



Instituto de Estadística y Cartografía de Andalucía
CONSEJERÍA DE ECONOMÍA, HACIENDA Y ADMINISTRACIÓN PÚBLICA

Estadísticas Longitudinales de Biografías Reproductivas

Autores:

Víctor Montañés Cobo
Rosa Cánovas Balboa
Francisco Viciano Fernández
Juan de Dios Ramos Poyatos

Organismo: Instituto de Estadística y Cartografía de Andalucía

Palabras clave: fecundidad; longitudinal; demografía; indicadores

Resumen

El IECA publicó por primera vez en 2017 una explotación sobre biografías reproductivas derivada de la infraestructura de información denominada “Base de Datos Longitudinal de Población de Andalucía” (BDLPA), de la que ya se había realizado un estudio sobre mortalidad.

Su objetivo ha sido suministrar información para analizar los diferentes comportamientos de la población en relación al fenómeno de la fecundidad combinando datos de nacimientos con características personales, familiares y del entorno de sus progenitores recogidas en los Censos de Población y Viviendas de 2001 y 2011.

El colectivo bajo estudio lo conforman las personas nacidas entre 1950 y 1999 de forma que se contemplan cinco décadas al completo y se puede estudiar el fenómeno en todas las edades que conforman el rango biológico clásico (15-49 años), estudiándose tanto mujeres como hombres.

Para cada individuo se han calculado dos tipos de episodios temporales, uno residencial y otro reproductivo. Para los episodios residenciales se ha utilizado fundamentalmente la información de las variaciones padronales, mientras que para los reproductivos se ha tenido en cuenta información de MNP (desde 1975), de los censos desde 1991 y de la convivencia padronal. La residencia en cada episodio se geocodificó utilizando el Callejero Digital Unificado de Andalucía y el Censo de Edificios de 2011 para poder representar los datos en la malla cuadrada (250m de lado) de referencia del IECA. Finalmente, se realizó la intersección de estos dos tipos de episodios.

Esta publicación plantea dos tipos de diseños longitudinales. El primero es de tipo prospectivo, fijando las características de las personas (salvo edad y residencia) al inicio del estudio a través de la información recogida en el censo de 2001. El segundo es (semi)retrospectivo, observándose las características de las personas cerca de la fecha de fin de estudio (recopilada por el censo de 2011).

La publicación final consta de tres productos para difusión:

- Un conjunto de tablas estadísticas (hipercubos) sobre el conjunto de la población con seguimiento efectivo en base a criterios temáticos y territoriales.
- Ficheros muestrales de microdatos para usuarios especializados.
- Mapas con indicadores inframunicipales.

Entre otras muchas cosas, los resultados han permitido constatar la evolución del porcentaje de personas sin hijos a las distintas edades, el desplazamiento progresivo de las curvas de fecundidad o la diferenciación por nivel de estudios. Además, como subproducto de uso futuro, se ha creado una tabla de relaciones familiares de filialidad.

1 Introducción

El presente documento está basado en los trabajos realizados para la difusión de las “Estadísticas Longitudinales de Biografías Reproductivas en Andalucía, 2002-2013”. Esta publicación supone la continuación de la línea de producción estadística basada en la integración y reutilización de información recopilada en otras fuentes y operaciones estadísticas. Del mismo modo se avanza en el conocimiento demográfico desde una perspectiva longitudinal que ya se inició con las “Estadísticas Longitudinales de Supervivencia y Longevidad” pero centrando ahora la atención sobre el fenómeno

de la fecundidad, un suceso con características muy diferentes para su análisis. Teniendo en cuenta que la última encuesta sociodemográfica con datos para Andalucía se realizó en 1991, parecía más que adecuado ahondar en el conocimiento de la fecundidad, básico a la hora de planificar tanto para la administración como para los ciudadanos.

Existen antecedentes en este sentido en algunas CCAA como encuesta demográfica realizada por IDESCAT en el año 2007 o la encuesta demográfica elaborada por EUSTAT de periodicidad quinquenal desde 1997)

En el Instituto de Estadística y Cartografía de Andalucía (IECA) se decidió realizar la operación estadística con base en la información ya existente en las diversas fuentes de datos disponibles. El sistema de información que se está construyendo para realizar la integración de estas fuentes se ha denominado Base de Datos Longitudinal de Población (BDLPA). La Ley 3/2013 del Plan Estadístico de Andalucía 2013-2017 (ahora ampliado hasta el año 2020) recogía la importancia de esta nueva fuente al incluirla entre los “*ámbitos prioritarios de actuación en infraestructuras de información*” (Art 17.1.b) a desarrollar durante el periodo de vigencia del Plan.

El objetivo principal es suministrar información adecuada para analizar los diferentes comportamientos de la población en relación con el fenómeno demográfico de la fecundidad, combinando la información del nacimiento de los hijos (MNP fundamentalmente) con la información sobre las características personales (de la madre y el padre), familiares y del entorno recogidas en los Censos de Población y Viviendas de 2001 y 2011 y los movimientos residenciales recogidos en los padrones municipales.

El periodo del estudio es el mismo que el que se determinó en su día para el estudio de la longevidad, abarcando desde el año 2002 a 2013. La fecundidad se considera en un sentido amplio, recogiendo el comportamiento de las mujeres respecto al fenómeno, pero también el de los hombres. Con estas explotaciones de la BDLPA se generan nuevas estadísticas de fecundidad y dinámicas familiares hasta ahora no disponibles y que se vertebran en una serie de hilos conductores:

- Suministrar información a detalle geográfico infra-municipal.
- Reconstruir los episodios de paridad con objeto de poder estimar tasas de fecundidad de primera categoría, propensiones de ampliación familiar y tablas de fecundidad.
- Investigar el efecto en la fecundidad y la dinámica familiar de las características socio-económicas de los progenitores.
- Estudiar las diferencias que se producen entre generaciones, edades y sexo.

Este estudio plantea diferencias relevantes y lógicas a la hora de tratar y analizar los datos. Como ejemplos, la maternidad o paternidad son eventos renovables, no exhaustivos, pueden tener una planificación personal y hay limitaciones biológicas sobre el rango de edades en la que se producen los eventos.

Con el objetivo de poder estimar la diferencia en fecundidad en diferentes colectivos poblacionales definidos tanto por su lugar de residencia como por sus características socio-económicas, se ha diseñado un “estudio longitudinal de cohortes” basado en la información histórica almacenada en la

BDLPA relativa a los padrones municipales, los Censos de Población y Viviendas de 2001 y 2011 y los Boletines Estadísticos de Parto.

Esta publicación plantea dos tipos de diseño longitudinal. El primero de ellos es de tipo **prospectivo**, fijando las características de las personas (salvo edad y residencia¹) al inicio del estudio a través de la información recogida en el censo de población y viviendas de 2001. El segundo es de tipo (semi)**retrospectivo**, es decir, observa las características de las personas cerca del final del estudio (con la información recopilada por el censo de 2011). En ambos casos (**Ilustración 1**) se investiga cómo evoluciona la historia reproductiva de los individuos incluidos en cada estudio, durante la ventana de observación 2002-2013, pero mientras en el diseño prospectivo los eventos reproductivos ocurren después de observar las características personales recogidas en el censo de 2001, en el diseño retrospectivo los eventos ocurren, mayoritariamente, antes de observar las características personales recogidas en el censo de 2011. Estas diferencias de diseño tienen importantes implicaciones en cuanto a cómo analizar, estimar indicadores y plantear hipótesis causales en cada uno de los dos tipos de estudio.

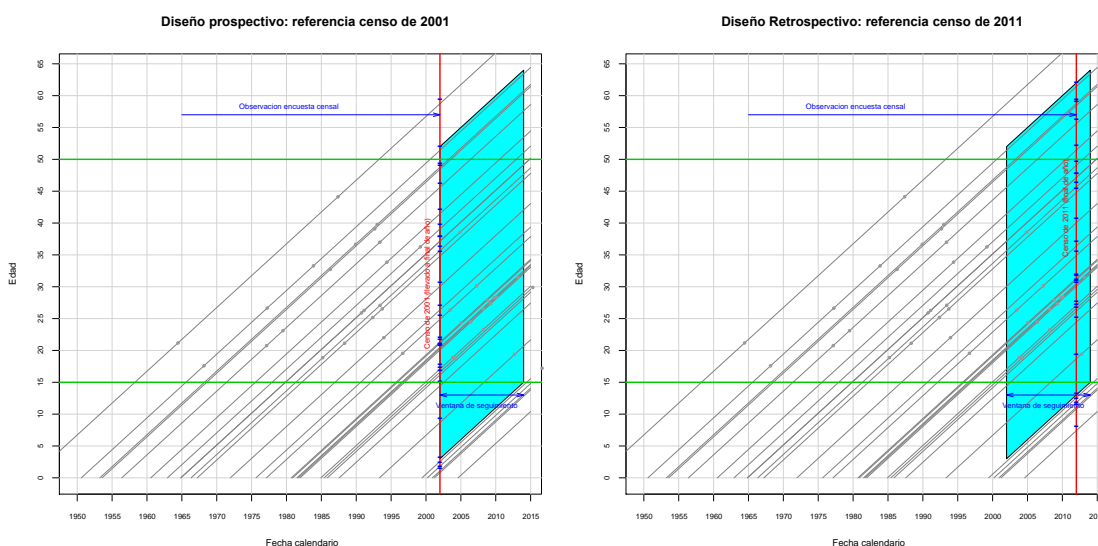


Ilustración 1. Esquema comparativo de estudio prospectivo y retrospectivo

2 Fuentes y población de estudio

La fuente de información a partir de la cual se ha realizado este producto de difusión es una infraestructura de información estadística denominada “Base de Datos Longitudinal de Población de Andalucía” (BDLPA), que está desarrollando el IECA. Las fuentes principales que la nutren casi desde sus inicios provienen de las inscripciones en los padrones municipales de Andalucía, los boletines estadísticos del Movimiento Natural de Población y los Censos de Población y Viviendas. La integración de esta información permite la incorporación de variables recopiladas en una fuente a las de otra, que o bien no disponía de dichas variables o bien las tenía recogidas con un menor nivel de detalle. Este proceso de integración y reutilización de información ya recopilada está comenzando a

¹ Edad y Residencia son las dos únicas variables personales tratadas como tiempo-dependientes en las explotaciones de esta publicación, mientras que el resto de las variables son observadas en una fecha temporal fijada: la censal.

generar una nueva gama de productos estadísticos que amplían y complementan la producción estadística clásica, con un coste menor que el que supondría diseñar nuevos circuitos de recogida de información o la realización de procesos de encuestación periódicos. Actualmente, en el marco vigente de planificación estadística en Andalucía, la BDLPA está considerada como infraestructura de información dentro del ámbito prioritario de la población.

La BDLPA es la nueva fuente estadística donde se realiza la integración y reutilización de la información ya recopilada en otros registros estadísticos y administrativos de base poblacional. Esta integración permite, por un lado, la reutilización de la información recogida en una determinada actividad para comprobar, mejorar la calidad o ampliar la información de otras estadísticas sectoriales. Por otro lado, permite diseñar nuevos productos de difusión, como es el caso de esta publicación, a partir de la información surgida de las interrelaciones de las fuentes originales.

A grandes rasgos, la información relevante para el estudio que se tiene de las distintas fuentes es la siguiente:

- Variaciones padronales desde 1996. Se incluyen todas las altas y bajas padronales comunicadas en municipios de Andalucía. Son proporcionadas por el INE semestralmente.
- Inscripciones en el MNP. Se incluyen mensualmente todos los nacimientos, partos, matrimonios desde 1998 y las defunciones desde 1992.
- Fichero histórico de partos desde 1975 (sin nombres y apellidos).
- Censo de Personas y Viviendas de 1991.
- Censo de Personas y viviendas de 2001.
- Censo de Personas y viviendas de 2011, muestra y marco muestral.
- Callejero Digital Unificado de Andalucía (CDAU), creado y mantenido por el IECA, que permite la geocodificación de los eventos.

Todas estas fuentes han sido utilizadas de una u otra forma en esta publicación.

El punto de partida de la publicación es el año 2002, ya que es el año completo más próximo a la operación censal de 2001, de forma que se pueden conocer las condiciones de los individuos bajo estudio al inicio del periodo analizado. A diferencia de la anterior publicación de supervivencia y longevidad en Andalucía, para la que se seleccionó toda la población censada en 2001, en esta ocasión sólo se han seleccionado a las personas de las generaciones nacidas entre los años 1950 y 1999 (ambos inclusive), condicionada esta selección por dos motivos fundamentales. En primer lugar, se contemplan cinco décadas al completo, permitiendo divisiones por tramos quinquenales y decenales. En segundo lugar, por razones relativas al rango biológico en el que suceden la mayoría de los partos y el rango de edad habitualmente usado en demografía para el cálculo de los indicadores de fecundidad entre 15 y 49 años. Fuera de este rango, los partos suelen ser eventos poco frecuentes, suponiendo menos de un 0,1% del total de maternidades y un 3% de las paternidades. La elección de las fechas 1950 y 1999 se traduce finalmente en que las personas nacidas en 1950 tenían 51 años al iniciarse el estudio (63 al final) y los nacidos en 1999 tenían 14 años al finalizar el periodo de observación.

3 Preparación de la información

Es sobradamente conocido que el uso de los registros administrativos para su uso estadístico conlleva una serie de limitaciones, restricciones y, cuando se combinan varios, incluso contradicciones².

El primer paso es establecer entre las distintas fuentes qué individuos son los mismos, teniendo en cuenta que en una fuente concreta un mismo individuo puede aparecer inscrito más de una vez con datos diferentes, ya sea por errores de escritura del individuo, del funcionario que transcribe, de lectura OCR...o de cambios reales. Después hay que darles coherencia temporal y comprobar que no suceden combinaciones de variables que resultan imposibles.

Para obtener finalmente la información utilizada en la publicación es necesario construir una tabla de episodios residenciales y otra de episodios reproductivos, obteniéndose por intersección temporal la tabla final. Por otro lado, al tratar el lugar de residencia, se geocodificó la dirección postal asociada y se agrupó la nube de puntos espaciales resultante de acuerdo con la rejilla cuadrada de 250m de lado que publica el IECA.

3.1 Construcción de episodios

Los episodios contruidos para la publicación constan fundamentalmente de los datos demográficos del individuo (sexo, fecha de nacimiento, lugar de nacimiento y nacionalidad), los datos residenciales, los datos relacionados con fechas de permanencia (entrada, salida, causas) y el número de hijos al inicio del episodio. Su construcción, como ya indicamos anteriormente, pasa por la elaboración de una tabla de episodios residenciales y otra de episodios reproductivos.

3.1.1 Episodios residenciales

La tabla de episodios residenciales se deriva de los trabajos realizados para la publicación “Distribución espacial de la población en Andalucía”. Si bien en la publicación se acaba dando la población a 1 de Enero de un año concreto, el fichero que la origina dispone de todos los episodios residenciales desde 1996. En las XVIII Jecas, en Oviedo, se presentó el trabajo sobre cálculo de poblaciones en el que se apoya esta fase³. A continuación se da un ejemplo simplificado de la estructura de esta tabla.

² Este tema se aborda en la comunicación “Integración y conciliación de ficheros administrativos a nivel individual para la elaboración de estadísticas públicas en Andalucía”, presentado en estas mismas JECAS.

³ “Estimación de los efectivos y estructuras poblacionales de Andalucía a partir de la integración de fuentes estadísticas y administrativas”. Montañés, Cánovas, Poza y Viciana. XVIII JECAS Oviedo, 2014.

IDPersona	Municipio	Edificio	CausaAlta	FechaAlta	CausaBaja	FechaBaja
10064491	23050	648758	-	19960501	CR	20090219
10064491	41091	544427	CR	20090219	CR	20090519
10064491	21007	59535	CR	20090519	CR	20091001
10064491	41091	85220	CR	20091001	CD	20101007
10064491	41091	544427	CD	20101007	CD	20120511
10064491	41091	630476	CD	20120511	CD	20150305
10064491	41091	523202	CD	20150305	XX	20990101

Tabla 1. Episodios residenciales del individuo (ficticio) 10064491. Periodo 1996-2016.

La persona ficticia a la que se refieren los episodios residenciales de la tabla ha vivido en un total de 3 municipios diferentes, representándose los cambios de residencia intermunicipales por CR, y ha vivido en un total de 6 edificios diferentes, representándose los cambios de domicilio intramunicipales por CD.

3.1.2 Episodios reproductivos

El cálculo de los episodios reproductivos es algo más complejo. Si bien a través de los boletines de parto de MNP se pueden obtener las fechas de los nacimientos y quiénes son los progenitores implicados con una cobertura muy amplia, el gran problema reside en establecer cuántos hijos tenía una persona al inicio del periodo de estudio. Para llevar a buen puerto este reto, se decidió construir una tabla de relaciones familiares de primer orden. Esta tabla, además de indicar para cada persona quiénes son su padre y madre a través de sus claves personales (si es que se puede establecer), puede permitir establecer relaciones de órdenes superiores (quiénes son sus tíos, primos, abuelos, cuñados...).

La información para confeccionar esta tabla se ha obtenido de las siguientes fuentes:

- MNP partos 1998-2013. Es sin duda la fuente más fiable, ya que se trata del propio boletín de parto. Además de los datos del nacido se incluyen campos (para la madre) como el número de hijos previos, la fecha del parto anterior o la fecha de matrimonio o unión estable.
- Censo de 1991. Se determinaban en el cuestionario la relación con la persona principal, si el padre, madre o cónyuge residían en la misma vivienda, el número de hijos nacidos con vida y el año de boda...aunque sólo para las mujeres.
- Censo de 2001. La información del cuestionario es algo más pobre que la del censo precedente, pero se dispone de relación con la persona de referencia. No obstante, en el fichero final se podían establecer las relaciones de parentesco.
- Censo de 2011. La información era similar a la de 2001, aunque se concretaba más a la hora de señalar quiénes eran padre, madre, cónyuge o incluso hermano. Por el contrario, se trata de una muestra cercana al 10%. Se retoma la pregunta sobre el número de hijos nacidos con vida de la mujer...sólo de la mujer.
- MNP partos 1975-1997. Se dispone de la información de los partos (aunque no de manera exhaustiva, pero sí con una gran cobertura) y con pocas variables. No obstante, se tiene la

fecha de matrimonio y de nacimiento de los progenitores y la fecha del hijo anterior, lo que permite afinar la búsqueda de relaciones.

- Padrón. A través de convivencia en la misma hoja padronal o vivienda y compatibilidad de apellidos y edad.
- MNP matrimonios 1998-2013. Permite conocer quiénes son pareja y afinar en la búsqueda de hijos en la vivienda que compartan apellidos.

Como podemos observar, parte de la información nos da directamente las relaciones de parentesco y otra nos permite delimitar candidatos que no han sido asignados por diversos motivos, como haber emigrado entre 1991 y 1996, haber nacido el hijo fuera de la comunidad autónoma...

De esta forma, a cada uno de los empadronados se le asignó, si se pudo, la clave de su padre, su madre o ambas. Además, se creó un fichero de no empadronados que eran hijos de empadronados. Tras la creación de estas relaciones se establecieron relaciones de consistencia como, por ejemplo, que una mujer no podía tener un parto 7 meses después del anterior parto, que debía existir una edad mínima a la paternidad o maternidad, que un individuo no podía ser padre y madre a la vez... Toda una serie de pequeñas reglas, en algunos casos aparentemente absurdas, pero que podían tirar por tierra todo el trabajo realizado por unos pocos casos residuales. Para algo más del 90% de los individuos se pudo establecer la secuencia del número de hijos y el número de hijos iniciales. De los restantes, en su mayoría se trataba de población inmigrante que no entraba en la delimitación temporal del estudio y que no había tenido hijos desde su entrada a Andalucía. Como ejemplo, se muestran los episodios reproductivos del mismo individuo del apartado anterior

IDPersona	fecha_inicio	fecha_fin	Hijos
10064491	19910301	20100702	0
10064491	20100702	20120629	1
10064491	20120629	20150101	2

Tabla 2. Itinerario reproductivo del individuo (ficticio) 10064491.

Es decir, esta persona ha tenido dos hijos en partos simples cuyas fechas de nacimiento han sido el 02/07/2010 y el 29/06/2012. Combinando las dos tablas, junto con la limitación temporal, el resultado sería el siguiente:

IDPersona	Municipio	Edificio	CausaAlta	FechaAlta	CausaBaja	FechaBaja	Hijos
10064491	23050	648758	-	20020101	CR	20090219	0
10064491	41091	544427	CR	20090219	CR	20090519	0
10064491	21007	59535	CR	20090519	CR	20091001	0
10064491	41091	85220	CR	20091001	PA	20100702	0
10064491	41091	85220	PA	20100702	CD	20101007	1
10064491	41091	544427	CD	20101007	CD	20120511	1
10064491	41091	630476	CD	20120511	PA	20120629	1
10064491	41091	630476	PA	20120629	XX	20131231	2

Tabla 3. Episodios residenciales y reproductivos combinados del individuo (ficticio) 10064491 para el periodo de estudio 2002-2013.

Por ejemplo ahora puede verse cómo en el edificio 85220 esta persona ha vivido una etapa sin hijos y otra con un hijo. La causa PA representa en este caso que el inicio o fin de episodios ha sido la paternidad. Con la información estructurada de esta forma y conocida la fecha de nacimiento del individuo en cuestión, podemos obtener para cada individuo su edad, generación y año de calendario. Por acumulación, lo que en realidad estamos teniendo son los triángulos (el mayor nivel de detalle) de un diagrama de Lexis, con la ventaja de que estamos teniendo en cuenta personas-tiempo, no personas, y el individuo puede aportar “tiempo” en distintas categorías (municipio, número de hijos...) a lo largo de un mismo año y edad. De esta estructura parten todos los datos publicados.

3.2 Construcción de la malla estadística

Como ya se ha indicado en la estructura básica de datos, se asigna una clave de edificio a cada episodio. Dicha clave se ha enlazado (está en fase de sincronización) con el Callejero Digital de Andalucía Unificado (CDAU), lo que ha permitido asignar coordenadas (directas o aproximadas) a cada uno de los tramos temporales o episodios. Estas coordenadas se han agregado en celdas cuadradas de 250m de lado que conforman una rejilla ficticia que cubre todo el territorio andaluz y que se utiliza en la publicación “Distribución espacial de la población en Andalucía”, permitiendo ofrecer un alto nivel de detalle territorial con una morfología zonal fija en el periodo de estudio y al margen del clásico sistema de secciones censales, no continuas en el tiempo. Este grid se ha utilizado como soporte para el cálculo de indicadores en área pequeña.

3.3 Información de otras fuentes

Con la información sobre el número de hijos el aprovechamiento estadístico es bastante limitado. Si bien permite el estudio del calendario por sexo, edad, año de nacimiento, año de calendario y un gran detalle territorial (que no es poco), su paralelismo con la publicación de longevidad llamaba a incluir otro gran grupo de variables contextuales que se podían obtener de la información censal.

Se plantearon, como ya hemos dicho, dos tipos de estudios, uno con base en el censo de 2001 y otro apoyándonos en el censo de 2011. El estudio basado en el censo de 2001 es puramente prospectivo, tomando las características del individuo y su entorno al inicio del estudio⁴ y estudiando su comportamiento reproductivo a partir de ese momento. La limitación reside en que los valores censales permanecen invariables en los 12 años de estudio. Por ejemplo, si al inicio (2002) una mujer de 19 años tuviera como estudios “Bachillerato” acabaría con 31 años y ese mismo nivel de “Bachillerato”, pese a que con la edad inicial fuera imposible que tuviera un nivel más alto. Por lo tanto, los resultados en determinados grupos han de ser interpretados con cautela.

⁴ El estudio se inicia el 1/1/2002 y la fecha de referencia es de dos meses antes. No obstante, los trabajos del recorrido censal se alargaron en el tiempo, por lo que el desfase temporal no tiene mucha importancia.

Del lado contrario tenemos el estudio retrospectivo, si bien no es puramente retrospectivo. Se ofrece la información cerca del final del periodo de estudio (censo 2011 vs año 2013) para analizar el comportamiento reproductivo una vez conocida la situación de las personas cerca del final del estudio. La limitación en esta ocasión es la contraria, una persona que tenía 23 años en el censo de 2011 y estaba trabajando, al inicio del estudio era imposible que estuviera trabajando por no alcanzar la edad legal y, sin embargo, estará clasificada como trabajadora en el estudio.

4 Contenido de la publicación

A partir de las tablas extraídas de la BDLPA, tras su depuración estadística para suprimir inconsistencias, se han realizado tres tipos de productos para difusión: hipercubos para tabulación, muestras de microdatos e indicadores para área pequeña.

4.1 Hiper cubos

Los hipercubos para la obtención de tablas estadísticas utilizan la información del conjunto de la población con seguimiento efectivo. Estas tablas combinan la información recogida en los censos (2001 y 2011) con los sucesos (fallecimiento, censura, parto) y tiempos de exposición experimentados por esta población durante el periodo 2002-2013. Se han realizado tablas por separado para hombres y mujeres.

Se han definido una serie de tabulaciones que permiten estudiar la relación de la fecundidad con un amplio abanico de características personales socio-demográficas, relativas a la forma de convivencia o a la relación con la actividad que fueron recogidas por los Censos de Población y Viviendas de 2001 y 2011.

La tabulación de los ficheros estadísticos preparados para este trabajo se realiza cruzando un grupo de variables censales (tiempo-independientes) con otras relativas a la experiencia de seguimiento de los miembros de la cohorte. Las variables asociadas al seguimiento son tiempo-dependientes y pueden cambiar a lo largo de la duración del seguimiento. En este trabajo estas variables son el número de hijos (paridad y multiplicidad), la edad y periodo en el que se ha producido el suceso o ha habido una residencia efectiva. Los cruces de estas variables generan un hipercubo multidimensional que el usuario puede configurar en BADEA (Banco de Datos Estadístico de Andalucía) para la visualización de las variables que se seleccionen. A partir de BADEA el número de posibles tablas estadísticas extraíbles de los hipercubos originales es enorme.

La información almacenada es básicamente el número de sucesos (partos simples y múltiples y nacimientos) y el tiempo de seguimiento acumulado por todos los individuos que han transitado por el espacio de estados delimitado por las variables que definen la celdilla. A la suma de los tiempos de exposición a riesgo de todos los individuos que han transitado por ella lo denominamos "Personas-Año" (de exposición). En lo sucesivo, se hablará en este documento sobre nacimientos, si bien se hace extensible al caso de partos con al menos un nacido vivo.

Además, se ofrece la tasa de fecundidad según número de hijos previos (no ajustada por estructura de sexo y edad) expresada en tanto por mil calculada para cada cruce, definida como el cociente entre el número de nacimientos y el tiempo de exposición en riesgo de la subpoblación.

Los hipercubos originales también se suministran en formato PC-Axis para aquellos usuarios que prefieran usar sus propias herramientas análisis.

Este plan de tabulación se ha diseñado limitando cada una de las tablas en base a tres criterios:

1. Criterio temático: determina el objetivo y secundariamente las variables (aunque no su nivel de detalle) a incluir en las tablas. Las *características demográficas sexo, edad y paridad*, que se incluyen en todos los hipercubos, son combinadas con algunas variables correspondientes a otras categorías temáticas: *características socio-económicas personales, estructuras de convivencia o características relativas a la actividad económica*.

2. Criterio territorial: se han definido dos agrupaciones con diferente nivel de detalle territorial: Andalucía y sus provincias, y las unidades territoriales del Plan de Ordenación Territorial de Andalucía (POTA)⁵.

3. Colectivo sub-poblacional: determina el subconjunto de la población a la que hace referencia una tabulación concreta. La información se ofrece para *población residente en viviendas familiares*, *permitiendo los distintos cruces establecer filtros posteriores*.

4.2 Microdatos

Para el estudio *prospectivo*, se han extraído inicialmente dos muestras del fichero censal: una de personas y otra de viviendas familiares a nivel territorial POTA. Adicionalmente, se ha extraído una tercera muestra de microdatos relativos a viviendas a nivel de la comunidad autónoma.

Las dos primeras muestras se distribuyen previa solicitud expresa y tras la firma de un compromiso de uso y confidencialidad. La tercera muestra se distribuye sin restricciones especiales y permite a los usuarios especializados familiarizarse con la información del estudio y realizar los análisis que precisen para el conjunto de Andalucía.

Continuando con el caso del estudio prospectivo, cuyas variables de referencia son las del Censo de 2001, se ha realizado mediante muestreo estratificado utilizando como estratos las secciones censales en la muestra de personas y las unidades territoriales del POTA en la muestra de viviendas. En cada uno de los estratos se ha realizado un muestreo aleatorio simple de las respectivas unidades muestrales, con una fracción de muestreo fija del 10%. En caso de no existir información para el elemento muestreado durante el periodo de referencia, se sustituyó dicha unidad por otra, de forma que se tenga una muestra efectiva del 10%.

De cada uno de los registros seleccionados (personas y viviendas) se ha extraído la información relativa a su seguimiento durante la ventana de observación de 1-1-2002 al 31-12-2013. La información de seguimiento de cada una de las muestras incluye los eventos fallecimiento, cambio de residencia, cambio en el número de hijos y censura por fin del estudio el 31-12-2013. El nivel de

⁵ www.juntadeandalucia.es/medioambiente/pota

detalle de la variación residencial es el de una salida o reentrada en el conjunto de la Comunidad Autónoma de Andalucía en la muestra de personas y las variaciones residenciales que ocurrían entre dos unidades territoriales de las definidas en el POTA en la muestra de viviendas.

Para el estudio *retrospectivo*, y debido a que el censo de 2011 es en sí mismo un proceso muestral, se distribuyen como microdatos todos los individuos muestreados que cumplen las condiciones necesarias para formar parte de este estudio de fecundidad (año de nacimiento, residencia, conocimiento de la trayectoria reproductiva...). La muestra es de alrededor del 97% de la muestra original y las variables han sido recodificadas para preservar la confidencialidad estadística. Se distribuye mediante descarga directa la muestra sin detalle territorial y con las categorías de las variables más agrupadas respecto a la original con objeto de evitar el posible riesgo de revelación de datos personales. Bajo petición expresa y compromiso de uso, se proporciona una segunda muestra con detalle territorial a nivel de comarca POTA, tanto la residencia en la fecha censal como las variaciones residenciales ocurridas durante el periodo de seguimiento.

En el caso de las fechas de inicio y fin de los episodios, éstas han sido truncadas a un decimal para dificultar la identificación de un evento vital de forma más o menos directa. Con este truncado a un dígito se indica que el evento ha sucedido en un rango de 36,5 días.

4.3 Indicadores para área pequeña

La ampliación del detalle territorial de las estadísticas clásicas es uno de los añadidos más interesantes que la integración de fuentes realizada en la BDLPA conlleva. Esta parte de la publicación se desarrolla con detalle en la comunicación “Indicadores de fecundidad en área pequeña” presentada en estas mismas JECAS, por lo que resumiremos brevemente qué se ha publicado.

Con objeto de mostrar el potencial de la BDLPA se estimó inicialmente un indicador poblacional de la intensidad de la fecundidad al primer nacimiento, tanto para hombres como para mujeres. La elección de este indicador se debe a su analogía con la “Razón de Mortalidad Estandarizada Suavizada” (RMES) que se empleó en la publicación de longevidad. Para su cálculo se utilizan los efectivos de población sin hijos previos y los eventos primeros nacimientos.

Posteriormente se calcularon la edad media a la maternidad/paternidad y el número medio de hijos, siendo el ámbito territorial de análisis sendas mallas de celdas regulares de 250 metros y 1 kilómetro de lado.

5 Resultados

Pese a las limitaciones indicadas, hay una gran riqueza en la información aportada. La principal ventaja de estos estudios longitudinales siempre ha sido que a la hora de calcular indicadores el numerador y el denominador provienen de la misma fuente. Por ejemplo, se puede establecer cuál era la propensión de tener un hijo en 2003 si ya se tenía un hijo y un trabajo en el sector servicios. Se trata de hacer preguntas a las bases de datos: ¿Influye el nivel de estudios en el número de hijos o el calendario? ¿Tienen más hijos los autónomos o los asalariados? ¿Cómo influye que una persona

esté en situación de paro? ¿Influye igual si el parado es el hombre que si es la mujer? ¿Cómo ha evolucionado el porcentaje de personas que no tienen hijos a las distintas edades o generaciones?

Aquí presentaremos una serie de resultados como ejemplo, pero evidentemente se pueden obtener muchísimos más indicadores. Básicamente presentaremos probabilidades de ampliación de familia, número de hijos y porcentaje de personas sin hijos en función de diversas características.

Comenzaremos por estudiar las diferencias en el calendario entre hombres y mujeres utilizando la información de base del censo de 2001 que tiene un mayor volumen poblacional. En el Gráfico 1 puede observarse una diferenciación, ya conocida, entre ambos sexos, teniendo las mujeres un calendario anticipado unos dos años respecto a los hombres. En este gráfico también se aprecian diferencias entre los distintos años de calendario, existiendo unas tasas en 2007 claramente superiores en las edades de mayor incidencia. De hecho, en los años 2007 y 2008 se registró el mayor número de nacimientos en Andalucía en los últimos 30 años. No obstante, las tasas así presentadas mezclan información de varias generaciones tal y como se hace para construir las tablas de fecundidad o las de mortalidad. La ventaja de estos datos es que también podemos hacer el estudio por generación.

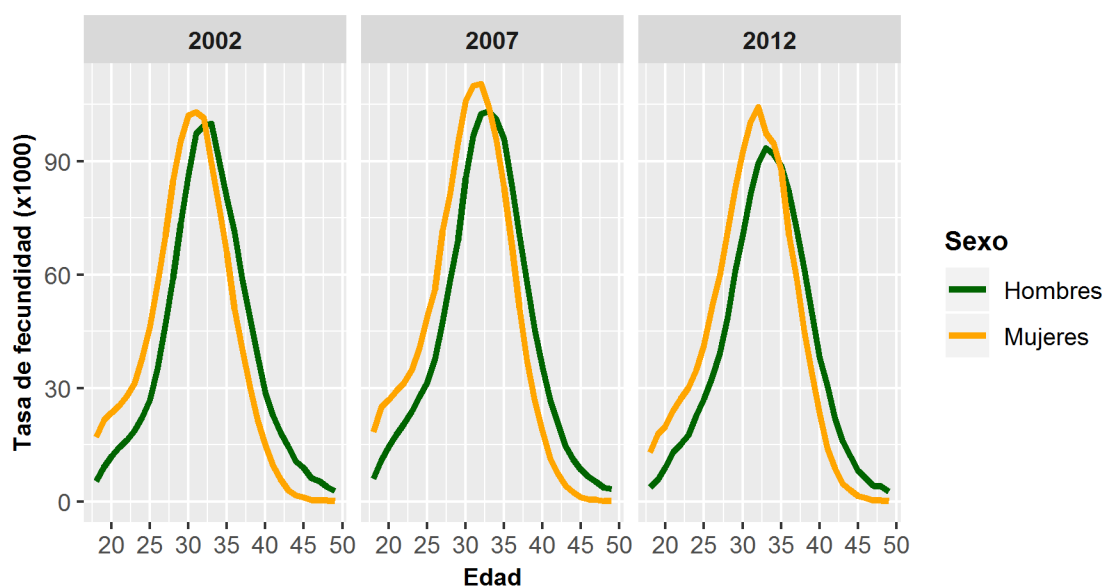


Gráfico 1. Tasa de fecundidad por sexo y edad. Años 2002, 2007 y 2012.

Para interpretar el Gráfico 2 tomaremos como referencia la generación de nacidos en 1975 (color amarillo). Para esta generación se tiene información entre los 27 y los 40 años. Si comparamos en las edades tardías, las tasas son más altas que las de las generaciones previas, es decir, en edades tardías la generación de 1975 tiene una mayor propensión a tener hijos que la de 1970 y esta a su vez que la de 1965. Sin embargo, si comparamos con las generaciones posteriores en edades más jóvenes, vuelve a suceder lo mismo, es decir, las tasas son mayores que las de la generación de 1980 y éstas mayores que las de 1985. Por lo tanto, parece que la generación de 1975 supone un punto de inflexión en el comportamiento reproductivo de los residentes en Andalucía. Son

especialmente llamativas estas tasas tan altas en edades tradicionalmente tardías teniendo en cuenta que a los 33 años sufrieron el inicio de la crisis económica.

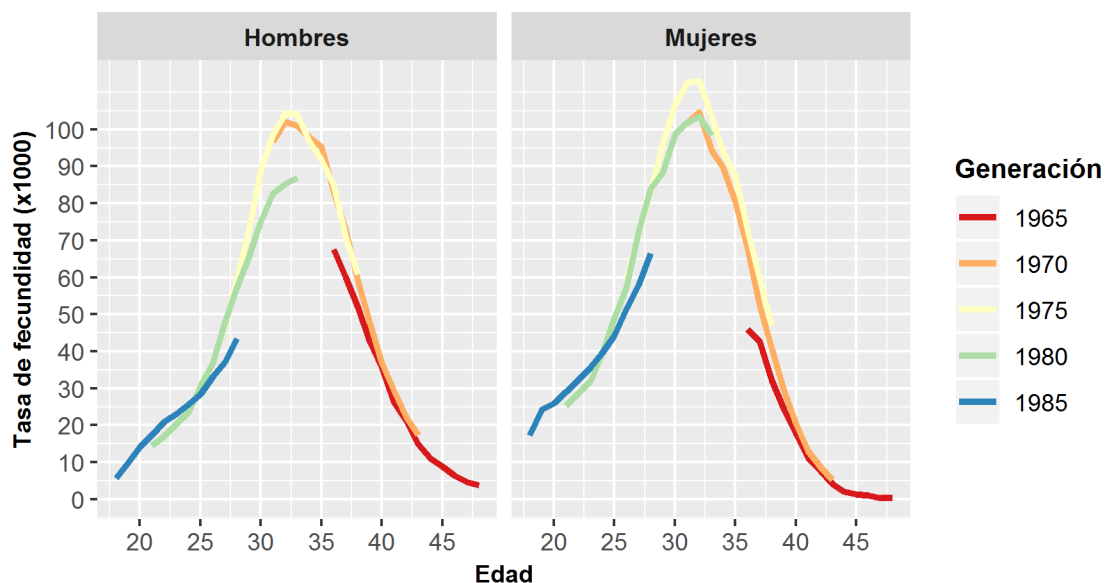


Gráfico 2. Tasa de fecundidad por generación, edad y sexo. Período 2002-2013.

En el Gráfico 3 se muestra el porcentaje de personas sin hijos en tres años (2002, 2007, 2012). Hasta los 27 años no se aprecian diferencias, para a partir de esa edad comienza a observarse un comportamiento diferenciado. Para ambos sexos el porcentaje de personas que no tienen hijos se ha ido incrementando, más si cabe en el caso de los hombres. Además de este cambio, se vuelve a observar, ahora de manera mucho más clara, el retraso en el comportamiento en relación a la fecundidad. La edad a la que la mitad de los hombres había tenido ya su primer hijo era de 32,2 en el año 2002, mientras que en 2012 era de 33,8. En el caso de las mujeres pasó de 30,0 a 31,2.

Si nos desplazamos hacia el final de la curva, el porcentaje de mujeres sin hijos a los 49 años en 2012 era del 12,7%, mientras que en 2002 se situaba en 11,7%. En el caso de los hombres el incremento es mucho mayor, pasando de un 14,7% a un 18%.

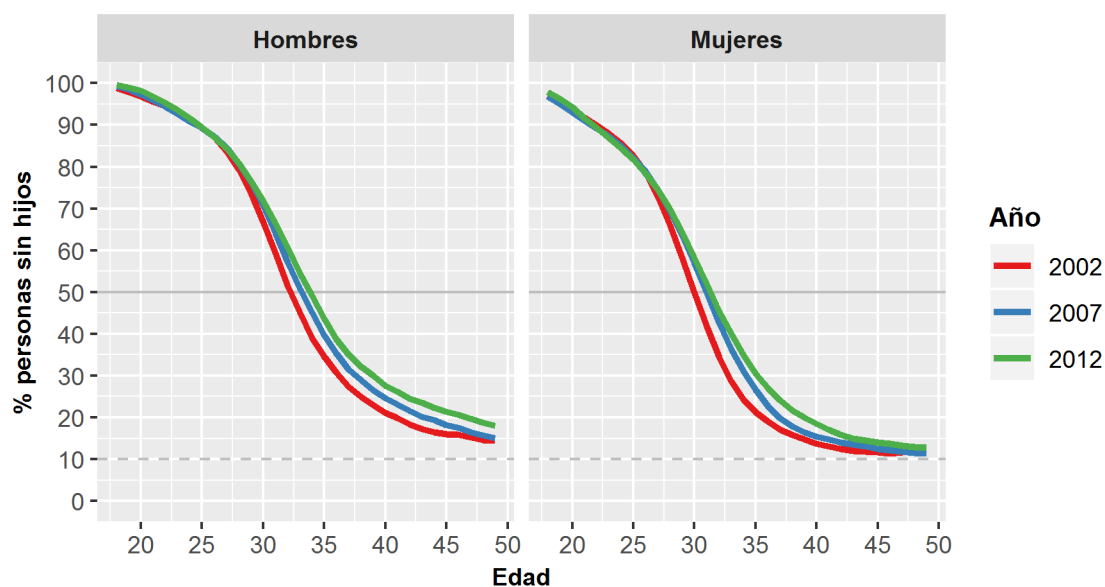


Gráfico 3. Porcentaje de personas sin hijos por sexo y edad. Años 2002, 2007 y 2012

Al margen de problemas como la infertilidad, posteriormente veremos cómo hay características sociodemográficas diferenciadoras a la hora de tomar la decisión de no tener hijos.

Otro gráfico (Gráfico 4) en el que se comparan varias generaciones, puede ayudar a explicar el progresivo cambio de las personas que no tienen hijos. Vemos cómo las curvas están separadas y que por una parte el calendario a la hora de tener el primer hijo se va retrasando a la vez que va aumentando el número de mujeres sin hijos.

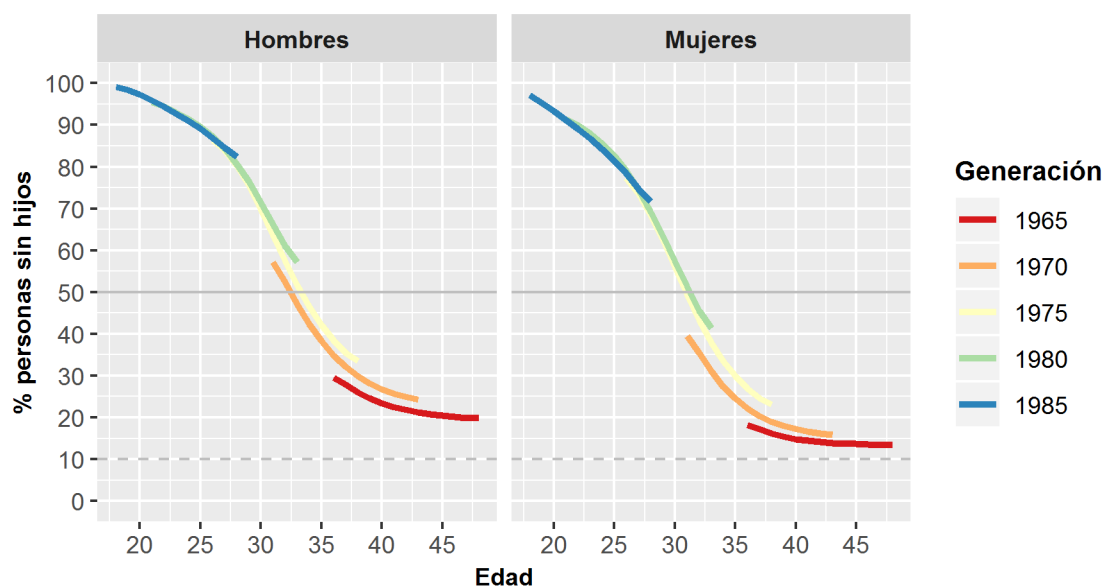


Gráfico 4. Porcentaje de personas sin hijos por sexo, edad y generación.

Ahora bien, ¿cómo afectan estos cambios a la descendencia final? La descendencia de una mujer de 49 años pasó de 2,36 hijos en 2002 (generaciones de 1962 y 1963) a 1,91 hijos en 2012 (generaciones de 1972 y 1973). En el siguiente Gráfico 5 podemos ver que además del número de

personas con hijos también ha bajado el número medio de hijos y la clara diferenciación entre las generaciones previas y posteriores a 1975.

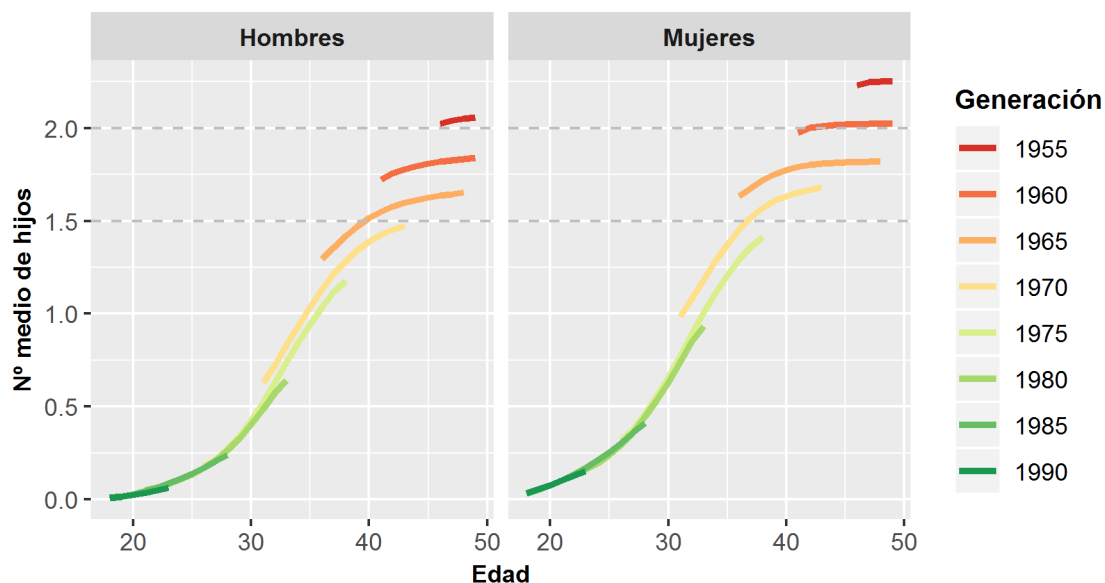


Gráfico 5. Número medio de hijos por sexo, edad y generación.

Como ejemplo, a los 33 años, el número medio de hijos de una mujer ha ido descendiendo desde 1,18 para aquellas nacidas en 1970, a 1,01 para las nacidas en 1975 y a 0,93 para las nacidas en 1980. En el caso de hombres, pasó de 0,83 a 0,74 y 0,64.

A los 38 años, en el caso de los hombres pasó de los 1,42 hijos para los padres de la generación de 1965, a 1,28 para padres de 1970 y 1,18 para los de 1975. En el caso de las mujeres el cambio fue desde 1,72 a 1,57 y a 1,41. Por lo tanto, este cambio ha sido más brusco para la generación de hombres que para la de mujeres y las diferencias son aún más notables a los 38 años que a los 33 años.

El hecho de que exista un comportamiento diferenciado entre generaciones hace que el clásico índice coyuntural de fecundidad pierda fuerza como indicador del número de hijos que pretende ser, ya que parte del supuesto de que el comportamiento de la fecundidad se mantiene en el tiempo o, al menos, sufre cambios graduales y podemos observar cómo los cambios han sido bastante bruscos para las generaciones analizadas.

Después de estudiar las diferencias por sexo, edad, año y generación, introduciremos más información de interés. Por los resultados obtenidos, nos vamos a centrar en el nivel formativo y en la relación con la actividad. Focalizaremos la atención en las generaciones nacidas antes de 1976, ya que para las posteriores el nivel formativo aún no estaba establecido (había muchas personas cursando estudios universitarios) o incorporándose al mundo laboral.

En el caso de la relación con la actividad (declarada en el censo de 2001), existe un comportamiento claramente diferenciado para ambos sexos, como puede apreciarse en el Gráfico 6.

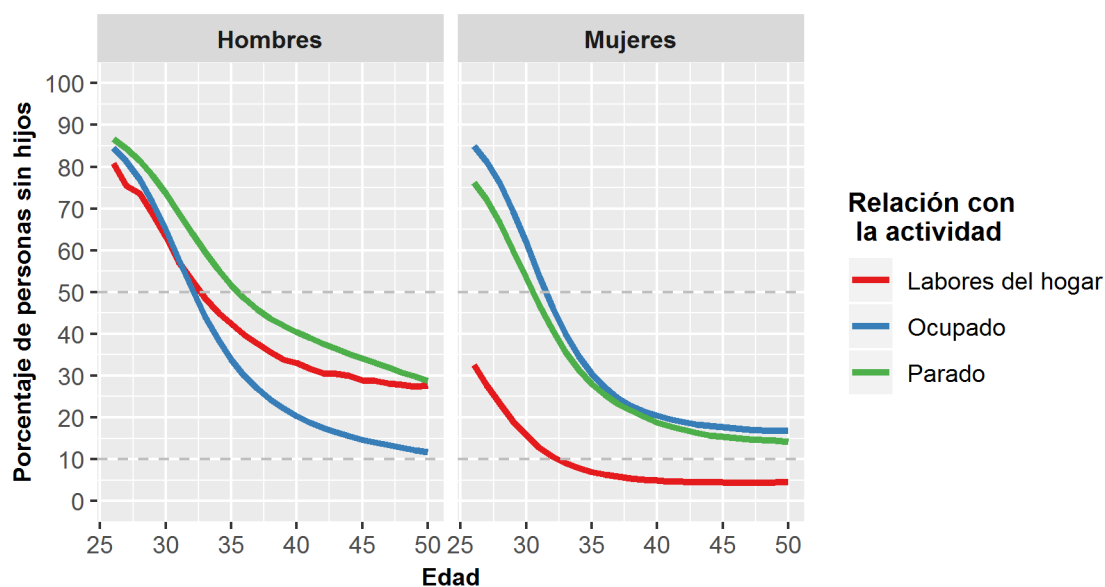


Gráfico 6. Porcentaje de personas (nacidas entre 1950 y 1975) sin hijos por sexo, edad y relación con la actividad.

El porcentaje de hombres parados sin hijos a los 49 años casi triplica al de hombres de la misma edad sin hijos que sí tenían empleo. Por el contrario, en el caso de las mujeres el tener o no tener empleo parece ser un factor diferenciador antes de los 35 años, pero no tanto después, afectando en mucha menor medida que a los hombres. De hecho, a los 49 años, el porcentaje de mujeres en paro con hijos es superior al de las mujeres ocupadas. Por otro lado cabe reseñar el comportamiento de la categoría “Labores del hogar”. Mientras en el caso de la clasificación ocupado/parado puede presuponerse una cierta causalidad, en el caso de las labores del hogar podría ser a la inversa. Es decir, es posible que el hecho de tener hijos acabe llevando a una persona a dedicarse al cuidado del hogar en lugar del suceso inverso, de ahí que el porcentaje de mujeres que no tienen hijos y se dedican a labores del hogar sea tan bajo. Probablemente en las generaciones venideras este gráfico cambiará sustancialmente. No obstante, al no disponer de información sobre el estado previo antes de tener descendencia, resulta imposible concluir a partir de estos datos cuál es la secuencia temporal.

En el caso del nivel de estudios también se repite el comportamiento diferenciado para hombres y mujeres. Podemos verlo a través del **Gráfico 7**.

En el caso de las mujeres puede apreciarse con nitidez que a mayor nivel de formación, más tardía es la decisión de tener hijos y mayor es el porcentaje de mujeres que finalmente no los tienen.

En el caso de los hombres, en las edades más jóvenes también se repite el mismo patrón, pero finalmente no acaba afectando al porcentaje de hombres que tienen hijos a pesar de tener un calendario diferenciado.

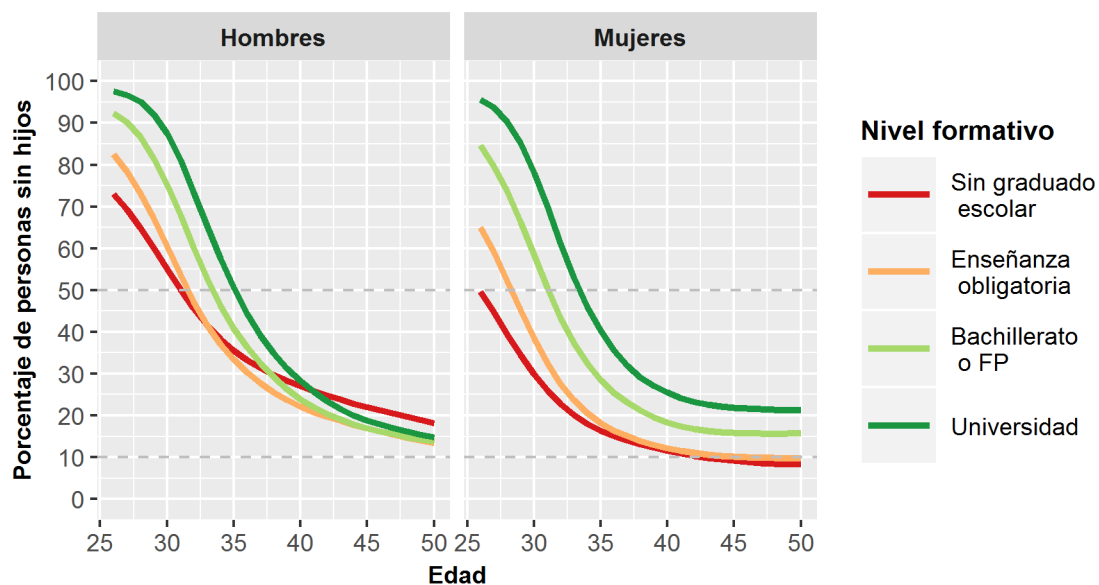


Gráfico 7. Porcentaje de personas (nacidas entre 1950 y 1975) sin hijos según sexo, edad y nivel de estudios.

Si analizamos las probabilidades de aumento de familia (**Gráfico 8**), se observa claramente el desplazamiento del calendario a la hora de tener el primer hijo con el aumento del nivel de estudios. Por otro lado, lo que llama la atención es la probabilidad de tener otro hijo cuando ya se tiene uno previamente. Las probabilidades de ampliación de familia de las personas con estudios universitarios a partir de los 33 años son superiores a las del resto de grupos, muy superiores. El paso de 2 hijos a 3 o más también presenta unas tasas superiores para el colectivo de personas con formación universitaria, aunque no se muestra en el gráfico por una simple cuestión de legibilidad del gráfico completo.

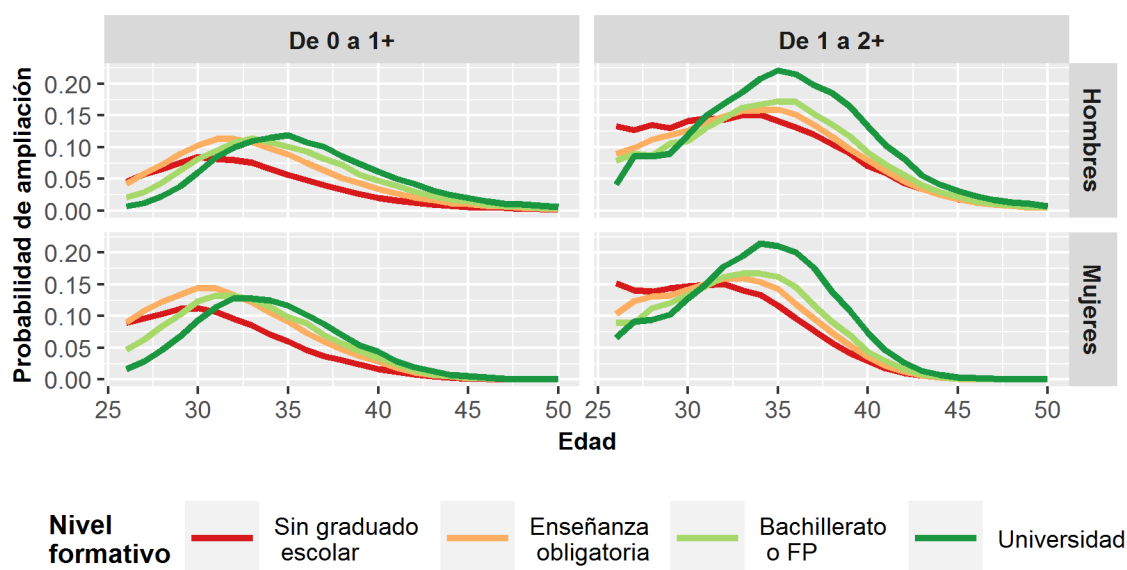


Gráfico 8. Probabilidad de ampliación de familia por sexo y edad, según nivel de estudios. Nacidos entre 1950 y 1975

¿Cómo se ve afectado el número de hijos por estas probabilidades y estas tasas? Para eso recurriremos al Gráfico 9. En él vemos de nuevo el comportamiento claramente diferenciado por nivel de estudios para las mujeres y el mismo patrón en el caso de los hombres, pero de una manera menos manifiesta. Para cualquiera de los niveles de estudios establecidos, el número medio de hijos de un hombre a los 40 años varía desde los 1,53 hijos de personas que no han completado los estudios obligatorios a los 1,37 hijos de los que tienen estudios universitarios. En el caso de las mujeres el rango es mucho más amplio y esta cifra oscila desde 2,03 a 1,43.

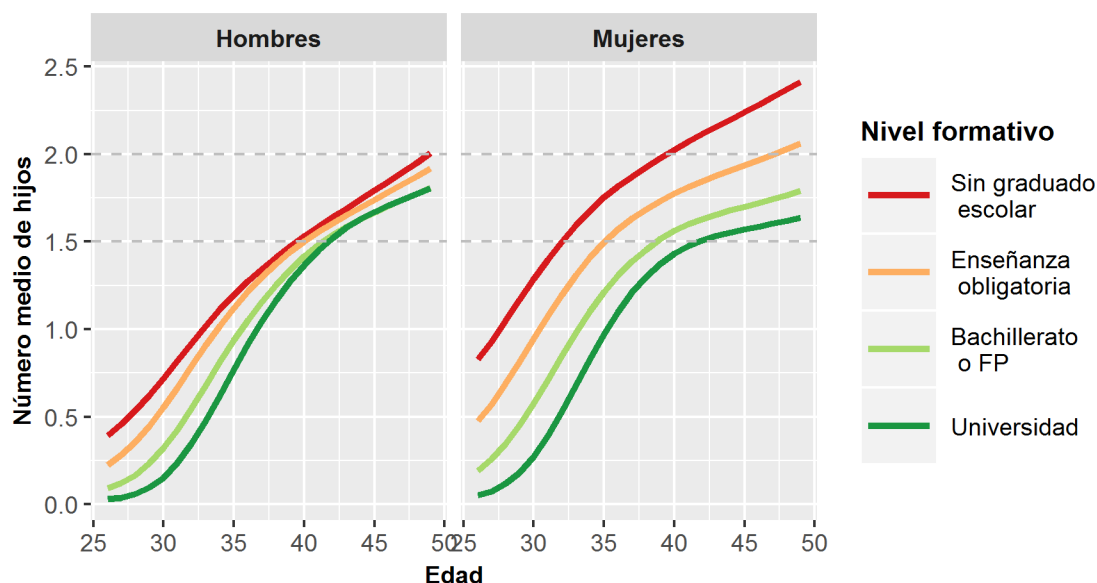


Gráfico 9. Número medio de hijos por sexo y edad, según nivel de estudios. Nacidos entre 1950 y 1975

Ahora bien, como vimos que el porcentaje de personas sin hijos estaba graduado según el nivel de estudios, podemos hacer la misma prueba sobre el número medio de hijos, pero sólo para las personas que ya han tenido hijos a las distintas edades (Gráfico 10). Desde esta perspectiva, la media de hijos de las personas con estudios universitarios (a los 40 años) se iguala o aproxima en ambos sexos, contrariamente a lo que ocurre en el caso de personas sin graduado escolar. Así, el número medio de hijos de personas con estudios universitarios se sitúa en 1,91 hijos para hombres y 1,92 para mujeres, mientras que para el nivel formativo más bajo estas cifras son 2,10 y 2,29 respectivamente.

Por tanto, para las personas con estudios universitarios hay un retraso en el calendario reproductivo, un mayor porcentaje de personas sin hijos (sobre todo en el caso de las mujeres), pero aquellas personas que tienen hijos acaban teniendo tantos hijos (o casi) como el resto de personas que han completado estudios reglados.

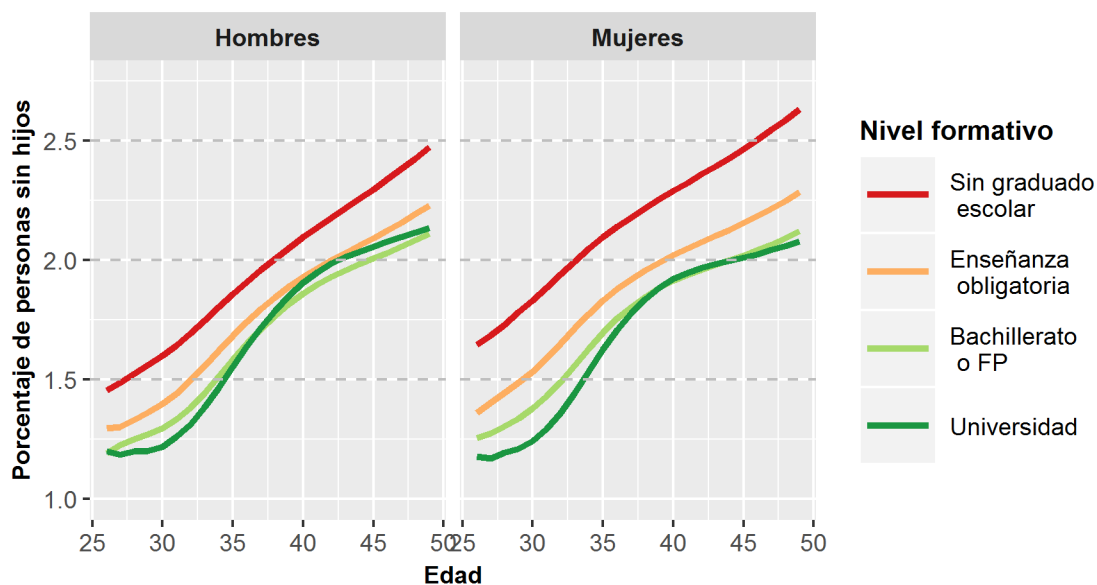


Gráfico 10. Número medio de hijos por sexo y edad, según nivel de estudios. Nacidos entre 1950 y 1975 que han tenido algún hijo.

Ahora bien, cabe señalar que el porcentaje de mujeres con estudios universitarios nacidas en 1950 era del 9,7% según el censo de 2001, mientras que para las nacidas en 1975 esta cifra ya era del 28,0%. Todos estos factores suman y ayudan a entender la caída de la fecundidad.

6 Futuro y conclusiones

Las limitaciones de este tipo de estudios son claras. Las variables tempo-dependientes son muy reducidas y hay que limitarse a tomar puntos de información al principio y al final del estudio. La integración de fuentes administrativas que está llevando el IECA hará que a medio plazo se puedan considerar como tiempo-dependientes la relación con la actividad, la percepción de ingresos regulares o el nivel educativo. Esto enriquecerá futuros estudios y facilitará la obtención de conclusiones sobre la verdadera influencia de los factores sobre la fecundidad. No obstante, datos como los que se muestran en la publicación ayudan a identificar los perfiles de las personas para comprender la denominada crisis de fecundidad.

Con el paso del tiempo también dispondremos del comportamiento completo de algunas generaciones, como las cercanas a 1975, de forma que se podrá corroborar si hay un descenso de la fecundidad, un desplazamiento del calendario o un aumento del número de personas que “deciden” no tener hijos.

Con el paso de los años también se podría observar si el desarrollo del teletrabajo ejerce una influencia sobre la fecundidad y en qué grado afecta.

Lo ideal puede ser ir mezclando resultados a medio plazo, como los publicados, con resultados más a largo plazo que recojan cambios estructurales. De momento ya hemos visto que hay una clara diferenciación, sobre todo en calendario, en el comportamiento de las mujeres respecto al fenómeno en función de su nivel de formación. También hemos visto que la relación con la actividad interacciona de manera muy diferente para hombres y para mujeres. En un futuro también se podrá

estudiar si el comportamiento reproductivo se ve modificado con la aplicación de nuevas políticas relativas a permisos de maternidad y paternidad.

Como ejemplo de estudios que se pueden hacer con base en la información ya disponible, podemos señalar “Parental Support and Transition to Second Birth in a Low-Fertility Setting: The Case of Andalusia in Spain”⁶, presentado por Albert Sabater, en el que se muestran las diferencias a la hora de tener un segundo hijo en función de la distancia a un familiar directo cercano.

En resumen, esta primera publicación facilita una gran cantidad de información con la que estudiar el fenómeno de la fecundidad y los factores que pueden influir en la fecundidad de las generaciones futuras.

⁶ https://risweb.st-andrews.ac.uk/portal/files/253160477/PAA2017_awarded_poster_final.pdf