

Proyecciones derivadas: proyección de los hogares y proyección de la población activa de Andalucía

Bermúdez Parrado, S.; Hernández Rodríguez, J.A.; Planelles Romero, J.

Instituto de Estadística y Cartografía de Andalucía

1. Introducción

En este trabajo se presentan dos ejercicios de proyección derivada realizados por el Instituto de Estadística y Cartografía de Andalucía. Concretamente se trata de la proyección de la población activa de Andalucía y de la proyección de los hogares de Andalucía. En ambos ejercicios se han incorporado novedades metodológicas que han sido objeto de publicación en documentos metodológicos y revistas especializadas ([1], [2] y [3]).

Ambos ejercicios comparten también las fuentes de información principales. Tratándose de proyecciones derivadas, se ha utilizado en ambos casos como población de base la proyección de población de Andalucía que elabora el propio Instituto de Estadística y Cartografía de Andalucía. Por otro lado, la Encuesta de Población Activa (EPA) ha proporcionado la información relativa a la relación con la actividad y la correspondiente a la composición de los hogares.

No obstante, las proyecciones de población y la EPA tienen fechas de referencia y población objetivo ligeramente distintas. Por ello, resulta necesario realizar dos ajustes que acomoden ambas.

- En primer lugar, hay que convertir la serie de poblaciones referenciada a 1 de enero, que proporciona la proyección de la población, en unas poblaciones medias para cada año que sean consistentes con los datos de población de la EPA. Esto se ha resuelto mediante la semisuma de las poblaciones a 1 de enero de cada año.
- En segundo lugar, hay una pequeña parte de la población que al no tener fijada su residencia habitual en una "vivienda familiar" queda fuera del ámbito de investigación de la EPA. Se trata de las personas que residen en "alojamientos" o en "viviendas colectivas"¹. En el caso de las proyecciones de hogares, se trata de un colectivo que también queda fuera del ámbito de estudio. Por ello, las proyecciones de población que se utilicen deben limitarse al segmento de la población total que efectivamente reside en viviendas familiares. En el caso de la proyección de la población activa, se

¹ Los "alojamientos" son viviendas móviles, semipermanentes o improvisadas, o bien que no han sido concebidas en principio con fines residenciales. Las "viviendas colectivas" son viviendas destinadas a ser habitadas por un grupo de personas sometidas a una autoridad o régimen común, no basados en lazos familiares y de convivencia (residencias, conventos, hospitales...). Alojamientos y viviendas colectivas suponían el 0,5% del total de la población residente en España en el último censo disponible (2011).

han aplicado a la población total las tasas de actividad de la población residente en viviendas familiares de la EPA.

El trabajo que aquí se presenta se ha estructurado en torno a dos grandes apartados: la proyección de hogares y la proyección de población activa, respectivamente. Dada la limitación de espacio, el texto se centra en los aspectos metodológicos. Los resultados están disponibles en la web del Instituto de Estadística y Cartografía de Andalucía¹.

2. Una metodología para la proyección de los hogares

Las transformaciones experimentadas por la sociedad española en las últimas décadas tienen un importante correlato en las estructuras familiares actuales. Los nuevos modelos de convivencia son fiel reflejo de un proceso de cambio que afecta a la composición de los hogares. El retraso en la formación de parejas, la caída de la fecundidad, el paulatino incremento de la edad de la primera maternidad, la disminución del número de hijos por pareja, las disoluciones de estas dando lugar a núcleos con un solo progenitor, la creciente autonomía de los mayores, son algunas de las manifestaciones más visibles de estos cambios.

De esta forma, el conocimiento de la evolución del número y de la estructura de los hogares se convierte en un elemento fundamental para determinar con mayor precisión las demandas sociales, así como para esclarecer algunos aspectos económicos relacionados con el sector de la vivienda, el consumo privado de recursos como el agua o la energía y otros.

2.1. Aspectos conceptuales y fuentes de información

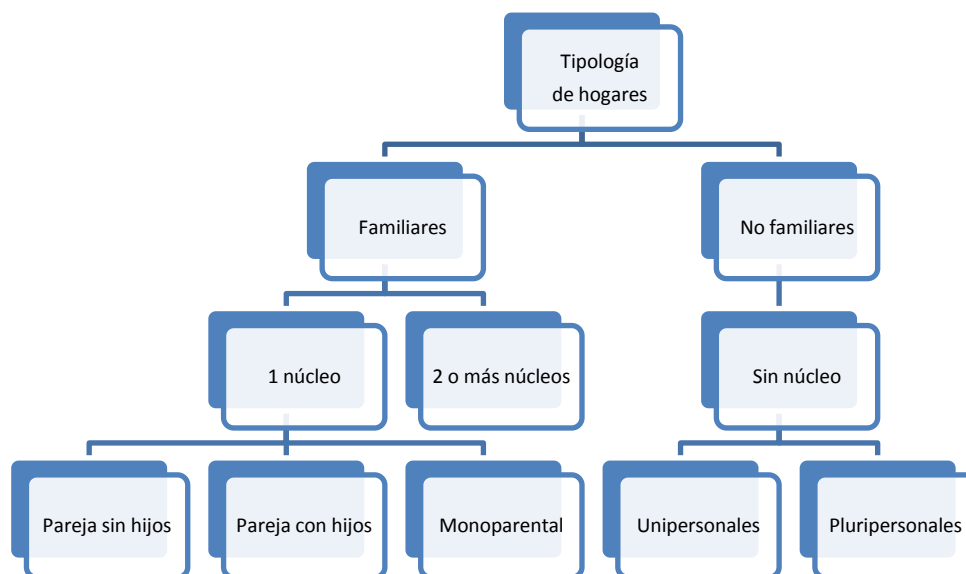
Antes de la presentación de la metodología más específica de la proyección, es necesario establecer algunas definiciones básicas empleadas en este apartado.

Las estadísticas y la investigación sociológica en el campo de la familia se apoyan en algunos conceptos básicos: núcleo familiar, familia, hogar y vivienda. Estos términos se pueden definir con precisión pero, en algunos casos, su sentido se confunde. En este caso se han seguido las definiciones utilizadas en el Censo de

Población y Viviendas 2011². Al mismo tiempo, las tipologías de hogares consideradas en el ejercicio de proyección están enumeradas en la Figura 1.

La fuente de información que nos permite analizar las dinámicas recientes en las relaciones de convivencia es la EPA. Dicha Operación la realiza el Instituto Nacional de Estadística (INE) desde 1964, es una investigación continua y de periodicidad trimestral dirigida a la población que reside en viviendas familiares principales³. En esta encuesta se recoge información acerca de las relaciones de parentesco que los miembros del hogar mantienen entre sí, sus edades y sexos. Es más, la EPA permite analizar la estructura de los hogares por dos caminos distintos. En primer lugar, una pregunta relativa a la relación de cada miembro con la "persona principal" (la persona que el propio hogar inscribe en primer lugar en el cuestionario). Esta pregunta ha sido explotada en numerosos análisis acerca de la estructura de los hogares e incluso para realizar proyecciones de hogares [12, 19, 22].

Figura 1. Tipologías de hogares.



Una segunda vía, menos explorada hasta la fecha, se abrió a partir del año 1999 gracias a la introducción de tres nuevas preguntas en el cuestionario de la EPA. En ellas se registra la presencia en el hogar del cónyuge o pareja de cada informante, su padre y su madre⁴. De este modo, es posible construir todas las relaciones de

² Documento metodológico en http://www.ine.es/censos2011/censos2011_proyecto.pdf

³ En la EPA se identifica hogar con vivienda o dirección postal.

Más información en el documento metodológico.

⁴ Variables NCONY, NPADRE y NMADRE, respectivamente.

parentesco de primer grado, esté o no implicada en ellas la persona principal. Esta nueva posibilidad resulta de lo más oportuna, pues permite completar la caracterización de las relaciones nucleares de parentesco dentro de los hogares, permitiendo por ejemplo identificar los hogares con más de un núcleo familiar, si bien hacerlo supone un trabajo previo de tratamiento de los micro-datos y de jerarquización de las relaciones de parentesco que no es sencillo.

1.2. Alternativas de proyección

El modelo de proyección de hogares más conocido y más ampliamente utilizado hasta la fecha en la mayoría de los organismos oficiales es el llamado *método de la tasa de jefatura* [26]. El método gira en torno a la selección de un único miembro del hogar, de modo que se puede establecer una correspondencia uno a uno entre hogares e individuos seleccionados⁵. De este modo, lo que se hace es estimar cuántas personas han ocupado en los últimos años esa posición única dentro del hogar. En segundo lugar, se estiman las llamadas tasas de jefatura, que se definen como la proporción de individuos que han resultado seleccionados dentro de cada hogar sobre el total de personas en la población, para cada edad y sexo. A partir de las dinámicas observadas en las tasas de jefatura, se obtiene una proyección de las mismas hacia el futuro. Finalmente se deriva la proyección del número de hogares multiplicando las tasas de jefatura proyectadas por la población proyectada de cada sexo y edad. Entre sus ventajas cabe destacar su baja dificultad técnica y los bajos requisitos de información necesarios, que se obtienen con facilidad de un censo o de una encuesta a hogares. Por ello, se trata de un modelo fácilmente aplicable a distintos periodos y lugares. No obstante, este modelo presenta algunos inconvenientes como el de no distinguir distintos tipos de hogares, ni las dinámicas que transforman la posición que ocupa cada individuo dentro del mismo.

El *método de propensiones* lo podemos considerar como una extensión de las tasas de jefatura. Como aquél, se trata de un modelo estático, pues no se consideran los flujos que transforman un tipo de hogar en otro. La diferencia principal con el método de la tasa de jefatura radica en que en el cálculo de la tasa no sólo entra el "cabeza de familia", sino todos los miembros del hogar, en función de la posición

⁵ Según la fuente de datos de la que se disponga o según los objetivos perseguidos en la investigación, será conveniente escoger un miembro concreto del hogar. A este respecto, en algunas encuestas se pide al hogar que identifique a un miembro del mismo como "cabeza de familia" o similar ("persona de referencia", "persona principal", etc.). En algunas ocasiones, pero no siempre, se sugiere al hogar un determinado criterio para su elección (por ejemplo, el miembro activo de mayor edad o el perceptor de mayores rentas). En otros casos, se utiliza el miembro que el hogar inscribe en primer lugar en el cuestionario, o directamente, el que el investigador considere más conveniente.

que ocupen en el mismo [1]. De este modo, se resuelven algunas de las limitaciones que presentaba el método de tasa de jefatura. En el debe, su principal inconveniente es que aumentan las restricciones de consistencia, pues hay que relacionar los distintos miembros del hogar entre sí, para obtener agrupaciones consistentes de individuos, una tarea no siempre sencilla.

Por último, alcanzando el máximo en el gradiente de complejidad técnica y requerimientos de información entre los modelos de proyección de hogares, se encuentran los *modelos dinámicos* de proyección de hogares [14] que modelizan explícitamente las transiciones entre unos estados (en nuestro caso tipos de hogar) y otros. Los posibles estados pueden ser más o menos detallados y por tanto, exigir datos más o menos desagregados. Dentro del gran paraguas de los modelos dinámicos existen dos enfoques distintos, los *modelos de macro-simulación* y los *modelos de micro-simulación*. En los primeros, las transiciones las experimentan conjuntamente subconjuntos de entre las unidades que pertenecen a un estado. Por su parte, los *modelos de micro-simulación*, cada vez más utilizados en las ciencias sociales, tratan al individuo como unidad analítica, completando el curso vital de cada biografía. La gran ventaja de los modelos dinámicos de proyección es que tienen en consideración los procesos de formación, transformación y disolución de los hogares. Sin embargo, sus grandes requerimientos de datos imposibilitan en la mayor parte de los casos su aplicación práctica.

1.1. Metodología empleada

Si bien todas las metodologías pueden ser, de una u otra forma, adecuadas para la realización de la proyección, en este caso se ha optado por el método de las propensiones, incorporando algunas novedades que han permitido modelizar el comportamiento de los distintos colectivos considerados. Las tasas de propensión se han descompuesto teniendo en cuenta la intensidad y el calendario facilitando la adopción de hipótesis independientes en cada caso.

El proceso seguido se detalla a continuación.

A partir de los microdatos de la EPA [21] para el periodo determinado y teniendo en cuenta las definiciones presentadas en el apartado 1.1., se obtienen las poblaciones desagregadas para el ámbito territorial fijado según tipo de hogar, posición de la persona dentro del mismo (lo que aquí se denomina rol), edad y sexo. Esta población se puede representar como:

$$P_{h,r,x,s,t}^{EPA} \quad \forall h, r, x, s, t \quad (1)$$

donde⁶

P^{EPA} : población estimada por la EPA

h : tipo de hogar, h = unipersonal, pluripersonal, pareja con hijos, pareja sin hijos, monoparental, dos o más núcleos

r : rol, r = pivote, hijo, otros

x : edad⁷, x = [0-4], [5-9], [10-15], [16-19],..., [60-64], 65 y más

s : sexo, s = hombre, mujer

t : tiempo⁸ en años

Con (1) se calculan las tasas de propensión, TP , cuya expresión viene dada por:

$$TP_{h,r,x,s,t} = \frac{P_{h,r,x,s,t}^{EPA}}{\sum_h \sum_r P_{h,r,x,s,t}^{EPA}} \quad \forall h, r, x, s, t \quad (2)$$

Estas tasas recogen, para cada periodo, edad y sexo, la proporción de personas en cada rol y tipo de hogar. Por definición, la formulación anterior verifica la siguiente condición:

$$\sum_h \sum_r TP_{h,r,x,s,t} = 1 \quad \forall x, s, t \quad (3)$$

Esto es, fijado un año, se asume que todas las personas de una determinada edad y sexo viven en alguna de las tipologías de hogar y tienen asignado un rol dentro del mismo.

Por la forma en la que se construyen las tasas en (2), éstas se pueden descomponer en dos factores: intensidad y calendario.

De esta forma, la expresión (2) se puede escribir como:

$$TP_{h,r,x,s,t} = \sum_x TP_{h,r,x,s,t} \frac{TP_{h,r,x,s,t}}{\sum_x TP_{h,r,x,s,t}} = ISP_{h,r,s,t} C_{h,r,x,s,t} \quad \forall h, r, x, s, t \quad (4)$$

donde

$ISP_{h,r,s,t}$: índice sintético de propensión fijado h, r, s y t

$C_{h,r,x,s,t}$: función calendario evaluada en h, r, x, s, t

⁶ Las variables y categorías que aquí se definen son extensibles al resto de expresiones que aparecen a lo largo del documento.

⁷ Se utilizan todos los grupos de edad disponibles en los microdatos de la Encuesta de Población Activa.

⁸ La EPA es una operación de periodicidad trimestral. Para la obtención de datos anuales se considera la media aritmética de los cuatro trimestres.

En base a la estabilidad que muestra el calendario (figuras 2 a 8), en particular en las tipologías de hogar más frecuentes, se puede asumir que éste permanece estable a lo largo del periodo de proyección. Para ello se construye un calendario más robusto, \bar{c} , a partir del promedio de los n último años:

$$\bar{c}_{h,r,x,s} = \frac{\sum_{t=1}^n c_{h,r,x,s,t}}{n} \quad \forall h, r, x, s \quad (5)$$

donde

\bar{c} = calendario medio de los n años fijados

Figura 2. Calendarios de propensión en Andalucía, 2006-2012. Parejas con hijos.

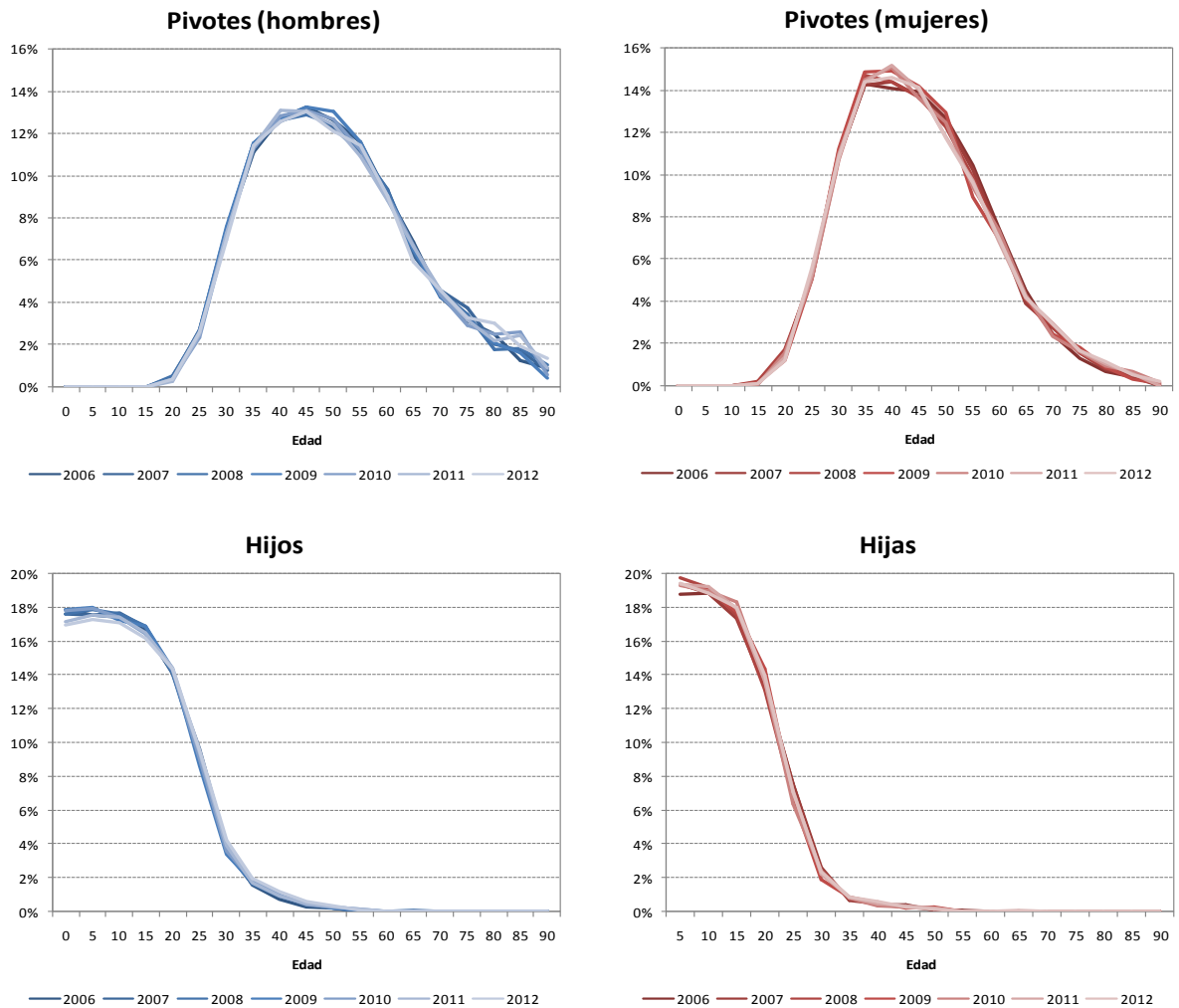


Figura 3. Calendarios de propensión. Parejas sin hijos.

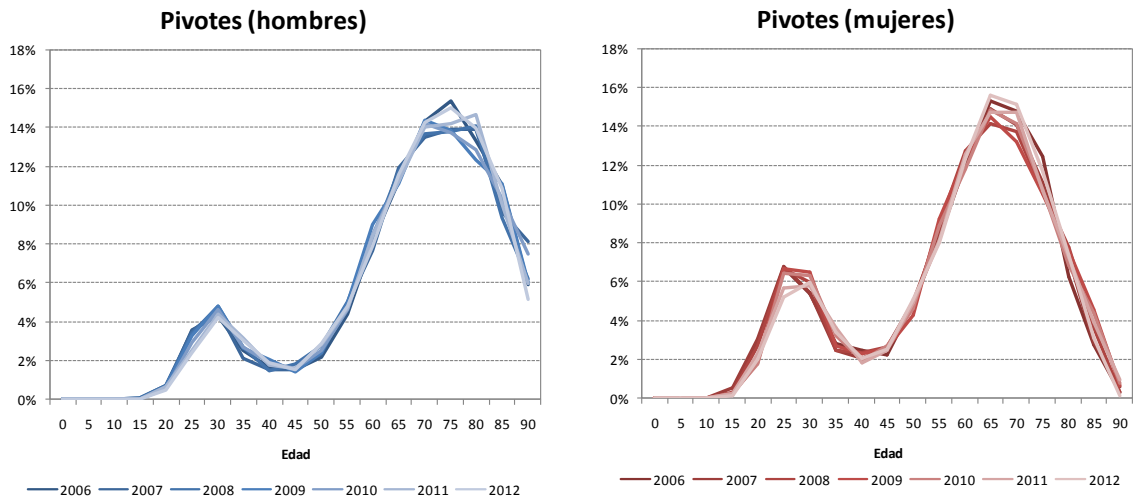


Figura 4. Calendarios de propensión. Hogares monoparentales.

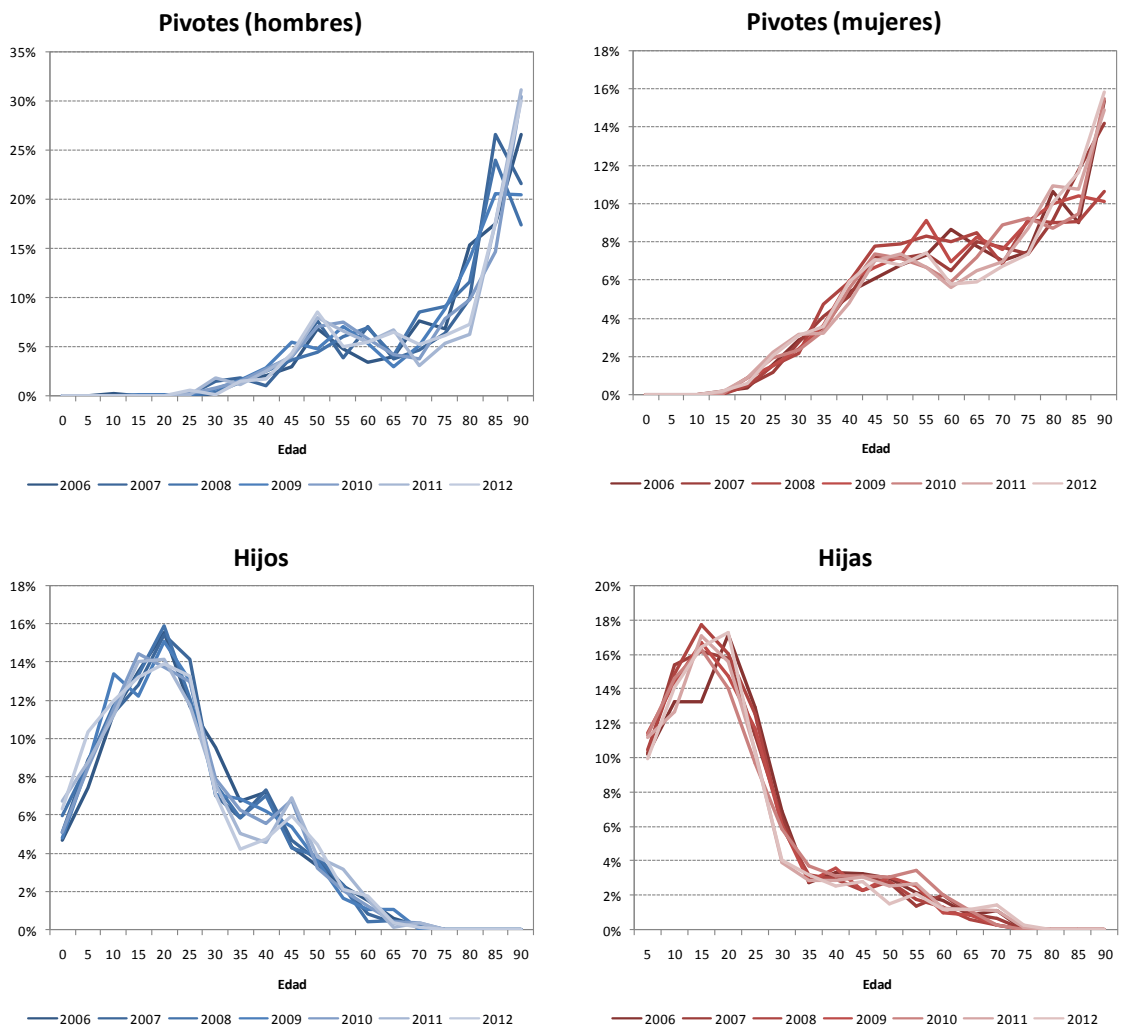


Figura 5. Calendarios de propensión. Hogares de dos o más núcleos

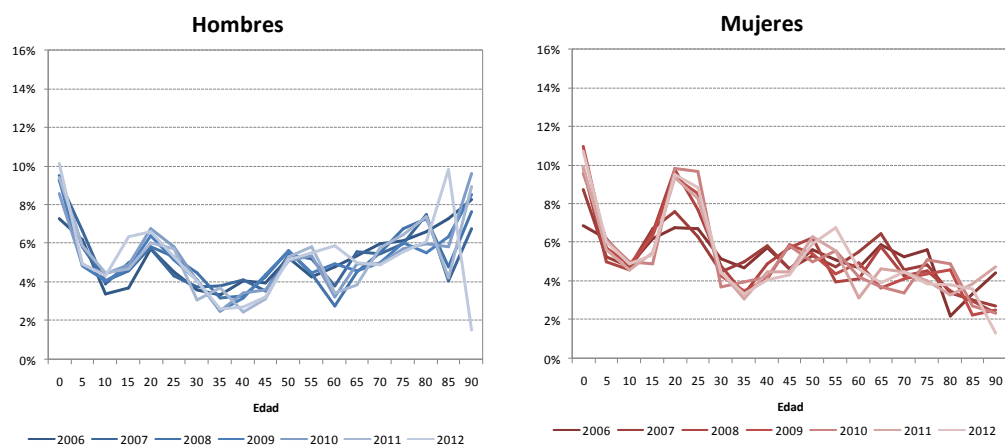


Figura 6. Calendarios de propensión. Otros en hogares familiares

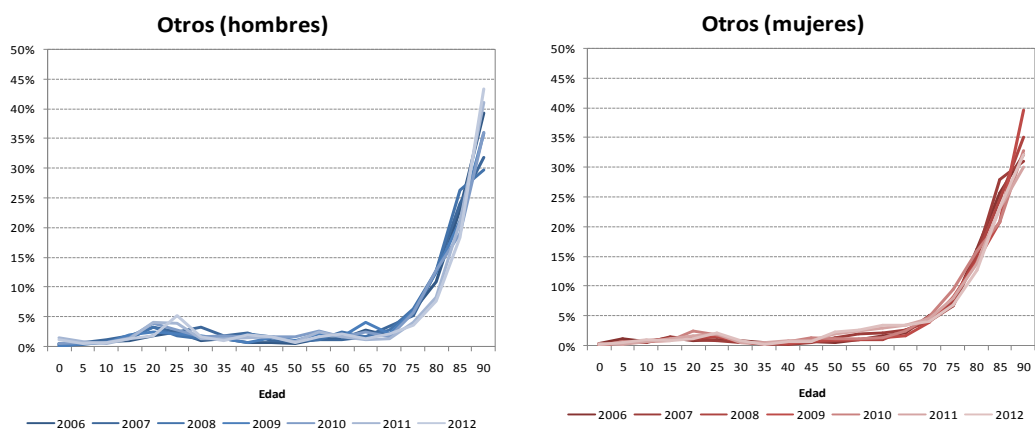


Figura 7. Calendarios de propensión. Hogares unipersonales

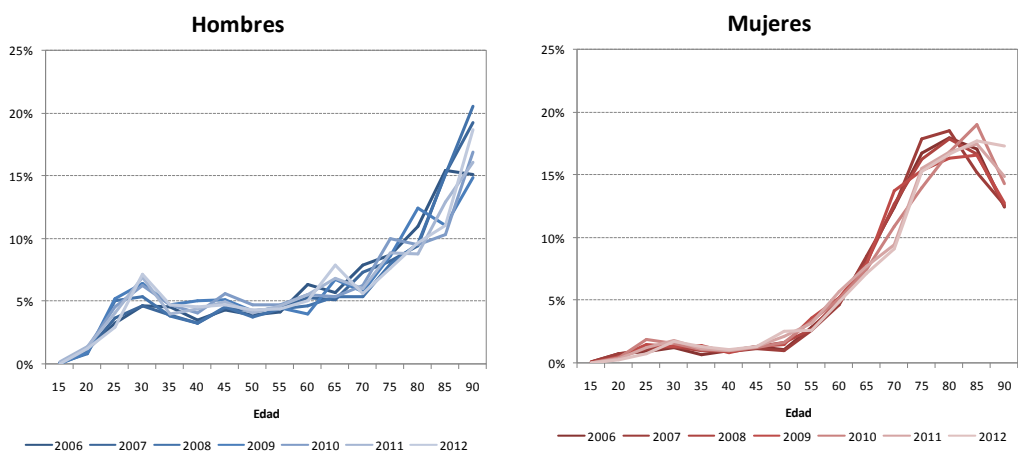
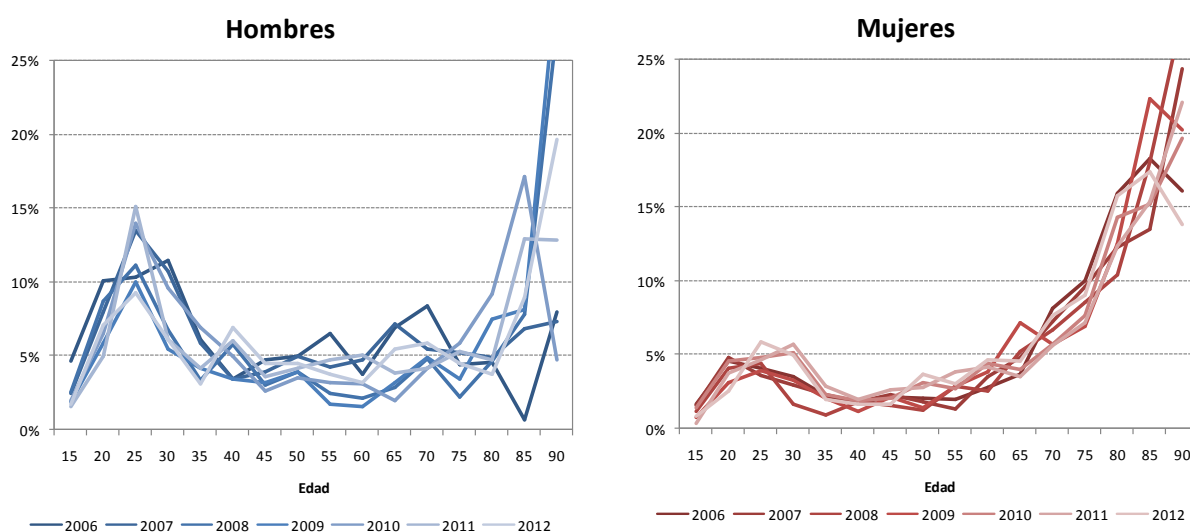


Figura 8. Calendarios de propensión. Hogares pluripersonales



En contraste con lo observado en el calendario, la intensidad sí ha experimentado pequeñas variaciones a lo largo del periodo analizado. Por ello, se opta por proyectar la intensidad a partir de una regresión que depende del logaritmo del tiempo. Este método mantiene las dinámicas de cambio observadas con tendencia a moderarse. La relación queda definida como sigue:

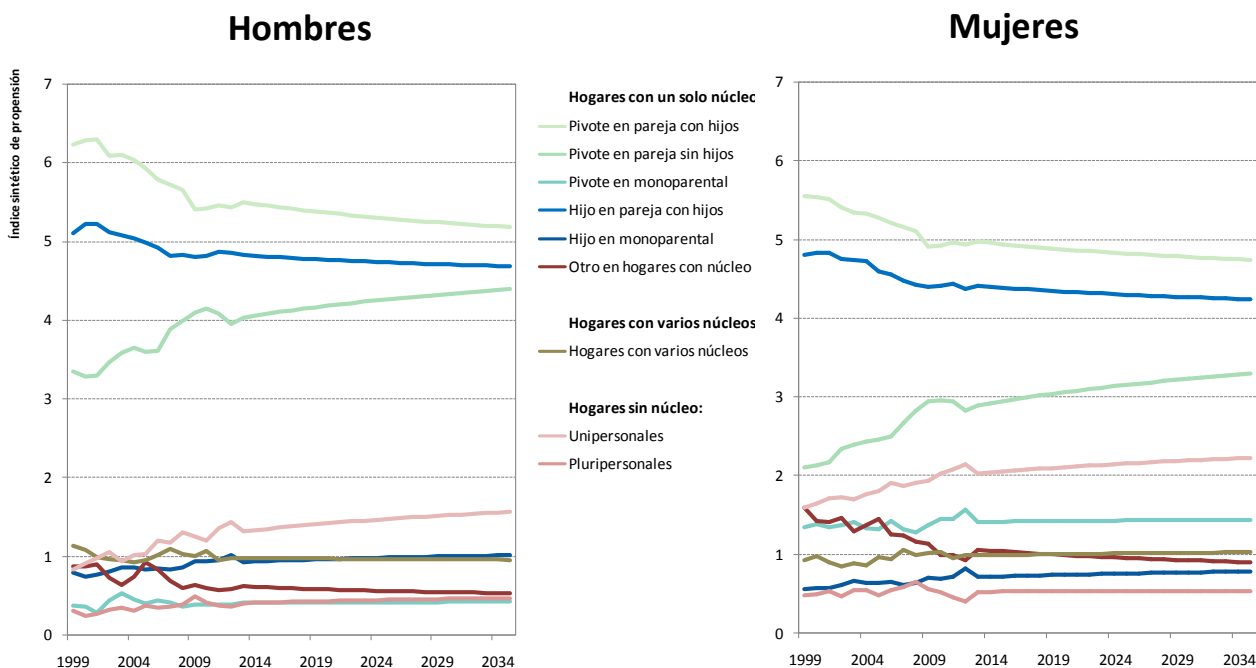
$$\ln(ISP_{h,r,s,t}) = \alpha_{h,r,s} + \beta_{h,r,s} \ln t + u_{h,r,s,t} \quad \forall h, r, s, t \quad (6)$$

Con los datos conocidos del periodo de observación, 2000-2012, se calculan las estimaciones de los parámetros $\hat{\alpha}_{h,r,s}$ y $\hat{\beta}_{h,r,s}$ para cada rol y sexo de la persona que convive en el hogar h.

Se proyecta la intensidad, Figura 9, sustituyendo en la expresión (6) los parámetros estimados. La expresión queda:

$$ISP_{h,r,s,t} = e^{\hat{\alpha}_{h,r,s} + \hat{\beta}_{h,r,s} \ln t} \quad \forall h, r, s \quad t > 2012 \quad (7)$$

Figura 9. Indicadores sintéticos de propensión, por tipo de hogar y posición dentro del mismo



Con la intensidad proyectada (7) y calendario (5) se obtiene la proyección de las tasas de propensión aplicando la ecuación (4). De esta forma queda:

$$TP_{h,r,x,s,t} = ISP_{h,r,s,t} \bar{C}_{h,r,s} \left[\frac{ISP_{h,r,s,t} \bar{C}_{h,r,x,s}}{\sum_h \sum_p ISP_{h,r,s,t} \bar{C}_{h,r,x,s}} \right] \quad \forall h, r, x, s \quad t > 2012 \quad (8)$$

donde el último término impone la condición (3).

Las tasas de propensión obtenidas a partir de (8) se multiplican por la población proyectada, P , correspondiente para obtener la población proyectada, 0P , que pertenece a cada una de las categorías consideradas.

$${}^0P_{h,r,x,s,t} = TP_{h,r,x,s,t} P_{x,s,t} \quad \forall h, r, x, s \quad t > 2012 \quad (9)$$

Al aplicar (9) a la categoría *pivote* de la variable *rol* se comprueba que el número de pivotes de cada sexo dentro de las parejas no tiene por qué coincidir ya que no es una condición impuesta en la metodología de proyección. De hecho, no coincide con exactitud en la serie histórica. Este resultado deriva de la presencia en la muestra utilizada en la EPA de parejas del mismo sexo. Como hemos comprobado que el número de parejas del mismo sexo en la muestra es reducido, no nos

permite realizar hipótesis diferenciadas para esta categoría en la proyección se ha optado por imponer que el número de pivotes de cada sexo ha de coincidir. Esto supone asumir que el verdadero número de pivotes de cada sexo en la población coincide, extremo perfectamente asumible con los datos disponibles.

Para este último ajuste la decisión tomada ha sido asignar la mitad de la diferencia observada dentro de las parejas, D , a cada uno de los sexos. Esto se ha hecho teniendo en cuenta la edad y el tipo de hogar al que pertenece mediante la siguiente expresión:

$$D_{h,r,s,t} = \sum_x ({}^0P_{h,r,x,s,t} - {}^0P_{h,r,x,\bar{s},t}) \quad h = \text{parejas con/sin hijos}; r = \text{pivote} \quad (10)$$

donde \bar{s} corresponde al sexo complementario considerado en la primera parte de la ecuación, s .

La población final, P , se obtiene aplicando (10). Y por tanto:

$$P_{h,r,x,s,t} = \begin{cases} {}^0P_{h,r,x,s,t} + \left[\frac{D_{h,r,s,t}}{2} \frac{{}^0P_{ob_{h,r,x,s,t}}}{\sum_x {}^0P_{h,p,x,s,t}} \right] & \text{si } h = \text{parejas con/sin hijos}; r = \text{pivote} \\ {}^0P_{h,r,x,s,t} - \left[\frac{D_{h,r,s,t}}{2} \frac{{}^0P_{h,r,x,s,t}}{\sum_x {}^0P_{h,r,x,s,t}} \right] \left[\frac{{}^0P_{ob_{h,r,x,s,t}}}{\sum_h \sum_r {}^0P_{ob_{h,r,x,s,t}}} \right] & \text{en caso contrario} \end{cases} \quad (11)$$

Obtenida la proyección de personas por tipo de hogar, ya sólo resta derivar el número de hogares a partir de ellas.

Para tal fin se analiza en primer lugar el caso de las personas en hogares con un único núcleo (hogar *pareja con/sin hijos* y hogar *monoparental*), las categorías posibles de la variable *rol* en este tipo de hogares son tres: pivote, hijo y otro. No obstante, de cara a obtener el número de hogares sólo nos interesa la posición de pivote. En los hogares *monoparentales* habrá tantos hogares como personas con el rol de pivote, y en las *parejas con/sin hijos* se tendrá la mitad de hogares que de pivotes (pues a ambos miembros de la pareja se le asigna el *rol pivote*).

En los *hogares sin núcleo* tenemos dos posibles casos: hogares en los que sólo vive un individuo (en los que el número de hogares coincide con el número de personas) y hogares sin núcleo *pluripersonales*. En este último caso, la opción elegida ha sido analizar y proyectar la evolución del indicador *tamaño medio del hogar*, τ . Este indicador se define como:

$$\tau_{h,t} = \frac{\sum_x \sum_s P_{h,x,s,t}}{H_{h,t}} \quad h = \text{hogar pluripersonal}, \quad t = 2000, \dots, 2012 \quad (12)$$

donde $H_{h,t}$ representa el número de hogares de la categoría *pluripersonal* en el año t .

Se ajusta la evolución del número de personas por hogar a una regresión de la forma: $\tau_{h,t} = \alpha_h + \beta_h \ln t + u_{h,t} \quad h = \text{pluripersonal}, \quad t = 2000, \dots, 2012 \quad (13)$

obteniéndose las estimaciones de los parámetros $\hat{\alpha}_h$ y $\hat{\beta}_h$ en el caso de $h = \text{pluripersonal}$.

Esto permite proyectar el número de personas en este tipo de hogares sustituyendo los parámetros estimados en la ecuación (13):

$$\tau_{h,t} = \hat{\alpha}_h + \hat{\beta}_h \ln t + u_{h,t} \quad h = \text{pluripersonal}, \quad t > 2012 \quad (14)$$

Se obtiene finalmente el número de hogares pluripersonales despejando la variable:

$$H_{h,t} = \frac{P_{h,t}}{\tau_{h,t}} \quad h = \text{pluripersonal}, \quad t > 2012$$

Por último, en el caso de los *hogares con más de un núcleo* se ha procedido de modo análogo al caso de los hogares pluripersonales: obtención del tamaño medio del hogar para el periodo observado, 2000-2012, y posterior proyección.

3. Una metodología para la proyección de la población económicamente activa

Los ejercicios prospectivos en el mercado de trabajo se pueden plantear desde dos grandes perspectivas, los enfoques de oferta y de demanda de trabajo. El enfoque de demanda se centra en analizar el empleo que será requerido en el futuro, lo que está íntimamente relacionado con la evolución prevista de la producción, de la estructura sectorial de la economía y de la productividad. Como se puede intuir, se trata de modelos de carácter económico, cuyo resultado final es la proyección del empleo disponible.

En cambio, lo que aquí se pretende no es obtener una proyección del empleo sino de la población activa, por lo que se ha utilizado un enfoque de oferta. La población activa está compuesta por el número total de personas que quieren trabajar, independientemente de que su deseo se encuentre "satisfecho" (estén en situación de empleo) o por el contrario, estén desempleadas y en búsqueda de un trabajo. Es esta una perspectiva de carácter demográfico, pues el número de personas que desean trabajar se encuentra muy condicionado por el volumen y estructura (edades y sexos) de la población. De hecho, se puede expresar el total de activos como la suma ponderada de los efectivos por edad y sexo, donde las ponderaciones vienen determinadas por la propensión de cada colectivo a ofertar su trabajo en el mercado (por las tasas específicas de actividad):

$$Activos_t = \sum_{x=16}^{75+} \sum_s Activos_{x,s,t} = \sum_{x=16}^{75+} \sum_s \left[\frac{Activos_{x,s,t}}{Pob_{x,s,t}} \right] Pop_{x,s,t} = \sum_{x=16}^{75+} \sum_s T_Act_{x,s,t} \overline{Pob_{x,s,t}}$$

donde:

x: edad, x = 16,17, ... ,74, 75+

s: sexo, s = hombre, mujer

t: año de calendario, t = 2009, 2010, ... , 2035

$\frac{R_Act_{x,s,t} T_Act_{x,s,t}}{Pop_{x,s,t}^t}$: Tasa de actividad en el año t, edad x y sexo s

$\overline{Pop_{x,s,t}^t}$: Población media en el año t, edad x y sexo s

Esta descomposición de la oferta de trabajo en colectivos de población y sus respectivas tasas de actividad permite derivar el número futuro de activos de unas proyecciones de población y de algunas hipótesis adicionales acerca de la evolución de las tasas específicas de actividad.

2.1 Principales tendencias en el mercado de trabajo

La Tabla 1 contiene las tasas de actividad para Andalucía, España y algunos países de nuestro entorno europeo.

Tabla 1. Tasas de actividad en Andalucía y países seleccionados. 1990-2012

Grupos de edad País/Región	Total			Hombres			Mujeres		
	1990	2000	2012	1990	2000	2012	1990	2000	2012
Entre 15 y 24 años									
ANDALUCÍA	46,9	42,5	39,1	51,1	47,0	40,6	42,4	37,9	37,5
EU - 15	47,2	47,5	45,6	50,4	50,9	48,1	44,0	44,1	43,1
Alemania	60,7	50,4	50,7	62,5	53,7	53,2	58,8	47,1	48,1
España	47,0	43,1	38,8	51,6	46,7	40,2	42,6	39,4	37,4
Francia	44,6	35,5	37,8	47,7	38,6	41,1	41,6	32,4	34,5
Italia	46,8	38,1	28,7	50,7	42,2	33,1	43,0	34,0	24,0
Países Bajos	59,6	72,2	69,9	60,0	73,4	68,5	59,2	70,9	71,4
Suecia	45,5	40,7	52,6	44,2	41,1	51,8	46,8	40,4	53,4
Reino Unido	71,8	63,4	59,3	76,7	67,0	61,7	66,7	59,8	56,8
Entre 25 y 49 años									
ANDALUCÍA	67,7	75,6	85,8	93,8	91,9	92,0	41,3	59,1	79,4
EU - 15	82,1	83,8	86,2	93,6	93,6	92,6	70,4	73,9	79,8
Alemania	80,9	86,1	87,9	94,0	94,1	93,3	67,4	77,8	82,3
España	72,1	80,1	88,0	95,0	93,9	93,3	49,8	66,2	82,5
Francia	85,2	87,4	88,9	96,6	95,2	94,2	74,0	79,7	83,8
Italia	75,3	76,6	78,4	95,4	91,9	89,3	55,7	61,1	67,4
Países Bajos	78,0	85,6	88,4	94,6	95,0	93,4	60,6	76,0	83,5
Suecia	89,9	86,5	90,7	92,3	88,4	93,8	87,3	84,6	87,5
Reino Unido	84,8	84,8	85,9	95,9	93,0	92,7	73,8	76,8	79,1
50 y más años									
ANDALUCÍA	23,7	22,5	31,4	41,7	35,4	40,0	8,9	11,7	24,0
EU - 15	28,6	29,5	35,9	39,6	39,1	42,9	19,7	21,5	29,9
Alemania	32,9	31,0	39,6	49,0	40,2	46,2	20,7	23,3	33,9
España	26,0	25,2	33,6	42,8	38,5	41,3	12,4	14,0	27,0
Francia	26,0	27,5	33,9	34,2	33,3	38,2	19,5	22,7	30,2
Italia	25,2	22,1	28,2	40,5	33,8	37,6	12,2	12,4	20,4
Países Bajos	25,2	31,8	42,0	38,9	42,8	50,5	13,5	21,9	34,2
Suecia	52,1	41,8	44,3	57,0	47,8	49,4	47,4	36,7	39,6
Reino Unido	34,1	36,2	40,8	45,8	44,6	47,1	24,6	28,9	35,1

Fuente: INE y EUROSTAT. (i) Dato relativo a 1995

Aun cuando existen ciertos rasgos específicos en cada país, las tendencias recientes y los determinantes principales son comunes a todos ellos:

1. Los hombres presentan un perfil de tasas de actividad muy estable, por encima del 90% en las edades centrales (25 a 49 años), siendo en todo momento el colectivo con tasas de actividad más altas. Por el contrario, los hombres de más edad no mostraron un comportamiento tan estable: sus tasas de actividad cayeron en los 90 y crecieron después.
2. Las tasas de actividad femeninas crecieron de forma ininterrumpida a lo largo de los últimos 20 años, acercando el comportamiento laboral de las mujeres al de los hombres.
3. La participación de los jóvenes tiene un comportamiento dispar entre países.

2.2 Proyección de las tasas de actividad masculinas

La metodología empleada para proyectar las tasas de actividad masculinas se basa en el "método de la cohorte" [8, 15]. No obstante, como se ha señalado en la

sección anterior, las tasas de actividad masculinas son bastante estables en el tiempo y en el espacio. Los mayores cambios se han registrado en los calendarios de entrada y salida de la población activa, regularidad que, como ahora se expondrá, se ha incorporado en el procedimiento de proyección.

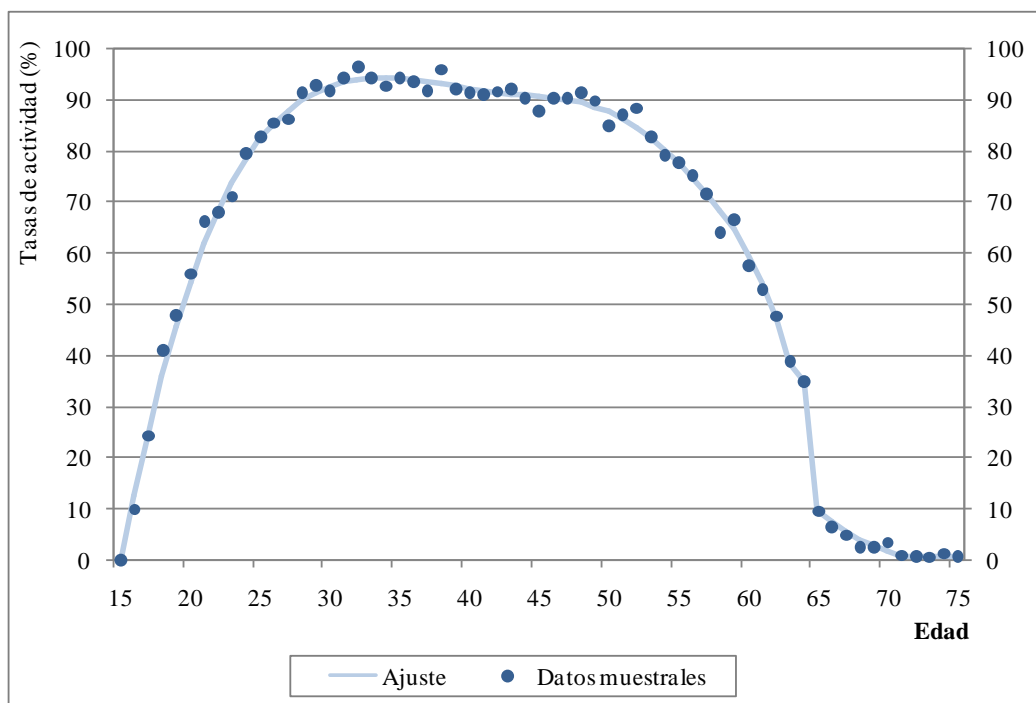
El primer paso consistió en ajustar a una curva las tasas de actividad muestrales, obtenidas a partir de la EPA (Figura 10). Para ello, se han utilizado técnicas de interpolación polinómicas⁹. No obstante, el salto que se produce a los 65 años se ha mantenido, dado que es un salto real, derivado del marco normativo que regula el acceso a los derechos de jubilación.

En un segundo paso, se estimaron las probabilidades de entrada y de salida del mercado de trabajo que se derivan de nuestro perfil robusto de tasas de actividad.

$$P_{Ent_{x,t_0}} = \frac{T_{Act_{x+1,t_0}} - T_{Act_{x,t_0}}}{1 - T_{Act_{x,t_0}}}$$

$$P_{Sal_{x,t_0}} = \frac{T_{Act_{x,t_0}} - T_{Act_{x+1,t_0}}}{T_{Act_{x,t_0}}}$$

Figura 10. Datos muestrales y modelo ajustado de tasas de actividad. Andalucía



Fuente: INE. Encuesta de Población Activa.

⁹ Concretamente, se han utilizado splines [9]. La interpolación por splines garantiza que se pase por un conjunto de puntos (en nuestro caso, las tasas quinquenales de actividad asignadas a la marca de clase) y lo hace mediante un ajuste polinómico (en nuestro caso, cúbico), que tiene aspecto suave.

No obstante, ya se ha mencionado que existen ciertas diferencias entre los calendarios de entrada y de salida del mercado de trabajo en Andalucía y en otros países de nuestro entorno. Por ello, hemos tomado como referencia dos países que tradicionalmente han sido referentes de una entrada temprana al mercado de trabajo (Países Bajos) y de una salida tardía del mismo (Suecia). Y hemos planteado la posibilidad de que las probabilidades de entrada y de salida, bien se queden como están, bien converjan en un horizonte temporal a las de esos países:

$$P_Ent_{x,t}^{And} = \theta_t P_Ent_{x,t_0}^{And} + (1 - \theta_t) P_Ent_{x,t_0}^{Ref}$$

$$P_Sal_{x,t}^{And} = \rho_t P_Sal_{x,t_0}^{And} + (1 - \rho_t) P_Sal_{x,t_0}^{Ref}$$

donde:

And, Ref: datos de Andalucía y modelo de referencia, respectivamente

$0 \leq \theta_t, \rho_t \leq 1$. En nuestro caso, hemos fijado un proceso de convergencia lineal: $\theta_t, \rho_t = \frac{T-t}{T-t_0}$

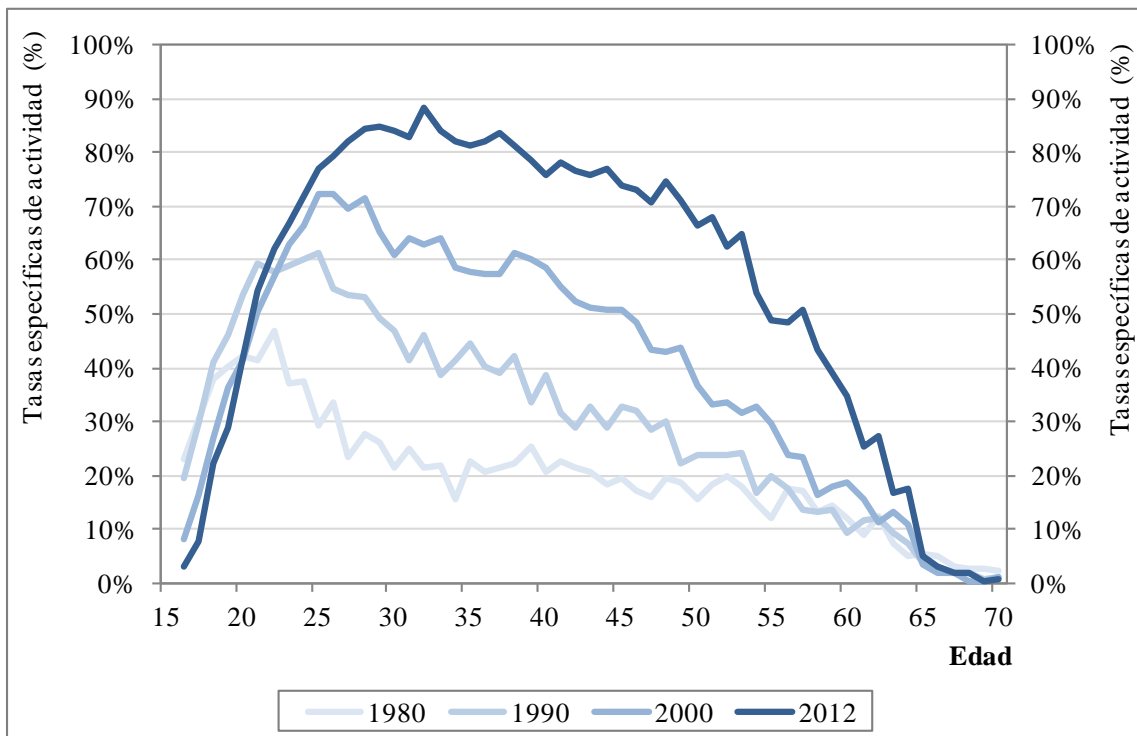
Se debe subrayar que la convergencia de las probabilidades en un determinado horizonte temporal no implica la convergencia de las tasas de actividad en ese mismo lapso. Las probabilidades se irán aplicando a generaciones que ya tienen su propia historia y por tanto, la convergencia de tasas de actividad se producirá más tarde.

$$T_Act_{x+1,t+1}^{And} = \begin{cases} T_Act_{x,t}^{And} + (1 - T_Act_{x,t}^{And}) P_Ent_{x,t}^{And} & \text{si } R_Act_{x+1,t+1}^{And} > R_Act_{x,t}^{And} \\ T_Act_{x,t}^{And} (1 - P_Exit_{x,t}^{And}) & \text{si } R_Act_{x+1,t+1}^{And} < R_Act_{x,t}^{And} \end{cases}$$

2.3 Proyección de las tasas de actividad femeninas

En contraste con la estabilidad que mostraban los hombres, las tasas de actividad femeninas han venido experimentando importantes variaciones en los últimos 30 años (Figura 11). En su rasgo más general, el comportamiento laboral de las mujeres se ha ido acercando progresivamente al de los hombres.

Figura 11. Evolución de las tasas de actividad por edad. Andalucía



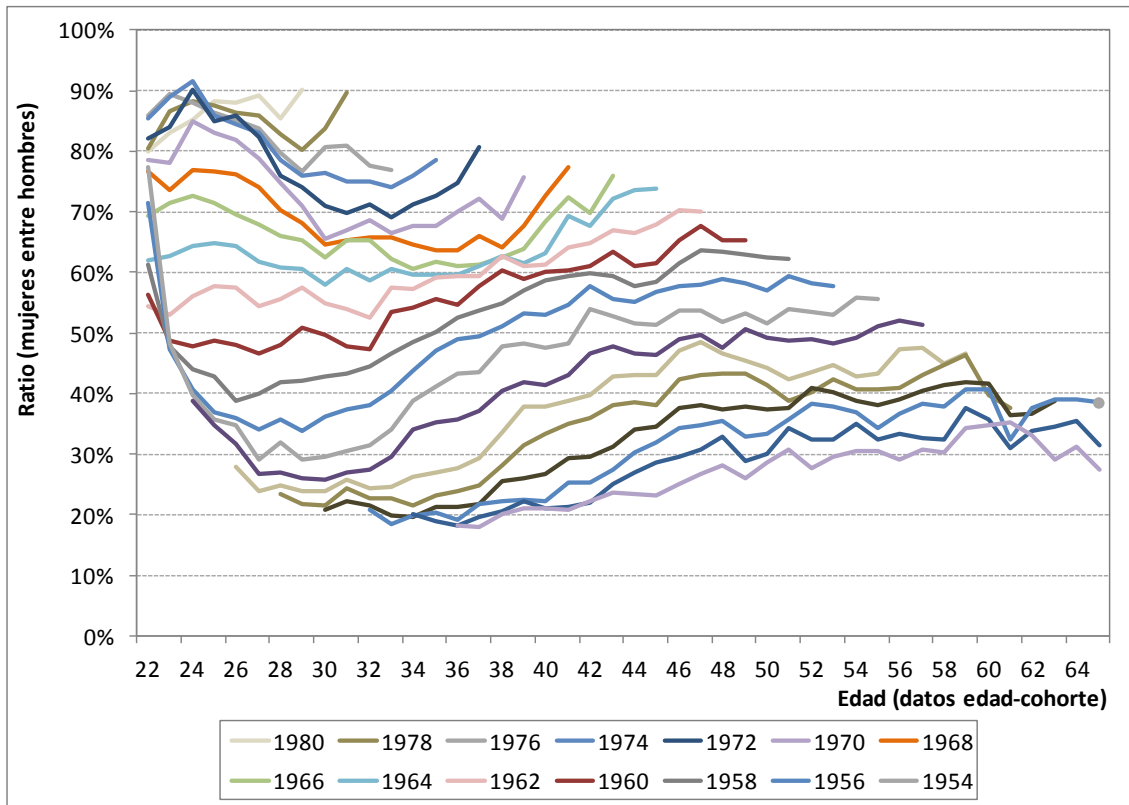
Fuente: INE. Encuesta de Población Activa.

Siendo esta la principal regularidad de largo plazo que han registrado las tasas de actividad femeninas, la metodología que se ha desarrollado para su proyección se basa en la modelización explícita del proceso de convergencia. Para ello, se comenzó estimado los ratios de tasas de actividad femenina y masculina. Se ha utilizado una media móvil centrada de orden 3 para que el resultado fuera todo lo robusto posible.

$$Ratio_{g,t}^{And} = \frac{\sum_{i=g-1}^{i=g+1} T_{Act_{i,Mujeres,t}}^{And}}{\sum_{i=g-1}^{i=g+1} T_{Act_{i,Hombres,t}}^{And}}$$

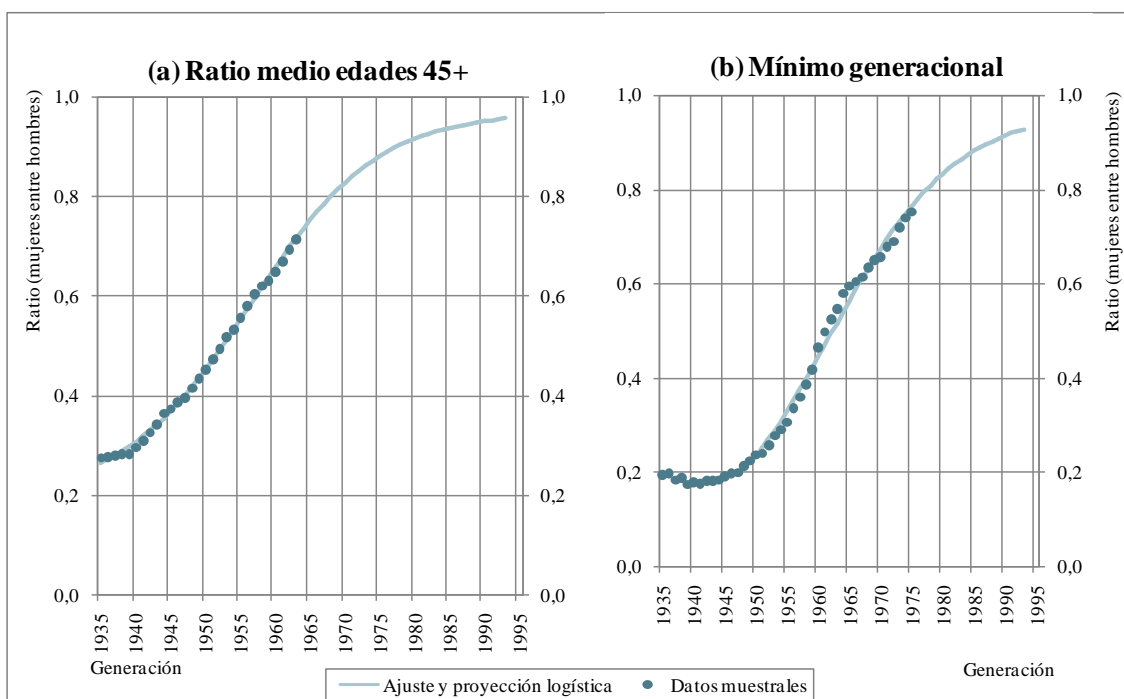
El proceso de convergencia en curso da lugar a que las generaciones más jóvenes registren valores cada vez más próximos a la unidad (Figura 12). Evidentemente, las generaciones más jóvenes no han cerrado su ciclo de actividad, por lo que el segmento de curva para el que se dispone de información se va haciendo progresivamente menor. Por otro lado, se puede observar también que dada una generación, se produce una estabilización de los valores del ratio en torno a los 45 años de edad.

Figura 12. Ratio de tasas de actividad. Generaciones 1940 a 1980



Como se muestra a continuación, estas regularidades se han incorporado al proceso de proyección mediante el uso de funciones logísticas. En primer lugar, la Figura 13a representa el ratio de tasas de actividad para los mayores de 45 años. Como se puede observar, la función logística se ajusta muy bien a esta evolución. Se ha empleado un procedimiento similar para proyectar el mínimo valor del ratio para cada generación (Figura 13b).

Figura 13. Ajuste logístico y proyección de la cola y el mínimo de los ratios...



Finalmente, disponemos ya de los dos elementos que nos van a permitir proyectar los ratios: por un lado, la información empírica hasta el último año disponible. En segundo lugar disponemos también de cierta información adicional que hemos proyectado con la ayuda de funciones logísticas. Esto nos permite completar las curvas de cada generación mediante el uso del siguiente modelo paramétrico:

Sea $R_g(x) = (r_g(x_1), \dots, r_g(x_N))$ el vector de valores del ratio de tasas de actividad para la generación g a la edad $x = \{x_1, \dots, x_N\}$. Dada una familia de funciones dependiente de un vector de parámetros λ , que representamos $f(x; \lambda)$, el ajuste del ratio de actividad para la generación g puede determinarse seleccionando el valor de λ que proporciona una mejor aproximación de $f_g(x; \lambda)$ al vector $R_g(x)$, según el criterio de mínimos cuadrados. Es decir, la función de ajuste se obtendría mediante la resolución del siguiente problema de optimización:

$$\min_{\lambda} \sum_{x=1}^N (r_g(x) - f_g(x, \lambda))^2$$

Concretamente, el modelo que se ha desarrollado hace uso de una mixtura de funciones:

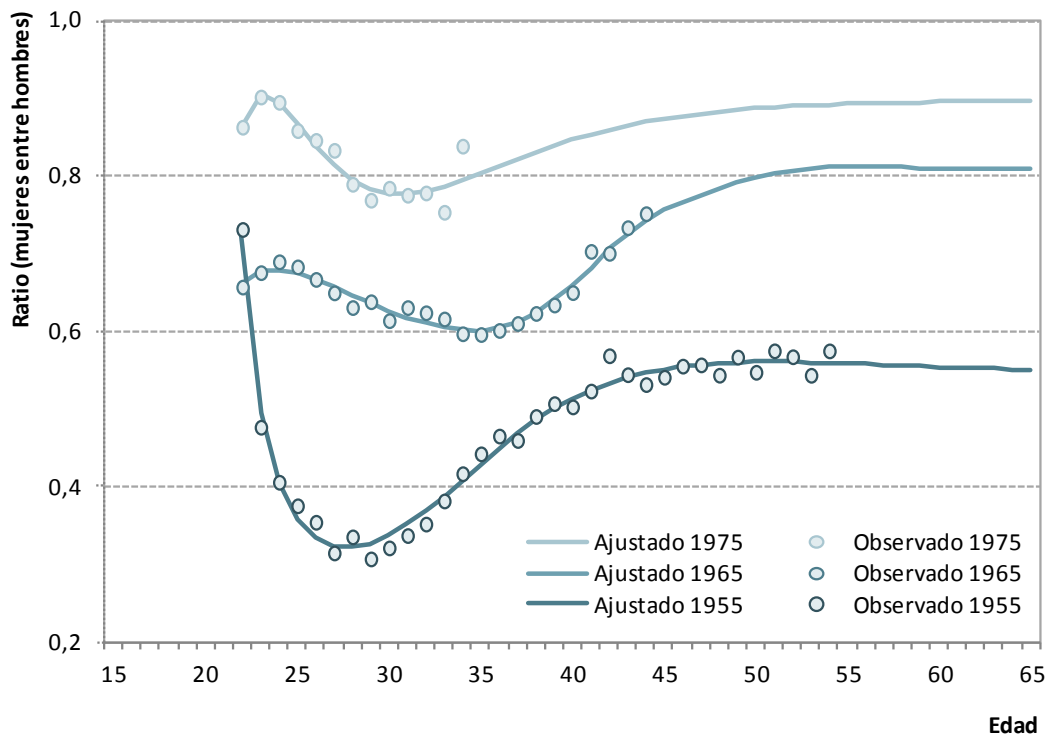
$$f_g(x; \lambda) = \frac{1}{a + be^{-cx}} + \alpha x^\beta e^{-\gamma x}$$

donde:

$\lambda = (a, b, c, \alpha, \beta, \gamma)$ son los parámetros a estimar, siendo
 x: edad
 t: año calendario
 g: generación

Los resultados obtenidos, tras aplicar este procedimiento a los datos correspondientes para Andalucía, se muestran en la Figura 14.

Figura 14. Ratios observados, ajustados y proyectados.
 Generaciones seleccionadas.



Finalmente, la proyección de las tasas de actividad femeninas se obtiene a partir de las tasas de actividad masculinas previamente proyectadas y de la proyección de los ratios de tasas de actividad (mujeres entre hombres).

Por su parte, la proyección de la población activa se obtiene multiplicando las tasas de actividad por las poblaciones medias, para lo que se usan las proyecciones de población.

4. Conclusiones

En este trabajo se han presentado dos metodologías desarrolladas por el Instituto de Estadística y Cartografía de Andalucía: la proyección de los hogares y la proyección de la población activa.

Por un lado, en el caso de la proyección de los hogares, se ha desarrollado un modelo que incorpora algunas variantes novedosas respecto al método de las propensiones, descomponiendo las tasas de propensión en las componentes intensidad y calendario. Cabe destacar que, dentro de los hogares, se han podido especificar todas las relaciones de parentesco de primer grado, aun cuando no esté implicada en ellas la persona principal, una posibilidad poco explorada hasta la fecha.

Por su parte, la segunda parte de este trabajo se centra en la metodología desarrollada para la proyección de la población económicamente activa. Se trata de un modelo paramétrico, basado en dos hechos empíricos. Uno, el perfil por edades de las tasas de actividad masculinas es mucho más estable en el tiempo y en el espacio que el de las femeninas. Los principales cambios que han tenido lugar en las tasas de actividad masculinas se refieren a los calendarios de entrada y salida del mercado laboral. Y dos, la participación de las mujeres en el mercado de trabajo se encuentra mucho más influida por un efecto cohorte, con las generaciones más jóvenes, mostrando un comportamiento laboral cada vez más próximo al de los hombres.

REFERENCIAS

1. Bell, M., Cooper, J.A. and Less, M. (1995), "Household and Family Forecasting Models - A Review". Commonwealth Department of Housing and Regional Development.
2. Bermúdez, S., Hernández, J.A. y Planelles, J (2014), "Una metodología para la proyección de los hogares utilizando datos de la Encuesta de Población Activa. Aplicación al caso de España". Estadística Española, pendiente de publicación.
3. Bermudez, S., Blanquero, R., Hernandez, J. A., Planelles, J. (2012), "A new parametric model for fitting fertility curves". Population Studies, 66-3, págs. 297-310.
4. Bermúdez, S., Hernández, J.A. y Planelles, J (2011), "Proyección de la población activa de Andalucía, 2009-2035". Documentos de trabajo, 3. IECA, Sevilla.
5. Bermúdez, S., Hernández, J.A. y Planelles, J (2010), "El Futuro de la Población en España. Distintos Escenarios". Estadística Española, 52, págs. 237 - 276.
6. Bermúdez, S., Hernández, J. y Planelles (2010). 'Sub-national and foreign-born population projections. The case of Andalusia'. In "Work Session on Demographic Projections, Lisbon 28-30 April", págs. 407-417. Eurostat Methodologies and Working Papers.
7. Bijak, J. et al (2007). "Population and labour force projections for 27 European countries, 2002-2052: impact of international migration on population ageing". European Journal of Population 23-1, págs. 1-31.
8. Carone, G. (2005). "Labour force projections: a set of data for assessing the long-run economic impact of ageing". Joint EUROSTAT-UNECE Work Session on Demographic projections.
9. De Boor, C. (1978). "A practical guide to splines". Springer-Verlag, New York.
10. EUROSTAT (2012), "European Union Statistics on Income and Living Conditions". European Commission.
11. Fernández-Cordón, J.A. y Tobio, C. (1999), "Las Familias Monoparentales en España". Madrid, Ministerio de Trabajo y Asuntos Sociales.
12. Garrido, L., Requena, M. y Toharia, L. (2000), "La Encuesta de Población Activa desde la perspectiva de los hogares". Estadística Española, 42, págs. 115 - 152.
13. General Register Office for Scotland (2007), "Household Projections for Scotland 2004 - based (revised)". Scottish Administration.

14. Hooimeijer, H. and M. Linde (1988), "Vergrijzing, individualisering en de woningmarkt, het WODYN-simulatiemodel". Rijksuniversiteit Utrecht, Utrecht.
15. ILO (2009). "Estimates and Projections of the Economically Active Population: 1980-2020 (fifth edition)". Geneva 2007, Update December 2009.
16. Instituto de Estadística de Andalucía (2012). "Proyección de los hogares de Andalucía 2012-2035". IECA, Sevilla.
17. Instituto de Estadística de Andalucía (2011). "Proyección de la población activa de Andalucía 2009-2035". IECA, Sevilla.
18. Instituto de Estadística de Andalucía (2010). "Proyección de la población de Andalucía 2009-2070". IECA, Sevilla.
19. Instituto de Estadística y Cartografía de Andalucía, IECA (2001). "Hogares y familias en Andalucía". IECA, Sevilla.
20. Instituto Nacional de Estadística, INE (2011). "Proyecto de los Censos Demográficos 2011". INE, Madrid.
21. Instituto Nacional de Estadística, INE (2008). "Encuesta de Población Activa 2005". INE, Madrid.
22. Recaño, J. y otros (2002), "Proyección de hogares de la Comunidad de Madrid, 2002-2017". Instituto de Estadística de la Comunidad de Madrid, Madrid.
23. Scherer (2002). "Age withdrawal from the labour force in the OECD countries". Labour Market and Social Policy Occasional Papers, 49-DELSA.
24. Taylor, S. and Rayner, L. (1999), "Projections of households, families and living arrangements". Australian Bureau of Statistics, Demography Section Australian Bureau of Statistics.
25. United Nations (1971). "Methods of projecting the economically active population".
26. United States National Resources Planning Committee (1938), "The problems of a changing population". Washington, DC, Government Printing Office.
27. Willekens, F., Drewe, P. (1984). "A multiregional model for regional demographic projection". Demographic Research and Spatial Policy. Academic Press. London.