales como el índice de Gini, el de Theil, o el de Atkinson.

En esta misma línea habría que situar el trabajo de Plaza Acero (1994), en el que se investiga la evolución del “coeficiente de variación” regional de las tasas de desempleo de las provincias españolas. Dicho coeficiente de variación tiene la ventaja de ser una medida de dispersión *relativa* y por tanto no se ve afectada por el valor absoluto de las tasas de desempleo provinciales; en este sentido se puede decir que constituye una medida de dispersión “más pura” que la simple varianza. Plaza Acero apunta la existencia de una cierta tendencia hacia la reducción de las diferencias provinciales de desempleo a la vez que subraya el carácter cíclico de las mismas (se hacen mayores en los auges y menores en las recesiones). Este mismo enfoque ha sido utilizado en el capítulo 2 de esta tesis como una primera aproximación intuitiva al problema de la convergencia de las tasas de paro de las Comunidades Autónomas en España.

Aunque el análisis de σ -convergencia pueda revelar una cierta tendencia a la reducción de las diferencias entre las tasas de paro regionales –como puede verse en el trabajo de Plaza Acero (1994) o en el capítulo 2 de esta tesis– existen razones suficientes, tanto teóricas como empíricas, para esperar que dichas tasas no lleguen nunca a igualarse del todo. Las razones teóricas fueron expuestas en el capítulo 3 y las empíricas también quedaron indicadas en el capítulo 2 de esta tesis. En el capítulo 5 veremos como el análisis econométrico corrobora, aunque con ciertos matices que se

verán en su momento, la hipótesis de convergencia condicionada (estados estacionarios distintos para las distintas regiones).

4.4.- Métodos de cointegración

Los métodos descritos en este apartado surgen de la necesidad de dar un tratamiento sistemático al componente temporal en los estudios de convergencia. En esta línea se encuentran los trabajos de Fernández-Jardón y Caneda (1999), Pallardo y Esteve (1997), Avilés *et al.* (1997) o Bernard and Durlauf (1994), que introducen series de tiempo en sus estudios sobre convergencia mediante la utilización de técnicas de cointegración.¹⁴

Al igual que sucedía con los datos de corte transversal, los métodos de convergencia basados en técnicas de cointegración permiten establecer varios estratos o niveles de convergencia (Camarero *et al.* 1995 y 2001, Olloqui *et al.* 2002, Pallardó y Esteve 1997).

El primer caso (el tipo de convergencia más fuerte), también denominado convergencia de largo plazo, se da cuando la diferencia entre dos series genera otra cuyo valor medio tiende a cero en el largo plazo. Desde un punto de vista más formal, diremos que existe convergencia a largo plazo entre el valor que toma la variable X en la región i y el que toma en la región j si:

¹⁴ En este apartado nos limitaremos a las ideas más esenciales sobre la teoría de la cointegración, aunque un desarrollo más amplio de tales cuestiones puede encontrarse en el apéndice 1. En dicho apéndice también aparecen recogidas las referencias bibliográficas pertinentes.

$$\lim_{k \rightarrow \infty} E(X_{i,t+k} - X_{j,t+k} / \xi_t) = 0$$

donde X es la variable estudiada en las regiones i y j en los momentos t y $t+k$, y ξ_t recoge la información relevante en el momento t .

En este primer caso debemos hacer un supuesto adicional para que tenga lugar la convergencia, y es que la varianza de esa diferencia sea finita; en caso contrario, podría existir divergencia entre las variables aunque se produzca un acercamiento entre sus esperanzas.

De acuerdo con la teoría de la cointegración, esta definición nos indica la ausencia de una raíz unitaria en la diferencia existente entre X_i y X_j . Si la variable X es integrada de orden 1 en las dos áreas geográficas, la existencia de este tipo de convergencia implica una relación de cointegración entre las dos series que vendrá dada por el vector $(1, -1)$. Además, ésta será la única relación a largo plazo posible entre ellas, porque solo hay dos variables, y por lo tanto, solo puede existir un vector de cointegración. Esto quiere decir que, para que tenga lugar la convergencia a largo plazo, la diferencia entre las dos variables no ha de presentar ni tendencia estocástica ni tampoco determinista; es decir, esa diferencia ha de ser estacionaria sin tendencia ni término independiente.

Este primer concepto nos estaría indicando que la diferencia entre los niveles que presenta la variable dentro de cada uno de los entes territoriales va disminuyendo a lo largo del tiempo, y por lo tanto, ambos niveles convergen hacia un mismo estado estacionario.

El segundo tipo de convergencia que podemos identificar a través del análisis de cointegración está asociado a un menor nivel de exigencia, y se designa a través del término *catching up*. En este caso lo que ocurre es que la brecha entre ambas series va disminuyendo a lo largo del tiempo pero cada una puede tender hacia su propio equilibrio de largo plazo. Este tipo de convergencia se da cuando

$$E(X_{i,T} - X_{j,T} / \xi_t) < X_{i,t} - X_{j,t}$$

donde t indica el momento actual y T algún momento futuro para el que se hace la proyección.

En este segundo caso seguimos manteniendo que la diferencia entre $X_{i,t}$ y $X_{j,t}$ debe ser estacionaria, pero se permite la existencia de una tendencia determinista; lo que no debe aparecer nunca es una tendencia estocástica en esa diferencia de variables. Nuevamente $X_{i,t}$ y $X_{j,t}$ han de ser cointegradas, y como en el caso de la convergencia de largo plazo, $(1, -1)$ es el único vector de cointegración que puede existir entre ellas.

Esta definición estaría muy relacionada con la definición de β -convergencia, en el sentido de que, si ha habido *catching up* entre los momentos t y T , la tasa de crecimiento tiene que haber sido mayor en la región que inicialmente presentaba menor nivel.

El tercer caso de convergencia – el menos estricto de todos – se da cuando $X_{i,t}$ y $X_{j,t}$ tienen tendencias diferentes, pero es posible establecer una relación funcional entre ambas tendencias. La representación formal de este caso sería la siguiente:

$$\lim_{k \rightarrow \infty} E(X_{i,t+k} - \beta X_{j,t+k} / \xi_t) = 0$$

siendo $\beta > 0$

En este caso las series $X_{i,t}$ y $X_{j,t}$, también deben estar cointegradas pero no es preciso la relación de cointegración sea la que refleja el vector (1, -1), sino que viene dada por el vector (1, $-\beta$), con $\beta > 0$. Estamos pues admitiendo que se corrigen las tendencias estocásticas de las series, pero que éstas afectan con distinta intensidad a cada ente territorial, y podemos saber cual es la diferencia existente entre esas series a largo plazo.

Este último caso admite interpretaciones muy distintas dependiendo de qué valor tome β (que sea mayor o menor que uno), de cuál sea el valor inicial de cada serie, y de qué tipo de evolución (ascendente o descendente) siga cada una. Podría ocurrir que las series se alejaran entre sí a lo largo del tiempo; y también puede ocurrir que dos series con tendencias distintas se estén acercando durante un período determinado y al cabo de cierto tiempo acaben convergiendo (aunque, si esas tendencias distintas se mantienen, después de ese momento de convergencia, las dos series empezarían a divergir).

El análisis de convergencia a través de series temporales, aunque corrige parte de los inconvenientes de los métodos con base en datos de corte transversal, sigue manteniendo algunas deficiencias; y la más importante de estas deficiencias tiene que ver con la posibilidad de que existan “saltos” o cambios estructurales en las series. Estos cambios estructurales pueden falsear los resultados de los tests de raíces unitarias y hacer, por ejemplo, que aceptemos la presencia de raíces unitarias cuando éstas no

existenten. En tales casos puede ocurrir que se acepte, o que se rechace, erróneamente la hipótesis de convergencia.¹⁵

Desde el punto de vista empírico existen bastantes trabajos, algunos de ellos ya citados, donde se analizan variables como la renta per capita, el nivel de precios o la tasa de desempleo y donde se utiliza este tipo de metodología.

Para países europeos, y utilizando el PIB per cápita, Greasley y Oxley (1997) comprueban que solo existe convergencia si se analizan los países dos a dos, pero nunca de forma conjunta. En este mismo sentido Pallardó y Esteve (1997) se vuelve a centrar en los países europeos y solamente encuentra evidencia de acercamiento entre las rentas per cápita de Suecia y Dinamarca. Camarero *et al.* (2001) examinan los precios y la inflación de algunos países europeos y rechazan la convergencia de largo plazo en todos los casos, aunque si encuentran *catching up* de algunos países con Alemania y con la media europea. Usando esta misma metodología para el ámbito regional, Mcguinness y Sheehan (1998) no encuentran una clara evidencia de convergencia entre los ingresos per capita de las distintas regiones del Reino Unido.

Para el caso español Olloqui *et al.* (2002) analizan la convergencia de los precios de las distintas provincias y solo encuentran evidencia de un acercamiento en seis de las cincuenta. Suriñach *et al.* (1995) se centran en las comunidades autónomas y analizan la convergencia en las tasas de desempleo encontrando evidencia de acercamiento con el agregado nacional en cinco de ellas. Avilés *et al.* (1997) también se centran en el

¹⁵ En este sentido, Rappoport y Reinlich (1989) y Perrón (1989) mantienen que los test de raíces unitarias están sesgados a favor de la aceptación de dichas raíces cuando aparecen cambios estructurales en las series estudiadas.

mercado de trabajo y aunque encuentra evidencia de acercamiento entre la población activa de Andalucía y España, ese acercamiento desaparece cuando se analizan las tasas de paro y ocupación. Finalmente Fernández-Jardón y Caneda (1999), tampoco encuentran evidencia de convergencia entre Galicia y España, cuando analizan variables como la producción o la población.

4.5.- Método de descomposición de la varianza

Una forma alternativa de abordar el análisis de convergencia podría ser a través del método de descomposición de la varianza. Aunque este método se ha utilizado con profusión en multitud de estudios econométricos, parece ser que aún no se han explotado sus posibilidades en relación con el tema de la convergencia.

Para investigar la existencia o inexistencia de convergencia en términos de renta per cápita, tasa de desempleo o cualquier otra variable entre varias áreas geográficas se puede proceder del siguiente modo. En primer lugar procederíamos a “explicar” la variable en cuya convergencia estamos interesados en función de otras, como podrían ser el sexo, la edad, o cualquier otro aspecto que pueda considerarse relevante, además de –y esto resulta fundamental- una variable o conjunto de variables referidas al ámbito territorial. Una vez realizada esta regresión (que no sería más que una regresión simple por mínimos cuadrados ordinarios), el tema de la convergencia puede ser directamente abordado: habría que calcular qué porcentaje de la varianza de la variable en cuestión se explica a través del componente territorial, y cómo evoluciona ese porcentaje a lo largo del tiempo.

La descomposición de la varianza se puede realizar a través de métodos distintos, dos de los cuales pasamos a describir a continuación. En el primero de ellos el porcentaje de la varianza explicado por una variable cualquiera (o conjunto de variables) es el que resulta de calcular el porcentaje en que se reduce el término residual de la regresión cuando se añade una (o varias) variable(s) explicativa(s) a un conjunto dado de regresores¹⁶.

El segundo método utilizado para realizar la descomposición de la varianza -que es el que vamos a utilizar en esta tesis- emplea para su cálculo, en lugar del término residual, el valor del coeficiente de determinación de la regresión¹⁷. Este método permite calcular dos efectos distintos: un *efecto total* y un *efecto marginal*, que pasamos a explicar a continuación, y que son los que determinan el porcentaje de la varianza explicado por cada variable (o variables).

El *efecto total* coincide con el valor que toma el estadístico R^2 (multiplicado por cien), cuando solamente se tiene en cuenta una variable explicativa (o un conjunto de variables), además del término independiente, y se ignoran todos los demás regresores del modelo.

El *efecto marginal* se mide a través del incremento porcentual del coeficiente de determinación de la regresión cuando se añade una variable explicativa (o un conjunto de variables) al conjunto formado por el resto de regresores del modelo. Ese conjunto de

¹⁶ Como ejemplo de este tipo de procedimiento podemos el *OCDE employment outlook*, del año 2000. En este informe se analiza la composición de la varianza del desempleo en función de factores regionales, poblacionales o de nivel de estudios.

¹⁷ Este procedimiento se puede encontrar en algunos trabajos como el de Simón Pérez (2001), donde se analizan los factores que influyen en la determinación de los salarios pactados en la negociación colectiva de ámbito sectorial en España.

variables añadidas son las que se utilizan para reflejar la cualidad estudiada en ese momento.

Una vez calculados estos dos efectos, el porcentaje de la varianza que explica cada variable (o variables) se situará dentro del intervalo que marcan estos dos efectos. Hay que observar que la suma de los efectos marginales de todas las variables (o conjuntos de variables) explicativas tiene que coincidir con el estadístico R^2 de la regresión.

El principal inconveniente que introduce el análisis de descomposición de la varianza se debe al cálculo del efecto marginal de cada variable. Este efecto puede verse afectado por las correlaciones –ya sean positivas o negativas– que puedan existir entre las distintas variables explicativas que aparezcan en la regresión. Estas correlaciones pueden provocar tanto aumentos como disminuciones en el mencionado efecto marginal, y en consecuencia, reducir la fiabilidad del intervalo de variación del porcentaje de la varianza.

4.6.- Recapitulación y observaciones finales

En este capítulo hemos visto qué tipo de métodos pueden utilizarse para investigar la convergencia de las tasas de paro regionales en España. Hemos observado que los métodos basados en datos de corte transversal –fundamentalmente los basados en la noción de β -convergencia– presentan algunas deficiencias importantes, sobre todo cuando las tasas de variación de las variables objeto de estudio sufren oscilaciones más o menos bruscas, como es el caso de las tasas de paro regionales. Como alternativa

tenemos el método de σ -convergencia, pero hemos considerado que las ventajas que pueda tener este método se obtienen también con la descomposición de la varianza, que es uno de los métodos seguidos en esta tesis. El otro método que seguiremos aquí para investigar la posible convergencia de las tasas de paro regionales se basa en el análisis de cointegración de las series correspondientes a las distintas áreas geográficas. Al seleccionar estos dos métodos tenemos una doble perspectiva que combina los datos de corte transversal con el análisis de series temporales.

En general, todos los métodos estudiados tienen sus inconvenientes que hemos procurado destacar a lo largo de las páginas precedentes. No obstante, creemos que los métodos seleccionados son más adecuados para el problema que nos ocupa. En cualquier caso, hemos procurado corregir las deficiencias de los métodos seleccionados a medida que éstas se van manifestando en el proceso de aplicación de los mismos.

BIBLIOTECA VIRTUAL

MIGUEL D
SERVANTES

CAPÍTULO 5:
ANÁLISIS EMPÍRICO

1.- Introducción

Como ya hemos justificado debidamente en el capítulo anterior, nuestro propósito es realizar un análisis de la convergencia real entre las regiones españolas en términos de tasas de paro. Para ello vamos a servirnos de dos metodologías diferentes: el análisis de cointegración y la descomposición de la varianza.

Concretando un poco más, los objetivos perseguidos con este capítulo son fundamentalmente dos: por un lado trataremos de ver si las tasas de desempleo regionales siguen, o no, un proceso de acercamiento a las tasas medias de la nación; por otro, intentaremos ver que influencia ejerce en ese proceso de convergencia el grado de homogeneidad del grupo al cual se refieren las tasas de desempleo. Ese grado de homogeneidad será medido a través de variables tales como la edad, el sexo o el nivel de estudios.

El capítulo se organiza de la siguiente manera. En el epígrafe 2 se discuten los datos empleados, así como las fuentes utilizadas para la obtención de los mismos, y se hace especial referencia a los problemas que nos hemos encontrado a la hora de elaborar las distintas variables. El epígrafe 3 se dedica a estudiar la convergencia desde el punto de vista de la cointegración, intentando corroborar la idea de que el grado de convergencia es mayor cuando utilizamos grupos poblacionales cada vez más homogéneos. En el epígrafe 4 se vuelve a estudiar la convergencia, pero esta vez a través de la descomposición de la varianza de la variable dependiente. Finalmente, en el epígrafe 5, se recogen todas las conclusiones del análisis empírico precedente.

5.2.- Fuentes Estadísticas

La totalidad de los datos utilizados en este trabajo provienen de la Encuesta de Población Activa (*EPA*) y de la base de datos *TEMPUS*, ambas proporcionadas por el Instituto Nacional de Estadística (*INE*). La tasa de paro de los distintos grupos, como variable fundamental de nuestro estudio, se ha elaborado a partir de las series de activos y ocupados de la *EPA*, todas ellas desagregadas por Comunidades Autónomas, sexo, grupo de edad y nivel de estudios. Sin embargo, la serie del Producto Interior Bruto (*PIB*), proviene de la contabilidad regional que proporciona la base de datos *TEMPUS*.

El hecho de usar dos fuentes estadísticas diferentes explica por qué el periodo analizado es distinto cuando realizamos el estudio desde el punto de vista de la cointegración y cuando éste tiene lugar mediante la descomposición de la varianza. El análisis de cointegración estudia la relación existente entre dos tasas de paro, la regional y la nacional, pero no introducimos ninguna variable adicional, por lo tanto, todos los datos provienen de la *EPA*, que proporciona datos desde 1977 hasta 2001. Sin embargo, al descomponer la varianza, introducimos el *PIB* de cada región como variable explicativa de la tasa de paro, y esta variable proviene de la contabilidad regional, la cual únicamente suministra datos desde 1980 hasta 1999.

Esa misma diversidad de fuentes también da lugar a diferencias en cuanto a la periodicidad de los datos utilizados. En el análisis de cointegración usamos datos de periodicidad trimestral y en el análisis de descomposición de la varianza los datos utilizados son de periodicidad anual.

Cuando se usa el enfoque de cointegración, todas las variables empleadas provienen de la misma fuente estadística: la *EPA*. Por lo tanto, tenemos datos trimestrales para todas ellas y eso nos permite llevar a cabo un análisis de periodicidad trimestral. Este hecho, además de proporcionarnos una gran cantidad de observaciones, nos permite contar con la información más completa posible. En cuanto a la posible existencia de problemas de estacionalidad, hemos procurado corregirlos y para ello hemos depurando los datos de la manera indicada en el epígrafe siguiente.

Cuando se recurre al método de descomposición de la varianza, tenemos variables procedentes de dos fuentes distintas. Una de estas fuentes –la *EPA*– proporciona datos trimestrales, mientras que la otra –la contabilidad regional– proporciona datos anuales. Por esta razón hemos tenido que homogeneizar las observaciones dándole a todas una periodicidad anual. El número de observaciones se nos reduce considerablemente, pero sigue siendo alto (180 observaciones en el caso de menos desagregación), y la razón de esto es que, al analizar simultáneamente las tasas de paro de varios grupos de población, diferenciados además por regiones y por años, se acumula un número bastante grande de observaciones.

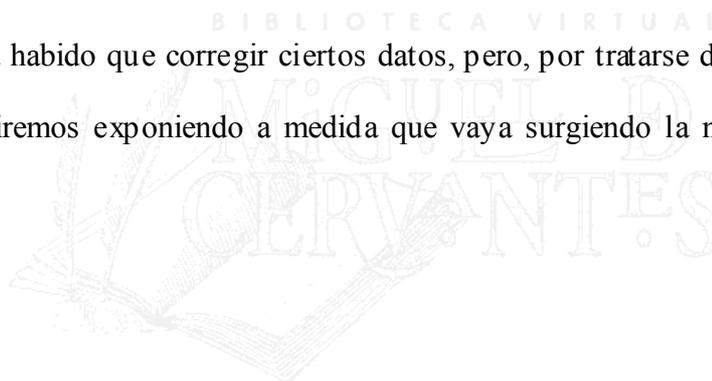
Para llevar a cabo el cambio de datos trimestrales a anuales en las series desagregadas de actividad y ocupación, se ha utilizado un recurso muy común en la literatura económica que consiste en identificar el dato anual con el correspondiente al segundo trimestre del año. Para justificar esto se suele aducir que dicho trimestre es el más representativo del año, ya que se cree que es el menos sujeto a fluctuaciones, lo cual parece ser verdad al menos en lo relacionado con el mercado de trabajo español (Toharia *et al.*, 1998).

Otro problema que nos hemos encontrado es el de la representatividad de las muestras. Dicho problema surge principalmente en relación con el análisis de descomposición de la varianza y tiene que ver con el alto grado de desagregación de los datos. En nuestro análisis de descomposición de la varianza, la variable dependiente utilizada consiste en una “macroserie” dentro de la cual se incluyen las tasas de paro de todos los grupos de población correspondientes a todas las regiones. Por esta razón, cuando fraccionamos mucho los grupos a fin de hacerlos cada vez más homogéneos, el número de personas que integran la muestra correspondiente a cada grupo se hace cada vez menor; y esto puede generar dudas sobre la representatividad de dichas muestras. El criterio seguido en estos casos ha sido el del *INE*, el cual considera que, cuando la muestra tiene menos de 5000 datos, dicha muestra deja de ser representativa. Para evitar este problema hemos intentado reducir el nivel de desagregación del análisis, diferenciando solo dos niveles de estudios (individuos con estudios medios o superiores e individuos por debajo de ese nivel), y solo dos grupos de edad (mayores de 25 años y menores de esa edad). A pesar de todo, este problema nos ha obligado a eliminar de nuestro análisis algunas Comunidades Autónomas, cuyos datos no resultaban suficientemente representativos, quedándonos únicamente con nueve Comunidades. En el epígrafe siguiente se darán más detalles sobre esta cuestión.

Este problema de representatividad de las muestras de tamaño reducido también se da, aunque en menor grado, en el análisis de cointegración. Para solucionarlo hemos optado por centrar nuestro análisis en el grupo de varones mayores de 25 años, ya que dicho grupo es el más numeroso y proporciona datos suficientes. En las regiones donde existían problemas estadísticos con los datos, se han “reconstruido” los valores ausentes de acuerdo con el procedimiento descrito en el epígrafe siguiente.

Finalmente, también nos hemos encontrado con problemas a la hora de utilizar el *PIB* regional. En los datos proporcionados por el *INE* aparecen dos series, una desde 1980 hasta 1995 y otra desde 1995 hasta 1999. Esta duplicidad de fuentes nos obliga a realizar una homogeneización de los datos, y para ello se han reconstruido los valores del *PIB* entre 1995 y 1999. Para reconstruir estos datos se calculan los incrementos que sufre el *PIB* en cada año según la segunda serie, y se aplican esos incrementos al *PIB* de 1995 que proporciona la primera serie. De esta manera se van consiguiendo todos los datos hasta 1999.

Existen otras consideraciones relevantes acerca de cómo se han construido las variables y cómo ha habido que corregir ciertos datos, pero, por tratarse de cuestiones más puntuales, las iremos exponiendo a medida que vaya surgiendo la necesidad de aclaraciones.



5.3.- Análisis de Cointegración

El primer método que vamos a emplear se basa en las técnicas de cointegración. Con este método se pretende descubrir la existencia de relaciones de largo plazo entre las variables; y serán esas relaciones las que nos indiquen si existe convergencia o divergencia entre las variables estudiadas.

Desde un punto de vista práctico, nuestro estudio se estructurará en varias fases. En primer lugar, se llevará a cabo el análisis de convergencia entre la tasa de paro de cada región y la media nacional, tomando como base la tasa de paro de toda la

población. Posteriormente, se repetirá el análisis para grupos de población cada vez más homogéneos, intentando extraer las conclusiones pertinentes.

En lo referente a los datos también hemos de hacer algunas consideraciones previas. En primer lugar, y como ya hemos mencionado en el epígrafe anterior, el período de análisis es el comprendido entre 1977 y 2001, y se van a emplear datos trimestrales. En segundo lugar, el estudio de la convergencia se realiza entre cada Comunidad Autónoma y el correspondiente agregado nacional, entendiendo por agregado nacional al resto de regiones que componen España. Por lo tanto, a la hora de elaborar la tasa de paro del agregado nacional, se eliminan los datos de esa región y de este modo se evita la posible relación que pudiera existir entre la Comunidad Autónoma y España por el mero hecho de que la primera esté incluida dentro de la segunda.

En cuanto al criterio de elección de grupos cada vez más homogéneos, también está íntimamente relacionado con la disponibilidad de datos. El primer grupo estudiado es el conjunto de la población trabajadora. Después, nos hemos centrado en el grupo formado por los varones con más de 25 años; y la razón de esta selección es que se trata de un grupo suficientemente amplio, que admite desagregaciones posteriores sin plantear problemas serios de disponibilidad de datos. Finalmente, hemos centrado nuestro análisis en un subgrupo del anterior: los varones con más de 25 años que, además de reunir esas características, poseen estudios superiores. Este grupo, aunque no es demasiado amplio, presenta un comportamiento muy regular y esto facilita su tratamiento estadístico.

Nuestro análisis econométrico constará de dos fases: una primera, donde se verá si todas las series presentan una raíz unitaria; y una segunda, donde se procederá a

estudiar si la diferencia existente entre la tasa de paro de cada región y el correspondiente agregado nacional sigue un proceso estacionario, y si dicho proceso está o no está condicionado por tendencias deterministas. De esta manera vamos a poder concluir si existe convergencia real, y también si ésta se debe a un proceso de *catching up*, o si, por el contrario, el equilibrio de largo plazo es el mismo para ambas series.

Una vez hechas estas consideraciones, el primer paso de nuestro trabajo debe de ser la elaboración de las series que vamos a utilizar. Las tasas de paro que utilizaremos se definen de la siguiente manera:

$$TP^i = \frac{P^i}{A^i}$$

$$TPE^i = \frac{PE - P^i}{AE - A^i}$$



donde TP^i es la tasa de paro de la región i calculada como cociente entre parados de la región i (P^i) y activos de esa región (A^i). Por su parte, TPE^i es la tasa de paro del agregado nacional correspondiente a esa región, y se define como el cociente entre el total de parados de España (PE) menos los que aporta esa región (P^i), y el total de activos de España (AE) menos los correspondientes a la región i .

El siguiente paso, una vez definidas nuestras variables, consiste en determinar el grado de integración que tienen las series, y fundamentalmente, demostrar que las tasas de paro de cada región y las de su correspondiente agregado nacional presentan una raíz

unitaria. De esta manera se puede establecer si existe o no existe una relación de cointegración entre ellas, que además será única.

Para determinar el número de raíces unitarias, vamos a utilizar dos métodos distintos: por un lado, la metodología propuesta por Dickey y Fuller (1979, 1981); y por otro lado, la recogida en los trabajos de Phillips (1987) y Phillips y Perron (1988). El primer método da lugar al test que se denomina *augmented Dickey-Fuller (ADF)*, y el segundo método al test *Phillips-Perron (PP)*.

Como ya hemos dicho en el capítulo anterior, el principal problema que tienen los trabajos de convergencia realizados a través de técnicas de cointegración proviene del hecho de que puedan existir cambios estructurales en las series utilizadas. Estos cambios pueden llevar a conclusiones erróneas sobre el grado de integración de las series. Para intentar solucionar este problema hemos procurado homogeneizar las series. Para ello se han eliminando tanto los posibles efectos estacionales que causa el estar trabajando con datos trimestrales, como los posibles cambios estructurales que pueda provocar en una serie la presencia de datos atípicos especialmente relevantes.

Existen varios métodos a la hora de homogeneizar las series. Uno de ellos consiste en filtrar la serie una vez que se han determinado los efectos estacionales y los datos atípicos que son especialmente representativos (Fernández-Jardón y Caneda, 1999). Existe una segunda opción, que es la que hemos utilizado en esta tesis, y que consiste en aplicar conjuntamente dos procedimientos distintos: en primer lugar se aplica el método *TRAMO (Time series Regression with ARIMA noise, Missing values, and Outliers)* que elimina el efecto de los datos atípicos más influyentes, y en segundo lugar se aplica el método *SEATS (Signal Extraction in ARIMA Time Series)* que elimina

la estacionalidad de las series. La aplicación de esos dos métodos se puede conseguir de forma simultánea a través del programa *TSW (TRAMO-SEATS para WINDOWS)* (Maravall y Gómez, 1993, 1994, 2001 y Maravall *et al.* 1999).

Este segundo procedimiento corrige los efectos de series con periodicidad inferior a la anual y nos proporciona una serie desestacionalizada. Además, nos permite depurar las series eliminando el efecto de los “outliers” (datos atípicos) especialmente relevantes. Otra ventaja que nos proporciona el uso de este método, es que nos permite estimar el valor de observaciones perdidas o ausentes. De esta manera, en regiones donde el alto grado de desagregación aplicado nos impide tener una muestra fiable, este procedimiento nos permite disponer de un dato verosímil que podemos utilizar en el análisis empírico. Esta última aplicación es importante en las regiones con menor tamaño en términos poblacionales; en especial, La Rioja, Murcia, Navarra, Baleares y Extremadura.

La aplicación de la técnica de homogeneización, la corrección de datos atípicos corregidos, y la interpolación de datos ausentes se recoge en el Apéndice 3; más concretamente, en las TABLAS A3.1, A3.2 y A3.3.

Una vez que poseemos las series depuradas, ya podemos estudiar cual es su grado de integración. La aplicación de los tests de raíces unitarias sobre las distintas tasas de paro utilizadas aparecen recogidos en las tablas que se incluyen dentro del Apéndice 2 (A2.1...A2.6). El resultado más relevante de la aplicación de estos tests nos permite comprobar que las series de tasas de paro utilizadas se pueden considerar integradas de orden 1

Después de haber homogeneizado las series y comprobado que todas ellas poseen una raíz unitaria, el siguiente paso es consiste en investigar si existe una relación de cointegración entre ellas, y más concretamente, si esta relación viene dada por el vector (1, -1). Para saber si esto es verdad, analizamos lo que ocurre con la diferencia existente entre la tasa de paro de cada región y la del correspondiente agregado nacional. Dependiendo de cómo sea esa diferencia (estacionaria o no estacionaria), podremos deducir si existe o no existe convergencia entre cada región y España. Además, si esa diferencia resulta ser estacionaria, sin término independiente ni tendencia, podremos afirmar que existe convergencia a largo plazo y que la Comunidad Autónoma y el agregado nacional tienden a un mismo estado estacionario. Sin embargo, si aparecen tendencias deterministas, la convergencia será del tipo “catching-up” y podremos tener estados estacionarios diferentes.

El análisis de cointegración se realiza en tres fases distintas, a lo largo de las cuales intentamos conseguir grupos poblacionales cada vez más homogéneos. En primer lugar, se estudia la existencia de relaciones de largo plazo entre las tasas de paro de cada región con España, referidas dichas tasas a la totalidad de la población. En segundo lugar, se repite el mismo análisis pero utilizando las tasas de paro de los varones mayores de 25 años; finalmente, se buscará un grupo aún más homogéneo, y se vuelven a estudiar las relaciones de largo plazo entre cada región y el agregado nacional, utilizando las tasas de paro de los varones mayores de 25 años con estudios superiores.

En el Apéndice 2, y más concretamente en las TABLAS A2.7, A2.8 y A2.9, se exponen los resultados de todos los tests de raíces unitarias correspondientes a las diferencias entre las tasas de paro regionales y los agregados nacionales pertinentes. Para todas las regiones se han aplicado los dos tests a los que ya nos hemos referido (*PP*

y *ADF*) con el fin de obtener resultados más concluyentes. En este epígrafe únicamente presentaremos una tabla -TABLA 5.1- que nos resumirá los resultados obtenidos y que nos proporcionará un soporte visual a la hora de aportar conclusiones sobre la existencia de convergencia real entre las tasas de paro de las regiones españolas.

El resultado que cabe esperar de todo este análisis sería el aumento progresivo del número de regiones que convergen con el agregado nacional, según vamos avanzando en el grado de homogeneidad del grupo poblacional estudiado. Este resultado estaría en concordancia con el hecho de que cuanto mejor definido esté el grupo estudiado menos son las fuentes de divergencia existentes a nivel regional.

Los resultados recogidos en la TABLA 5.1 parecen estar en concordancia con nuestra hipótesis inicial: el grado de convergencia aumenta efectivamente a medida que el grupo poblacional se define mejor y se hace más homogéneo. Al menos esto es lo que parece indicar el aumento experimentado por el número de Comunidades Autónomas que presentan un resultado estacionario al ser comparadas con su correspondiente agregado nacional, según nos vamos centrando en un mercado laboral cada vez más homogéneo en términos de sexo, edad y nivel de estudios.

Profundizando un poco más en los resultados obtenidos, y haciendo un estudio separado para cada grupo de población, podemos extraer algunas conclusiones adicionales (todos los resultados aquí resumidos aparecen detallados en las TABLAS A2.7, A2.8 y A2.9).

TABLA 5.1.
Comunidades Autónomas donde se cumple que la diferencia con el agregado nacional es estacionaria.

	DIFERENCIA ESTACIONARIA
CONJUNTO DE LA POBLACION	ANDALUCIA, ARAGON, VALENCIA, MADRID*, MURCIA*, LA RIOJA
VARONES MAYORES DE 25 AÑOS	ANDALUCIA, ASTURIAS, BALEARES, CANTABRIA, VALENCIA, MURCIA*, PAIS VASCO, LA RIOJA
VARONES MAYORES DE 25 AÑOS CON ESTUDIOS SUPERIORES	ANDALUCIA*, ARAGON, ASTURIAS, CANARIAS, CANTABRIA, CASTILLA-LEON, CASTILLA LA MANCHA, CATALUÑA, VALENCIA, EXTREMADURA*, GALICIA, MADRID*, MURCIA*, PAIS VASCO, NAVARRA, LA RIOJA

*Estacionarias sin término independiente ni tendencia

Fuente: Elaboración propia a partir de datos EPA

En primer lugar, y refiriéndonos a la tasa de paro de toda la población, podemos apreciar que solo seis regiones presentan un resultado estacionario al compararlas con el agregado nacional (Andalucía, Aragón, Madrid, Murcia, La Rioja y Valencia). Además, Madrid solo muestra estacionariedad cuando bajamos el nivel de exigencia al 10%, y

Andalucía aún cumpliéndolo solo para uno de los test aplicados, también debe reducir su nivel de significación al 10%.

En cuanto al tipo de convergencia, podemos apreciar que solamente en dos regiones se produce convergencia a largo plazo (Murcia y Madrid). El resto alcanzan una convergencia del tipo *catching-up*, lo cual indica la existencia de componentes deterministas, ya sean tendencias o términos independientes. Estos componentes, aunque no impiden la existencia de un cierto equilibrio en la evolución seguida por las series, sí impiden que se alcance la convergencia completa de las tasas de desempleo. Esto estaría indicando la existencia de más de un estado estacionario en la evolución seguida por las tasas de paro regionales.

Pasemos ahora al análisis de la convergencia de las tasas de paro de los varones mayores de 25 años. En este caso se produce un ligero aumento del número de regiones que presentan un cierto grado de convergencia (Andalucía, Asturias, Baleares, Cantabria, Murcia, País Vasco, La Rioja y Valencia), aunque algunas de ellas, como ya ocurría en el caso anterior, solo cumplen la condición de estacionariedad para uno de los dos test de raíces unitarias aplicados, y otras vuelven a exigir un nivel de confianza del 10%.

Resulta un poco sorprendente el hecho de que dos regiones que antes convergían (Aragón y Madrid) no sigan haciéndolo ahora. No obstante, este resultado deja de parecer tan anormal si se tiene en cuenta que el grupo de los mayores de 25 años sigue siendo un grupo muy amplio y por lo tanto muy heterogéneo. Al agrupar a la gente por edades de esta manera podrían aflorar discrepancias que antes permanecían ocultas o diluidas en el conjunto de la población.

En cuanto al grado de convergencia, vuelve a ser Murcia la única que mantiene el equilibrio a largo plazo sin término independiente ni tendencia (recordemos que Madrid ya no cumple la condición de convergencia). El resto sigue manteniendo componentes deterministas que impiden la convergencia de las tasas de paro de todas las regiones hacia un único estado estacionario.

El paso final que damos con este estudio se centra en un grupo mucho más definido: el de los varones mayores de 25 años con estudios superiores. En este caso, la situación ya es mucho más clara, y el aumento de regiones que cumplen la condición de estacionariedad es más que considerable. Ahora, la única Comunidad Autónoma que no presenta relación de equilibrio con su correspondiente agregado nacional es Baleares. Este resultado, que podría sorprendernos si tenemos en cuenta que Baleares aparecía como estacionaria en el caso anterior, no es demasiado extraño después de todo: los datos de Baleares presentaban graves problemas de muestreo y han tenido que ser reconstruidos a través del programa TSW. Aparte de Baleares, la única Comunidad que presenta problemas es el País Vasco. La condición de estacionariedad solo se cumple para uno de los tests, y para esto debemos exigir un nivel de significación del 10%. No obstante, este resultado es coherente con la evolución seguida por esta región a lo largo de nuestro estudio.

Finalmente, habría que destacar que la convergencia estricta a largo plazo solo se da en cuatro regiones -Andalucía, Extremadura, Madrid y Murcia-. En el resto se produce convergencia del tipo *catching-up*; es decir, el equilibrio a largo plazo de cada región está condicionado por la existencia de términos independientes o tendencias, lo

cual pone de manifiesto que no existe un único estado estacionario al que converjan las tasas de paro de todas las comunidades autónomas.

5.4.-Análisis de descomposición de la varianza

El siguiente método que vamos a emplear para profundizar un poco más en la idea de convergencia de las tasas de paro regionales es el de la descomposición de la varianza. Con este método de análisis perseguimos dos objetivos: por un lado intentaremos demostrar que la importancia que tiene el componente regional en la explicación de las tasas de paro disminuye conforme nos vamos centrando en un grupo cada vez más homogéneo de individuos; por otro lado, este método nos permitirá estudiar la evolución temporal que sigue el porcentaje de la varianza explicado por ese componente regional.

La metodología, como ya hemos explicado en el capítulo anterior, consiste en la realización de una batería de regresiones, una para cada nivel de desagregación de los datos. Este enfoque se basa en “regresar” la tasa de paro correspondiente a cada grupo de población, región y año contra una serie de variables categóricas y contra el *PIB*. Las variables categóricas incluidas se refieren a las regiones, los años, el género, los grupos de edad y el nivel educativo¹⁸.

Una vez realizada cada una de las regresiones, a la hora de calcular el porcentaje de la varianza que explica cada variable, o grupo de variables, recurriremos al método

¹⁸ El conjunto de variables que se introduzcan en el modelo dependerá del nivel de desagregación que tengamos en cada momento

descrito en el capítulo anterior, el cual nos permite acotar el porcentaje de la varianza atribuible a cada conjunto de regresores dentro de un intervalo que viene delimitado por el efecto total y el efecto marginal (ver capítulo 4)

El conjunto de regresiones se realizará mediante la estimación por mínimos cuadrados ordinarios de funciones del siguiente tipo:

$$TP = \Phi(C, Y, R, T, D)$$

donde TP es la tasa de paro de un conjunto de individuos clasificados según edad, estudios, sexo, región y año; C es el término independiente; Y es el Producto Interior Bruto de cada región en cada uno de los años considerados; R es un conjunto de nueve variables ficticias que recogen el efecto de estar en una u otra región; T es otro conjunto de variables ficticias, una por cada año estudiado, que reflejan la influencia del componente temporal; y D es un tercer conjunto de variables cualitativas que recogen características tales como el sexo, la edad o el nivel de estudios.

La variable R solo refleja la existencia de nueve regiones debido a un problema de datos. Ya hemos mencionado anteriormente que la desagregación de los datos a ciertos niveles no es fiable; eso se debe a que las muestras disponibles son demasiado pequeñas, por lo que no resultan suficientemente significativas. Este problema es particularmente grave en las regiones de escasa población, por lo que dichas regiones han sido excluidas del análisis. A pesar de todo, podemos considerar que las regiones incluidas en este apartado (Andalucía, Aragón, Castilla y León, Castilla la Mancha, Cataluña, Comunidad Valenciana, Galicia, Comunidad de Madrid y País Vasco) son lo suficientemente representativas como para dar validez al estudio realizado.

Por su parte, el vector D está formado por las tres variables siguientes: *SEXO*, que distingue entre hombres y mujeres; *EDAD*, que diferencia entre mayores de 25 años y menores de esa edad; y *ESTUDIOS*, que establece la distinción entre individuos con estudios medios o superiores e individuos por debajo de ese nivel.

Habría que aclarar también que en las regresiones se ha eliminado tanto la variable ficticia correspondiente a 1980 como el término independiente. Se ha hecho esto con el fin de evitar el problema de multicolinealidad perfecta que de otro modo se presentaría¹⁹. También hay que aclarar que el periodo al que se refiere el análisis de este apartado (1980 – 1999) ha venido impuesto por las características de la base de datos utilizada. El hecho de que los datos de la contabilidad regional que proporciona el INE tengan una periodicidad anual y la ausencia de datos previos a 1980 ha condicionado la amplitud de nuestro estudio.

Este tipo de análisis puede presentar otros problemas. El primero de ellos estaría relacionado con el uso de series temporales, lo que exigiría llevar a cabo algún test de raíces unitarias para comprobar el orden de integración de dichas series. En relación con este punto, podría ser que la tasa de paro y el *PIB* tuviesen tendencias más o menos relacionadas. No obstante, la forma en que hay que organizar los datos para construir la serie de tasas de paro –que en realidad es una mezcla de varias series correspondientes a regiones distintas– hace que desaparezca el aspecto tendencial de dichos datos. Por eso, no parece que sea necesario comprobar el grado de integración de dicha serie. En realidad, lo único que se podría comprobar es si, para cada región, las tasas de paro

¹⁹ Debemos tener en cuenta que en las regresiones se incluyen tantas ficticias regionales como regiones se estudian. Por eso no pueden introducirse a la vez todas las variables ficticias anuales; y por eso también debemos eliminar el término independiente.

tienen el mismo grado de integración que el *PIB*. Las tasas de paro ya hemos comprobado que en general son integradas de orden uno; en cambio el *PIB* no presenta un mismo grado de integración para todas las series (véase TABLA A2.10 del apéndice 2). Por estas razones se ha decidido no tomar diferencias y mantener la mayor información posible en los datos.

El segundo problema aludido al principio del párrafo anterior tiene que ver con la posible correlación entre las variables. No obstante, y como puede comprobarse en la tabla recogida en el Apéndice 4, los coeficientes de correlación entre las distintas variables son suficientemente bajos como para admitir la hipótesis de que no existe correlación entre ellas. Únicamente aparece una cierta correlación entre el *PIB* y las variables ficticias regionales. En principio, esta correlación podría afectar al cálculo de los efectos marginales²⁰ –la contribución de cada conjunto de variables al R^2 global de la regresión– pero de ninguna manera al cálculo de los efectos totales. Por tanto, podemos concluir que el intervalo sigue siendo representativo, y el problema, aunque existe, no parece ser demasiado grave.

Una vez realizadas estas consideraciones previas, ya podemos centrarnos en el desarrollo econométrico propiamente dicho y pasar a describir las regresiones realizadas. El objetivo de estas regresiones es ver cómo a medida que nos vamos centrando en grupos cada vez más homogéneos, la importancia del componente regional –como variable explicativa de la tasa de paro– va disminuyendo.

²⁰ La varianza total coincide con la suma de las varianzas individuales únicamente si las covarianzas son nulas; en caso contrario, la existencia de correlaciones positivas o negativas puede afectar al cálculo del efecto marginal, causando tanto su aumento como su disminución.

La primera regresión que vamos a realizar incluye únicamente, como variables explicativas de la tasa de paro, las que reflejan el componente regional, las referidas al aspecto temporal y el *PIB* de cada región. A partir de aquí vamos introduciendo nuevas variables con el fin de obtener grupos cada vez mejor definidos. Así la segunda regresión incluye, además de las anteriores, la variable *SEXO*. La tercera añade la variable *EDAD*, y la cuarta introduce el nivel de formación a través de la variable *ESTUDIOS*.

Hay que observar que cada una de estas regresiones exige una redefinición de la tasa de paro que cada vez se va desagregando más, además de modificar el número de datos empleados. En la primera regresión, la tasa de paro es simplemente la media de cada región, sin embargo, en la segunda regresión ya tenemos tasas de desempleo masculinas y femeninas. En la tercera tenemos tasas masculinas y femeninas diferenciadas por edad, y en la cuarta, la diferenciación se extiende también al nivel de estudios.

Como ya hemos dicho anteriormente, la primera regresión que vamos a realizar es la que menor nivel de desagregación presenta. Las 180 observaciones empleadas en esta primera regresión resultarían de multiplicar las tasas de las 9 regiones por los 20 años recogidos en nuestro análisis. Por otra parte, tenemos 20 variables dicotómicas que recogen el efecto temporal, una para cada año del período analizado, que toman el valor 1 cuando la observación se corresponde con ese año y 0 en el resto de los años. Representamos estas variables por *A1980*, *A1981*... *A1999*. También tenemos nueve variables que recogen el componente regional. Éstas toman el valor 1 cuando estamos en la región correspondiente y 0 en el resto de los casos. Las variables son *ANDA*, *ARAG*, *CASL*, *CASM*, *CATA*, *CVAL*, *GALI*, *MADR* y *PVAS*, y se corresponden con

Andalucía, Aragón, Castilla y León, Castilla la Mancha, Cataluña, Comunidad Valenciana, Galicia, Comunidad de Madrid y País Vasco, respectivamente. Por último aparece la variable PIB que representa el producto interior bruto correspondiente a cada región y cada año estudiados.

Los resultados de esta primera regresión aparecen recogidos en la TABLA 5.2. De una primera observación podemos extraer ya una serie de conclusiones relevantes: en primer lugar, el coeficiente de determinación presenta un valor superior al 88%. Esto nos muestra la importancia que tienen las variables utilizadas para explicar la variabilidad de la tasa de desempleo. En segundo lugar, y a la vista de los valores que toma el estadístico t , podemos decir que casi todas las variables introducidas en el modelo son significativas a un nivel del 1%. La única excepción es el *PIB* que exige un nivel de significación del 5%.

Si nos fijamos en los valores de los coeficientes estimados también podemos extraer conclusiones relevantes. Los coeficientes referidos a las variables anuales presentan todos valores positivos, siendo especialmente elevados los referidos a los años centrales, tanto de la década de los ochenta como de los noventa, los cuales se corresponden con los años más recesivos desde el punto de vista del desempleo en España.

Las variables referentes al componente regional también presentan coeficientes positivos, siendo especialmente alto y significativo el valor correspondiente a Andalucía. Este resultado no debería sorprendernos, si tenemos en cuenta los elevados niveles que presenta la tasa de paro en esta comunidad.

TABLA 5.2

	COEFICIENTES	T ²	% DE LA VARIANZA	
			Efecto total	Efecto marginal
A1980	-	-	32.73	31.42
A1981	0.026089	2.730249		
A1982	0.041210	4.314014		
A1983	0.058100	6.080841		
A1984	0.083034	8.684519		
A1985	0.101255	10.56630		
A1986	0.096578	10.05137		
A1987	0.086548	8.919779		
A1988	0.084015	8.538012		
A1989	0.062434	6.251950		
A1990	0.050552	4.967815		
A1991	0.049155	4.784977		
A1992	0.067978	6.588173		
A1993	0.111359	10.86448		
A1994	0.131485	12.69240		
A1995	0.117898	11.21960		
A1996	0.113907	10.67900		
A1997	0.103688	9.518514		
A1998	0.083009	7.453794		
A1999	0.055216	4.825958		
ANDA	0.238755	12.05975	55.48	47.81
ARAG	0.093498	4.722673		
CASL	0.093960	11.16901		
CASM	0.097979	11.64684		
CATA	0.106620	12.51059		
CVAL	0.113616	13.33146		
GALI	0.075871	7.217789		
MADR	0.106701	10.15077		
PVAS	0.178572	6.388652		
PIB	-7.80E-09	-1.852729	10.17	0.26
R²	0.884772			

Fuente: Elaboración propia a partir de datos procedentes del INE

El *PIB* presenta un signo negativo que refleja una relación inversa entre la producción y el desempleo. Este signo podría considerarse adecuado porque nos indica que los aumentos en el *PIB* se asocian con reducciones de la tasa de desempleo. El bajo valor de este coeficiente es probablemente consecuencia de la magnitud en que viene expresada la producción.

Las dos últimas columnas de esta TABLA 5.2 indican el peso relativo de cada conjunto de variables explicativas en la explicación de la variable dependiente. Como ya se ha indicado anteriormente, se han calculado dos efectos, el total y el marginal. Estos dos efectos marcan los límites entre los que se mueve el porcentaje correspondiente a cada grupo de variables explicativas.

A la vista de los resultados obtenidos, podemos destacar la importancia que tienen las variables regionales –entre un 47% y un 55%– de cara a la determinación de las tasas de paro. A continuación, y con porcentajes en torno al treinta por ciento, aparecen las variables que recogen el componente temporal. Todas ellas pueden ser consideradas importantes a la hora de explicar la variable dependiente. Finalmente, aparece la variable referida al nivel de producción. Esta presenta un intervalo de variación bastante amplio, y el valor tan bajo que presenta su efecto marginal puede ser consecuencia de la correlación existente entre esta variable y las regionales.

En relación con la hipótesis de la convergencia de las tasas de paro regionales, es de destacar que el componente regional explica en torno al 50% de la variable dependiente. Este resultado indica ya un cierto grado de divergencia entre las tasas de

paro regionales, cuando nos fijamos en grupos de población muy agregados en los que no distinguimos ni por sexo, ni por edad, ni por nivel de estudios.

El segundo paso que vamos a dar en el análisis de descomposición de la varianza nos lleva a estimar una segunda versión del modelo, donde añadimos la desagregación por sexo, manteniendo el resto de variables igual que antes. El único cambio que conlleva esta nueva versión pasa por redefinir la tasa de paro, que ahora tiene que distinguir entre hombres y mujeres. Hay que señalar que en esta regresión los datos se duplican – ahora tenemos 360 – y que se introduce una nueva variable en el modelo. Esta variable es *SEXO*, que toma el valor 1 cuando se refiere a hombres y 0 cuando se trata de mujeres. Todas las demás variables han sido construidas de la misma manera y tienen la misma interpretación que la versión anterior del modelo.

Los resultados de esta segunda regresión aparecen recogidos en la TABLA 5.3. Al igual que ocurría anteriormente, el nivel de significación es muy elevado, superando el 80%, y todas las variables explicativas son significativas al 1%, incluso el PIB. Estos primeros resultados corroboran nuevamente la adecuación de nuestro modelo, y la importancia que tienen las variables utilizadas como factores explicativos de las tasas de paro.

Por lo que se refiere a los coeficientes estimados, los resultados son una continuación de los que ya aparecían en la TABLA 5.2, y por lo tanto las explicaciones previas siguen siendo igualmente válidas. Todos los años mantienen el signo positivo, y siguen siendo los periodos del 84 al 88 y del 93 al 98 los que presentan coeficientes mayores. Todas las variables que recogen el componente regional siguen con el signo positivo, y vuelve a ser Andalucía la que más peso tiene como variable explicativa del

desempleo. El *PIB* también mantiene su signo y sigue teniendo un coeficiente muy reducido, pero ahora aumenta su importancia como variable explicativa. En cuanto a la variable *SEXO*, que es la que aparece nueva, presenta un signo negativo que nos está indicando las mayores tasas de paro que sufren las mujeres en el mercado de trabajo.

Las dos últimas columnas de la TABLA 5.3 vuelven a reflejar el efecto total y el efecto marginal, que nos dicen dentro de qué márgenes se encuentra el porcentaje de la varianza de las tasas de paro explicado por cada grupo de variables. La primera apreciación que podemos hacer es que todas las variables, tanto las anuales como las poblacionales, reducen sus porcentajes respectivos en torno a la mitad.

También se puede apreciar la importancia de la variable *SEXO*, que pasa a ser la que más cuenta dentro de nuestro modelo, explicando más del 37% de la varianza de la tasa de desempleo. En cuanto al *PIB*, vemos que sigue siendo la menos explicativa, aunque también ha reducido su intervalo de variación, ya que ha disminuido su efecto total y ha aumentado su efecto marginal.

Por último, podemos apreciar que las variables referidas al componente regional han sufrido una fuerte reducción en el porcentaje de la varianza de las tasas de paro explicado por las mismas. Este porcentaje ha pasado de estar entre el 47.81 y el 55.48, a situarse en niveles comprendidos entre 24.92 y 26.73. Por lo tanto vemos que la influencia regional es menor cuando pasamos a este segundo nivel de desagregación, y también es menor la divergencia regional de las tasas de desempleo.

TABLA 5.3

	COEFICIENTES	T ²	% DE LA VARIANZA	
			Efecto total	Efecto marginal
A1980	-	-		
A1981	0.028377	2.454346		
A1982	0.046573	4.027872		
A1983	0.064197	5.548729		
A1984	0.088760	7.666335		
A1985	0.109878	9.470992		
A1986	0.106967	9.183553		
A1987	0.104327	8.861830		
A1988	0.105444	8.825492		
A1989	0.085552	7.033436	17.23	17.32
A1990	0.074515	6.019310		
A1991	0.072126	5.748676		
A1992	0.090601	7.204075		
A1993	0.131055	10.49909		
A1994	0.153015	12.11722		
A1995	0.140240	10.91532		
A1996	0.136646	10.47393		
A1997	0.127644	9.528298		
A1998	0.108875	7.883214		
A1999	0.081327	5.707397		
ANDA	0.333708	14.57004		
ARAG	0.131876	12.92770		
CASL	0.162574	15.76969		
CASM	0.181457	14.61973		
CATA	0.265360	8.215658	26.73	24.92
CVAL	0.214606	12.28769		
GALI	0.130886	10.92653		
MADR	0.231349	8.961621		
PVAS	0.217658	16.79618		
PIB	-1.78E-08	-3.643172	3.75	0.70
SEXO	-0.097613	-26.69828	37.82	37.82
R²	0.824907			

Fuente: Elaboración propia a partir de datos procedentes del INE

El tercer modelo que vamos a estudiar, además de todas las características ya incluidas en el anterior, añade una nueva desagregación por edad. Esto hace que aumente del número de datos, que ahora pasan a ser 720. La nueva variable –*EDAD*– toma el valor 1 cuando nos referimos a personas menores de 25 años, y 0 en el caso contrario.

Los resultados de esta nueva regresión aparecen recogidos en la TABLA 5.4. En ella podemos ver nuevamente que el estadístico R^2 sigue siendo superior al 80%. También se siguen manteniendo, como ya ocurría en los modelos anteriores, unos valores del estadístico t superiores al 3.5 en la gran mayoría de los casos. Esto nos indica que todas las variables explicativas son significativas al 1%, y que nuestro modelo sigue manteniéndose bastante ajustado a la realidad.

En cuanto al valor de los coeficientes y a los signos que poseen, se mantienen las explicaciones que hemos dado para los modelos anteriores. Se confirman los valores positivos para todas las variables temporales y regionales, siendo especialmente significativos los periodos de finales de los ochenta y mediados de los noventa, así como la referencia andaluza, para la determinación de las tasas de paro.

También se sigue manteniendo el efecto negativo del *PIB* sobre la tasa de paro y sigue siendo muy importante el sexo en la determinación de las tasas de paro. En cuanto a la variable *EDAD*, vemos que es muy significativa, y que presenta un coeficiente positivo. Este resultado también es coherente con la situación del mercado de trabajo español. No en vano podemos observar que son los estratos más jóvenes de población los que sufren mayores tasas de desempleo.

TABLA 5.4

	COEFICIENTES	T ²	% DE LA VARIANZA	
			Efecto total	Efecto marginal
A1980	-	-	8.09	8.50
A1981	0.036513	2.690649		
A1982	0.065894	4.855406		
A1983	0.091938	6.770420		
A1984	0.122004	8.978137		
A1985	0.145597	10.69250		
A1986	0.144639	10.58007		
A1987	0.130196	9.422500		
A1988	0.131122	9.350528		
A1989	0.099331	6.957700		
A1990	0.086163	5.930199		
A1991	0.082580	5.607865		
A1992	0.107108	7.256216		
A1993	0.165954	11.32737		
A1994	0.195159	13.16745		
A1995	0.174846	11.59484		
A1996	0.179351	11.71272		
A1997	0.164984	10.49303		
A1998	0.140704	8.680140		
A1999	0.101827	6.088511		
ANDA	0.289052	10.71853	8.88	8.72
ARAG	0.049392	4.060613		
CASL	0.064572	5.254534		
CASM	0.133995	9.099950		
CATA	0.241948	6.372011		
CVAL	0.152848	7.415919		
GALI	0.073531	5.170161		
MADR	0.201294	6.626839		
PVAS	0.186387	12.13435		
PIB	-2.71E-08	-4.709983		
SEXO	-0.082109	-19.13421	7.51	7.51
EDAD	0.233897	54.50607	60.93	60.93
R²	0.858682			

Fuente: Elaboración propia a partir de datos procedentes del INE

En lo referente al porcentaje de la varianza que explica cada variable o grupo de variables, la situación sí que ha cambiado. Esto se debe fundamentalmente a que hemos introducido como fuente de variación la edad, que ha pasado a ser la variable explicativa con mayor peso sobre la variable dependiente (más del 60%). Esto ha provocado que tanto la variable *SEXO*, como los componentes regional y temporal hayan disminuido mucho sus porcentajes correspondientes (todas en torno al 8%). Finalmente, el *PIB* sigue explicando un porcentaje muy pequeño de la varianza de las tasas de paro.

Con respecto al componente regional, que como ya hemos dicho es el más relevante desde el punto de vista de la convergencia, podemos apreciar que vuelve a experimentar una fuerte disminución: se reduce a más de la tercera parte. Esto se puede interpretar como signo de una menor divergencia entre las tasas de paro regionales: cuanto menor sea el porcentaje explicado por las variables categóricas regionales, mayor es el grado de convergencia entre las mismas.

El cuarto y último paso que vamos a dar en nuestro análisis de descomposición de la varianza nos sitúa en el modelo con mayor nivel de desagregación. Ahora introducimos una nueva distinción en función del nivel de estudios. Por lo tanto, además de las variables que recogen los componentes regional y espacial, y además de las variables *PIB*, *SEXO* y *EDAD*, aparece la variable *ESTUDIOS*. Este nuevo regresor tomará el valor 1 para los que tienen estudios medios o superiores y 0 en el caso contrario. Con esta operación se vuelve a duplicar el número de observaciones, que pasan a ser 1440.

Los resultados de esta última estimación aparecen recogidos en la TABLA 5.5. Las conclusiones iniciales de esta nueva estimación vuelven a coincidir con las de los modelos anteriores. Todos los regresores vuelven a ser significativos a un nivel del 1%; y por lo que se refiere al coeficiente de determinación, éste vuelve a presentar un valor superior al ochenta por ciento.

Si nos fijamos en el valor de los coeficientes, tampoco hay grandes modificaciones con respecto a los modelos anteriores. Todos los coeficientes siguen manteniendo el mismo signo, positivo para las variables regionales, las temporales y la *EDAD* y negativo para el *PIB* y para la variable *SEXO*. Además, las variables ligadas a los años centrales de las décadas de los ochenta y los noventa, así como la ligada a Andalucía, siguen apareciendo como las que más influyen en la tasa de desempleo regional.

En cuanto a la variable *ESTUDIOS*, que no aparecía en los modelos anteriores, tenemos que destacar que es bastante significativa, y que su signo resulta positivo, como era de esperar. Este signo nos está indicando que los trabajadores con estudios medios o superiores sufren mayores tasas de desempleo que los demás (en condiciones *ceteris paribus*). Un resultado como éste podría parecer sorprendente; pero no lo es tanto, si tenemos en cuenta que se han englobado dentro de un mismo grupo a los individuos con estudios superiores y a los que solo tienen estudios medios (lo cual ha venido impuesto por las características de los datos disponibles); y ocurre que los individuos con estudios medios son los que presentan tasas de desempleo más altas, además de ser más numerosos que los de estudios superiores. Por qué los individuos con estudios medios tienen tasas de desempleo superiores a las de los otros dos grupos de estudios (estudios primarios y estudios universitarios) no está muy claro, pero tampoco es demasiado

sorprendente. Cabe pensar que, por una parte, los individuos con estudios medios no tienen la formación suficiente para acceder a los puestos de trabajo más cualificados, para los que se suele preferir gente con educación universitaria; pero, por otra parte, los individuos con estudios medios se ven a sí mismos como gente educada y no están tan dispuestos a aceptar los empleos menos cualificados como lo está la gente de educación más baja.

En cuanto al porcentaje de la varianza que explica cada grupo de variables, hay que destacar no se modifica mucho en relación con el modelo anterior. Esto quiere decir que la nueva variable introducida en este modelo –*ESTUDIOS*– aporta muy poco a la explicación de la variable dependiente. La variable *EDAD*, que explica más del 50% de la varianza, sigue siendo la más relevante, seguida muy de lejos por las variables temporales, las regionales y la variable *SEXO*, todas ellas en torno al 7%. Los porcentajes correspondientes al *PIB* y a la variable *ESTUDIOS* son inferiores al 1%.

Estos resultados vienen a corroborar nuestra idea inicial de que la convergencia regional es mayor cuando nos centramos en el estudio de grupos de población más agregados. No obstante, también podemos apreciar que sigue existiendo un porcentaje de la varianza explicado por las variables regionales (un 7% en el último modelo) que impide la completa igualación de las tasas de paro regionales para grupos de población homogéneos. Este residuo puede deberse, simplemente, al hecho de que las variables que hemos utilizado como controles de las características de la fuerza de trabajo (*EDAD*, *SEXO*, *ESTUDIOS*) son medidas inevitablemente toscas y no captan demasiado bien toda la heterogeneidad de la mano de obra.

TABLA 5.4

	COEFICIENTES	t ²	% DE LA VARIANZA	
			Efecto total	Efecto marginal
A1980	-	-		
A1981	0.0312	2.67		
A1982	0.0585	5.02		
A1983	0.0846	7.25		
A1984	0.1129	9.67		
A1985	0.1392	11.89		
A1986	0.1368	11.64		
A1987	0.1204	10.14		
A1988	0.1188	9.85		
A1989	0.0910	7.42	7.86	7.92
A1990	0.0789	6.32		
A1991	0.0727	5.75		
A1992	0.0989	7.79		
A1993	0.1587	12.60		
A1994	0.1883	14.78		
A1995	0.1697	13.06		
A1996	0.1792	13.62		
A1997	0.1557	11.52		
A1998	0.1342	9.63		
A1999	0.0923	6.42		
ANDA	0.2559	11.01	7.52	7.16
ARAG	0.0450	4.24		
CASL	0.1172	9.16		
CASM	0.0566	5.28		
CATA	0.1993	6.10		
CVAL	0.1331	7.47		
GALI	0.0597	4.83		
MADR	0.1668	6.37		
PVAS	0.1631	12.23		
PIB	-2.08E-08	-4.21		
SEXO	-0.0816	-22.13	6.99	7.00
EDAD	0.2326	63.04	56.83	56.86
ESTUDIOS	0.0183	4.97	0.35	0.35
R²		0.80		

Fuente: Elaboración propia a partir de datos procedentes del INE

De todos modos, se puede apreciar una considerable disminución del porcentaje de la varianza que explican las variables regionales según nos vamos centrando en grupos de población cada vez menos desagregados. Así pasamos de un porcentaje algo superior al 50% en el caso de menor desagregación a otro en torno al 25% cuando se desagrega en función de la edad, porcentaje que luego se reduce hasta un 8-9 % cuando se desagrega en función de la edad y el sexo, y hasta la un 7-8% cuando la desagregación se hace máxima, es decir, cuando se introducen los estudios como fuente de variación.

El análisis precedente tiene un inconveniente en la medida en que da la impresión de que los efectos regionales son de magnitud fija a lo largo del tiempo. Con el fin de ver si dichos efectos varían con los años, vamos a proceder de la siguiente manera. En primer lugar tomaremos como punto de partida el último modelo, que es el más desagregado y también el que incluye un mayor número de datos, y a partir de él, fraccionaremos la regresión en veinte regresiones distintas, una para cada año de la muestra.

Vamos a disponer pues de veinte estimaciones, cada una ellas con 72 observaciones, en las que se regresará la tasa de paro contra todas las variables regionales, el *SEXO*, la *EDAD* y el nivel de *ESTUDIOS*. Nuestro propósito no es otro que calcular el porcentaje de la varianza que representa el componente regional en cada una de ellas, y ver si dicho porcentaje aumenta o disminuye a lo largo del tiempo.

TABLA 5.6

PERIODO	R ²	% DE LA VARIANZA EXPLICADO POR EL COMPONENTE REGIONAL	
		Efecto Total	Efecto Marginal
1980	0,82	8,12	6,28
1981	0,87	6,64	5,40
1982	0,86	6,96	4,70
1983	0,87	8,41	5,92
1984	0,87	6,47	4,14
1985	0,88	7,51	4,50
1986	0,89	9,57	6,75
1987	0,87	13,58	9,78
1988	0,91	10,06	8,14
1989	0,86	14,02	13,89
1990	0,86	19,88	19,70
1991	0,85	19,68	19,62
1992	0,88	18,73	18,71
1993	0,89	12,40	12,05
1994	0,90	8,56	7,64
1995	0,89	10,28	10,19
1996	0,88	6,58	6,46
1997	0,89	11,76	11,72
1998	0,86	12,06	11,15
1999	0,77	18,07	17,96

Fuente: Elaboración propia a partir de datos proporcionados por el INE

GRÁFICO 5.1
Evolución de los efectos total y marginal entre los que varía la importancia del componente regional



Fuente: Elaboración propia a partir de TABLA 5

Para calcular el porcentaje en cuestión seguiremos el mismo camino de antes; es decir, calcularemos los efectos total y marginal, cuyos significados ya conocemos. Los resultados de estas nuevas estimaciones aparecen recogidos en la TABLA 5.5 y en el GRAFICO 5.1. En dicha tabla aparece el coeficiente de determinación correspondiente a cada año estudiado, así como el porcentaje de la varianza explicado por las variables regionales en cada caso. Más concretamente, aparecen los dos valores entre los que se espera que fluctúe este porcentaje y que son los efectos total y marginal.

A la vista de los resultados expuestos en esta última tabla podemos apreciar que los coeficientes de terminación siguen presentando valores muy elevados superando holgadamente el 80%, y en muchos de los casos incluso el 85%. En lo referente a la evolución temporal que sigue el porcentaje de la varianza explicado por las variables regionales, podemos apreciar, que tras empezar con un nivel muy bajo en los primeros años de la muestra, este porcentaje sufre un aumento progresivo y alcanza su máximo a principios de los 90. Posteriormente, sigue una senda descendiente, que se ve truncada en los últimos años de la muestra.

Este resultado es muy similar al que obteníamos en el segundo capítulo de esta tesis al analizar la evolución del coeficiente de variación de las tasas de desempleo entre regiones. La diferencia principal está en el fuerte descenso inicial que sufría el coeficiente de variación, y que no tiene paralelo en este GRÁFICO 5.1. Esto, sin embargo, no es nada extraño, ya que el coeficiente de variación se calculaba desde 1977 y entre este año y 1980 (año en que comienza el análisis de descomposición de la varianza) es donde experimentaba ese brusco descenso del que hemos hablado.

Desde el punto de vista de la convergencia podemos decir que tras una primer periodo de divergencia, que se produce durante los ochenta, durante la década de los noventa sí que se produce una cierta convergencia o acercamiento de las tasas de desempleo regionales. Dicho acercamiento se refleja en la disminución que sufre el componente regional durante este periodo, y se ve bruscamente truncado en los últimos tres años de la muestra.

Como ya hemos indicado antes, esta evolución es muy similar a la que seguía el coeficiente de variación de las tasas de paro regionales en el epígrafe 6 del capítulo

segundo, y, al igual que allí, podemos ver la clara influencia que tiene el ciclo económico en dicha evolución. En este sentido, podemos concluir que la importancia del componente regional aumenta en las fases recesivas del ciclo económico, y disminuye en las expansivas.

Los resultados principales que se obtienen de este análisis de descomposición de la varianza son los siguientes.

En primer lugar, podemos apreciar la importancia que tiene el grado de homogeneidad de grupo poblacional estudiado a la hora de medir la convergencia, y en consecuencia, la gran influencia de variables como la edad, el sexo o el nivel de estudios. En este sentido, aparece como un resultado bastante claro el hecho de que, cuando mejor definido está el grupo estudiado menor es la importancia del componente regional en nuestro estudio. Este resultado está en la misma línea que el que se obtuvo anteriormente en el análisis de cointegración, donde vimos que a medida que aumenta el grado de homogeneización del grupo tratado, es mayor el número de regiones que convergen.

En segundo lugar, hay que destacar la existencia de diferencias regionales en las tasas de desempleo que no se explican a través de la edad, el sexo, el nivel de estudios y el *PIB* (además del tiempo). Este “residuo regional no explicado” sería simplemente una consecuencia del hecho de que los grupos poblacionales considerados son todavía demasiado heterogéneos. Si pudiésemos disponer de variables más desagregadas (en términos de características personales y sectoriales) sería de esperar que este residuo se redujese aún más hasta llegar eventualmente a desaparecer.

BIBLIOTECA VIRTUAL

CAPÍTULO 6:

RESUMEN Y CONCLUSIONES

Esta tesis se ha planteado como objetivo el estudio de la convergencia de las tasas de desempleo de las regiones españolas. El primer intento que hemos hecho para abordar dicho objetivo (capítulo 2) no ha sobrepasado mucho el nivel puramente intuitivo. Hemos comenzado con un análisis descriptivo de los datos de desempleo regional, así como de los datos correspondientes a los posibles factores explicativos, y eso nos ha permitido llegar a una serie de conclusiones:

- En primer lugar, podemos apreciar un fuerte aumento en las tasas de desempleo y en las diferencias que existen entre las tasas de unas regiones y otras.
- En segundo lugar, hemos podido ver que hay ciertas características de la población, como son la edad, el nivel de estudios, o el peso de cada sexo dentro del conjunto de los activos, que influyen –o parecen influir– decisivamente sobre las diferencias regionales de desempleo.
- En tercer lugar, hemos podido apreciar también la importancia que tienen, de cara a la determinación de las diferencias regionales de desempleo, factores tales como la composición sectorial de las distintas regiones o el grado de presión que ejercen los sindicatos a través de los convenios colectivos.
- En cuarto lugar, también hemos podido observar, a través de un análisis *shift-share*, que los factores más específicamente relacionados con cada región tienen mucha más importancia que los factores poblacionales en la explicación de las diferencias interregionales de tasas de paro.
- Finalmente, una primera exploración de la dinámica de las tasas de paro regionales nos ha permitido ver que, aunque se ha producido una disminución en la dispersión de dichas tasas entre el año 1977 y el 2000, la

evolución de los años intermedios no permite concluir la existencia de un proceso de convergencia.

En el capítulo 3 de la tesis hemos tratado de aclarar los criterios teóricos sobre los que tendría que basarse un análisis de convergencia de las tasas de paro regionales que quiera ser consecuente con los presupuestos de la microeconomía convencional. Para ello nos hemos fijado en dos elementos fundamentales: el grado de homogeneidad de los trabajadores y el mecanismo de fijación de los salarios. A partir de estos elementos hemos construido cuatro escenarios alternativos. En el primer escenario suponemos que todos los trabajadores son homogéneos y que los salarios se determinan competitivamente. En el segundo escenario mantenemos el supuesto de homogeneidad de la mano de obra pero introducimos un supuesto alternativo acerca de cómo se fijan los salarios: suponemos que éstos se determinan a través de la negociación colectiva. El tercer escenario supone mano de obra heterogénea y salarios competitivos. Y el cuarto escenario supone mano de obra heterogénea y salarios negociados en convenios colectivos.

El análisis de estos cuatro escenarios nos ha llevado a la conclusión de que siempre va a haber un cierto nivel de desempleo en cada región, con independencia del escenario elegido, y que dicho desempleo es una consecuencia de la falta de sincronización entre las oferta y las demandas de trabajo. Este desempleo, que hemos denominado friccional, podrá ser mayor o menor dependiendo de las particularidades de cada región (la estructura sectorial, lo que hemos denominado “factores hedónicos”, como el clima, las infraestructuras, etc., y la mezcla de cualificaciones de los trabajadores), pero nunca va a desaparecer del todo.

Hemos visto también que, además de este paro friccional, básicamente voluntario, puede haber otra clase de paro, de carácter esencialmente involuntario, que surge cuando la negociación colectiva sustituye a las negociaciones competitivas o atomizadas. Esta clase de paro puede variar dependiendo de las características de cada región (estructura sectorial, etc.), además de depender, lógicamente, del grado de presión que ejerzan los sindicatos en los convenios colectivos de cada zona.

Todo esto nos lleva a la conclusión de que no podemos esperar que haya un proceso de convergencia, sin más cualificaciones, entre las tasas de paro de las distintas regiones. No obstante, también llegamos a la conclusión de que cuanto más homogéneo sea el grupo de población estudiado tanto más verosímil resulta la convergencia de las tasas de paro correspondientes a las distintas regiones.

Una vez establecidos los criterios teóricos anteriores, el paso siguiente consiste en verificar, desde el punto de vista econométrico, la conclusión básica de nuestro análisis: cuanto más homogéneo es el grupo de población tanto más probable resulta la convergencia de las tasas de desempleo. Con este fin procedemos en primer lugar a una revisión de los procedimientos más habitualmente empleados en los análisis de convergencia. Esta revisión se lleva a cabo en el capítulo 4 y nos sirve para situar esta tesis dentro del contexto de la literatura existente sobre el tema de la convergencia. Al final de esta discusión llegamos a la conclusión de que los métodos más convenientes, dado el objetivo de esta tesis, son dos: el análisis de cointegración y el método de descomposición de la varianza. Los resultados de estos dos métodos se encuentran en la misma línea y vienen a corroborar nuestra hipótesis básica, como se comprueba a través del análisis desarrollado en el capítulo 5.

Las conclusiones obtenidas a través del análisis de cointegración se podrían concretar en lo siguiente:

- Por un lado, podemos confirmar la ausencia de convergencia absoluta entre las regiones; cuando usamos toda la población son pocas las regiones que presentan cointegración, pero cuando nos centramos en grupos más homogéneos, es necesario introducir componentes deterministas (es decir, aspectos diferenciadores) en la mayoría de las regiones para admitir la convergencia.
- Por otro lado, se corrobora que las características de la población son sumamente importantes en este tipo de análisis; según vamos concretando el grupo en términos de edad, sexo y nivel de estudios, se aprecia un incremento considerable en número de regiones que convergen con su agregado nacional correspondiente.

Por lo que respecta al análisis de descomposición de la varianza, las conclusiones principales son las siguientes:

- Se confirma la ausencia de un claro proceso de convergencia, lo cual se concreta en dos resultados: por un lado, el porcentaje de la varianza explicado por el componente regional alcanza valores relativamente altos; y, por otro lado, dicho porcentaje no tiene una clara evolución decreciente, como sería de esperar si la hipótesis de la convergencia fuese cierta.
- Se vuelve a confirmar la importancia del sexo, la edad y el nivel de estudios, y se comprueba que cuando se tienen en cuenta dichas características de la

población el componente regional pierde importancia como factor explicativo del desempleo regional.

En suma, se puede concluir que no existe un proceso de convergencia real generalizado, en términos de tasa de desempleo, entre las distintas Comunidades Autónomas españolas. No obstante, sí que se puede admitir que, a medida que nos vamos centrando en grupos de población con características más y más homogéneas, se percibe un cierto acercamiento entre ellas. A pesar de todo, es preciso destacar la persistencia de un componente residual no explicado en las tasas de desempleo regionales. Este “factor regional” podría reflejar dos hechos: (1) los grupos de población caracterizados por edad, sexo y nivel de estudios no son todavía suficientemente homogéneos; y (2) las características regionales que pueden dar lugar a diferencias de desempleo –los “factores hedónicos” antes mencionados o el poder de los sindicatos de cada región– no han sido adecuadamente medidos. Si pudiésemos afinar más en la definición de los grupos, y si pudiésemos medir mejor todas esas características regionales, sería de esperar que el residuo acabase por desaparecer.

BIBLIOTECA VIRTUAL

MIGUEL DE
SERVANTES

APÉNDICE 1

TEORÍA DE LA COINTEGRACIÓN

El hecho de trabajar con series temporales nos obliga a realizar un análisis previo para determinar si los procesos que representan dichas series son estacionarios o no. La existencia de tendencias estocásticas, o procesos no estacionarios, como ocurre con las series de tasa de paro, hace que los contrastes de convergencia más adecuados sean los propuestos por Bernard y Durlauf (1996). Estos tests están basados en la utilización de técnicas de cointegración.

De acuerdo con la teoría econométrica, cuando tenemos series que no son estacionarias (series con un determinado número de raíces unitarias), la relación que pueda existir entre ellas, lejos de ser una relación de causalidad, puede ser espuria y deberse a simple casualidad.

Este problema fue tratado por primera vez en el trabajo de Yule (1926). Posteriormente en el trabajo de Phillips (1986) encontramos cuales son los principales problemas que resultan de aplicar mínimos cuadrados ordinarios entre variables no estacionarias: en primer lugar los estimadores dejan de ser consistentes, en segundo lugar los estadísticos t y R^2 no siguen las distribuciones estándar, y finalmente el estadístico Durbin Watson tiende hacia cero (este último resultado nos sirve de indicador para detectar la presencia de relaciones espurias entre las variables). Ante este tipo de situaciones, la teoría presenta dos opciones alternativas para solucionar el problema y garantizar que la relación considerada -si existe- es la adecuada:

La primera opción consiste en diferenciar las series con el fin de eliminar todas las raíces unitarias que posean. La teoría econométrica establece (Suriñach et al, 1995) que si una variable x_t es integrada de orden n , $x_t \rightarrow I(n)$, al tomar su primera diferencia la variable se transforma en una variable de un orden de integrabilidad menor, $\Delta x_t \rightarrow I(n-1)$. De esta manera, diferenciando tantas veces como raíces unitarias aparezcan en la variable, es posible conseguir siempre series estacionarias. El principal problema que presenta este método es que, cuando tomamos diferencias solo estamos considerando la relación entre los cambios de las variables, y por lo tanto, estamos prescindiendo de toda la información que proporcionan los niveles de la variable utilizada.

La segunda opción consiste en utilizar técnicas de cointegración. Siguiendo a Engle y Granger (1987): “Los componentes de un vector X_t se dice que están cointegrados de orden (d,b) y se denota por $X_t \rightarrow CI(d,b)$, si todas las componentes de X_t son integradas del mismo orden d , $I(d)$, y además existe un vector α , no nulo, tal que $\alpha'X_t = z_t \rightarrow I(d-b)$, con $b > 0$ ”. Es decir, existe una combinación lineal entre los componentes del vector X_t que presenta un orden de integración inferior.

Al vector α también se le denomina vector de cointegración, y bastaría con multiplicar ese vector por un escalar no nulo para obtener otro distinto. No obstante, lo que si podemos decir es que solo puede haber $(K-1)$ vectores de cointegración linealmente independientes, donde K es el número de variables incluidas en X_t .

El caso más común, y al que nos vamos a referir en este trabajo, es aquel en que $d=b=1$. Por lo tanto, todas las variables que componen el vector X son integradas de orden uno (poseen una única raíz unitaria), y la combinación lineal de ellas, z_t , sigue un proceso estacionario. Además, como el estudio que nosotros realizamos tiene lugar

entre cada región y el agregado nacional, solo vamos a tener dos variables, (K tomará el valor 2) y por lo tanto, solo vamos a poder tener una relación de cointegración en cada uno de los casos.

El hecho de que exista una relación de cointegración entre dos variables $I(1)$ implica que existe una combinación lineal de dichas variables (la que refleja el vector de cointegración) con residuos estacionarios. Este resultado, lo que nos está diciendo en el fondo, es que hay una relación de largo plazo entre las variables, y que cualquier perturbación que se produzca sobre esa relación tiene un carácter transitorio, aunque el efecto sea permanente sobre cada una de las variables.

Para terminar este apéndice podemos decir que esta relación de equilibrio es la que nos va a permitir extraer conclusiones sobre la existencia de convergencia. De esta manera, diremos que hay convergencia entre dos variables siempre que el único vector de cointegración que exista entre ellas sea el vector $(1, -1)$, o dicho en otras palabras, siempre que la diferencia entre dichas variables sea estacionaria²¹. Posteriormente, si para que esa diferencia sea estacionaria debemos introducir constantes o tendencias deterministas, vamos a poder diferenciar entre distintos grados de convergencia entre dichas variables.

²¹ Debemos tener en cuenta que, para que exista cointegración, como nuestras variables son $I(1)$, la combinación lineal debe tener un grado de integración menor, y por lo tanto, necesariamente debe ser $I(0)$ (estacionaria).

BIBLIOTECA VIRTUAL

MIGUEL DE
SERVANTES

APÉNDICE 2

TESTS DE RAÍCES UNITARIAS

En este apéndice vamos a recoger los resultados de aplicar los tests de raíces unitarias sobre las distintas series utilizadas en el análisis de cointegración. Las seis primeras tablas se refieren a los tests de raíces unitarias sobre las distintas tasas de paro de cada región y grupo poblacional utilizado. Las tres siguientes recogen esos mismos tests para la diferencia de tasas de paro entre cada región y el agregado nacional. Estas últimas tablas, además de informarnos sobre si las diferencias entre variables son estacionarias o no, también nos dicen si es necesaria la introducción de términos independientes o tendencias deterministas que son importantes a la hora de analizar el grado de convergencia existente. Finalmente, la décima tabla se refiere al número de raíces unitarias que se han encontrado en la serie del PIB de cada región estudiada en la descomposición de la varianza.

Desde el punto de vista práctico, y con el fin de hacer más comprensibles los resultados que aparecen recogidos en las siguientes tablas, se va a intentar aportar una explicación previa del funcionamiento de los distintos tests de raíces unitarias. De esta manera, será más fácil dar una interpretación adecuada a los resultados.

En esta tesis vamos a aplicar dos tests de raíces unitarias: en primer lugar, el propuesto por Dickey y Fuller (1981), también conocido como *Augmented Dickey-Fuller (ADF)*, y en segundo lugar, el propuesto por Phillips y Perron (1988) y conocido como *Phillips y Perron (PP)*. Estos tests nos van a permitir determinar cual es el grado de integración que presentan los niveles de las variables. Antes de comenzar con el

desarrollo de cada uno, sería conveniente decir que la aplicación de estos dos tests viene incluida dentro de varios programas estadísticos. El que nosotros hemos utilizado es el programa *e-views*.

Para el primero de los tests, si el valor que toma el estadístico *ADF* es menor que el de los valores tabulados por McKinnon (1991) para diferentes niveles de significación (número que aparece entre paréntesis), entonces se puede concluir que la variable es estacionaria. En caso contrario se toma la primera diferencia de la variable y se repite el proceso. Si tras diferenciar la variable, los estadísticos si que son menores que los valores tabulados, entonces diremos que la variable es integrada de orden 1; en caso contrario, se vuelven a tomar diferencias y se inicia nuevamente el test, y así sucesivamente.

La interpretación hecha para el segundo de los test es muy similar, cuando el valor del estadístico *PP* es menor que los valores tabulados por Fuller (1976) (número que aparece entre paréntesis), entonces podemos admitir la estacionariedad de la variable. En caso contrario se toman diferencias y se repite el proceso igual que ocurría con el *ADF*.

Sin entrar en más detalles, que pueden ser consultados en Phillips y Perron (1988) y Dickey y Fuller (1981), podemos decir que, a la vista de los resultados que presentamos a continuación, la mayoría de las variables utilizadas en este trabajo resultaron ser integradas de orden 1 en niveles, a un nivel de significación del 5% (véanse TABLAS A2.1..A2.6.).

En el caso de las diferencias entre variables (TABLAS A2.7... A2.9), podemos apreciar cuáles son las regiones que presentan cierto grado de convergencia (aquellas que proporcionan un resultado estacionario). En todas estas regiones se puede admitir la existencia de una relación de largo plazo con el agregado nacional. También vemos cuáles de esas regiones necesita de factores deterministas (términos independientes o tendencias) para conseguir la convergencia, y como el número de regiones con resultado estacionario aumenta según avanzamos en el nivel de homogeneización del grupo estudiado.



TABLA A2.1
Test ADF. Tasas de paro de toda la población

	REGION		AGREGADO NACIONAL	
	Niveles	Diferencias	Niveles	Diferencias
ANDALUCIA	-0.3146 (-3.4561)	-10.7765 (-3.4566)	-2.7485 (-2.8915)	-2.1096 (-1.9434)
ARAGON	-2.3824 (-2.8918)	-5.1138 (-3.4571)	-2.8314 (-2.8918)	-2.1108 (-1.9434)
ASTURIAS	-2.6573 (-2.8912)	-9.0722 (-3.4566)	-2.8992* (-2.8918)	-2.2250 (-1.9434)
BALEARES	-1.9937 (-2.8912)	-3.8393 (-1.9435)	-2.7742 (-2.8918)	-2.0689 (-1.9435)
CANARIAS	-0.0038 (-3.4561)	-9.6459 (-3.4566)	-2.7354 (-2.8918)	-1.9574 (-1.9435)
CANTABRIA	-2.0272 (-2.8912)	-9.1619 (-3.4566)	-2.7206 (-2.8918)	-3.3326* (-3.4566)
CASTILLA-LEON	-2.3291 (-2.8918)	-3.9857 (-3.4571)	-2.8904 (-2.8918)	-2.3171 (-1.9434)
CASTILLA-LA MANCHA	-2.7072 (-2.8918)	-3.7435 (-3.4571)	-2.6433 (-2.8918)	-2.1569 (-1.9435)
CATALUÑA	-2.6505 (-2.8918)	-2.3210 (-1.9435)	-2.6003 (-2.8918)	-1.9969 (-1.9435)
EXTREMADURA	-2.0409 (-3.4566)	-13.6049 (-3.4566)	-2.7424 (-2.8918)	-2.4117 (-1.9435)
GALICIA	-2.0849 (-2.8918)	-4.6369 (-3.4571)	-2.6636 (-2.8918)	-2.2956 (-1.9435)
MADRID	-2.8685 (-2.8918)	-4.3933 (-3.4571)	-2.6850 (-2.8918)	-2.0610 (-1.9435)
MURCIA	-1.4952 (-2.8912)	-9.9440 (-3.4566)	-2.7865 (-2.8918)	-2.3451 (-1.9434)
NAVARRA	-1.2516 (-3.4561)	-10.8900 (-3.4566)	-2.6970 (-2.8918)	-2.5523 (-1.9434)
PAIS VASCO	-0.6488 (-3.4561)	-5.5007 (-3.4571)	-2.7150 (-2.8918)	-2.3353 (-1.9435)
RIOJA	-2.1528 (-2.8918)	-8.1057 (-3.4566)	-2.8277 (-2.8918)	-2.2459 (-1.9434)
VALENCIA	-2.1703 (-2.8918)	-7.4103 (-3.4566)	-2.6900 (-2.8918)	-2.1263 (-1.9435)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%

*Para valores críticos distintos del 5% puede admitirse la existencia de una raíz unitaria

Fuente: Elaboración propia a partir de datos del INE

TABLA A2.2
Test ADF. Tasas de paro de varones mayores de 25 años

	REGION		AGREGADO NACIONAL	
	Niveles	Diferencias	Niveles	Diferencias
ANDALUCIA	-1.7383 (-2.8912)	-96725 (-3.4566)	-3.0901 (-2.8918)	-2.1449 (-1.9535)
ARAGON	-1.8855 (-2.8912)	-5.6419 (-1.9435)	-2.7515 (-2.8918)	-2.6429 (-1.9535)
ASTURIAS	-19583 (-2.8912)	-10.4836 (-1.9434)	-2.9883* (-2.8918)	-2.1106 (-1.9535)
BALEARES	-1.1480 (-2.8912)	-10.7293 (-1.9434)	-2.8180 (-2.8918)	-2.5560 (-1.9535)
CANARIAS	-0.3356 (-3.4561)	-10.7363 (-3.4566)	-2.8192 (-2.8918)	-2.6519 (-1.9535)
CANTABRIA	-2.0874 (-2.8912)	-11.5617 (-3.4566)	-2.8248 (-2.8918)	-2.5801 (-1.9535)
CASTILLA-LEON	-1.0768 (-3.4571)	-8.5783 (-3.4566)	-2.6980 (-2.8918)	-2.7842 (-1.9535)
CASTILLA-LA MANCHA	-1.7028 (-2.8912)	-9.1078 (-3.4566)	-2.7685 (-2.8918)	-2.5991 (-1.9535)
CATALUÑA	-0.6075 (-1.9434)	-5.1977 (-1.9434)	-2.4989 (-2.8918)	-2.3643 (-1.9535)
EXTREMADURA	-3.6225* (-3.4561)	-10.3045 (-1.9434)	-2.8797 (-2.8918)	-2.5576 (-1.9535)
GALICIA	-2.2195 (-2.8918)	-5.6226 (-3.4571)	-2.7635 (-2.8918)	-2.6187 (-1.9535)
MADRID	-2.3498 (-2.8918)	-3.9785 (-3.4576)	-2.6964 (-2.8918)	-2.3808 (-1.9535)
MURCIA	-1.6698 (-2.8912)	-4.1130 (-1.9435)	-2.7821 (-2.8918)	-2.6641 (-1.9535)
NAVARRA	-1.2645 (-3.4561)	-10.9932 (-3.4566)	-2.7923 (-2.8918)	-2.6219 (-1.9535)
PAIS VASCO	-0.3378 (-3.4561)	-8.4112 (-3.4566)	-2.8144 (-2.8918)	-2.6048 (-1.9535)
RIOJA	-1.9568 (-2.8912)	-10.6234 (-1.9434)	-2.7986 (-2.8918)	-2.6017 (-1.9535)
VALENCIA	-2.0713 (-2.8918)	-3.0594 (-1.9435)	-2.6837 (-2.8918)	-2.7467 (-1.9535)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%

*Para valores críticos distintos del 5% puede admitirse la existencia de una raíz unitaria

Fuente: Elaboración propia a partir de datos del INE

TABLA A2.3

Test ADF. Tasas de paro de varones mayores de 25 años con estudios superiores.

	REGION		AGREGADO NACIONAL	
	Niveles	Diferencias	Niveles	Diferencias
ANDALUCIA	-2.1881 (-2.8912)	-11.1477 (-3.4566)	-2.1857 (-2.8918)	-5.9356 (-3.4571)
ARAGON	-2.6865 (-2.8912)	-11.0249 (-1.9434)	-2.1741 (-2.8918)	-9.1720 (-3.4566)
ASTURIAS	-2.5145 (-2.8915)	-12.6042 (-1.9434)	-2.2254 (-2.8912)	-10.1349 (-3.4566)
BALEARES	-3.1431* (-2.8912)	-9.4314 (-1.9434)	-1.8081 (-3.4566)	-12.2699 (-3.4566)
CANARIAS	-2.5000 (-2.8915)	-12.8012 (-1.9434)	-2.1771 (-2.8918)	-9.4053 (-3.4566)
CANTABRIA	-2.0014 (-2.8915)	-13.8626 (-1.9434)	-2.1559 (-2.8918)	-9.0628 (-3.4566)
CASTILLA-LEON	-2.9586* (-2.8912)	-7.3579 (-1.9435)	-2.2062 (-2.8912)	-9.1735 (-3.4566)
CASTILLA-LA MANCHA	-3.3169* (-2.8912)	-10.6232 (-1.9434)	-2.1777 (-2.8918)	-9.4098 (-3.4566)
CATALUÑA	-2.7100 (-2.8912)	-11.6758 (-1.9434)	-2.2992 (-2.8912)	-9.5432 (-3.4566)
EXTREMADURA	-2.9557* (-2.8915)	-15.6535 (-1.9434)	-2.1948 (-2.8918)	-8.7711 (-3.4566)
GALICIA	-2.6934 (-2.8912)	-11.6758 (-1.9434)	-2.1654 (-2.8918)	-9.2449 (-3.4566)
MADRID	-2.5544 (-2.8918)	-10.4648 (-3.4571)	-2.3140 (-2.8918)	-9.5199 (-3.4566)
MURCIA	-2.3169 (-2.8922)	-13.9710 (-1.9434)	-2.2280 (-2.8918)	-9.2306 (-3.4566)
NAVARRA	-2.0317 (-2.8915)	-13.1068 (-1.9434)	-2.1377 (-2.8915)	-8.1118 (-3.4566)
PAIS VASCO	-0.3378 (-3.4566)	-8.4112 (-3.4566)	-2.8144 (-2.8918)	-4.6223 (-3.4566)
RIOJA	-2.8144 (-2.8912)	-10.9721 (-1.9434)	-2.2220 (-2.8918)	-5.4192 (-3.4571)
VALENCIA	-2.3549 (-2.8912)	-11.1812 (-1.9434)	-2.2408 (-2.8912)	-9.6755 (-3.4566)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%

*Para valores críticos distintos del 5% puede admitirse la existencia de una raíz unitaria

Fuente: Elaboración propia a partir de datos del INE

TABLA A2.4
Test PP. Tasas de paro de toda la población

	REGION		AGREGADO NACIONAL	
	Niveles	Diferencias	Niveles	Diferencias
ANDALUCIA	-0.3190 (-3.4561)	-10.7342 (-3.4566)	-0.6092 (-3.4561)	-2.2522 (-1.9434)
ARAGON	-1.0089 (-3.4561)	-8.8626 (-3.4566)	-0.1122 (-3.4561)	-2.1046 (-1.9434)
ASTURIAS	-2.5191 (-2.8912)	-9.1000 (-3.4566)	-0.1864 (-3.4561)	-2.1880 (-1.9434)
BALEARES	-2.0219 (-2.8912)	-9.6815 (-3.4566)	-0.2966 (-3.4561)	-3.5209 (-3.4566)
CANARIAS	-0.0379 (-3.4561)	-9.0512 (-3.4566)	-0.3147 (-3.4561)	-3.4061* (-3.4566)
CANTABRIA	-0.1047 (-3.4561)	-9.1708 (-3.4566)	-0.0687 (-3.4561)	-3.2850* (-3.4566)
CASTILLA-LEON	0.0366 (-3.4561)	-8.1278 (-3.4566)	-0.2010 (-3.4561)	-2.2605 (-1.9434)
CASTILLA-LA MANCHA	-0.7236 (-3.4561)	-8.3494 (-3.4566)	-0.2758 (-3.4561)	-3.6130 (-3.4566)
CATALUÑA	-1.0701 (-3.4561)	-3.9845 (-3.4566)	-0.1482 (-3.4561)	-3.8677 (-3.4566)
EXTREMADURA	-2.2065 (-3.4561)	-13.5793 (-3.4566)	-0.0427 (-3.4561)	-2.3175 (-1.9434)
GALICIA	-1.9078 (-2.8912)	-8.4915 (-3.4566)	-0.3649 (-3.4561)	-3.9764 (-3.4566)
MADRID	-1.0604 (-3.4561)	-7.4387 (-3.4566)	-0.1501 (-3.4561)	-3.5930 (-3.4566)
MURCIA	-1.4640 (-2.8912)	-9.9430 (-3.4566)	-0.1877 (-3.4561)	-2.2426 (-1.9434)
NAVARRA	-1.2302 (-3.4561)	-10.8490 (-3.4566)	-0.0424 (-3.4561)	-2.4285 (-1.9434)
PAIS VASCO	-0.7155 (-3.4561)	-9.5523 (-3.4566)	-0.2962 (-3.4561)	-3.7517 (-3.4566)
RIOJA	-0.6151 (-3.4561)	-8.2736 (-3.4566)	-0.1392 (-3.4561)	-2.1074 (-1.9434)
VALENCIA	-0.4149 (-3.4561)	-7.5383 (-3.4566)	-0.3130 (-3.4561)	-3.8640 (-3.4566)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%

*Para valores críticos distintos del 5% puede admitirse la existencia de una raíz unitaria

Fuente: Elaboración propia a partir de datos del INE

TABLA A2.5
Test PP. Tasas de paro de los varones mayores de 25 años

	REGION		AGREGADO NACIONAL	
	Niveles	Diferencias	Niveles	Diferencias
ANDALUCIA	-1.8515 (-2.8912)	-9.7373 (-3.4566)	-0.6134 (-3.4561)	-3.6804 (-3.4566)
ARAGON	-2.0082 (-2.8912)	-10.3878 (-3.4566)	-0.5399 (-3.4561)	-4.8442 (-3.4566)
ASTURIAS	-1.9580 (-2.8912)	-10.4608 (-1.9434)	-0.5241 (-3.4561)	-3.8133 (-3.4566)
BALEARES	-2.0780 (-2.8912)	-10.7814 (-1.9434)	-0.5638 (-3.4561)	-4.8105 (-3.4566)
CANARIAS	-0.2905 (-3.4561)	-10.6989 (-3.4566)	-0.6134 (-3.4561)	-4.8673 (-3.4566)
CANTABRIA	-2.0419 (-2.8912)	-11.5602 (-3.4566)	-0.5705 (-3.4561)	-4.7865 (-3.4566)
CASTILLA-LEON	-0.6760 (-3.4561)	-8.7071 (-3.4566)	-0.5941 (-3.4561)	-4.9517 (-3.4566)
CASTILLA-LA MANCHA	-1.8056 (-2.8912)	-9.9062 (-3.4566)	-0.5641 (-3.4561)	-4.8545 (-3.4566)
CATALUÑA	-1.2018 (-3.4561)	-5.5037 (-3.4566)	-0.4529 (-3.4561)	-5.6137 (-3.4566)
EXTREMADURA	-3.6207* (-3.4561)	-10.3005 (-1.9434)	-0.5414 (-3.4561)	-4.6290 (-3.4566)
GALICIA	-1.8526 (-2.8912)	-9.9109 (-3.4566)	-0.6123 (-3.4561)	-5.1586 (-3.4566)
MADRID	-1.2629 (-3.4561)	-9.3927 (-3.4566)	-0.4603 (-3.4561)	-4.5388 (-3.4566)
MURCIA	-1.7483 (-2.8912)	-9.7092 (-3.4566)	-0.5805 (-3.4561)	-5.0118 (-3.4566)
NAVARRA	-1.2072 (-3.4561)	-10.9751 (-3.4566)	-0.5601 (-3.4561)	-4.8796 (-3.4566)
PAIS VASCO	-0.6022 (-3.4561)	-8.4622 (-3.4566)	-0.5736 (-3.4561)	-4.5250 (-3.4566)
RIOJA	-1.9572 (-2.8912)	-10.0617 (-1.9434)	-0.5564 (-3.4561)	-4.8410 (-3.4566)
VALENCIA	-0.6000 (-3.4561)	-7.9498 (-3.4566)	-0.5767 (-3.4561)	-5.2873 (-3.4566)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%

*Para valores críticos distintos del 5% puede admitirse la existencia de una raíz unitaria

Fuente: Elaboración propia a partir de datos del INE

TABLA A2.6

Test PP. Tasas de paro de los varones mayores de 25 años con estudios superiores.

	REGION		AGREGADO NACIONAL	
	Niveles	Diferencias	Niveles	Diferencias
ANDALUCIA	-2.1040 (-2.8912)	-11.2972 (-3.4566)	-2.2002 (-2.8912)	-10.7518 (-2.8912)
ARAGON	-2.6850 (-2.8912)	-11.0588 (-1.9434)	-1.2948 (-3.4561)	-9.2476 (-2.8912)
ASTURIAS	-2.6947 (-2.8912)	-13.1329 (-1.9434)	-2.2460 (-2.8912)	-10.1324 (-2.8912)
BALEARES	-3.2460 (-2.8912)	-9.4508 (-1.9434)	-1.8802 (-3.4561)	-12.2374 (-2.8912)
CANARIAS	-2.8801 (-2.8912)	-12.8059 (-1.9434)	-1.3554 (-3.4561)	-9.4724 (-2.8912)
CANTABRIA	-3.4480 (-3.4561)	-14.2163 (-1.9434)	-1.3548 (-3.4561)	-9.1473 (-2.8912)
CASTILLA-LEON	-2.7973 (-2.8912)	-11.2003 (-1.9434)	-2.2254 (-2.8912)	-9.2282 (-2.8912)
CASTILLA-LA MANCHA	-3.2166* (-2.8912)	-10.6007 (-1.9434)	-2.2405 (-2.8912)	-9.4697 (-2.8912)
CATALUÑA	-2.6835 (-2.8912)	-11.0386 (-1.9434)	-2.3036 (-2.8912)	-9.5564 (-2.8912)
EXTREMADURA	-3.8827** (-2.8912)	-16.0846 (-1.9434)	-1.3202 (-3.4561)	-8.8790 (-2.8912)
GALICIA	-2.5754 (-2.8912)	-11.7241 (-1.9434)	-1.3359 (-3.4561)	-9.3314 (-2.8912)
MADRID	-2.6453 (-2.8912)	-12.5604 (-1.9434)	-2.2137 (-2.8912)	-9.5735 (-2.8912)
MURCIA	-4.7125** (-3.4561)	-14.6125 (-1.9434)	-1.3957 (-3.4561)	-9.3270 (-2.8912)
NAVARRA	-2.4745 (-2.8912)	-13.3480 (-1.9434)	-0.3045 (-3.4561)	-8.1902 (-2.8912)
PAIS VASCO	-0.6022 (-3.4561)	-8.4622 (-3.4566)	-0.5736 (-3.4561)	-4.5250 (-2.8912)
RIOJA	-2.6978 (-2.8912)	-11.0893 (-1.9434)	-1.2778 (-3.4561)	-9.2308 (-2.8912)
VALENCIA	-2.2448 (-2.8912)	-11.2472 (-1.9434)	-2.2656 (-2.8912)	-9.6986 (-2.8912)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%

*Para valores críticos distintos del 5% puede admitirse la existencia de una raíz unitaria

**Series estacionarios en niveles para valores críticos de hasta el 1%.

Fuente: Elaboración propia a partir de datos del INE

**TABLA A2.7. Tasas de paro de toda la población.
Test de raíces unitarias sobre diferencias entre variables**

	ADF		P-P	
	NIVELES	DIFERENCIAS	NIVELES	DIFERENCIAS
ANDALUCIA	-2.2394 (-2.8915)	-12.5132 (-1.9434)	-2.5413* (-2.8912)	C -
ARAGON	-4.3177 (-3.4561)	-	-4.3727 (-3.4561)	C T -
ASTURIAS	-2.0172 (-3.4561)	-9.5639 (-1.9434)	-2.0836 (-3.4561)	-9.5600 (-1.9434)
BALEARES	-1.7878 (-2.8918)	-4.9274 (-3.4576)	-2.4587 (-2.8912)	-11.1598 (-3.4566)
CANARIAS	-1.2112 (-1.9434)	-4.9184 (-1.9435)	-1.1411 (-1.9434)	-11.0948 (-1.9434)
CANTABRIA	-2.4954 (-2.8912)	-6.5562 (-1.9435)	-2.5287 (-2.8912)	-10.4411 (-1.9434)
CASTILLA-LEON	-1.8876 (-3.4561)	-10.1458 (-1.9434)	-1.9330 (-3.4561)	-10.1404 (-1.9434)
CASTILLA-LA MANCHA	-2.0092 (-2.8912)	-10.9515 (-1.9434)	-1.9911 (-2.8912)	-10.9436 (-1.9434)
CATALUÑA	-2.2749 (-3.4571)	-4.4060 (-1.9435)	-1.8711 (-3.4561)	-7.8769 (-1.9434)
EXTREMADURA	-0.9276 (-1.9434)	-13.2742 (-1.9434)	-2.9980 (-3.4561)	-13.1345 (-1.9434)
GALICIA	-1.9327 (-3.4571)	-5.2464 (-3.4571)	-1.7253 (-3.4561)	-9.7226 (-3.4566)
MADRID	-1.6663* (-1.9434)	-	-1.7746* (-1.9434)	-
MURCIA	-3.4355 (-1.9434)	-	-3.3863 (-1.9434)	-
NAVARRA	0.2794 (-1.9434)	-11.5900 (-1.9434)	-2.0577 (-3.4561)	-11.5547 (-1.9434)
PAIS VASCO	-1.9410 (-3.4566)	-12.1187 (-3.4566)	-1.9451 (-3.4561)	-12.2704 (-3.4566)
RIOJA	-3.4208 (-2.8918)	-	-3.1320 (-2.8912)	C -
VALENCIA	-3.1315 (-2.8912)	-	-2.9787 (-2.8912)	C -

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%.

C y T, indican la existencia de término independiente o tendencia respectivamente.

*Se admite la existencia de estacionariedad para un nivel de significación del 10%.

Fuente: Elaboración propia a partir de datos del INE

**TABLA A2.8. Tasas de paro de los varones mayores de 25 años.
Test de raíces unitarias sobre diferencias entre variables**

	ADF		P-P	
	NIVELES	DIFERENCIAS	NIVELES	DIFERENCIAS
ANDALUCIA	-3.1320* (-2.8912)	C	-	-
ARAGON	-0.1259 (-3.4561)		-10.4811 (-3.4566)	-10.4519 (-3.4566)
ASTURIAS	-3.2392* (-3.4561)	C T	-	-12.7363 (-1.9434)
BALEARES	-3.3195 (-2.8912)	C	-	-
CANARIAS	-1.2112 (-1.9434)		-11.0692 (-1.9434)	-11.0948 (-1.9434)
CANTABRIA	-2.6721* (-2.8912)	C	-	-
CASTILLA-LEON	-0.4484 (-1.9434)		-8.9328 (-1.9434)	-8.9383 (-1.9434)
CASTILLA-LA MANCHA	-2.1989 (-2.8912)		-9.7675 (-3.4566)	-9.7809 (-3.4566)
CATALUÑA	-2.0238 (-3.4561)		-8.5703 (-1.9434)	-8.5066 (-1.9434)
EXTREMADURA	-0.4586 (-1.9434)		-9.6377 (-1.9434)	-9.6428 (-1.9434)
GALICIA	-2.3011 (-3.4566)		-12.8804 (-1.9434)	-12.6484 (-1.9434)
MADRID	-1.1288 (-1.9434)		-12.4439 (-1.9434)	-12.5548 (-1.9434)
MURCIA	-2.5492 (-1.9434)		-	-
NAVARRA	-1.4760 (-2.8912)		-10.6678 (-1.9434)	-10.6350 (-1.9434)
PAIS VASCO	-2.2929 (-2.8918)		-9.0671 (-1.9435)	-
RIOJA	-3.3572 (-2.8912)	C	-	-
VALENCIA	-2.5730* (-2.8912)	C	-	-11.5744 (-1.9434)

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%
C y T, indican la existencia de término independiente o tendencia respectivamente.
*Se admite la existencia de estacionariedad para un nivel de significación del 10%..
Fuente: Elaboración propia a partir de datos del INE

TABLA A2.9. Tasas de paro de varones mayores de 25 años con estudios superiores. Test de raíces unitarias sobre diferencias entre variables

	ADF		P-P	
	NIVELES	DIFERENCIAS	NIVELES	DIFERENCIAS
ANDALUCIA	-2.0809 (-1.9434)	-	-2.7150 (-1.9434)	-
ARAGON	-4.6291 (-3.4561)	C T	-4.6657 (-3.4561)	C T
ASTURIAS	-5.7702 (-3.4561)	C T	-5.7584 (-3.4561)	C T
BALEARES	-2.9265 (-3.4561)	-6.7393 (-1.9435)	-2.9024 (-3.4561)	-10.2073 (-1.9434)
CANARIAS	-5.3437 (-2.8912)	C	-5.3966 (-2.8912)	C
CANTABRIA	-3.2876* (-3.4566)	C T	-4.2915 (-3.4561)	C T
CASTILLA-LEON	-4.5781 (-2.8912)	C	-4.5019 (-2.8912)	C
CASTILLA-LA MANCHA	-3.6981 (-2.8912)	C	-3.9094 (-2.8912)	C
CATALUÑA	-3.3595 (-2.8912)	C	-3.2727 (-2.8912)	C
EXTREMADURA	-3.2573 (-1.9434)	-	-5.5411 (-2.8912)	C
GALICIA	-5.1266 (-3.4561)	C T	-5.1695 (-3.4561)	C T
MADRID	-2.4194 (-1.9435)	-	-3.7476 (-1.9434)	-
MURCIA	-4.7169 (-1.9435)	-	-5.0145 (-1.9434)	-
NAVARRA	-5.2086 (-3.4561)	C T	-5.2803 (-3.4561)	C T
PAIS VASCO	-2.3148 (-2.8915)	-12.6276 (-1.9434)	-2.8440* (-2.8912)	C
RIOJA	-3.7541 (-2.8912)	C	-3.6449 (-2.8912)	C
VALENCIA	-4.8255 (-2.8912)	C	-4.9074 (-2.8912)	C

Entre paréntesis aparecen los valores críticos para un nivel de significación del 5%
 C y T, indican la existencia de término independiente o tendencia respectivamente.
 *Se admite la existencia de estacionariedad para un nivel de significación del 10%.
 Fuente: Elaboración propia a partir de datos del INE

**TABLA A2.10. Producto Interior Bruto.
Número de raíces unitarias que tiene cada serie**

	ADF	P-P
ANDALUCIA	0*	1
ARAGON	0*	1
CASTILLA-LEON	2	1
CASTILLA-LA MANCHA	2	1*
CATALUÑA	1	2
GALICIA	0	2
MADRID	1	1
PAIS VASCO	2	2
VALENCIA	1	2

*Número de raíces para un nivel de significación del 10%.
Fuente: Elaboración propia a partir de datos del INE

BIBLIOTECA VIRTUAL

MIGUEL DE
SERVANTES

APÉNDICE 3

HOMOGENEIZACIÓN DE DATOS

En este apéndice recogemos el proceso de homogeneización que tiene lugar mediante la aplicación del programa TSW. Este programa cumple tres funciones fundamentales, por un lado nos proporciona una serie desestacionalizada, en segundo lugar, identifica los datos atípicos corrigiendo su influencia en nuestra serie, y en tercer lugar, proporciona un valor para los datos ausentes de nuestra muestra.

Los datos que aparecen en las TABLAS A3.1, A3.2 y A3.3 se refieren tanto a los datos atípicos corregidos como a los ausentes interpolados, y se identifican con el año y trimestre al que corresponden.

En esas mismas tablas que ya hemos dicho, y que son las que aparecen a continuación, nos encontramos con tres tipos de datos atípicos:

OA (Outlyers Aditivos): Hacen referencia a datos atípicos que solo afectan al dato concreto y luego no vuelven a afectar.

CT (Cambios Transitorios): Son alteraciones, que tras aparecer en el modelo, permanecen en él durante algunos periodos y luego desaparecen.

CN (Cambios de Nivel): Son modificaciones, que una vez que aparecen en el modelo, permanecen en él hasta el final.

Otra consideración que sería conveniente hacer en cuanto a la homogeneización de datos, proviene de los resultados que aparecen en la última de las tablas de este primer apéndice. En esa TABLA A3.3 podemos apreciar que existen muchos datos atípicos en determinadas regiones, en especial para Baleares, eso se debe a que en esos atípicos están incluidos los datos ausentes que estima el programa.

Debemos tener en cuenta que según avanzamos en el nivel de desagregación, los grupos poblacionales son cada vez más pequeños, y por lo tanto en regiones donde la población es baja, al recurrir a grupos demasiado concretos (en nuestro caso varones parados mayores de 25 años con estudios superiores de regiones como Baleares, Murcia o La Rioja) puede haber problemas muestrales que nos impidan disponer de datos fiables desde el punto de vista estadístico.

En concreto, el problema aparece cuando buscamos una muestra de grupos poblacionales reducidos, en estos casos es posible que los datos que obtengamos no sean lo suficientemente fiables como para hacer inferencia a partir de ellos. En términos del Instituto Nacional de Estadística (INE), el número mínimo de individuos que harían fiable una muestra estaría en torno a 5000.

Por otra parte, el hecho de corregir todos esos datos tiene la finalidad de buscar una serie suficientemente suavizada e impedir que obtengamos un orden de integración incorrecto para nuestras series. Por lo tanto prevalece el hecho de obtener una serie suficientemente homogénea, aunque no exista una justificación adecuada para todos los atípicos encontrados.

Para cualquier consideración adicional sobre la técnica y su desarrollo más concreto, se puede recurrir a la bibliografía referida a este programa que hemos detallado dentro del apartado referido a la cointegración.



TABLA A3.1
Datos atípicos detectados y corregidos por el programa TSW en las series de población total

	REGION			AGREGADO NACIONAL		
	OA	CT	CN	OA	CT	CN
ANDALUCIA	-	-	1980.1 1984.1	-	-	1999.1
ARAGON	-	1987.2	-	1977.4 1980.2	-	1984.1 1988.4
ASTURIAS	1973.3	1980.1	-	-	-	1984.1 1988.4
BALEARES	1980.2	2000.2	1999.2	-	-	-
CANARIAS	-	-	-	1977.2	-	1978.1
CANTABRIA	1978.4	-	-	1977.4	-	1988.4
CASTILLA-LEON	1976.3 1977.1	-	-	-	-	1984.1 1988.4
CASTILLA-LA MANCHA	1987.1	-	-	1977.2	-	1978.1
CATALUÑA	1992.3	1988.1	-	-	-	1984.1 1988.4
EXTREMADURA	-	-	1984.1 1993.1	1977.2	-	1978.1 1988.4
GALICIA	1978.2	-	1979.3	1977.2	-	1978.1
MADRID	-	1985.4	1988.4	-	-	-
MURCIA	1977.4	-	-	1977.4	1977.2	1984.1 1988.4
NAVARRA	-	1985.2	-	1977.4	-	1988.4
PAIS VASCO	1987.1	-	1992.4	1977.2	-	-
RIOJA	1981.3 1992.3	-	1986.3	1977.4	-	1984.1 1988.4
VALENCIA	-	-	1977.3	-	-	-

Fuente: Elaboración propia a partir de datos del INE

TABLA A3.2
Datos atípicos detectados y corregidos por el programa TSW en las series de
varones mayores de 25 años

	REGION			AGREGADO NACIONAL		
	OA	CT	CN	OA	CT	CN
ANDALUCIA	1990.3	1978.4	1980.1 1984.1	-	-	-
ARAGON	-	-	1977.3	-	-	-
ASTURIAS	-	-	-	1978.2 1982.2	-	-
BALEARES	-	-	-	-	-	-
CANARIAS	-	-	-	-	-	-
CANTABRIA	1994.1	1984.1	1983.3	-	-	-
CASTILLA-LEON	1977.1	-	1992.1	-	-	-
CASTILLA-LA MANCHA	-	-	1979.3	-	-	-
CATALUÑA	-	1988.1	-	-	-	-
EXTREMADURA	1984.2 1987.2 1989.1	-	1984.1 1993.1 1999.2	-	-	-
GALICIA	-	-	-	-	-	-
MADRID	-	-	-	-	-	-
MURCIA	-	1977.4	-	-	-	-
NAVARRA	1988.1	-	1993.1	-	-	-
PAIS VASCO	-	-	1978.2	-	-	-
RIOJA	1981.2	-	-	1982.2	-	-
VALENCIA	1978.4 1984.4 1999.1	1984.4	1979.2	-	-	-

Fuente: Elaboración propia a partir de datos del INE

TABLA A3.3
Datos atípicos detectados y corregidos por el programa TSW en las series de
varones mayores de 25 años con estudios superiores

	REGION			AGREGADO NACIONAL		
	OA	CT	CN	OA	CT	CN
ANDALUCIA	-	-	1992.4	1984.3	-	-
ARAGON	1976.3 1981.4	1982.4	-	1980.3 1984.3	-	-
ASTURIAS	-	1985.3	-	-	-	-
BALEARES	1985.2 1991.2 1994.3	1979.2 1979.3 1979.4 1985.1 1985.3 1991.4 1992.1 1992.2	-	1991.2	1979.2 1979.4 1991.4	1985.1 1985.3 1992.2
CANARIAS	-	-	-	1984.3	-	-
CANTABRIA	-	-	1988.3	1980.3 1984.3	-	-
CASTILLA-LEON	-	-	-	1984.3	-	-
CASTILLA-LA MANCHA	-	1983.2	-	1984.3	-	-
CATALUÑA	1984.3	-	-	-	1980.3	-
EXTREMADURA	1976.4	-	-	1976.4 1984.3 1988.3	-	-
GALICIA	-	-	-	1980.3 1984.3	-	-
MADRID	-	1980.3	-	-	-	-
MURCIA	1978.1 1980.3 1987.2	-	-	1978.3 1980.3 1984.3	-	-
NAVARRA	1977.4 1985.3	1983.4 1988.1 1993.1	-	1977.4 1984.3 1987.4 1988.3	-	-
PAIS VASCO	-	-	-	1980.3 1982.4 1984.3 1986.3	-	1981.2
RIOJA	1977.3 1983.2	-	1979.3 1980.3	1977.3 1984.3	-	1979.3 1980.3
VALENCIA	1984.3 1985.3	-	-	-	-	-

Fuente: Elaboración propia a partir de datos del INE

BIBLIOTECA VIRTUAL

MIGUEL DE
SERVANTES

APÉNDICE 4

COEFICIENTES DE CORRELACIÓN ENTRE LAS VARIABLES

En este último apéndice incluimos los coeficientes de correlación entre las distintas variables que aparecen en la regresión utilizada para la descomposición de la varianza. Estos coeficientes son importantes para determinar la fiabilidad del efecto marginal de cada variable o conjunto de variables.

También es conveniente añadir que estas correlaciones se han calculado para el modelo más completo y desagregado. Esta elección se ha debido a que es el modelo que más información introduce y en el que más variables aparecen, por lo tanto, también es el modelo en que más problemas de correlación pueden aparecer.

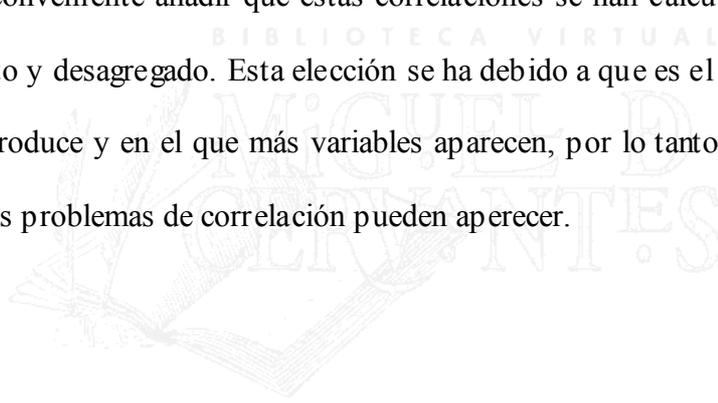


TABLA A4.1
Coefficientes de correlación entre las variables

	A1980	A1981	A1982	A1983	A1984	A1985	A1986	A1987	A1988	A1989
A1980	1.000	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1981	-0.053	1.000	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1982	-0.053	-0.053	1.000	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1983	-0.053	-0.053	-0.053	1.000	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1984	-0.053	-0.053	-0.053	-0.053	1.000	-0.053	-0.053	-0.053	-0.053	-0.053
A1985	-0.053	-0.053	-0.053	-0.053	-0.053	1.000	-0.053	-0.053	-0.053	-0.053
A1986	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	1.000	-0.053	-0.053	-0.053
A1987	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	1.000	-0.053	-0.053
A1988	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	1.000	-0.053
A1989	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	1.000
A1990	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1991	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1992	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1993	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1994	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1995	-0.052	-0.052	-0.052	-0.052	-0.052	-0.052	-0.052	-0.052	-0.052	-0.052
A1996	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1997	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1998	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
A1999	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053	-0.053
ANDA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ARAG	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CASL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CASM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CATA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CVAL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GALI	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MADR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PVAS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SEXO	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EDAD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ESTUDIOS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PIB	-0.082	-0.084	-0.078	-0.072	-0.068	-0.060	-0.049	-0.031	-0.012	0.006

TABLA A4.1 (continuación)

	A1990	A1991	A1992	A1993	A1994	A1995	A1996	A1997	A1998	A1999
A1980	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1981	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1982	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1983	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1984	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1985	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1986	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1987	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1988	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1989	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1990	1.000	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1991	-0.053	1.000	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1992	-0.053	-0.053	1.000	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1993	-0.053	-0.053	-0.053	1.000	-0.053	-0.052	-0.053	-0.053	-0.053	-0.053
A1994	-0.053	-0.053	-0.053	-0.053	1.000	-0.052	-0.053	-0.053	-0.053	-0.053
A1995	-0.052	-0.052	-0.052	-0.052	-0.052	1.000	-0.052	-0.052	-0.052	-0.052
A1996	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	1.000	-0.053	-0.053	-0.053
A1997	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	1.000	-0.053	-0.053
A1998	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	1.000	-0.053
A1999	-0.053	-0.053	-0.053	-0.053	-0.053	-0.052	-0.053	-0.053	-0.053	1.000
ANDA	0.000	0.000	0.000	0.000	0.000	-0.009	0.000	0.000	0.000	0.000
ARAG	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
CASL	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
CASM	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
CATA	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
CVAL	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
GALI	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
MADR	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
PVAS	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
SEXO	0.000	0.000	0.000	0.000	0.000	-0.003	0.000	0.000	0.000	0.000
EDAD	0.000	0.000	0.000	0.000	0.000	-0.003	0.000	0.000	0.000	0.000
ESTUDIOS	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000
PIB	0.021	0.031	0.033	0.027	0.036	0.045	0.058	0.075	0.093	0.111

TABLA A4.1 (continuación)

	ANDA	ARAG	CASL	CASM	CATA	CVAL	GALI	MADR	PVAS
A1980	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1981	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1982	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1983	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1984	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1985	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1986	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1987	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1988	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1989	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1990	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1991	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1992	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1993	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1994	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1995	-0.009	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
A1996	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1997	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1999	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ANDA	1.000	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125
ARAG	-0.125	1.000	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125
CASL	-0.125	-0.125	1.000	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125
CASM	-0.125	-0.125	-0.125	1.000	-0.125	-0.125	-0.125	-0.125	-0.125
CATA	-0.125	-0.125	-0.125	-0.125	1.000	-0.125	-0.125	-0.125	-0.125
CVAL	-0.125	-0.125	-0.125	-0.125	-0.125	1.000	-0.125	-0.125	-0.125
GALI	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	1.000	-0.125	-0.125
MADR	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	1.000	-0.125
PVAS	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	-0.125	1.000
SEXO	-0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EDAD	-0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ESTUDIOS	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PIB	0.268	-0.374	-0.199	-0.360	0.631	0.046	-0.226	0.384	-0.168

TABLA A4.1 (continuación)

	SEXO	EDAD	ESTUDIOS	PIB
A1980	0.000	0.000	0.000	-0.082
A1981	0.000	0.000	0.000	-0.084
A1982	0.000	0.000	0.000	-0.078
A1983	0.000	0.000	0.000	-0.072
A1984	0.000	0.000	0.000	-0.068
A1985	0.000	0.000	0.000	-0.060
A1986	0.000	0.000	0.000	-0.049
A1987	0.000	0.000	0.000	-0.031
A1988	0.000	0.000	0.000	-0.012
A1989	0.000	0.000	0.000	0.006
A1990	0.000	0.000	0.000	0.021
A1991	0.000	0.000	0.000	0.031
A1992	0.000	0.000	0.000	0.033
A1993	0.000	0.000	0.000	0.027
A1994	0.000	0.000	0.000	0.036
A1995	-0.003	-0.003	0.003	0.045
A1996	0.000	0.000	0.000	0.058
A1997	0.000	0.000	0.000	0.075
A1998	0.000	0.000	0.000	0.093
A1999	0.000	0.000	0.000	0.111
ANDA	-0.002	-0.002	0.002	0.268
ARAG	0.000	0.000	0.000	-0.374
CASL	0.000	0.000	0.000	-0.199
CASM	0.000	0.000	0.000	-0.360
CATA	0.000	0.000	0.000	0.631
CVAL	0.000	0.000	0.000	0.046
GALI	0.000	0.000	0.000	-0.226
MADR	0.000	0.000	0.000	0.384
PVAS	0.000	0.000	0.000	-0.168
SEXO	1.000	-0.001	0.001	-0.001
EDAD	-0.001	1.000	0.001	-0.001
ESTUDIOS	0.001	0.001	1.000	0.001
PIB	-0.001	-0.001	0.001	1.000

Fuente: Elaboración propia a partir de datos del INE



Ashenfelter, O. C. y Layard, R. (1986): *Manual de economía del trabajo. II*. Madrid, Ministerio de Trabajo y Seguridad Social.

Avilés, S., Gámez, C. y Torres, J. L. (1997): “La convergencia real de Andalucía: Un análisis de cointegración del mercado de trabajo”. *Revista de Estudios Regionales*, nº 47, pp. 15-36

Barro, R. J. Y Sala i Martín, X. (1990): “Economic Growth and Convergence across the United States”. *NBER Working Paper*, nº 3419.

Barro, R. J. Y Sala i Martín, X. (1991): “Convergence across States and Regions”. *Brooking Papers on Economic Activity*, nº 1, pp. 107-182.

Barro, R. J. Y Sala i Martín, X. (1992): “Convergence”. *Journal of Political Economy*, nº 2, pp. 223-251.

Bernard, A. B. y Durlauf, S. N. (1996): “Interpreting test of the convergence hipótesis”. *Journal of Applied Econometrics*, Vol. 10, nº 2, pp. 97-108.

Brechelling, T. (1967): “Trends and cycles in British regional unemployment”. *Oxford Economic Papers*, nº 19, pp. 1-21

Brown, J. N. y Ashenfelter, O. (1986): “Testing the efficiency of employment contracts”. *Journal of Political Economy*, nº 94, pp. 540-587.

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, pp. 62-76.

Chatterjee, K. (1982): “Incentive compatibility in bargaining under uncertainty”. *Quarterly Journal of Economics*, nº 97, pp. 717-726.

De la Fuente, A.(1993): “Expenditure Levels and Incentives in Public R&D”. Mimeo, Instituto de Análisis Económico (FEDER III).

Dickey, D. A. y Fuller, W. A. (1979): “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”. *Journal of Business and Economics Statistics*, nº 74, pp. 427-431.

Dickey, D. A. y Fuller, W. A. (1981): “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root”. *Econometrica*, nº 49, pp. 1057-1072.

Dolado, J., González-Páramo, J. M. y Roldán, J. M.(1993): “Convergencia Económica entre las Provincias Españolas: Evidencia Empírica (1955-1989)”. *Moneda y crédito*, nº198, pp. 81-132.

Engle, R. y Granger, C. W. J. (1987): “Cointegration and error correction: representation, estimation and testing”. *Econometrica*, nº 55, pp. 251-276.

Esteban, J. M. (1994): “La desigualdad interregional en Europa y en España: descripción y análisis”. *Crecimiento y convergencia regional en España y Europa*, vol. 2, pp. 13-82. Instituto de Análisis Económico-CSIC y Fundación de Economía Aplicada, Barcelona.

Fernandez-Jardón, C. M. y Caneda, A. (1999): *El nivel educativo de los trabajadores como factor de crecimiento económico: Un análisis comparativo entre Galicia y España*. Tesis doctoral.

Fuller, W. A. (1976): *Introduction to Statistical Time Series*. John Wiley and Sons, New York.

Galton, F. (1989): *Natural Inheritance*. Macmillan, Londres

Galton, F. (1992): *Hereditary Genius*. Macmillan, Londres

García, A., Raymond, J. L. y Villaverde, J. (1995): “La convergencia de las provincias españolas”. *Papeles de Economía Española*, nº 64, pp. 38-53.

Greasley, D. y Oxley, L. (1997): “Time-series based test of convergence hypothesis: Some positive results”. *Economic letters*, vol. 56, pp. 143-147.

Grossman, S. J. y Hart, O. D. (1983): “Implicit contracts under asymmetric information”. *Quarterly Journal of Economics*, nº98 (suplemento), pp. 123-156.

Hart, O. D. (1983): “Optimal labor contracts under asymmetric information: an introduction”. *Review of Economic Studies*, nº50 (2), pp. 3-35.

Hofler, R. y Murphy, K. (1989): “Using a composed error model to estimate the frictional and excess supply components of unemployment”. *Journal of Regional Science*, vol. 29, nº 2, pp. 213-228.

Hyclak, T y Lynch, G. (1980): “An empirical analysis of state unemployment rates in the 1970's”. *Journal of Regional Science*, vol. 20, nº 3, pp. 377-386

Jimeno, J. F. (1992): “Las implicaciones macroeconómicas de la negociación colectiva, el caso español”. *Moneda y Crédito*, nº 195, pp. 223-281.

Jimeno, J. F. (1997): “Los factores específicos del paro en Andalucía”. *Documento de trabajo FEDEA*.

Jimeno, J. F. y Toharia, L. (1993): “Spanish labour markets: institutions and outcomes”. En J. Hartog y J. Theeuwes (comps.), *Labour Markets Contracts and Institutions, A Cross-national Comparison*, Amsterdam, North-Holland.

Leontief, W.(1946): “The pure theory of the guaranteed annual wage contract”. *Journal of Political Economy*, nº 54, pp. 99-150.

Lindbeck, A. y Snower, D. (1986): “Wage setting, unemployment, and insider-outsider relations”. *American Economic Review*, nº 76, pp. 235-239.

Lindbeck, A. y Snower, D. (1988): *The insider-outsider theory*. MIT Press, Cambridge, MA.

MaCurdy, T. E. y Pencavel, J. (1986): "Testing between competing models of wage and employment determination in unionized labor markets". *Journal of Political Economy*, nº 94, pp. 3-39.

Mas, M., Maudos, J., Pérez, F. Y Uriel, E. (1994): "Disparidades regionales y convergencia en las CC.AA. españolas". *Revista de Economía Aplicada*, vol. II (4), pp. 37-58.

Mankiw, G., Romel, D. Weil, D.(1992): "A Contribution to the Empirics of Economic Growth". *Quarterly Journal of Economics*, vol 107(2), pp. 407-437.

Maravall, A. y Gomez, V. (1993): Initializing the Kalman Filter with incompletely specified initial conditions in approximate Kalman Filter (Series on Approximation and Decomposition), Chen, R. (ed), London, World Scientific Publ. Co., pp. 39-62.

Maravall, A. y Gomez, V. (1994): "Estimation, prediction and interpolation for nonstationary series with the Kalman Filter". *Journal of the American Statistical Association*, nº 89, pp. 611-624

Maravall, A. y Gomez, V. (2001): *Automatic modeling methods for univariate series*, en *A Course in Time Series Analysis*, Peña, D., Tiao, G.C. and Tsay, R.S. (ed), New York: Wiley, J. and Sons Ch. 7, pp 170-201.

Maravall, A., Gómez, V. Y. y Peña, D. (1999): “!Missing observations in ARIMA models: Skipping Approach versus Additive Outlier Approach”. *Journal of Econometrics*, n° 88, pp. 341-363

McDonald, I. M. y Solow, R. M. (1981): “Wage bargaining and unemployment”. *American Economic Review*, n° 71, pp. 896-908.

McGuinness, S. y Sheehan, M. (1998): “Regional convergence in U.K.”. *Applied Economic Letters*, vol. 5, pp. 653-658.

McKinnon, J. (1991): “Critical Values for Cointegration Test”. En Engle, R. y Granger, C. W. J. editores (1991).

Metcalf, D. (1975): “Urban employment in England”. *The Economic Journal*, n° 85 (september), pp. 578-589

Milner, S. y Metcalf, D. (1995): “Relaciones laborales y comportamiento macroeconómico”. *Moneda y Crédito*, n° 201, pp. 11-48.

Milner, S. y Nombela, G. (1994): “Trade union strength, organisation and impact in Spain”. *Discussion Paper*, n° 258, Centre for Economic Performance, LSE.

Mortensen, D. T. (1986): “Job search and labor market analysis”. En Ashenfelter, O. C. Y Layard, R. (eds), *Handbook of labor economics*, Vol. 2, North-Holland, Amsterdam.

OECD (2000): *Employment Outlook*: junio 2000, (OECD, Paris)

Olloqui, I., Sosvilla, S. y Alonso, J. (2002): “Convergencia en precios en las provincias españolas”. *ICE Tribuna de Economía*, nº 797, febrero, pp. 160-178.

Oswald, A. J. (1984): “Efficient contracts are on the labor demand curve: theory and facts”. Documento nº 178, Princeton University, Industrial Relations Section.

Pallardó, V. J. Y Esteve, V. (1997): “Convergencia real en la Unión Europea”. *Revista de Economía Aplicada*, nº 14 (vol. V), pp. 25-49

Perron, P. (1989): “The Great Crash, the Oil Price Shock, and the Unit Roots Hypothesis”. *Econometrica*, nº 57, pp. 1361- 1401.

Phillips, P. C. B. (1986): “Understanding Spurious Regressions in Econometrics”. *Journal of Econometrics*, nº 33, pp. 311-340.

Phillips, P. C. B. y Perron, P. (1988): “Testing for a Unit Root in Time Series Regression”. *Biometrika*, nº 75, pp. 335-346.

Plaza, R. A. (1994): *Desempleo y salarios en España. Diferencias interprovinciales*. Secretariado de Publicaciones, Universidad de Valladolid.

Quah, D.(1993a): “Empirical Cross-Section Dynamics in Economic Growth”. *European Economic Review*, vol. 37, pp. 426-434.

Quah, D.(1993b): “Galton’s Fallacy and Test of the Convergence Hypothesis”. *Scandinavian Journal of Economics*, vol. 95, nº 4, pp. 427-443.

Rappoport, P. Y Richlin, L. (1989): “Segmented Trends and Non-Stationary Time Series”. *The Economic Journal*, nº 99, pp. 168-177

Raymond, J. L. y García, A. (1994): “Las disparidades en el PIB per cápita entre las Comunidades Autónomas y la hipótesis de convergencia”. *Papeles de Economía Española*, nº 59, pp. 37-58.

Sala i Martín, X. (1994): “La riqueza de las regiones. Evidencia y teorías sobre crecimiento regional y convergencia”. *Moneda y Crédito*, nº 198, pp. 13-80.

Sala i Martín, X. (1996a): “Regional cohesion. evidence and theories of regional growth and convergence”. *Regional Economic Review*, vol. 40, nº 6, pp. 1325-1352.

Sala i Martín, X. (1996b): “The Classical Approach to Convergence Analysis”. *The Economic Journal*, nº 106, julio, pp. 1019-1036.

Sanchez Molinero, J. M. (1999): “ La negociación colectiva en una sociedad plural y abierta: España año 2000. La visión del economista”. *Papeles de la fundación para el análisis y los estudios sociales*, nº 50, pp 127-144.

Simón Pérez, H. J. (2001): “La estructura de salarios pactados en España”. *Estudios de Economía Aplicada*, nº 19 (diciembre), pags. 171-187.

Suriñach, J., Artís, M., López, E. y Sanso, A. (1995): *Análisis Económico Regional: Nociones Básicas de la Teoría de la Cointegración*. Antoni Bosch, editor, y Bosch i Gimpera, Fundació.

Toharia, L., Albert, C., Cebrián, I., García Serrano, C., García Mainar, I., Malo, M. A., Moreno, G. y Villagomez, E. (1998): *El mercado de trabajo en España*. McGraw-Hill/ Interamerica de España, S. A. U., Madrid.

Villaverde, J.(1999): *Diferencias regionales en España y Unión Monetaria Europea*. Colección “economía XXI”. Ediciones Pirámide, S. A. Madrid.

Yule, G. U. (1926): “Why Do We Sometimes Get Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Series”. *Journal of the Royal Statistical Society*, nº 89, pp. 1-64.