



Universidade de Vigo

Trabajo Fin de Máster

Análisis temporal y geográfico de la mortalidad evitable en España 1990-2011

Andrés Juanatey Martínez

Máster en Técnicas Estadísticas

Curso 2014-2015

Propuesta de Trabajo Fin de Máster

<p>Título en galego: Análise temporal e xeográfico da mortalidade evitable en España 1990-2011</p>
<p>Título en español: Análisis temporal y geográfico de la mortalidad evitable en España 1990-2011</p>
<p>English title: Temporal and geographical analysis of avoidable mortality in Spain 1990-2011</p>
<p>Modalidad: Modalidad B</p>
<p>Autor/a: Andrés Juanatey Martínez, Universidad de Santiago de Compostela</p>
<p>Director/a: César Andrés Sánchez Sellero, Profesor titular de universidad del Departamento de Estadística e Investigación Operativa de la Universidad de Santiago de Compostela.</p>
<p>Tutor/a: María Isolina Santiago Pérez, Dirección Xeral de Innovación e Xestión da Saúde Pública da Consellería de Sanidade</p>
<p>Breve resumen del trabajo:</p> <p>El objetivo de este TFM es analizar el patrón geográfico y temporal de la mortalidad evitable en España, a través de dos indicadores: tasas de mortalidad evitable y años de esperanza de vida perdidos, para una lista de causas consideradas evitables desde el punto de vista sanitario. Estos indicadores están disponibles por causa de defunción, sexo y trienio, desde 1990 hasta 2011, y por Comunidad Autónoma.</p> <p>Se pretende, además, analizar las diferencias entre las Comunidades Autónomas ajustando el posible efecto de posibles variables confusoras, como pueden ser el gasto dedicado a sanidad.</p>
<p>Recomendaciones:</p>
<p>Otras observaciones:</p>

Don César Andrés Sánchez Sellero, Profesor titular de universidad del Departamento de Estadística e Investigación Operativa de la Universidad de Santiago de Compostela, doña María Isolina Santiago Pérez de la Dirección Xeral de Innovación e Xestión da Saúde Pública da Consellería de Sanidade, informan que el Trabajo Fin de Máster titulado

Análisis temporal y geográfico de la mortalidad evitable en España 1990-2011

fue realizado bajo su dirección por don Andrés Juanatey Martínez para el Máster en Técnicas Estadísticas. Estimando que el trabajo está terminado, dan su conformidad para su presentación y defensa ante un tribunal.

En Santiago de Compostela, a 3 de Julio de 2015.

El/la director/a:

La tutora:

Don César Andrés Sánchez Sellero

Doña María Isolina Santiago Pérez

El/la autor/a:

Don Andrés Juanatey Martínez



Dna. M^a Isolina Santiago Pérez, nomeada pola Dirección Xeral de Innovación e Xestión da Saúde Pública da Consellería de Sanidade para exercer a función de titora de Andrés Juanatey Martínez no Convenio de Cooperación Educativa entre a Consellería de Sanidade e as Universidades de Santiago de Compostela, A Coruña e Vigo para a realización do seu traballo Fin de Máster, certifican que Andrés Juanatey Martínez realizou satisfactoriamente prácticas no servizo de Epidemioloxía da Dirección Xeral de Innovación e Xestión da Saúde Pública entre setembro de 2014 e xaneiro de 2015 co obxectivo de desenvolver o traballo titulado "Análisis temporal y geográfico de la mortalidad evitable en España 1990-2011".

M^a Isolina Santiago Pérez

Santiago de Compostela, 30 de xuño de 2015

Índice general

Resumen	XI
Prefacio	XIII
1. Preliminares	1
1.1. Introducción histórica	1
1.2. Conceptos	2
1.3. Listas de causas de muerte evitables	4
1.3.1. Lista 1: Nolte y Mckee (2004)	4
1.3.2. Lista 2: Gispert y cols.(2006)	4
1.3.3. Lista 3: AMEIHS (2011)	5
1.3.4. Lista 4: Plug in (2012)	5
1.4. Antecedentes de resultados de mortalidad evitable en España	6
2. Métodos de análisis de datos de panel	9
2.1. Conceptos previos	9
2.2. Modelos estáticos	11
2.2.1. Modelo de efectos fijos	11
2.2.2. Modelo de efectos aleatorios	12
2.2.3. Efectos fijos vs Efectos aleatorios	13
2.2.4. Diagnósis y validación de los modelos de efectos fijos	13
2.3. Modelos dinámicos	14
2.3.1. Enfoque de variables instrumentales: estimador de Anderson-Hsiao	15
2.3.2. Método Generalizado de Momentos	16
2.3.3. Consideraciones adicionales: los instrumentos	20
2.3.4. Diagnósis y validación de los modelos dinámicos	21
3. Preparación de la base de datos	23
3.1. Indicadores de mortalidad evitable	23
3.1.1. Fuente de datos	23
3.1.2. Causas de muerte	24
3.1.3. Tasa de mortalidad estandarizada	24
3.1.4. Años de esperanza de vida perdidos (AEVP)	25
3.2. Variables adicionales del modelo	26
4. Resultados	29
4.1. Análisis de la situación inicial	29
4.2. Evolución temporal	33
4.3. Análisis temporal y geográfico	34
4.3.1. Ajuste de modelos de efectos fijos	35
4.3.2. Ajuste de modelos dinámicos de datos de panel	36

A. Listas de causas de mortalidad evitable	51
B. Comparación de las listas	57
C. Evolución temporal	63
Referencias	65

Resumen

Resumen en español

La mortalidad evitable propuesta a mediados de los 1970, es un indicador que permite medir la calidad y efectividad de los servicios sanitarios, así como su evaluación y monitorización. El presente texto profundiza en este concepto, repasando diferentes listas de causas de muerte evitables, así como los antecedentes bibliográficos de estudios de esta índole existentes hasta la fecha en España. Además, se lleva a cabo una revisión de los métodos de análisis de datos de panel, tanto de los modelos estáticos como otros más complejos como son los modelos dinámicos.

Mediante esta metodología, se propone un análisis temporal y geográfico de la mortalidad evitable en España desde 1990 hasta 2011 empleando como variable respuesta distintos indicadores de mortalidad evitable (tasas de mortalidad evitable estandarizadas y años de esperanza de vida perdidos), valorando la influencia de diferentes factores socio-económicos, entre los que destaca el gasto sanitario. Como resultado, se aprecia una gran reducción en este tipo de mortalidad en el periodo considerado, así como un efecto significativo del gasto sanitario y de otras variables, como la tasa de paro. Desde el punto de vista geográfico, Asturias y Canarias destacan como las comunidades que presentan mayor mortalidad evitable debido a características específicas de la zona, mientras que Castilla y León, Castilla-La-Mancha y Madrid son las comunidades autónomas donde se produce menor mortalidad evitable por estos motivos.

English abstract

Avoidable mortality proposed in the mid-1970s, is an indicator to measure the quality and effectiveness of health services, as well as for its evaluation and monitoring. This concept is explored, by reviewing different lists of avoidable causes of death as well as bibliographic records of studies of this kind in Spain. In addition, a review of methods for analyzing panel data, including static and more complex models such as dynamic models is carried out.

Using this methodology, a temporal and geographical analysis of avoidable mortality in Spain from 1990-2011 is proposed, using as the response variable different indicators of avoidable mortality (standardized rates of avoidable mortality and years of life expectancy lost) and evaluating the influence of different socio-economic variables, including health expenditure stands. As a result, a large reduction in this type of mortality in the period considered, as well as a significant effect of health expenditure and other variables such as the unemployment rate is seen. From the geographical point of view, Asturias and Canarias stand out as the regions that have higher avoidable mortality due to specific characteristics of the area, while Castilla y Leon, Castilla-La-Mancha and Madrid are the regions where lower avoidable mortality occurs for these reasons.

Prefacio

Este trabajo corresponde a una memoria sobre las prácticas realizadas en el servicio de epidemiología de la Dirección Xeral de Innovación e Xestión da Saúde Pública (DXIXSP) da Consellería de Sanidade con el objetivo de que se considere como trabajo fin de máster, del "Máster en Técnicas Estadísticas".

En este trabajo se lleva a cabo un análisis temporal y geográfico de la mortalidad evitable en España desde 1990 hasta 2011, empleando métodos de análisis de datos de panel. Para ello, en un primer capítulo se profundiza en el concepto de mortalidad evitable y se repasan los antecedentes bibliográficos existentes hasta la fecha en nuestro país.

Seguidamente, en el Capítulo 2, se expone la metodología estadística empleada para la realización de nuestro estudio, lo que se conoce como modelos de datos de panel. En una primera parte se tratan los modelos más sencillos, los estáticos, para posteriormente hablar sobre modelos dinámicos de datos de panel, cuya principal característica reside en la inclusión de la variable respuesta rezagada como variable explicativa.

El tercer capítulo centra la atención en la descripción de la base de datos a utilizar, así como en el proceso seguido para su obtención. Éste se divide en dos partes, haciendo referencia a los indicadores de mortalidad evitable tratados y al resto de variables involucradas en el estudio respectivamente.

Por último, el Capítulo 4 tiene como finalidad la ilustración de los resultados obtenidos al ajustar los métodos de análisis de datos de panel sobre la base de datos considerada.

Como información adicional, en los apéndices A, B y C se presentan las 4 listas de causas de muerte evitables que se han tenido en cuenta en el presente trabajo, los gráficos relativos a la comparación de las tasas de mortalidad evitable estandarizada de estas listas y la evolución temporal de las variables involucradas en el estudio para cada comunidad autónoma respectivamente.

Capítulo 1

Preliminares

El objetivo principal de los sistemas de salud es mejorar la salud de la población, y para comprobar que se cumple este objetivo es necesario llevar a cabo una evaluación del rendimiento de los mismos. Esta tarea, sin embargo, presenta un gran número de dificultades, algunas de las cuales se exponen a continuación:

- ↯ La evaluación de los niveles y cambios en el estado de salud de las poblaciones es muy compleja, ya que hay multitud de dimensiones dentro de la salud de la población, algunas de ellas muy subjetivas.
- ↯ El estado de salud no depende de manera unilateral de las intervenciones del sistema de salud, existen determinantes no médicos, como riqueza, estilo de vida, factores ambientales, etc, que pueden jugar un papel fundamental.
- ↯ El impacto provocado en el estado de salud por las intervenciones de los sistemas de salud no siempre se sabe con certeza.

En cambio, la información sobre la mortalidad puede ser fácilmente obtenida en la mayoría de países. Por ello, los analistas han utilizado con mayor frecuencia indicadores de mortalidad general, como tasas de mortalidad o esperanza de vida para llevar a cabo la evaluación del rendimiento de los sistemas de salud. Lamentablemente no hay acuerdo unánime acerca de cuál de estos indicadores sería el adecuado para tal fin.

1.1. Introducción histórica

La utilización de la mortalidad por causas específicas para monitorizar la calidad asistencial nace en el siglo XX. Comienza en el Reino Unido en 1928, donde se realizaron investigaciones sobre la mortalidad materna con el objetivo de identificar los errores y las áreas en las que se podrían realizar mejoras, y así poder evitar muertes innecesarias. En los Estados Unidos, se llevaron a cabo estudios similares en 1930, lo que conllevó reducciones importantes en las tasas de mortalidad materna.

En la década de los 70, nace el concepto de mortalidad evitable, introducido por un grupo de clínicos liderado por David Rutstein (Working Group on Preventable and Manageable diseases). Este grupo esbozó un método de medición de la calidad de los cuidados médicos, basado en contar los casos de enfermedades innecesarias, discapacidad y muertes prematuras. Construyeron una lista de 90 condiciones que denominaron 'eventos centinela de la salud', pues suponían que su presencia era una señal de alerta para el seguimiento de la atención médica, ya que los casos de enfermedad, discapacidad o muerte en estas condiciones se consideran prevenibles y/o tratables mediante atención médica

adecuada y oportuna (Rutstein y cols, 1976).

A partir de este trabajo, numerosos autores de diferentes países investigaron la posibilidad de utilizar la mortalidad evitable como indicador para monitorizar el proceso y los resultados de la atención sanitaria. Charlton y colaboradores (Charlton, Harley, Silvera, y Holland, 1983) estrecharon el concepto de mortalidad evitable mediante la exclusión de condiciones que consideraban fuera del ámbito del cuidado médico, como accidentes de tráfico, tabaquismo, etc. Fueron los primeros en emplear la mortalidad evitable en un estudio empírico con datos procedentes de Inglaterra y Gales y en examinar las tendencias nacionales e internacionales, así como la importancia de la incidencia de la enfermedad y de los factores sociales. Al mismo tiempo, introdujeron un límite máximo de edad para algunas condiciones, que se impuso en 65 años.

Posteriormente, en 1986, surge el interés en estudiar la distribución geográfica de la mortalidad evitable con el inicio de un gran proyecto en la Comunidad Europea, que se tradujo en la publicación del Atlas de Mortalidad Evitable de la Comunidad Europea (Holland, 1988, 1991, 1993, 1997). Este proyecto extiende el trabajo de Charlton y colaboradores mediante la presentación de una lista de causas de mortalidad evitable agrupadas en 'enfermedades fundamentalmente sensibles a prevención secundaria o tratamiento médico' y 'enfermedades fundamentalmente sensibles a prevención primaria' y su propuesta como indicadores de asistencia médica y de política nacional de salud.

En esta misma década (la de 1980), Mackenbach y colaboradores analizaron la posible contribución de las innovaciones desarrolladas en la atención médica a los cambios producidos en la mortalidad mediante el análisis de tendencias en la mortalidad por enfermedades específicas en Holanda. El resultado fue que aunque dicha contribución no se puede determinar, el impacto en la mortalidad podría haber sido sustancial (Mackenbach, Looman, Kunst, Habbema, & van der Maas, 1988). Emplearon una definición más estricta de la atención médica y se centraron explícitamente en la diferenciación y comparación de los niveles de mortalidad evitable atribuible al sistema de salud y a las políticas de salud más amplias (por lo general no incluidas dentro del control directo de los servicios de salud).

Más adelante, Tobias y Jackson, después de un ejercicio de consenso de expertos en Nueva Zelanda, repartieron el relativo carácter evitable de la muerte por las diferentes condiciones en proporciones, dependiendo de si son evitables por acción primaria, secundaria o terciaria (Tobias y Jackson, 2001). El concepto 'evitable' se amplió para no cubrir únicamente causas de muerte susceptibles de intervención terapéutica sino también las que responden a una intervención preventiva individual y basada en la población. En 2004, Nolte y McKee examinan la mortalidad 'evitable' y el cambio de la esperanza de vida en la Unión Europea entre los años 1980 y 1990 utilizando una lista actualizada de las condiciones, teniendo en cuenta los avances en el conocimiento médico y en la tecnología (Nolte y McKee, 2004).

En España, la mortalidad evitable se introdujo en 1988 bajo el acrónimo de MIPSE (mortalidad innecesariamente prematura y sanitariamente evitable)(Ortún y Gispert, 1988). Con la introducción de las técnicas de representación y análisis espacial en el ámbito de la epidemiología, durante el siglo XX, como herramienta para explicar los patrones geográficos de distribución de enfermedades y detección de factores causales, en 1989 el Ministerio de Sanidad y Consumo publicó el primer Atlas de Mortalidad Evitable en España, con su distribución por comunidades autónomas.

1.2. Conceptos

Una vez realizada esta pequeña introducción histórica, conviene profundizar en el concepto de mortalidad evitable, pues será el tema principal del texto que nos ocupa.

La mortalidad evitable es un importante problema de salud pública, y ha sido estudiado en un gran

número de países. Este concepto se puede emplear para supervisar y evaluar la calidad de la asistencia sanitaria y las políticas de salud, tanto geográficamente como en el tiempo.

Los investigadores de esta rama han conceptualizado este tipo de mortalidad de varias maneras ligeramente diferentes, teniéndose de forma común ciertas consideraciones:

- La mortalidad evitable designa muertes por condiciones particulares a ciertas edades, las cuales no deberían ocurrir en la presencia de atención sanitaria oportuna y eficaz u otras intervenciones apropiadas.
- Se divide en dos sub-categorías bien diferenciadas, las conocidas como mortalidad tratable y mortalidad prevenible. La suma de estas abarca la mortalidad evitable.

⇒ **Mortalidad tratable:** se refiere a muertes que podrían evitarse mediante atención sanitaria oportuna y eficaz, es decir, a través de los esfuerzos de los servicios de atención de la salud. Esto no implica que el cuidado sanitario pueda impedir necesariamente la aparición de las enfermedades pertinentes, sino que puede prevenir la muerte en la mayoría de los casos.

⇒ **Mortalidad prevenible:** abarca las muertes que podrían impedirse mediante intervenciones más amplias, como podría ser la prevención de accidentes. Wheller et al. (2007) definió esta mortalidad en términos de causas que se consideraron prevenibles a través de:

- (a) el comportamiento individual, o bien
- (b) las medidas de salud pública que limitan la exposición individual a sustancias o condiciones peligrosas.

Las muertes por estas causas son evitables mediante la prevención de la enfermedad, o de un evento externo.

Sin embargo, no todas las muertes por causas consideradas evitables se pueden evitar en realidad. Por ejemplo, las provocadas por causas 'tratables' pueden ser inevitables debido a un diagnóstico tardío o a problemas de salud concurrentes, mientras que algunas muertes por causas 'prevenibles' se dan como consecuencia de eventos impredecibles y contra los que no podrían haber sido previstas medidas de protección posibles.

En resumen, existen multitud de definiciones diferentes tanto de mortalidad evitable, como de sus categorías. Por ello, y con la finalidad de aclarar estos conceptos, la ONS (*Office for National Statistics*) propone las definiciones expuestas a continuación:

- **Mortalidad tratable**

Una muerte es tratable si, a la luz de los conocimientos médicos actuales y la tecnología, la totalidad o la mayoría de las muertes por esa causa (sujeto a los límites de edad en su caso) podrían evitarse a través de una asistencia sanitaria de calidad.

- **Mortalidad prevenible**

Una muerte es prevenible si, a la luz de los conocimientos actuales de los determinantes de la salud, la totalidad o la mayoría de las muertes por esa causa (sujeto a los límites de edad en su caso) podría evitarse mediante intervenciones de salud pública en el sentido más amplio.

- **Mortalidad evitable**

Muertes evitables son todas aquellas definidas como prevenibles, tratables, o ambas. Cuando una causa de muerte se inscribe tanto en la definición de prevenible como en la de tratable, las muertes por esa causa se contabilizan en ambas categorías.

1.3. Listas de causas de muerte evitables

Llegados a este punto, procede abordar el tema relativo a las listas de causas de muerte seleccionadas como evitables, donde no se ha hallado un acuerdo unánime. A la hora de construir este tipo de listas, se deben tener en cuenta varios factores relevantes.

Se debe valorar qué categorías de mortalidad evitable abarcarán las causas de muerte que componen la lista. En algunos casos se consideran ambas subcategorías de mortalidad evitable, llegando incluso a su diferenciación dentro de la misma lista, mientras que en otras listas solo están incluidas las causas consideradas susceptibles a la intervención médica (tratables).

Otra consideración cuando se elabora una lista de esta índole es el límite de edad (si los hay) a valorar para cada causa. Si se realiza la selección de un límite máximo de edad para todas las causas de muerte, hay que tener en cuenta que es esencialmente arbitrario y no debe interpretarse en el sentido de que algunas muertes en personas de mayor edad que este límite no podrían también ser evitadas. Por ejemplo, las muertes por accidentes de tráfico deben ser evitables en todas las edades. En nuestro caso trataremos con 4 listas de causas de muerte evitable diferentes.

Lista 1: Nolte y Mckee (2004)

Lista 2: Gispert y cols. (2006)

Lista 3: AMEIHs (2011)

Lista 4: Plug in (2012)

A continuación, describiremos de forma somera cada una de las listas.

1.3.1. Lista 1: Nolte y Mckee (2004)

La selección de las 34 causas de muerte evitables que incluye esta lista, todas ellas susceptibles al cuidado sanitario, se deriva de una revisión previa de mortalidad evitable. La información detallada de esta selección está en la citada revisión, pero la lista final empleaba una modificación del trabajo de Tobias y Jackson (2001), quienes actualizaron un trabajo anterior de Mackenbach y Charlton y cols. (1988).

Se introdujo un límite superior de edad de 75 años, exceptuando la edad límite para la diabetes mellitus (donde se fijó una edad límite de 50 años) debido a que el carácter de las muertes a edades más avanzadas sigue generando controversia. Para otras causas, la edad límite fijada fue de 15 años, ya que las muertes relacionadas con estas causas, de no producirse en la infancia, es probable que reflejen la existencia de alguna otra enfermedad. El límite de edad de 75 años es arbitrario, pero este valor es consistente con la esperanza de vida al nacer en muchos países industrializados.

1.3.2. Lista 2: Gispert y cols.(2006)

Esta lista es el resultado de un proceso de consenso entre profesionales médicos con tareas asistenciales, elaboradores de las estadísticas de mortalidad y usuarios de los indicadores sanitarios. El objetivo de este proceso era la obtención de una lista de causas de mortalidad evitable que pudiera ser de uso común en España.

Se pretendía obtener un indicador para monitorizar el proceso de atención en los servicios sanitarios utilizable en la comparación entre distintas áreas geográficas, con sistemas sanitarios y organizaciones de salud diversas.

Para ello, se llevó a cabo una búsqueda bibliográfica de diferentes fuentes de datos. Los principales investigadores analizaron las distintas listas de causas de muerte evitables presentes en la bibliografía y seleccionaron las más consistentes entre sí, quedándose finalmente con las categorías de causas que aparecían con mayor frecuencia.

El resultado de este proceso fue una lista de partida con 37 categorías de enfermedades y causas externas. A partir de las opiniones de los expertos, y tras tres rondas de consulta en la que se mantuvieron algunas categorías, se modificaron otras y se introdujeron otras, se alcanzó el consenso. La lista final consta de 34 categorías:

27 susceptibles a intervenciones de política sanitaria intersectorial

7 susceptibles a intervenciones de los servicios de asistencia sanitaria

Resumiendo, esta lista es fruto de un procedimiento de consenso entre diversos expertos clínicos y de salud pública, con el objetivo de seleccionar una lista de enfermedades que, de acuerdo con los avances de la medicina actual, no deberían ser causa de muerte.

1.3.3. Lista 3: AMEIHS (2011)

Se trata de una lista de 45 causas seleccionadas por el proyecto europeo AMIEHS (*Amenable Mortality in the European Union: towards better Indicators for the Effectiveness of Health Systems*). Este proyecto elaboró un atlas electrónico de mortalidad evitable en el que se presentan tasas ajustadas de mortalidad por las 45 causas de la lista en 31 países europeos, entre los que se encuentra España, y en el período 2001-2009.

Este proyecto tenía como objetivo generar una lista de causas o condiciones de mortalidad evitable (en concreto tratable) que puedan estar vinculadas a una determinada intervención, la cual es esperable que reduzca la mortalidad en un 30 % o más.

Como primer paso se realizó una búsqueda sistemática para identificar las causas de muerte que habían disminuido en un 30 % o más (seleccionado como reducción significativa de la mortalidad) entre 1979 y 2000, en Inglaterra y Gales (ya que esto coincidió con el período en que la CIE-9 estaba en uso). Se excluyeron las causas de muerte para las que el número de muertes fue menos de 100 en el año 2000, ya que en las comparaciones realizadas en países mucho más pequeños carecería de utilidad.

De estas, se seleccionaron aquellas para las cuales es posible encontrar intervenciones que pueden reducir las muertes y para las que la variación en las tasas de mortalidad (en el tiempo o entre poblaciones) refleja la variación en la cobertura y calidad de la atención sanitaria (servicios preventivos o terapéuticos prestados a las personas). Sin embargo, ya que estos criterios podrían considerarse demasiado estrictos, se incluyeron otras condiciones que no cumplían algunas de estas premisas, como por ejemplo causas para las que la mortalidad había disminuido entre 1979 y 2000 entre un 20 y un 30 %.

Finalmente, y tras este proceso, se obtuvo una lista de 24 causas, que junto a otras 21 causas de muerte extra empleadas en listas anteriores, forman la lista final considerada.

1.3.4. Lista 4: Plug in (2012)

Se trata de una lista de causas de muerte susceptibles a la intervención médica basada en la lista original desarrollada por Rutstein y cols. (1976) a la que se añadieron unas pocas causas adicionales basándose en mejoras recientes.

La lista final consta de 15 causas de muerte evitable por intervención médica.

1.4. Antecedentes de resultados de mortalidad evitable en España

En vistas a finalizar este primer capítulo, se expondrá de forma somera una revisión de los principales estudios acerca de la mortalidad evitable llevados a cabo en España hasta la fecha. Los Cuadros 1.1 y 1.2 resumen, para cada uno de estos estudios, algunas de sus principales características:

- El periodo de estudio.
- El ámbito geográfico.
- La lista de causas de mortalidad evitable considerada.
- El indicador de mortalidad evitable empleado.

Como se evidencia en los cuadros 1.1 y 1.2, la lista de causas de mortalidad evitable empleada difiere de unos estudios a otros. En consonancia con esto, Albert (1996) realiza la comparación de dos listas distintas, concluyendo en que la elección de la misma tiene una influencia notable en los resultados. Esta se debe, en gran parte, a la presencia de causas prevenibles en únicamente una de las listas. En cuanto al ámbito geográfico, a lo largo de los últimos años se han llevado a cabo estudios relativos a la totalidad del territorio español, así como otros centrados en una comunidad autónoma, concretamente en la Comunidad Valenciana y Cataluña.

A la hora de analizar la evolución temporal de la mortalidad evitable, todos los estudios coinciden en la presencia de un descenso significativo en el número de muertes producidas por causas evitables. Además, la gran mayoría de los mismos, resaltan el hecho de que el citado descenso es considerablemente mayor que el producido en la mortalidad general (en cada periodo de estudio), así como la mayor incidencia de casos de muertes por causas evitables en hombres que en mujeres.

En relación a la distribución geográfica, Albert (1996) y Barés (2005) indican que la mortalidad evitable se distribuye de manera bastante homogénea en general (en la Comunidad Valenciana y Cataluña respectivamente). Albert (1996) añade además que la mortalidad evitable es mayor en los núcleos urbanos, encontrándose los picos de la misma en Valencia y Alicante, y relaciona los mismos a una posible presencia de bolsas de pobreza y marginación social.

Por otra parte, Gispert (2006), Vergara (2008) y Barrasa (2013) analizaron la distribución geográfica de la mortalidad evitable en el territorio español por provincias, municipios y comunidades autónomas respectivamente. El resultado de estos estudios, pone de manifiesto una mayor mortalidad evitable en las islas y en el suroeste peninsular. En concreto, Barrasa (2013) resalta las comunidades autónomas de Andalucía y Canarias como las comunidades donde tuvieron lugar un mayor número de muertes provocadas por causas evitables, y a Madrid y Navarra como las comunidades donde se produjeron menos muertes de este corte.

Por último, Nolasco (2009) valoró la relación existente entre el nivel de riqueza de la población y la mortalidad evitable. Con este objetivo, dividió la población de Alicante, Castellón y Valencia en 3 niveles de bienestar, y estudió la mortalidad evitable en cada uno de ellos. El resultado fue la existencia de una relación inversa entre la riqueza de la población y la mortalidad evitable en los hombres; sin embargo, dicha relación no se encontró para las mujeres, en donde no se identificó patrón alguno.

Referencia	Periodo	Ambito Geográfico	Lista	Indicador de Mortalidad Evitable
Albert (1996)	1975-1990 Análisis anual	Comunidad Valenciana.	Comparan las listas de Holland (1991) y Charlton (1983).	Tasas de mortalidad ajustadas en el grupo de edad 5-64 años.
Barés (2005)	1986-2001	Comunidad de Cataluña, por sectores sanitarios (46)	Lista del Registro de Mortalidad de Cataluña, basada en la de Holland (1991).	Tasas de mortalidad ajustadas para analizar la evolución temporal. Índices de mortalidad estandarizados (método indirecto) y la razón de mortalidad comparativa para los dos grupos de causas (tratables y prevenibles) para estudiar la distribución geográfica.
Gispert (2006)	Dos periodos: 1986-1993, 1994-2001	Provincias españolas (52)	Gispert et al(2006)	Tasas de mortalidad ajustadas por edad para la evolución temporal. Índices de mortalidad estandarizados (IME) por el método indirecto, y la razón de mortalidad comparativa (estandarizada por el método directo) para analizar el cambio temporal en la distribución geográfica.
Gispert (2008)	Tres periodos: 1987-91, 1992-96, 1997-2001	España	Gispert(2006)	Descomposición de diferencias en la esperanza de vida al nacer, con 85+ como último grupo de edad, por el método de Arriaga.

Cuadro 1.1: Resumen de las principales características de los principales estudios sobre mortalidad evitable en España existentes hasta la fecha (I).

Referencia	Periodo	Ambito Geográfico	Lista	Indicador de Mortalidad Evitable
Melchor (2008)	1990-2004	Comunidad Valenciana: por departamentos	Holland (1991) modificada	Tasa de mortalidad ajustada por edad y la razón comparativa de mortalidad, para cada sexo, periodo y departamento, tomando como población estándar la del periodo 1990-1994 de cada ámbito geográfico y sexo. Razón de mortalidad estandarizada (método indirecto), tomando como estándar la mortalidad del conjunto de la Comunidad Valenciana para cada periodo y sexo
Nolasco (2009)	Tres trienios (1996-1998, 1999-2001 y 2002-2003)	Alicante, Castellón y Valencia	Nolte y McKee (2004)	Tasas de mortalidad ajustadas por edad.
Vergara (2009)	1990-2001	2.218 áreas pequeñas formadas por municipios o municipios agregados de todo el territorio español	Gispert y cols (2006)	Riesgo relativo de mortalidad ajustado por edad en cada área pequeña
Barrasa (2013)	Dos trienios: 1999-2001 y 2006-2008	España de manera global y por comunidades autónomas	Nolte y McKee (2004)	Tasas de mortalidad ajustadas por edad y sexo.
Pérez (2014)	1981-2007	España	Lista basada en las utilizadas en Simonato (1998), Tobias (2001) y James (2006)	Tasas de mortalidad ajustadas por edad, para cada sexo y el total de causas evitables.

Cuadro 1.2: Resumen de las principales características de los principales estudios sobre mortalidad evitable en España existentes hasta la fecha (II).

Capítulo 2

Métodos de análisis de datos de panel

En este capítulo se expondrá la metodología estadística que se empleará en la realización del presente proyecto, el ajuste de lo que se conocen como modelos de datos de panel. Para ello, conviene introducir en primera instancia el concepto de datos de panel.

Se dice que estamos en presencia de un conjunto de datos de panel cuando se dispone, simultáneamente, de información de corte transversal y de serie temporal. Esta situación se da cuando se tienen observaciones sobre determinadas características de un conjunto de agentes (individuos, países, empresas, etc.) a lo largo de un período continuado de tiempo. La recogida de información se lleva a cabo, por tanto, en dos dimensiones y de este modo se generan múltiples observaciones puntuales para cada unidad. En consonancia con esto, el término modelo de datos de panel se aplica en este contexto a aquel modelo de regresión que utiliza, para la estimación de los parámetros de interés, la variabilidad temporal y transversal de los datos.

El ajuste de este tipo de modelos será el tema principal del presente capítulo. En un primer acercamiento, y a modo de introducción, repasaremos los casos más básicos de este tipo de modelos, denominados modelos de datos de panel estáticos. Dentro de los mismos valoraremos los dos principales enfoques de estimación existentes, efectos fijos y efectos aleatorios y la dicotomía existente entre ellos. Este tipo de técnica presenta una serie de carencias, dentro de las cuales es resaltable la incapacidad de tratar adecuadamente el problema de endogeneidad, por lo que no es posible analizar desde una perspectiva evolucionista la dependencia del pasado o el proceso acumulativo. Para subsanar este problema, pasaremos a tratar el ajuste de lo que se conoce como modelos de panel dinámicos, los cuales permiten incorporar en el modelo una estructura endógena, mediante la integración de efectos pasados, lo que exigirá la utilización de nuevos métodos de estimación. Este tipo de modelos serán el centro de atención del presente capítulo, pues la base de datos valorada en el presente texto requerirá de la utilización de los mismos.

2.1. Conceptos previos

En la presente sección, se introducirán algunos conceptos necesarios para abordar el análisis de este tipo de modelos, y que ayudarán a la mejor comprensión de los mismos.

En primer lugar, conviene hablar del concepto de endogeneidad, que se define como la existencia de correlación entre las variables y el término de error del modelo. Concretamente, la endogeneidad causada por la relación bidireccional entre el fenómeno a explicar y las variables explicativas consideradas es el principal problema a resolver mediante la introducción de los modelos dinámicos de datos de panel. A efectos prácticos podemos considerar dos tipos de endogeneidad:

Endogeneidad del modelo

Endogeneidad de las variables independientes

La primera de ellas es provocada por el efecto de la variable dependiente sobre sí misma, es decir, el modelo está determinado por su pasado. Por otra parte, la endogeneidad de las variables independientes es producida por el hecho de que las variables estén determinadas por su condición pasada. Además, también podría ocurrir que exista relación entre las variables independientes, lo que se reflejaría mediante un coeficiente de correlación alto entre ellas, que provocaría multicolinealidad.

Seguidamente, pasaremos a definir los diferentes tipos de variables a considerar en los diferentes modelos de datos de panel que introduciremos a lo largo del presente capítulo.

- Variable exógena: es aquella que no guarda relación con los errores del modelo, es decir,

$$Cov(X_1, \varepsilon) = 0$$

siendo X_1 la variable exógena y ε los errores del modelo.

- Variable endógena: se encuentra determinada dentro del modelo, es decir existe causalidad en ambos sentidos, ya que son influidas por, y pueden influir en, otras variables. Cumplen

$$Cov(X_1, \varepsilon) \neq 0$$

siendo X_1 la variable endógena y ε los errores del modelo.

- Variable predeterminada: no depende de valores presentes o futuros del error del modelo, pero puede estar relacionada con valores retardados de este, es decir,

$$cov(X_{1t}, \varepsilon_{t+l}) = 0 \quad \forall l \geq 0$$

siendo X_1 la variable predeterminada, ε los errores del modelo.

- Variable instrumental: son proxy de variables que no podemos observar o que no podemos utilizar directamente por diferentes problemas, como la falta de datos, mala calidad de la información, imposibilidad de cuantificarlas o series incompletas.

Una variable instrumental puede ser una proxy o la misma variable retardada, expresada en diferencias o niveles. Este tipo de variables tendrán un papel fundamental a la hora de estimar modelos dinámicos, pues serán utilizadas como instrumentos en el proceso de estimación, esto es, sustituirán en el citado proceso a aquellas variables endógenas o que causan endogeneidad en el modelo, con la finalidad de corregir los problemas derivados de la utilización de las mismas. Un instrumento o conjunto de instrumentos (Z) deben cumplir las siguientes propiedades:

Debe estar correlacionado con la variable independiente a instrumentalizar (X_i)

$$Cov(X_1, Z) \neq 0$$

No puede estar correlacionados con los errores del modelo

$$Cov(Z, \varepsilon) = 0$$

Cabe comentar, que una variable X_i presente en el modelo puede tener varios instrumentos válidos.

2.2. Modelos estáticos

En esta sección, introduciremos de forma somera el primer tipo de modelos de datos de panel que trataremos, conocidos como estáticos. El contexto de este tipo de modelos de datos de panel reside en una regresión lineal con efectos individuales de la forma

$$y_{it} = \mu_i + x'_{it}\beta + \varepsilon_{it} \quad i = 1, \dots, N; \quad t = 1, \dots, T \quad (2.1)$$

siendo

i y t subíndices que indican la unidad de sección cruzada y los diferentes periodos de tiempo considerados, respectivamente

y_{it} la variable dependiente o explicada

x_{it} un vector de $k \times 1$ variables independientes o explicativas

β el vector de coeficientes a estimar

μ_i los efectos individuales

ε_{it} el término de perturbación.

Nótese que, si el modelo incluye efectos temporales, éstos se supondrán incluidos en β , en cuyo caso x_{it} contiene las correspondientes variables ficticias de tiempo.

Existen dos grandes enfoques a la hora de estimar este tipo de modelos: el enfoque de efectos fijos y el de efectos aleatorios. La diferencia entre ambos no reside en la morfología del modelo, ya que esta es siempre la misma, sino que radica en el tratamiento de los efectos individuales μ_i . A continuación, expondremos brevemente ambas metodologías.

2.2.1. Modelo de efectos fijos

El denominado modelo de efectos fijos considera los efectos individuales μ_i como términos constantes específicos de cada grupo. En la formulación de este modelo se supone que las diferencias entre unidades pueden ser captadas a través de las diferencias en dicho término constante. Por tanto, a la hora de llevar a cabo la estimación, los μ_i serán tratados como un conjunto de N coeficientes adicionales que se pueden estimar junto con β .

Sea el modelo (2.1), con $\mu_i = \mu + v_i$, por lo que:

$$y_{it} = \mu + x'_{it}\beta + v_i + \varepsilon_{it} \quad (2.2)$$

Mediante esta operación, se está suponiendo que el error puede descomponerse en una parte fija, constante para cada individuo (v_i) y otra aleatoria (ε_{it}), lo que es equivalente a obtener una tendencia general en regresión dando a cada individuo un punto de origen (ordenadas) distinto. En este punto, existen dos prácticas habituales de estimación de esta índole.

La primera de ellas consiste en la introducción de una variable dummy por cada individuo, y estimar mediante Mínimos Cuadrados Ordinarios. La otra, más habitual, parte del cálculo de la media muestral de todas las variables para cada individuo. Una vez calculado, se tiene que:

$$\bar{y}_i = \mu + \bar{x}'_i\beta + v_i + \bar{\varepsilon}_i \quad (2.3)$$

Restando esta ecuación a la de partida (2.1), se obtiene

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)' \beta + \varepsilon_{it} - \bar{\varepsilon}_i \quad (2.4)$$

A partir de (2.4), y mediante Mínimos Cuadrados Ordinarios, se obtiene la estimación de los coeficientes β , y posteriormente se calcula el estimador de μ_i como sigue:

$$\hat{\mu}_i = \bar{y}_i - \bar{x}_i' \hat{\beta}$$

Este tipo de estimación, conocida como intragrupos, presenta la importante limitación de que las interpretaciones obtenidas pueden ser aplicables a los grupos considerados en el estudio, pero no a otros adicionales no incluidos en la muestra. Por ello, cuando los grupos considerados en un estudio de esta índole son extracciones de una población más grande, puede resultar más apropiado considerar que los términos constantes específicos de cada unidad están aleatoriamente distribuidos entre los grupos. Esto es lo que se conoce como el enfoque de efectos aleatorios, que introduciremos a continuación.

2.2.2. Modelo de efectos aleatorios

El modelo de efectos aleatorios tiene la misma especificación que el de efectos fijos con la salvedad de que en este caso los efectos individuales son variables aleatorias de media v_i y varianza $Var(v_i) \neq 0$. Además, estas estarán integradas en la perturbación aleatoria (v_{it}) de modo que:

$$\tau_{it} = v_i + \epsilon_{it}$$

Por esta razón a estos modelos se les llama también modelos de componentes del error (*error components models*).

La estimación de los parámetros de un modelo de efectos aleatorios no requiere, como en el caso de efectos fijos, diferencias, desviaciones o cualquier otra transformación que elimine la presencia del efecto fijo μ_i . Al contrario, junto con la estimación de los parámetros μ y β interesa la estimación diferenciada de la varianza de los efectos aleatorios σ_v dentro de la estimación global de la varianza de la perturbación aleatoria, por lo que no conviene que el efecto v_i sea obviado en el procedimiento de estimación. Llegados a este punto, podría pensarse en utilizar el estimador simple por Mínimos Cuadrados Ordinarios sobre el modelo en niveles:

$$y_{it} = \mu + x_{it}'\beta + (v_i + \epsilon_{it}) \quad (2.5)$$

Sin embargo, esta estrategia de estimación no resulta válida ya que la presencia de un efecto temporalmente constante en la perturbación aleatoria v_i provoca autocorrelación residual y modifica la expresión tradicional de la varianza de la perturbación aleatoria. Así, la matriz de varianzas y covarianzas de la perturbación queda:

$$\Omega = \begin{pmatrix} \sigma_v^2 + \sigma_\epsilon^2 & \sigma_v^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma_\epsilon^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \ddots & \vdots \\ \sigma_v^2 & \sigma_v^2 & \dots & \sigma_v^2 + \sigma_\epsilon^2 \end{pmatrix} = \sigma_v^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{pmatrix}$$

con

$$\rho = \frac{\sigma_v^2}{\sigma_\tau^2}$$

Por tanto, la estrategia correcta consiste en usar Mínimos Cuadrados Generalizados (MCG) sobre el modelo en niveles utilizando una estimación adecuada de la anterior matriz Ω . A este estimador se le denomina estimador Balestra- Nerlove o Estimador Entre Grupos, nombre este último que proviene de la equivalencia entre estos resultados y los que se obtendrían planteando la estimación MCO entre

las medias grupales de y y x .

La estimación Balestra - Nerlove requerirá por tanto necesariamente dos etapas:

1. Estimación de un modelo previo del que utilizar los residuos para aproximar σ_v^2 y σ_τ^2 . Normalmente suele utilizarse, o bien una estimación MCO simple del modelo, con término independiente μ común, o bien una estimación del modelo con efectos fijos.
2. Estimación de las varianzas σ_v^2 y σ_τ^2 y aplicación del estimador MCG ¹.

2.2.3. Efectos fijos vs Efectos aleatorios

La elección de un modelo de efectos aleatorios frente a uno de efectos fijos no siempre se nos muestra de forma obvia. De hecho, una costumbre muy extendida en el trabajo aplicado consiste en estimar ambos modelos para a continuación contrastar si los efectos son fijos o aleatorios. Sin embargo, el que los efectos se supongan fijos o aleatorios no representa una cualidad intrínseca de la especificación. La distinción crucial reside en si los efectos están correlacionados o no con las variables observables x_{it} . Si μ_i está correlacionado con x_{it} es conveniente hacer inferencia condicional sobre las realizaciones de los μ_i en la muestra (efectos fijos), mientras que si no están correlacionados, es natural realizar inferencia incondicional como se describe en el modelo de efectos aleatorios.

En consonancia con esto, procede introducir el denominado contraste de hipótesis de Wu-Hausman (1978), que se utiliza para analizar la posible relación de los μ_i con los regresores y poder así tomar una decisión entre efectos fijos y aleatorios. Bajo hipótesis nula $H_0 : E[\mu_i|x_{it}] = 0$, el estadístico

$$Q_{FE,RE} = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' (\hat{\sigma}_{\hat{\beta}_{FE}}^2 - \hat{\sigma}_{\hat{\beta}_{RE}}^2)^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \quad (2.6)$$

converge a una distribución χ_{NT}^2 .

En el presente texto, nos centraremos en modelos de efectos fijos, pues la base de datos de la que se dispone requiere del empleo de un enfoque de esta índole. Por ello, en la siguiente sección trataremos la diagnosis y validación de este tipo de modelos.

2.2.4. Diagnosis y validación de los modelos de efectos fijos

Cuando afrontamos el ajuste de un modelo de datos de panel mediante la utilización del estimador intragrupos o de efectos fijos, estamos asumiendo que los errores ε_{it} sean independientes e idénticamente distribuidos. Desafortunadamente, tal asunción será violada con frecuencia en este tipo de ajustes. El incumplimiento de esta condición puede estar derivado de tres posibles factores:

Correlación contemporánea: los errores de diferentes unidades se encuentran correlacionados.

Correlación serial: los errores dentro de cada unidad se correlacionan.

Heterocedasticidad: varianza no constante.

¹Mínimos Cuadrados Generalizados (MCG) es una técnica para la estimación de los parámetros desconocidos en un modelo de regresión lineal. El MCG se aplica cuando las varianzas de las observaciones son desiguales, es decir, cuando hay presencia de heterocedasticidad, o cuando existe un cierto grado de correlación entre las observaciones. Partiendo de un modelo de la forma $Y = X\beta + \varepsilon$ y siendo $E[X|\varepsilon] = 0$ y $Var[X|\varepsilon] = \Omega$, se tiene que β es un vector de coeficientes de regresión desconocidos que deben ser estimados a partir de los datos. Sea b un candidato para la estimación β , el residual de vector para b será $Y - Xb$. El método MCG estima β reduciendo al mínimo el cuadrado de Mahalanobis longitud de este vector residual, es decir, $\hat{\beta} = \arg_b \min(Y - Xb)\Omega^{-1}(Y - Xb)$. Dado que el objetivo es una forma cuadrática en b , el estimador tiene la forma explícita $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$

En esta línea, procede introducir distintos contrastes de hipótesis que nos permitan comprobar la incurrancia estos problemas. En primer lugar, introduciremos el denominado test de Wooldrige (2002), quien desarrolló una prueba muy flexible para detectar la existencia de correlación serial en el error, cuya hipótesis nula es la no existencia de autocorrelación en el error ε_{it} . El método de Wooldrige utiliza los residuos de una regresión en primeras diferencias, observando que si ε_{it} no está serialmente correlacionado, entonces

$$\text{Corr}(\Delta\varepsilon_{it}, \Delta\varepsilon_{i,t-1}) = -0,5$$

En realidad esta prueba consiste en verificar esta igualdad.

Para identificar otro de los problemas antes mencionados, la presencia de autocorrelación contemporánea en los errores ε_{it} , Breusch y Pagan (1980) desarrollaron un test donde la hipótesis nula es que los errores entre las unidades sean independientes entre sí. Concretamente, se prueba la hipótesis de que la matriz de correlación de los residuos, obtenida sobre las observaciones comunes a todas las unidades transversales, es una matriz de orden N , donde N es el número de unidades transversales. El correspondiente estadístico de contraste λ_{LM} es de la forma

$$\lambda_{ML} = T \sum_2^N \sum_j^{i-1} r_{ij}^2$$

siendo r_{ij}^2 el ij -ésimo coeficiente de correlación de los residuos. Este estadístico denominado estadístico del Multiplicador de Lagrange de Breusch y Pagan, bajo hipótesis nula, sigue una χ_d^2 donde $d = N(N-1)/2$.

Por último, para dilucidar posibles problemas de heterocedasticidad en el término del error ε_{it} , introducimos la prueba modificada de Wald, con hipótesis nula $\sigma_i^2 = \sigma^2$ para $i = 1, \dots, N$. Sea

$$\hat{\sigma}_i^2 = T_i^{-1} \sum_{t=1}^{T_i} e_{it}^2$$

el estimador de σ_i^2 , basado en los T_i residuos e_{it} disponibles para esa unidad i , se define

$$V_i = T_i^{-1} (T_i - 1)^{-1} \sum_{t=1}^{T_i} (e_{it}^2 - \hat{\sigma}_i^2)^2$$

como la varianza estimada de $\hat{\sigma}_i^2$. El estadístico modificado de Wald, definido como

$$W = \sum_{i=1}^N \frac{(\hat{\sigma}_i^2 - \hat{\sigma}^2)^2}{V_i}$$

sigue una χ_N^2 bajo la hipótesis nula.

2.3. Modelos dinámicos

Una vez finalizada la introducción de los modelos básicos, nos centraremos en el estudio de los modelos dinámicos, los cuales han sido desarrollados con el propósito de incorporar en la estimación las relaciones de causalidad que se generan en el interior del modelo, como una forma de tratar los problemas de endogeneidad presentes en el mismo. En general, incluyen en su análisis valores rezagados de la variable dependiente. A la hora de estimar este tipo de modelos con los métodos anteriores, surgen una serie de problemas. Para analizar los mismos, introducimos la especificación más simple de un modelo dinámico de datos de panel, una autorregresión de primer orden con efectos individuales:

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \varepsilon_{it} \quad (2.7)$$

y_{it} la variable dependiente o explicada, con $t = 2, \dots, T$ e $i = 1, \dots, N$

$y_{i,t-1}$ la variable dependiente retardada que actúa como regresora, con $t = 1, \dots, T-1$ e $i = 1, \dots, N$

α el vector de coeficientes asociado a $y_{i,t-1}$

μ_i los efectos individuales

ε_{it} el término de perturbación.

El estimador intragrupos (sección 2.2.1) de $\hat{\alpha}$ resulta de la realización de MCO sobre la ecuación transformada:

$$y_{it} - \bar{y}_i = \alpha[y_{i,t-1} - \bar{y}_{i(-1)}] + [\varepsilon_{it} - \bar{\varepsilon}_i] \quad (2.8)$$

en donde $\bar{y}_i = \sum_{t=2}^T y_{it}/(T-1)$, $\bar{y}_{i(-1)} = \sum_{t=1}^{T-1} y_{it}/(T-1)$ y $\bar{\varepsilon}_i = \sum_{t=2}^T \varepsilon_{it}/(T-1)$. Sin embargo, se tiene que $[y_{i,t-1} - \bar{y}_{i(-1)}]$ y $[\varepsilon_{it} - \bar{\varepsilon}_i]$ están correlacionados por lo que $\hat{\alpha}$ estará sesgado cuando T es pequeño, incluso para valores grandes de N . Nickell (1981) derivó la expresión exacta de este sesgo para este caso, la cual omitiremos por brevedad. Únicamente cabe comentar que este sesgo, conocido como sesgo de Nickell, es de orden $1/T$, y que para valores positivos de α siempre es negativo.

En el caso de un modelo con un retado de una unidad de tiempo con presencia de variables explicativas exógenas, el sesgo en valor absoluto presente en la estimación del parámetro autorregresivo α sería aún mayor que de omitir estos regresores. Asimismo, la estimación del vector de coeficientes asociados a las variables explicativas (β) será también sesgada, siendo este sesgo más amplio cuanto mayor sea la relación entre las variables regresoras x_{it} y el retardo de la variable dependiente $y_{i,t-1}$.

La aparición de este sesgo al emplear los métodos de estimación expuestos anteriormente sobre modelos dinámicos, sugieren la utilización de otros estimadores, que se ajusten mejor a las características de este tipo de modelos. La primera alternativa que introduciremos a continuación, es el enfoque de variables instrumentales de Anderson-Hsiao (1981).

2.3.1. Enfoque de variables instrumentales: estimador de Anderson-Hsiao

Un método alternativo para evitar los problemas de sesgo previamente citados presentes en la estimación de modelos dinámicos es el denominado enfoque de variables instrumentales. Con objeto de ilustrar en qué consiste este modo de proceder con más detalle, introducimos uno de los estimadores más utilizados y que con mayor sencillez ilustran este tipo de procedimiento, el denominado estimador de Anderson-Hsiao, introducido por Anderson y Hsiao en 1981.

Para exponer su morfología, supondremos un panel de datos con $T = 3$ que permita reducir el sistema dinámico a dos ecuaciones en niveles,

$$y_{i2} = \mu_i + \alpha y_{i1} + \varepsilon_{i2} \quad (2.9)$$

$$y_{i3} = \mu_i + \alpha y_{i2} + \varepsilon_{i3}$$

Partiendo de esta especificación en niveles, y con el objetivo de eludir la existencia de heterogeneidad transversal, se plantea la forma en diferencias como sigue:

$$\Delta y_{i3} = \alpha \Delta y_{i2} + \Delta \varepsilon_{i3} \quad (2.10)$$

donde $\Delta y_{it} = y_{it} - y_{it-1}$ y $\Delta \varepsilon_{it} = \varepsilon_{it} - \varepsilon_{it-1}$. Como se expuso con anterioridad, la existencia de correlación entre Δy_{i2} y $\Delta \varepsilon_{i3}$ ocasiona problemas de sesgo en la estimación del parámetro autorregresivo α

por Mínimos Cuadrados Ordinarios. La estimación de variables instrumentales, exigirá encontrar un instrumento incorrelacionado con $\Delta\varepsilon_{i3}$, y, sin embargo, correlacionado con la variable a la que deberá sustituir, Δy_{i2} . En el caso que nos ocupa, el instrumento seleccionado será el valor del nivel y_{i1} , que asumiendo ausencia de autocorrelación en la perturbación aleatoria, verifica las condiciones para ser un instrumento adecuado. Por tanto, el modelo estaría exactamente identificado al contar con una variable instrumental para la estimación de un único parámetro, siendo el estimador de Anderson-Hsiao de la forma:

$$\hat{\alpha}_{AH} = \frac{\sum_{i=1}^N y_{i1} [y_{i3} - y_{i2}]}{\sum_{i=1}^N y_{i1} [y_{i2} - y_{i1}]} \quad (2.11)$$

Cuando el panel tiene más de 3 observaciones seriales ($T > 3$), existen posibles instrumentos adicionales. De este modo, mientras que y_{i1} es el único instrumento posible en la ecuación de primeras diferencias para el periodo $t = 3$, y_{i1} e y_{i2} pueden ser ambos empleados para el periodo $t = 4$, y de modo análogo el vector $(y_{i1}, y_{i2}, \dots, y_{iT-2})$ para el periodo $t = T$. El modelo está sobreidentificado para $T > 3$, y se tiene que el estimador no será asintóticamente eficiente aunque se empleen para cada ecuación el conjunto completo de los instrumentos disponibles. En este punto, Hansen (1982), desarrolla el denominado Método Generalizado de Momentos (GMM) por el que se obtienen estimadores asintóticamente eficientes en este contexto.

Años más tarde, Arellano y Bond (1991), proponen este procedimiento como alternativa generalmente más eficiente a la aproximación simple de Anderson-Hsiao para modelos de esta índole.

2.3.2. Método Generalizado de Momentos

Este método de estimación propuesto por Hansen (1982) es una generalización del conocido método de los momentos que corrige las limitaciones del mismo ante la existencia de sobreidentificación, es decir, la presencia de un número de condiciones de momentos mayor que el número de parámetros a estimar. El objetivo de estos procedimientos es estimar un vector de parámetros $\theta \in \mathbb{R}^p$ con valor real θ_0 de un modelo basado en $q \times 1$ momentos incondicionales

$$E[g(\theta_0, x_i)] = 0$$

con x_i un vector de datos. Cuando existe un mismo número de parámetros que de condiciones de momentos ($p = q$) se tiene un sistema exactamente identificado, y mediante el método de los momentos se obtiene la estimación deseada. Sin embargo, cuando $q > p$, se da una sobreidentificación, por lo que no hay solución para el sistema de ecuaciones:

$$\bar{g}(\theta) = \frac{1}{n} \sum_{i=1}^n g(\theta, x_i) = 0$$

Aquí es donde se introduce el denominado método generalizado de los momentos, que trata este problema mediante la minimización de la siguiente función cuadrática:

$$Q(\theta) = \bar{g}(\theta)^T W \bar{g}(\theta)$$

siendo W una matriz de pesos $q \times q$ definida positiva.

A continuación se expondrá la adaptación de Arellano y Bond (1991) de este método para la estimación de modelos dinámicos, el conocido como estimador GMM en diferencias.

Estimador GMM en diferencias

Este método de estimación es una generalización del método de variables instrumentales expuesto previamente que proporciona mejores resultados en este tipo de situaciones. La idea consiste en afrontar la estimación combinando diversos instrumentos entorno a un único vector numérico de coeficientes, que logre correlaciones muestrales mínimas entre el término de error y cada uno de los instrumentos. Teniendo en consideración el modelo tratado en la sección anterior,

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \varepsilon_{it} \quad |\alpha| < 1 \quad (2.12)$$

Se asume que el término de error ε_{it} verifica $E[\varepsilon_{it}] = 0$, que es serialmente incorrelado

$$E[\varepsilon_{it}\varepsilon_{is}] = 0 \text{ para } i = 1, \dots, N \text{ y } t \neq s \quad (2.13)$$

y que se verifican las siguientes condiciones iniciales sobre y_{i1}

$$E[y_{i1}\varepsilon_{it}] = 0 \text{ para } i = 1, \dots, N \text{ y } t = 2, \dots, T \quad (2.14)$$

Bajo estas hipótesis Arellano y Bond proponen transformar el modelo inicial tomando primeras diferencias para eliminar los efectos fijos inobservados e instrumentar las variables explicativas con problemas de endogeneidad mediante una matriz de instrumentos de la forma

$$Z_i = \begin{pmatrix} y_{i1} & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & y_{i1} & y_{i2} & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & y_{i1} & \cdots & y_{iT-2} \end{pmatrix}, \quad (2.15)$$

donde las filas se corresponden con las ecuaciones de primeras diferencias para los periodos $t = 3, 4, \dots, T$ para cada individuo i , que cumple las siguientes condiciones de momentos impuestas:

$$E[Z_i' \Delta \varepsilon_i] = 0 \text{ para } i = 1, 2, \dots, N \quad (2.16)$$

donde $\Delta \varepsilon_i = (\Delta \varepsilon_{i3}, \dots, \Delta \varepsilon_{iT})'$. En general, el estimador *GMM* asintóticamente eficiente basado en este conjunto de condiciones minimiza el siguiente criterio

$$J_N = \left(\frac{1}{N} \sum_{i=1}^N \Delta \varepsilon_i' Z_i \right) W_N \left(\frac{1}{N} \sum_{i=1}^N Z_i' \Delta \varepsilon_i \right) \quad (2.17)$$

usando la matriz de pesos

$$W_N = \left[\frac{1}{N} \sum_{i=1}^N (Z_i' \widehat{\Delta \varepsilon}_i \widehat{\Delta \varepsilon}_i' Z_i) \right]^{-1}, \quad (2.18)$$

donde $\widehat{\Delta \varepsilon}_i$ son estimaciones consistentes de las primeras diferencias residuales obtenidas de un estimador preliminar consistente. Por ello se conoce como el estimador GMM en dos etapas. Bajo homocedasticidad de los ε_{it} , la particularidad del modelo de primeras diferencias implica que un estimador asintóticamente eficiente se puede obtener en una etapa, usando la matriz de pesos

$$W_{1N} = \left[\frac{1}{N} \sum_{i=1}^N (Z_i' H Z_i) \right]^{-1}, \quad (2.19)$$

siendo

$$H = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{pmatrix} \quad (2.20)$$

Nótese que W_{1N} no depende de ningún parámetro estimado. Esto sugiere que W_{1N} es una opción razonable como estimador inicial consistente a emplear para calcular la matriz de pesos óptima W_N y de este modo obtener el estimador en dos etapas. Sin embargo, en muchas aplicaciones prácticas se han centrado en el estimador en una etapa frente al de dos etapas. Esto se debe en parte a que diversos estudios de simulación sugieren que la eficiencia ganada al emplear el estimador en dos etapas frente al de una etapa es muy pequeña, incluso en presencia de heterocedasticidad. La razón principal reside en que la influencia en la matriz de pesos del estimador en dos etapas de parámetros estimados convierte las aproximaciones de la distribución asintótica menos fiable para el caso del estimador en dos etapas. En esta línea, Windmeijer (2000) realizó un análisis de este hecho y propuso una corrección del estimador en dos etapas para muestras finitas muy útil en este tipo de modelos.

Este tipo de métodos de estimación se extienden de manera sencilla tanto a órdenes autorregresivos mayores, como a extensiones del presente modelo con variables explicativas adicionales (a parte de la variable respuesta retardada). Sea ahora el modelo de la forma

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \mu_i + \varepsilon_{it} \quad i = 1, \dots, N \quad t = 2, \dots, T \quad (2.21)$$

donde x_{it} puede ser un vector de valores de variables explicativas (retardadas o no), existirán diferentes condiciones sobre los momentos posibles dependiendo de la correlación existente entre x_{it} y el término del error del modelo. Hay tres posibles asunciones posibles acerca de cada variable explicativa, que son cruciales a la hora de su tratamiento en el proceso de estimación, en concreto a la hora de imponer condiciones sobre los momentos.

- Si x_{it} se supone endógena, entonces será tratada simétricamente con la variable dependiente y_{it} . En este caso, los valores retardados de la variable ($x_{i,t-2}, x_{i,t-3}, \dots$) serán instrumentos válidos en las ecuaciones de primeras diferencias para los periodos $t = 3, \dots, T$. Así, se impondrá que

$$E[Z_i' \Delta \varepsilon_i] = 0 \quad \text{para } i = 1, 2, \dots, N,$$

teniendo en cuenta que las filas de la matriz de instrumentos Z_i ahora se formarán a partir del vector $(y_{i1}, \dots, y_{i,t-2}, x_{i1}, \dots, x_{i,t-2})$. El procedimiento posterior de estimación se realiza de modo análogo al anteriormente expuesto.

- Si en cambio asumimos que una variable x_{it} es predeterminada, es decir, incorrelacionada con ε_{it} pero no con sus retardos, entonces $x_{i,t-1}$ es un posible instrumento adicional en la ecuación de primeras diferencias para el periodo t . En ese caso, el vector $(y_{i1}, \dots, y_{i,t-2})$ es reemplazado por $(y_{i1}, \dots, y_{i,t-2}, x_{i1}, \dots, x_{i,t-1})$ en la formación de la de la matriz de instrumentos Z_i .
- Por último, si suponemos que la variable explicativa es estrictamente exógena, entonces la serie completa x_i' será válida como variables instrumentales en cada una de las ecuaciones de primeras diferencias. En este caso, el vector $(y_{i1}, \dots, y_{i,t-2})$ es reemplazado por $(y_{i1}, \dots, y_{i,t-2}, x_i')$ en la formación de la matriz de instrumentos Z_i .

A pesar de la mejora de los resultados respecto al estimador de Anderson-Hsiao, diversos estudios de simulación han demostrado que el estimador GMM en diferencias se ve afectado por un considerable sesgo cuando se emplea con muestras finitas y en presencia de autocorrelación en los términos de error (Kiviet (1995), Blundell y Bond (1998), Hsiao, Pesaran y Tahmiscioglu (1999)). Además, también presentan sesgos cuando el coeficiente de la variable autorregresiva está muy cercano a 1, es decir, cuando la serie es altamente persistente o cercana a un proceso de raíz unitaria, por lo que el parámetro no puede ser identificado usando las condiciones de momentos para las ecuaciones de primeras diferencias. En estos casos, las simulaciones muestran que este estimador estará fuertemente sesgado por defecto (Blundell y Bond (1998)), en particular cuando T es pequeño.

Un estimador alternativo que evita estos problemas es el sugerido por Arellano y Bover (1995) y Blundell y Bond (1998), conocido con el nombre de estimador GMM extendido o de sistema.

Estimador GMM de sistema

A partir del modelo (2.12), pasamos a exponer ahora el denominado estimador GMM de sistema propuesto por Blundell y Bond (1998), que posee mejores propiedades cuando tratamos muestras finitas. Es una modificación del estimador GMM en diferencias, cuya principal característica reside en la adición de la suposición sobre el modelo de que

$$E[\mu_i \Delta y_{i2}] = 0 \text{ para } i = 1, \dots, N \quad (2.22)$$

Para que se verifique (2.22), las medias de las series y_{it} , aunque difiriendo entre individuos, deben de ser constantes a través de los periodos $t = 1, \dots, T$ para cada individuo. Esto combinado con el modelo $AR(1)$ expuesto anteriormente (2.12), desemboca en las siguientes $T - 2$ condiciones de momentos lineales

$$E[v_{it} \Delta y_{i,t-1}] = 0 \text{ para } i = 1, \dots, N \text{ y } t = 3, \dots, T. \quad (2.23)$$

siendo $v_{it} = \mu_i + \varepsilon_{it}$. Estas condiciones nos permitirán emplear primeras diferencias retardadas de las series como instrumentos para ecuaciones en niveles, siguiendo la idea introducida por Arellano y Bover (1995).

En este punto, podemos construir un estimador siguiendo la metodología GMM que explote las condiciones de momentos impuestas en (2.16) y (2.23). Este partirá de un sistema apilado de $T - 2$ ecuaciones en diferencias y $T - 2$ ecuaciones en niveles, correspondientes a los periodos $3, \dots, T$. Para el mismo, la matriz de instrumentos tendrá la forma

$$Z_i^+ = \begin{pmatrix} Z_i & 0 & 0 & \cdots & 0 \\ 0 & \Delta y_{i2} & 0 & \cdots & 0 \\ 0 & 0 & \Delta y_{i3} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \Delta y_{i,T-1} \end{pmatrix}, \quad (2.24)$$

donde Z_i es la matriz expuesta en (2.15). A partir de esta matriz de instrumentos, se procede de modo análogo a lo explicado con anterioridad para el caso de estimador GMM en diferencias.

Como se puede apreciar, el estimador GMM de sistema propuesto por Arellano y Bover (1995) y Blundell y Bond (1998) combina dos conjuntos de ecuaciones. El primero está formado por el sistema de ecuaciones en diferencias que se empleaban en el caso anterior, utilizando como instrumentos los

niveles rezagados de la variable autorregresiva $y_{i,t-1}$ y de las variables no exógenas x_{it} . El segundo, está formado por un sistema de ecuaciones en niveles, que permite añadir un conjunto de condiciones de momentos en diferencias.

Al igual que ocurría anteriormente, este estimador se extiende de manera sencilla a órdenes autorregresivos mayores y a extensiones del presente modelo con variables explicativas adicionales. A partir del modelo (2.21), las condiciones de momentos impuestas dependerán de la naturaleza de las variables explicativas (exógena, predeterminada o endógena). Respecto a las condiciones de momentos (2.16), el modo de proceder es el mismo que el expuesto en el caso del estimador GMM en diferencias. Por otra parte, en consonancia con las condiciones vistas en (2.23), se tiene que

- Si la variable x_{it} es endógena, se asume que las primeras diferencias de x_{it} están incorrelacionadas con los efectos individuales,

$$E[\mu_i \Delta x_{it}] = 0 \quad \text{para } i = 1, \dots, N \text{ y } t = 2, \dots, T \quad (2.25)$$

por lo que se tendrían, las siguientes condiciones de momentos adicionales

$$E[v_{it} \Delta x_{i,t-1}] = 0 \quad \text{para } i = 1, \dots, N \text{ y } t = 3, \dots, T$$

además de las introducidas en (2.23).

- Si la variable x_{it} es predeterminada, la suposición 2.25 conlleva las siguientes condiciones de momentos adicionales a las introducidas en (2.23)

$$E[v_{it} \Delta x_{i,t}] = 0 \quad \text{para } i = 1, \dots, N \text{ y } t = 3, \dots, T$$

- Análogamente, si la variable x_{it} fuese exógena, se tendría que

$$E[v_{it} \Delta x_{i,t}] = 0 \quad \text{para } i = 1, \dots, N \text{ y } t = 2, \dots, T$$

Como se puede apreciar, el tratamiento de las variables explicativas presentes en el modelo juega un papel fundamental a la hora de estimar un modelo dinámico de datos de panel. Sin embargo, existen otro tipo de consideraciones claves a tener en cuenta, como el número de instrumentos que se deben emplear. Los problemas derivados de esto último, así como el modo de solucionar o corregir los mismos, se expondrán en la siguiente sección.

2.3.3. Consideraciones adicionales: los instrumentos

La elección del número de instrumentos a utilizar es, si cabe, el paso más trascendental al estimar modelos de panel dinámicos mediante los estimadores tratados en la sección anterior. La utilización de un número elevado de instrumentos, sobre todo en muestras pequeñas, puede llevar al sobreajuste de las variables instrumentales, eliminar su componente endógena y sesgar los coeficientes estimados hacia los valores de los estimadores no instrumentados. Tauchen (1986) demostró mediante simulaciones con muestras muy pequeñas (50-75 observaciones) que el sesgo aumenta con el número de instrumentos basados en retardos de las variables. Ziliak (1997) obtuvo resultados similares. Además, el uso de muchos instrumentos en el proceso de estimación puede desembocar en resultados erróneos que asemejan ser válidos, lo que se debe a la debilidad de los test de validación, que trataremos en la sección siguiente.

La ausencia de test formales para controlar este aspecto obliga a emplear diferentes técnicas para reducir el número de instrumentos a usar en la estimación. Hasta la fecha, los investigadores han introducido dos técnicas principales para limitar el número de instrumentos en el proceso de estimación. La primera de ellas consiste en considerar únicamente ciertos retardos en lugar de todos los disponibles como instrumentos. Para cada periodo, se siguen generando instrumentos de manera separada, pero

se reduce el número de instrumentos considerados por periodo. La segunda, sugiere combinar los instrumentos disponibles en conjuntos más pequeños, lo que se conoce como colapsarlos. Esta práctica tiene como principal ventaja que retiene mayor información, ya que no se elimina ningún retardo. De este modo, en el caso de la estimación mediante el GMM en diferencias, la matriz de instrumentos Z_i anteriormente descrita en 2.15 pasaría a tener la forma siguiente

$$Z_i = \begin{pmatrix} y_{i1} & 0 & 0 & \cdots \\ y_{i2} & y_{i1} & 0 & \cdots \\ y_{i3} & y_{i2} & y_{i1} & \cdots \\ \cdots & \cdots & \cdots & \ddots \end{pmatrix} \quad (2.26)$$

imponiendo ahora la condición siguiente sobre los momentos

$$E[y_{i,t-l}\Delta\varepsilon_{it}] = 0 \text{ para cada } l \geq 2$$

En el caso de emplear el estimador GMM de sistema, la idea sería análoga, aplicando el colapsamiento sobre la matriz Z_i^+ .

Nótese que ambas técnicas pueden ser combinadas en el mismo proceso de estimación. Cabe comentar además, que Roodman (2007) comprobó que la aplicación de las mismas conducen a una mejora tanto en los resultados como en el funcionamiento de los test de validación.

2.3.4. Diagnósis y validación de los modelos dinámicos

Una vez ilustrado el método de estimación que emplearemos en la realización del presente proyecto (GMM), conviene introducir los diferentes métodos de diagnóstico y validación correspondientes. El primer objetivo será el estudio de la existencia de correlación serial en las perturbaciones, dado que ésta provoca la inconsistencia de los estimadores antes introducidos. Arellano y Bond (1991) sugieren un test de autocorrelación que trabaja bajo la hipótesis de no correlación serial de segundo orden en las perturbaciones del modelo en diferencias, es decir,

$$H_0 \equiv E[\Delta\varepsilon_{it}\Delta\varepsilon_{i,t-2}] = 0$$

El estadístico m_2 definido por Arellano y Bond (1991) se distribuirá según una $N(0, 1)$ bajo la hipótesis nula. Se trata de un estadístico flexible ya que está basado en cualquier estimador GMM no necesariamente eficiente. En cambio, la potencia de este test sí dependerá del estimador empleado.

Otra asunción crucial que debemos validar es la exogeneidad de los instrumentos. Para ello, conviene introducir el denominado test de sobre-identificación de Hansen (1982), cuyo estadístico de contraste es de la forma:

$$J = \left(\frac{1}{NT} Z' \hat{\varepsilon} \right) S^{-1} \left(\frac{1}{NT} Z' \hat{\varepsilon} \right)$$

siendo S una estimación de $Var[Z' \hat{\varepsilon}]$, $\hat{\varepsilon}'$ los residuos de la estimación GMM en dos etapas y Z la matriz de instrumentos. Sea p el número de columnas de la matriz de instrumentos Z , se tiene que bajo la hipótesis nula,

$$J \sim \chi_{p-k}^2$$

Nótese que si la estimación está exactamente identificada (mismo número de instrumentos que de regresores), la detección de instrumentos inválidos será imposible, ya que incluso cuando $E[Z\varepsilon] \neq 0$, el

estimador considerará coeficientes tales que $Z\hat{\varepsilon} = 0$, por lo que $J = 0$.

Sin embargo, Andersen y Sorensen (1996) y Bowsher(2002) demostraron que la proliferación de instrumentos vicia el test. Este último, obtuvo en sus simulaciones de Monte Carlo que el test nunca rechazaba la hipótesis nula a partir de un cierto número de instrumentos a un 0,05 de significación, cuando este debería rechazar tal hipótesis el 5% de las veces la misma. En consonancia con esto, Roodman (2009) aconseja imponer un p-valor mínimo de 0,25 en lugar del 0,05 habitual como límite para no rechazar la hipótesis nula. Sin embargo, se debe tener cuidado ante un valor crítico cercano a 1, lo cual no concluye en la validez de los instrumentos sino en que probablemente no se estén cumpliendo las propiedades del test (Roodman (2009)), en cuyo caso debemos rechazar la hipótesis nula.

De modo complementario al mencionado test de sobre-identificación de Hansen, procede introducir el denominado Difference-in-Hansen test, mediante el que se comprueba la validez de los subconjuntos de instrumentos. Este contraste, calcula el incremento en J al añadir un determinado subconjunto de instrumentos a la estimación. Bajo la hipótesis nula de exogeneidad de los instrumentos, el cambio en J sigue una χ^2 con grados de libertad igual al número de instrumentos añadidos.

Capítulo 3

Preparación de la base de datos

Una vez expuesta la metodología estadística que emplearemos para llevar a cabo el análisis temporal y geográfico de la mortalidad evitable en España, es necesario hablar de las diferentes variables que estarán involucradas en la realización del mismo. Este será uno de los objetivos del presente capítulo, en el que describiremos detalladamente el proceso de obtención y modificación de las variables que formarán los modelos de datos de panel que se considerarán en el citado análisis de la mortalidad evitable. En un primer acercamiento introduciremos los dos indicadores de mortalidad evitable que valoraremos, que serán las variables respuesta en los modelos que ajustaremos en nuestro estudio. Posteriormente, hablaremos de las diferentes variables explicativas o regresores que tendremos en cuenta.

3.1. Indicadores de mortalidad evitable

El primer paso en este tipo de análisis es el cálculo de los indicadores de mortalidad evitable, que serán las variables respuesta en nuestro estudio. En el caso que nos ocupa, se utilizaron dos indicadores: la tasa de mortalidad estandarizada y los años de esperanza de vida perdidos. Estos se calcularon para cada una de las 17 comunidades autónomas, en el periodo concurrido entre 1990-2011 para el total de muertes evitables de cada una de las cuatro listas de causas de mortalidad evitable.

3.1.1. Fuente de datos

Los datos necesarios para el cálculo de los indicadores antes mencionados se obtuvieron del Instituto Nacional de Estadística (INE). Se trata de tres bases de datos, que contienen para cada una de las 17 comunidades autónomas y por año:

- Las defunciones ocurridas en España, clasificadas por causas de muerte, en el periodo de estudio (1990-2011). Para realizar la clasificación de las causas de muerte, se utilizó la 9ª Clasificación Internacional de Enfermedades (CIE-9) hasta 1998, y desde 1999 la CIE-10.
- Las poblaciones de residentes en España en el citado periodo, las cuales son estimaciones intercensales, a fecha 1 de Enero, entre los censos de 1981-1991, 1991-2001 y 2001-2011.
- Los nacimientos ocurridos en España entre 1990-2011. Estos son datos del Movimiento Natural de Población, clasificados por lugar de residencia de la madre.

Las poblaciones y los nacimientos se obtuvieron de la web del INE (www.ine.es). Sin embargo, las defunciones fueron proporcionadas en CD por el INE, previa petición, para este estudio.

3.1.2. Causas de muerte

Se consideraron las cuatro listas de causas de muerte evitables descritas en la sección 1.3 para el cálculo de los distintos indicadores. No se analizaron las causas específicas de cada lista, sino la mortalidad evitable en conjunto.

Estas listas consideran la edad de 75 años como límite máximo de edad para la mayor parte de las causas, ya que este valor es consistente con la esperanza de vida al nacer en muchos países industrializados. Existen excepciones, causas de muerte que son evitables hasta una edad inferior a 75, que poseen un límite de edad menor. Es importante recordar que estos límites son arbitrarios y no debe interpretarse en el sentido de que algunas muertes en personas de mayor edad no podrían ser evitadas.

En el Apéndice 1 se presentan detalladamente de causas de muerte de cada lista con sus respectivos códigos CIE-9 y CIE-10.

3.1.3. Tasa de mortalidad estandarizada

El primero de los indicadores que introduciremos es la tasa de mortalidad estandarizada. Antes de profundizar en este concepto, conviene introducir lo que se conoce como tasa bruta de mortalidad, que es la proporción de personas que fallecen respecto al total de la población (usualmente expresada en tanto por cien mil).

Matemáticamente, se define la tasa bruta t como:

$$t = \frac{c}{p}$$

siendo

c es el número total de casos

p es la población total

Sin embargo, el empleo de estas tasas para nuestro análisis no es adecuado, ya que la distribución por grupos de edad de la población varía de unas comunidades autónomas a otras. Al no distribuirse de manera similar, cuando se compararan las tasas de mortalidad, estas estarán influenciadas por la proporción de sujetos en cada grupo de edad de cada población. Por tanto, el hecho de que una de las poblaciones consideradas presente una tasa de mortalidad superior a la del otro no quiere decir que tenga necesariamente un problema de salud mayor, ya que la diferencia observada podría ser efecto de diferencias en las estructuras de edad poblacionales, esto es, la edad estaría causando un efecto de confusión.

Para tratar de resolver este problema, se desarrollaron métodos de estandarización o ajuste de tasas. Uno de ellos es el método directo, que consiste en aplicar las tasas específicas por cada estrato de la variable de confusión (en este caso edad) de las poblaciones cuyas tasas se quieren comparar, a una población estándar dividida en los mismos estratos o categorías.

La tasa estandarizada por el método directo se define como:

$$t_{est} = \sum_{j=1}^m \frac{p_{ej}}{p_e} t_j$$

donde

m es el número de grupos de edad considerados para el ajuste

p_{ej} es la población estándar del j -ésimo grupo de edad, $j = 1, \dots, m$

$p_e = \sum_{j=1}^m p_{ej}$ es la población estándar total

c_j es el número de casos en el j -ésimo grupo de edad, $j = 1, \dots, m$

p_j es la población del j -ésimo grupo de edad, $j = 1, \dots, m$

$t_j = c_j/p_j$ es la tasa específica del j -ésimo grupo de edad, $j = 1, \dots, m$

En nuestro caso particular, se consideraron $m = 15$ grupos de edad quinquenales para el ajuste. Estas tasas se calcularon para cada comunidad autónoma (se excluyen Ceuta y Melilla), año (desde 1990 hasta 2011) y para el total de muertes evitables de cada una de las listas. Están calculadas en el grupo de 0 a 74 años y están multiplicadas por 100.000. En cuanto a la población estándar utilizada, inicialmente se valoraron tres opciones, con la finalidad de explorar cuál es la más adecuada en nuestro estudio. Estas fueron:

Población estándar Europea (Waterhouse et al., 1976)

Población estándar Europea Actualizada (2010)

Censo España 2001

La primera de las opciones ha sido descartada ya que representa una estructura poblacional muy joven, y la población española es más envejecida. Entre las otras dos opciones, se ha optado por seleccionar el Censo de España en el año 2001 ya que al tratarse de una población española en un año intermedio del periodo de estudio, será más adecuada en nuestro caso.

3.1.4. Años de esperanza de vida perdidos (AEVP)

El segundo indicador en nuestro análisis son los años de esperanza de vida perdidos. Este fue propuesto por Arriaga, y representa la diferencia entre el máximo posible de años que se pueden vivir entre dos edades (por ejemplo, en nuestro caso se considera entre 0 y 75) y los que realmente se viven (esperanza de vida temporaria en esas edades). Además, el total de AEVP atribuibles a la mortalidad en una población se puede descomponer por causa de muerte.

En el caso que nos ocupa, se calcularon los años de esperanza de vida perdidos por el método de Arriaga, basándonos en el supuesto de mortalidad nula entre dos edades para cada población definida por área geográfica (comunidad autónoma), año y para el total de muertes evitables de cada una de las listas.

Matemáticamente, los años de esperanza de vida perdidos entre las edades x y $x + n$ por la causa j se definen como:

$$AEVP_{x,j}^n = \frac{dt_{x,j}^n [(n - a_x^n) + S_{x+n}]}{100000}, \quad a \leq x < b$$

donde

a y b son los límites inferior y superior de edad, respectivamente, establecidos para el cálculo. En este trabajo $a = 0$ y $b = 75$.

x es el límite inferior de los grupos de edad ($x = 0, 1, 5, 10, \dots, 70$)

n es la amplitud de los grupos de edad

j representa la causa de defunción

$dt_{x,j}^n = dt_x^n \frac{d_{x,j}^n}{d_x^n}$ es el número de fallecidos de la tabla de mortalidad entre x y $x+n$ por la causa j

$d_{x,j}^n$ es el número de defunciones por la causa j entre las edades x y $x+n$

$d_x^n = \sum_j d_{x,j}^n$ es el número total de defunciones entre las edades x y $x+n$

dt_x^n es el número de fallecidos de la tabla de mortalidad entre x y $x+n$

a_x^n es el promedio de años vividos entre x y $x+n$ por los que mueren a esas edades.

$S_x = b - x$ (supuesto de mortalidad nula entre dos edades)

Simplificando, en nuestro caso particular, la fórmula queda de la siguiente manera:

$$AEVP_{x,j}^n = \frac{dt_{x,j}^n [75 - a_x^n - x]}{100000}, \quad a \leq x < b$$

3.2. Variables adicionales del modelo

Introducidas ya las diferentes variables respuesta que serán consideradas en los modelos que se emplearán para llevar a cabo el análisis de la mortalidad evitable, es conveniente exponer cuáles son las diferentes variables explicativas que formarán parte de los citados modelos.

En primer lugar, procedentes del Ministerio de Sanidad, Servicios Sociales e Igualdad, de la Estadística del Gasto Sanitario Público, se obtuvieron datos del gasto sanitario público de cada comunidad autónoma para el periodo 1995-2011. Concretamente, se consideró el gasto sanitario territorial consolidado en miles de euros, desglosado en lo que se conoce como su clasificación funcional, es decir, en 8 categorías:

Servicios hospitalarios y especializados

Servicios primarios de salud

Servicios de salud pública

Servicios colectivos de salud

Farmacia

Traslado, prótesis y aparatos terapéuticos

Gasto de capital.

Transferencias intersectoriales

Los datos proceden de dos bases de datos diferentes, de la *serie 1995-2005* se obtuvieron los datos correspondientes al periodo 1995-2001, ambos inclusive, mientras que los datos referentes al periodo 2002-2011 fueron extraídos de la *serie 2002-2012*. Para los años 1990 al 1994 se estimaron los valores del gasto a partir de los datos de gasto de años posteriores.

Para llevar a cabo dicha estimación, se calcularon las tasas de variación interanual del gasto sanitario territorial consolidado total en el periodo 1995-1999, y se calculó la tasa de variación media (TVM, en %). Mediante la misma, y de modo recursivo, se estimó el gasto sanitario territorial consolidado total desde 1990 a 1994 como sigue

$$\text{GastoTotal}_t = \frac{\text{GastoTotal}_{t+1}}{1 + \text{TVM}/100} \quad t = 1990, \dots, 1994$$

Finalmente, se calculó el porcentaje que representa cada una de las categorías de gasto respecto al total en el periodo 1995-1999 conjunto, y se aplicó este porcentaje a cada uno de los gastos totales antes estimados para obtener el gasto desglosado en los años 1990 a 1994.

Sin embargo, este gasto sanitario no es comparable, ya que está expresado en precios corrientes (no se está teniendo en cuenta el diferente valor del dinero en cada año). Para que la serie sea comparable, es necesario deflactarla, es decir, expresarla en precios constantes (fijando un año cero o base) con objeto de eliminar el efecto del cambio del valor monetario. Para ello, fue necesaria la obtención del IPC medio (IPC_m) en el periodo 1990-2011, mediante el cual se calcula la tasa de variación anual (en %) tomando como año base el 2011 (TV_{2011}), pues pasaremos los valores de gasto acorde al valor del dinero en este año. Para cada año $t \in (1990, 2011)$, se tiene que:

$$TV_{2011}^t = 100 \cdot \left(\frac{IPC_m^t - IPC_m^{2011}}{IPC_m^{2011}} \right)$$

A continuación, para deflactar, por ejemplo el gasto sanitario total, se realiza para cada año $t \in (1990, 2011)$,

$$\text{GastoTotal}_t = \frac{\text{GastoTotal}_t}{1 + TV_{2011}^t}$$

Este cálculo, se ha realizado de modo análogo para cada una de las categorías de gasto consideradas. Finalmente, se dividen los valores resultantes entre la población en cada año, obteniendo el correspondiente gasto per cápita, que será el que incluiremos en los correspondientes modelos en capítulos posteriores.

Además, se obtuvieron del Instituto Nacional de Estadística (INE) las siguientes variables de corte económico correspondientes al periodo de estudio y para cada una de las 17 comunidades autónomas.

- Tasa de paro. Se obtuvieron las tasas de paro trimestrales para el periodo de estudio de los resultados de la Encuesta de Población Activa (EPA), concretamente de dos series diferentes. Los datos trimestrales correspondientes al periodo de 1990-2001 provienen de la serie de 1876-2004, mientras que para los datos del 2002 al 2011 se utilizaron los datos de la EPA de 2002-2014. Con estas tasas trimestrales se estimó la tasa de paro anual mediante la media.
- Producto Interior Bruto per cápita (PIB_{pc}). Datos procedentes de la *Base 2000* de la contabilidad regional de España para los años 1995-2007 (ambos inclusive), y de la *Base 2008* para los años 2008-2011. No hay datos anteriores a 1995, por lo que se estimó el PIB per cápita de los años 1990 al 1994 utilizando datos de los años siguientes, de modo análogo al expuesto para el gasto sanitario total.

A partir de estas series de datos, se creó una serie homogénea (en *Base 2008*), para lo que fue necesario el cálculo de las tasas de variación interanual del PIB_{pc} para los datos anteriores a 2008. Con estas tasas de variación y los datos del PIB_{pc} del 2008, y para cada año $t \in (1990, 2007)$ se calcula de manera recursiva, el PIB_{pc} homogéneo en *Base 2008* de la manera que sigue:

$$PIB_{pc}^t = \frac{PIB_{pc}^{t+1}}{1 + TV_t^{t+1}}$$

Capítulo 4

Resultados

Como último paso del presente proyecto, este capítulo se centrará en la exposición de los diferentes resultados obtenidos, haciendo especial hincapié en el análisis temporal y geográfico de la mortalidad evitable llevado a cabo mediante el ajuste de modelos de datos de panel, siguiendo la metodología estadística introducida en el Capítulo 2, sobre la base de datos descrita en el Capítulo 3.

En un primer acercamiento, y antes del ajuste de los modelos, se llevarán a cabo un análisis de la situación inicial de la mortalidad evitable (1990) y un análisis de la evolución temporal de la misma. El objetivo del primero de estos análisis será de carácter meramente descriptivo, a fin de ilustrar la situación en el punto de partida de nuestro periodo de estudio. Por otro lado el objeto del análisis de la evolución temporal de la mortalidad evitable será dilucidar si la elección de la lista de causas de muerte evitables tiene influencia en los resultados, para lo que se realizará una comparativa del comportamiento de uno de los indicadores de mortalidad evitable a lo largo de nuestro periodo de estudio para cada una de las 4 listas de causas de muerte evitables.

Una vez finalizada ésta, procederemos a explicar de manera detallada el ajuste de diferentes modelos de datos de panel, tratando de encontrar el método de estimación que mejor se adapte a nuestra base de datos. En primer lugar, se valorará el ajuste de modelos estáticos de datos de panel, dada su mayor sencillez. Seguidamente, se considerarán modelos más complejos, de corte dinámico, con la finalidad de obtener mejores resultados.

4.1. Análisis de la situación inicial

Antes de analizar de manera temporal y geográfica la mortalidad evitable en España entre 1990-2011, procede describir la situación de la misma en el año inicial del periodo de estudio, 1990. Para ello, ilustraremos tanto analíticamente como gráficamente los resultados obtenidos para los dos indicadores considerados (tasa de mortalidad estandarizada y años de esperanza de vida perdidos).

En primer lugar, en el Cuadro 4.1 se presentan las tasas de mortalidad estandarizadas en el año 1990 en cada una de las 17 comunidades autónomas para el total de causas de muerte evitables, considerando las 4 listas de causas de mortalidad evitable.

Comunidad Autónoma	Lista 1	Lista 2	Lista 3	Lista 4
Andalucía	161 [5]	223,8 [6]	299,9 [1]	171,3 [3]
Aragón	123,8 [15]	188,8 [12]	226,5 [14]	125,9 [15]
Asturias (Principado de)	141,2 [7]	224,8 [5]	278,6 [7]	152 [7]
Balears (Illes)	163,1 [3]	248,8 [1]	295,9 [2]	174,7 [2]
Canarias	168,1 [1]	228,8 [3]	292,9 [3]	178 [1]
Cantabria	124,3 [14]	192,3 [10]	240,1 [11]	134,9 [14]
Castilla y León	117,9 [16]	171,7 [17]	215,9 [17]	122,6 [16]
Castilla-La Mancha	130,4 [11]	183,6 [14]	232 [13]	140,3 [11]
Catalunya	137,2 [8]	209,9 [9]	255,8 [9]	141 [10]
Comunidad Valenciana	162,6 [4]	228,8 [4]	284,2 [6]	164,7 [6]
Extremadura	153,3 [6]	215,5 [7]	285,4 [5]	167,3 [5]
Galicia	133,6 [10]	213,8 [8]	260 [8]	147,8 [8]
Madrid (Comunidad de)	130 [12]	177,6 [15]	226 [15]	144,1 [9]
Murcia (Región de)	164 [2]	234,1 [2]	288,4 [4]	169,2 [4]
Navarra (Comunidad Foral de)	108,4 [17]	174,6 [16]	218,6 [16]	112,8 [17]
País Vasco	124,4 [13]	185,3 [13]	240,8 [10]	138,1 [13]
Rioja (La)	134,4 [9]	191,4 [11]	239,9 [12]	138,4 [12]

Cuadro 4.1: Tasas de mortalidad ajustadas para el total de causas evitables por comunidad autónoma y para cada una de las listas consideradas en el año 1990. Entre corchetes, se ordenan las comunidades autónomas de mayor a menor tasa de mortalidad evitable en cada una de las listas.

Con el objetivo de describir de una manera más detallada la situación inicial de la mortalidad evitable en el periodo estudiado, se ilustran en las Figuras 4.1, 4.2, 4.3 y 4.4, para cada una de las 4 listas de causas de muerte evitables respectivamente, las tasas de mortalidad estandarizadas diferenciando por sexo y para cada comunidad autónoma, y el total de causas de muerte evitables. Se ha incluido también una línea indicando la tasa de mortalidad estandarizada para todo el territorio español (sin incluir Ceuta y Melilla).

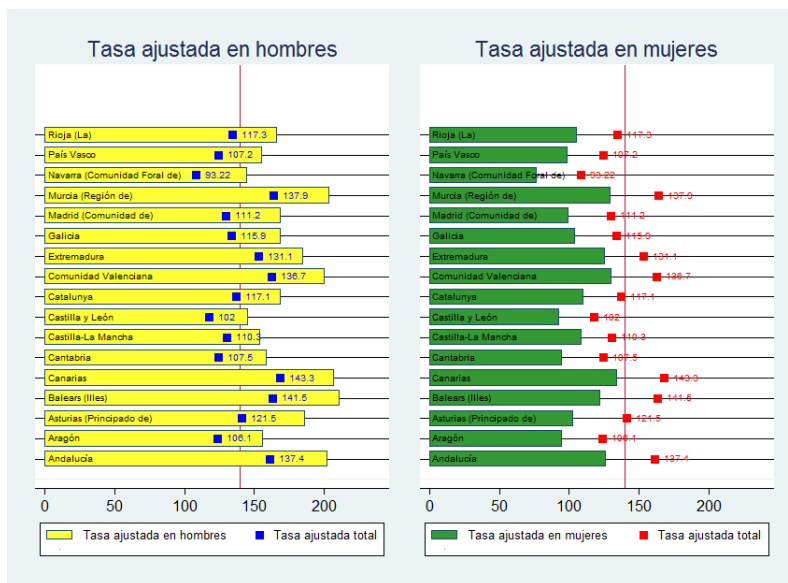


Figura 4.1: Tasas de mortalidad, por causas evitables, en hombres y mujeres y por Comunidad autónoma. Año 1990, Lista 1. En rojo, se presenta la tasa para el conjunto de España (línea vertical).

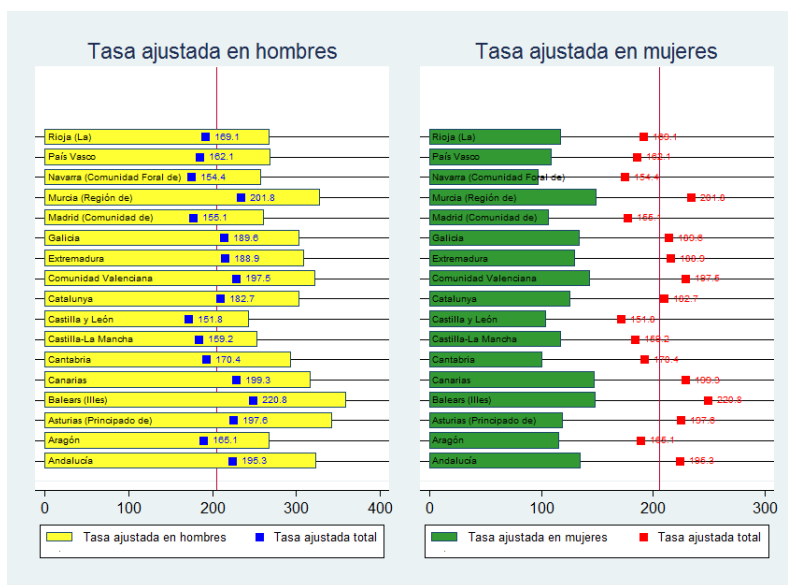


Figura 4.2: Tasas de mortalidad, por causas evitables, en hombres y mujeres y por Comunidad autónoma. Año 1990, Lista 2. En rojo, se presenta la tasa para el conjunto de España (línea vertical).

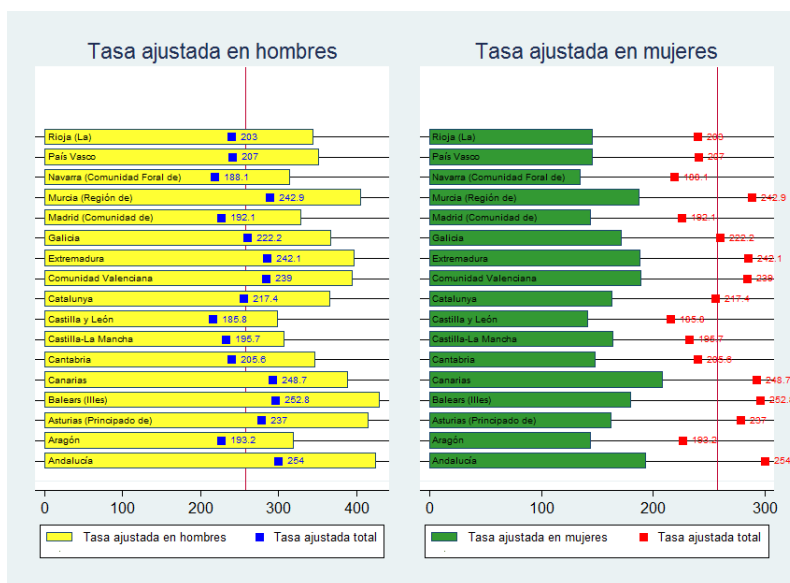


Figura 4.3: Tasas de mortalidad, por causas evitables, en hombres y mujeres y por Comunidad autónoma. Año 1990, Lista 3. En rojo, se presenta la tasa para el conjunto de España (línea vertical).

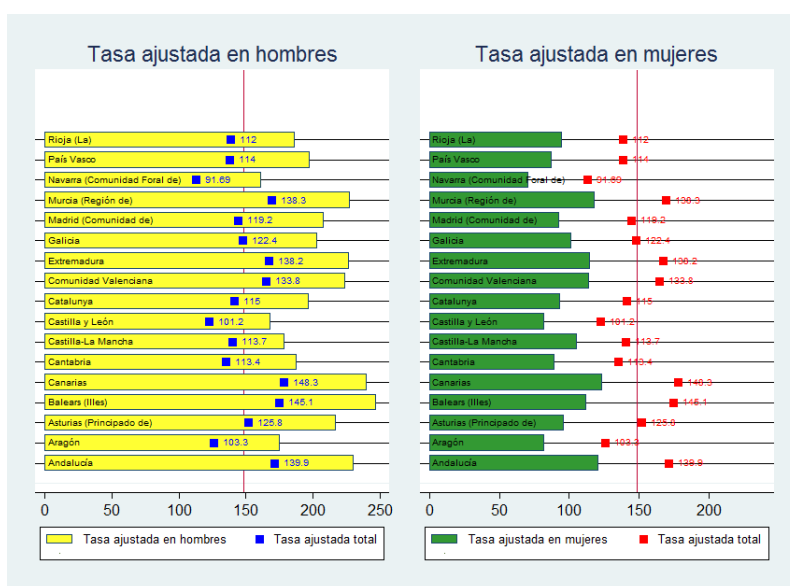


Figura 4.4: Tasas de mortalidad, por causas evitables, en hombres y mujeres y por Comunidad autónoma. Año 1990, Lista 4. En rojo, se presenta la tasa para el conjunto de España (línea vertical).

A la vista de los resultados, cabe comentar que las mayores tasas de mortalidad evitable se encuentran en las comunidades isleñas (Canarias e Illes Balears) seguidas por Andalucía y la Región de Murcia. En cambio, Navarra y Castilla y León destacan como las comunidades autónomas con menor mortalidad evitable en el año inicial del presente estudio. Asimismo, destacar que la tasa de mortalidad evitable es considerablemente mayor en la población masculina, encontrándose siempre por encima de la media española en todas las comunidades.

En cuanto al otro indicador de mortalidad evitable tratado, los años de esperanza de vida perdidos, se exponen en el Cuadro 4.2, de modo análogo al Cuadro 4.1, los resultados obtenidos en el año 1990.

Comunidad Autónoma	Lista 1	Lista 2	Lista 3	Lista 4
Andalucía	5,20 [3]	8,52 [7]	8,87 [1]	4,19 [4]
Aragón	4,11 [16]	7,73 [13]	7,20 [13]	3,20 [16]
Asturias [Principado de]	4,69 [9]	9,07 [3]	8,49 [3]	3,90 [8]
Balears [Illes]	5,27 [2]	10,05 [1]	8,84 [2]	4,44 [2]
Canarias	5,09 [4]	8,28 [9]	8,47 [4]	4,57 [1]
Cantabria	4,22 [13]	7,79 [11]	7,27 [12]	3,88 [9]
Castilla-La Mancha	4,14 [14]	7,12 [16]	6,92 [17]	3,49 [13]
Castilla y León	4,13 [15]	7,26 [15]	7,16 [14]	3,26 [15]
Catalunya	4,44 [10]	8,52 [6]	7,75 [11]	3,44 [14]
Comunidad Valenciana	4,77 [6]	8,71 [4]	8,16 [7]	3,92 [7]
Extremadura	4,84 [5]	8,21 [10]	8,45 [5]	4,34 [3]
Galicia	4,72 [8]	9,31 [2]	8,30 [6]	3,87 [10]
Madrid [Comunidad de]	4,34 [11]	7,43 [14]	7,00 [16]	4,12 [5]
Murcia [Región de]	4,75 [7]	8,60 [5]	8,11 [8]	4,09 [6]
Navarra [Comunidad Foral de]	3,59 [17]	7,06 [17]	7,03 [15]	2,81 [17]
País Vasco	4,24 [12]	7,74 [12]	7,77 [10]	3,68 [11]
Rioja [La]	5,37 [1]	8,40 [8]	7,88 [9]	3,56 [12]

Cuadro 4.2: AEVP para el total de causas evitables por comunidad autónoma y para cada una de las listas consideradas en el año 1990. Entre corchetes, se ordenan las comunidades autónomas de mayor a menor tasa de mortalidad evitable en cada una de las listas.

La representación gráfica de estos resultados, en cambio, no se hará del mismo modo que en el caso de las tasas de mortalidad. En este caso, se ilustran en las Figuras 4.5 4.6 los años de esperanza de vida perdidos para el total de causas de mortalidad evitable y no evitable, en cada una de las 17 comunidades autónomas y cada una de las 4 listas empleadas.

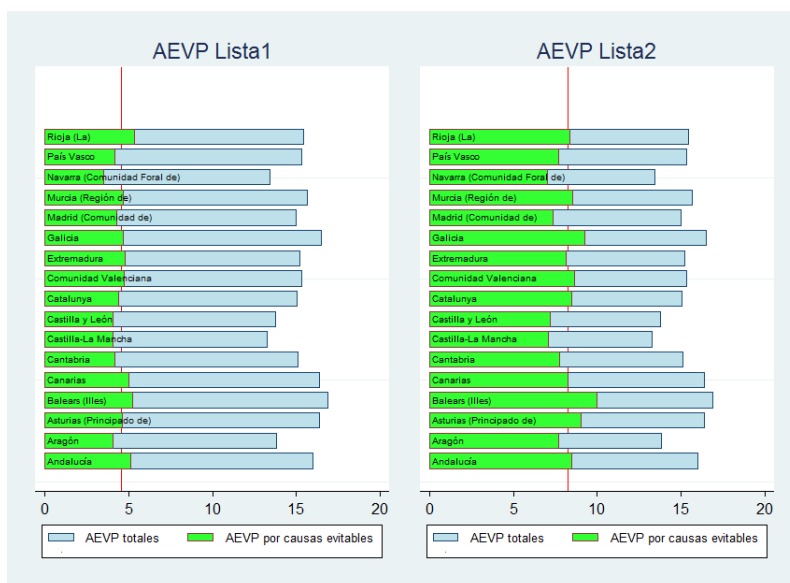


Figura 4.5: Años de esperanza de vida perdidos, por causas evitables, por Comunidad autónoma. Año 1990, Lista 1 y 2. En rojo, se presenta la tasa para el conjunto de España (línea vertical).



Figura 4.6: Años de esperanza de vida perdidos, por causas evitables, por Comunidad autónoma. Año 1990, Lista 3 y 4. En rojo, se presenta la tasa para el conjunto de España (línea vertical).

Los resultados obtenidos empleando este indicador son similares a los anteriormente expuestos respecto a las comunidades que presentan mayor y menor mortalidad evitable. Sin embargo, se aprecian diferencias considerables entre las cuatro listas a la hora de elaborar un ranking de mortalidad evitable en este año, ya que por ejemplo La Rioja pasa de ser la comunidad autónoma con mayor mortalidad evitable para la lista 1 a ocupar el puesto 12 al considerar la lista 4. Esto mismo ocurre en menor grado al emplear las tasas de mortalidad ajustadas por edad, pero parece que la elección de la lista de causas de muerte tendrá una influencia significativa en los resultados. Aún así, es conveniente realizar

una comparación de la evolución temporal de la mortalidad evitable de las cuatro listas de causas de muerte, para dilucidar si estas diferencias entre listas son un hecho puntual o si la hipótesis de que la lista de causas de mortalidad evitable influye en el resultado se verifica.

4.2. Evolución temporal. Comparación de las listas de causas de mortalidad evitable

El objetivo principal en esta sección será realizar una comparación de la evolución temporal de los indicadores de mortalidad evitable en cada una de las comunidades autónomas para cada una de las cuatro listas consideradas. De este modo analizaremos si la elección de una u otra lista cambia significativamente los resultados obtenidos o si en cambio la elección de una u otra no es relevante. Para llevar a cabo esta comparación, ilustraremos la evolución de uno de los indicadores de mortalidad evitable considerados, la tasa de mortalidad estandarizada, para cada comunidad y lista de causas de muerte evitables. Además, también se calculó el cambio relativo anual de las tasas en el periodo de estudio, el cual en 1990, será igual a cero ($CR_{1990} = 0$) y para el resto de años $t = 1991, \dots, 2011$, se define como:

$$CR_t = \frac{t_{est}^t - t_{est}^{t-1}}{t_{est}^{t-1}}$$

siendo t_{est}^t la tasa de mortalidad estandarizada en el año t .

En la Figura 4.7 representamos la evolución temporal de 1990-2011 de la tasa de mortalidad estandarizada (arriba) y de su cambio relativo anual (abajo) para la comunidad autónoma de Galicia. Los resultados para el resto de comunidades se adjuntan en el Anexo B.

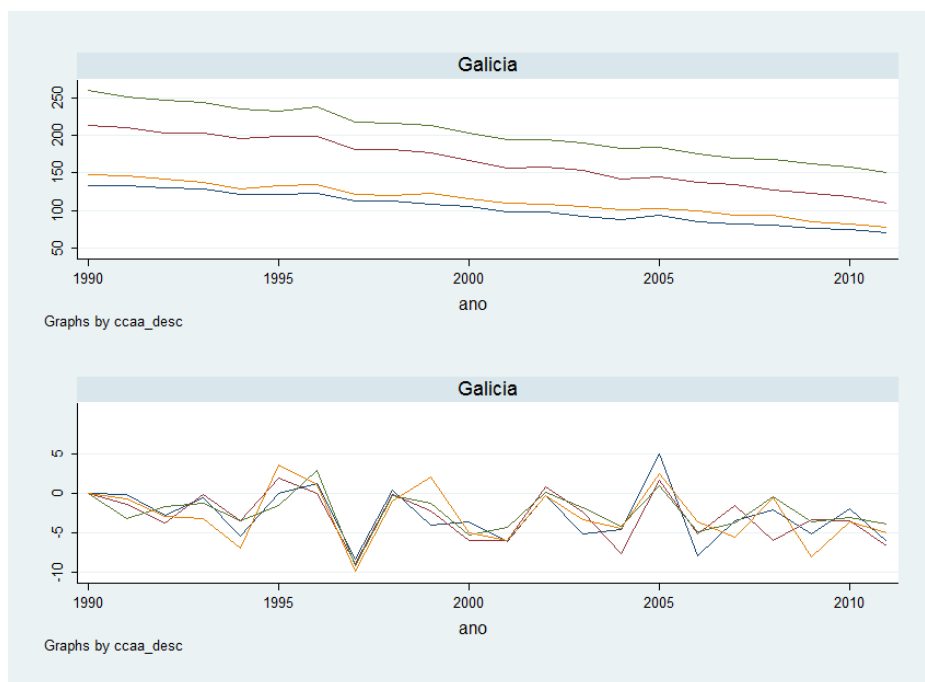


Figura 4.7: Tasas de mortalidad estandarizadas para el total de causas de muerte evitables (arriba) y su cambio relativo anual (abajo) para cada una de las cuatro listas en la comunidad autónoma de Galicia entre 1990 y 2011.

A la vista de los resultados, se observa un decrecimiento similar para las cuatro listas a lo largo del tiempo, sin embargo, observamos que las magnitudes de las tasas de mortalidad para cada lista son sustancialmente diferentes. En cuanto al cambio relativo, parece que evoluciona de manera similar para las cuatro listas, apreciándose leves diferencias en algunas comunidades autónomas que hacen pensar que la elección de la lista de causas de muerte tendrá alguna influencia sobre los resultados.

Ante esto, se ha decidido trabajar con la lista AMEIHs en lo sucesivo. Esta elección se sustenta en el hecho de que esta lista nace como resultado de un proyecto europeo (en el que se encontraba España) mediante el que se intentó poner en común una lista actualizada (2011) de causas de muerte evitables, lo que la convierte en la lista más adecuada de entre las cuatro que se han valorado para el estudio que nos ocupa.

4.3. Análisis temporal y geográfico de la mortalidad evitable en España (1990-2011)

En esta sección nos centraremos en la exposición y descripción del proceso seguido para la realización de un análisis temporal y geográfico de la mortalidad evitable en España desde 1990 hasta 2011. La técnica estadística empleada para efectuar el citado análisis es el ajuste de modelos de datos de panel, metodología expuesta en el Capítulo 2.

Para ello partimos de una base de datos de panel formada por las $N = 17$ comunidades autónomas (se excluye Ceuta y Melilla) y $T = 22$ años (1990-2011). Mediante la misma, es menester ajustar un modelo de datos de panel de la forma:

$$y_{it} = \mu_i + x'_{it}\beta + \varepsilon_{it} \quad i = 1, \dots, 17; \quad t = 1990, \dots, 2011 \quad (4.1)$$

siendo

i y t subíndices que indican respectivamente la unidad de sección cruzada (comunidad autónoma) y los diferentes periodos de tiempo (por año) considerados,

y_{it} el indicador de mortalidad evitable a considerar (tasa de mortalidad estandarizada o *AEVP*), que será nuestra variable respuesta o dependiente,

x_{it} un vector de $k \times 1$ variables independientes o explicativas,

β el vector de coeficientes a estimar,

μ_i los efectos individuales correspondientes a cada comunidad autónoma,

ε_{it} el término de perturbación.

En primera instancia, consideraremos distintos modelos considerando como variable respuesta cada uno de los indicadores de mortalidad evitable tratados (la tasa de mortalidad evitable estandarizada y los años de esperanza de vida perdidos) en los que el vector x_{it} estará formado por:

El gasto sanitario total consolidado per cápita ajustado a precios constantes

El Producto Interior Bruto (PIB) per cápita

La tasa de paro

Una vez realizado esto, trataremos de ajustar un modelo valorando el gasto de forma desglosada, con la finalidad de obtener una información más detallada del efecto que éste tiene sobre la mortalidad evitable en nuestro país.

4.3.1. Ajuste de modelos de efectos fijos

La metodología planteada para el ajuste de los modelos previamente descritos será la estimación intragrupos o de efectos fijos. Para llevar a cabo esta tarea, nos apoyaremos en el software estadístico Stata 12. De este modo, en primera instancia procederemos a ajustar un modelo donde la variable dependiente será la tasa de mortalidad evitable estandarizada, y se incluirán tres variables explicativas: el gasto sanitario total per cápita, el Producto Interior Bruto per cápita y la tasa de paro.

La estimación intragrupos de este modelo en Stata se realiza empleando el comando `xtreg`, que reporta los siguientes resultados:

```
>xtreg taxust3_t Gasto_Total pib_pc tasa_paro, fe

Fixed-effects (within) regression           Number of obs   =    374
Group variable: ccaa                       Number of groups =    17

R-sq:  within = 0.8760                     Obs per group:  min =    22
        between = 0.3602                    avg           =   22.0
        overall = 0.7072                    max           =    22

                                           F(3,354)       =   833.26
corr(u_i, Xb) = -0.1725                    Prob > F       =   0.0000

-----+-----
taxust3_t |      Coef.   Std. Err.   t    P>|t|   [95% Conf. Interval]
-----+-----
Gasto_Total | -.0651846   .0046285  -14.08  0.000   -.0742874   -.0560819
pib_pc      | -.0048894   .0005366   -9.11  0.000   -.0059447   -.0038341
tasa_paro   | -.9090521   .1770437   -5.13  0.000   -1.257242   -.5608623
_cons       | 399.5269    9.016823   44.31  0.000   381.7937    417.2602
-----+-----
sigma_u     | 18.960135
sigma_e     | 12.161781
rho         | .70849391   (fraction of variance due to u_i)
-----+-----
F test that all u_i=0:   F(16, 354) =   43.45           Prob > F = 0.0000
```

El ajuste de este primer modelo considerado pone de manifiesto una relación inversa entre todos los regresores respecto de la tasa de mortalidad evitable estandarizada, con unos p-valores asociados menores que los niveles de significación usuales. Por tanto, se tiene que para valores fijos del resto de regresores, el incremento tanto del gasto sanitario total per cápita, como del PIB per cápita o de la tasa de paro, provoca una bajada de la tasa de mortalidad evitable. En el caso del gasto sanitario y el PIB parece razonable dicha conclusión, pero resalta el hecho de que el incremento de la tasa de paro (bajo condiciones fijas de gasto sanitario y PIB ambos per cápita) se asocie a una reducción de la mortalidad evitable se refiere.

El término constante (`cons`) hace referencia a un valor común μ ¹ a todas las comunidades autónomas, el cual también resultó ser significativo. Asimismo, mediante lo que Stata denomina `rho`, se nos aporta información acerca de la fracción de variabilidad causada por diferencias entre paneles, la cual es de una magnitud considerable (0.7). Por último, se lleva a cabo la realización de un test F, que compara un modelo de esta índole con otro donde los efectos individuales se encuentren agrupados, es decir, sean un término constante común a todas las comunidades. El p-valor obtenido, menor que los niveles de significación usuales, nos lleva a poder afirmar que la consideración de los efectos individuales es adecuada en este caso. La estimación de estos efectos individuales $\mu_i = \mu + v_i$ no es reportada automáticamente por Stata y su cálculo se debe efectuar a posteriori como sigue

$$\hat{\mu}_i = \bar{y}_i - \bar{x}_i' \hat{\beta}$$

¹Se tiene que $\mu_i = \mu + v_i$.

Previo cálculo de los μ_i , incurriremos en la validación del presente modelo. Para ello, echaremos mano de los diferentes test introducidos en la sección 2.2.4, medio por el que se dilucidará la posible presencia de correlación serial, contemporánea y heterocedasticidad en el término de error del modelo. En caso de existir alguno de estos problemas, se estaría incumpliendo la asunción de que los errores ε_{it} sean independientes e idénticamente distribuidos, por lo que deberíamos afrontar el ajuste del modelo con otros estimadores, o de otro tipo de modelo.

En primer lugar, mediante la prueba de Wooldrige, implementada en Stata en la función `xtserial`, obtenemos un p-valor nulo que nos lleva a rechazar la hipótesis nula de incorrelación serial en el error. Seguidamente, realizamos el test de Breusch y Pagan mediante la sentencia `xttest2`, obteniendo otra vez un p-valor nulo, que evidencia la existencia de correlación contemporánea en el error. Por último, haciendo uso del comando `xttest3`, contrastamos la presencia de heterocedasticidad en el error ejecutando el test modificado de Wald. Como resultado se obtuvo un p-valor igual a cero que nos lleva a rechazar la hipótesis nula de homocedasticidad en el término del error ε_{it} .

En resumidas cuentas, el modelo ajustado empleando un estimador intragrupos no es adecuado en este caso, ya que no se verifican las asunciones que este tipo de estimador impone sobre los errores ε_{it} . Ante esta situación, procede valorar otro tipo de modelos de datos de panel más complejos, mediante los que seamos capaces de conseguir un mejor análisis de la mortalidad evitable en España. Se trata de los modelos dinámicos de datos de panel, los cuales serán el tema principal de la siguiente sección.

4.3.2. Ajuste de modelos dinámicos de datos de panel

Ante la imposibilidad de lograr explicar de manera adecuada la mortalidad evitable en España mediante un modelo estático de efectos fijos, pasamos a tratar un modelo de corte dinámico, cuya característica principal es la inclusión en el modelo de la variable dependiente rezagada como variable explicativa. Tal práctica no parece descabellada, pues el hecho de que la mortalidad evitable referente a un año tenga influencia sobre las muertes evitables del año siguiente asemeja ser un razonamiento adecuado.

Hechas tales consideraciones, pasamos a tratar en primera instancia un modelo introduciendo un retardo de la variable dependiente como variable explicativa, que se rige por la siguiente ecuación

$$y_{it} = \mu_i + \alpha y_{i,t-1} + x'_{it}\beta + \varepsilon_{it} \quad i = 1, \dots, 17; t = 1991, \dots, 2011 \quad (4.2)$$

siendo α el coeficiente asociado a la variable dependiente rezagada $y_{i,t-1}$. En cuanto a las variables involucradas en el mismo, volveremos a tratar de explicar tanto la tasa de mortalidad evitable estandarizada como los años de esperanza de vida perdidos (variable dependiente), empleando como regresores el gasto sanitario total per cápita, el PIB per cápita y la tasa de paro, acompañando a la tasa de mortalidad evitable estandarizada rezagada.

Para realizar esta tarea, utilizaremos el estimador GMM de sistema en dos etapas de Blundell y Bond (1998), implementado en Stata con la función `xtabond2` desarrollada por Roodman (2007). Como ya se expuso en el Capítulo 2, a la hora de estimar esta clase de modelos se debe prestar atención a dos cuestiones principales:

El tratamiento de las variables explicativas

El número de instrumentos utilizados

En primer lugar, se debe valorar la consideración de cada variable explicativa según su naturaleza dentro del modelo, es decir, según sean exógenas, predeterminadas o endógenas, ya que juega un papel fundamental en el proceso de estimación del modelo. Además, el tratamiento de una variable de uno u otro modo tiene una relación directa con el número de instrumentos usados en la estimación, como

se comentó en el Capítulo 2.

En relación al número de instrumentos a emplear, cabe comentar que la base de datos involucrada en nuestro estudio se caracteriza por tener un número pequeño de individuos N (comunidades autónomas) y un número elevado de periodos de tiempo T (años), lo que genera una tendencia a la proliferación de instrumentos. De hecho, la relación N/T en nuestro caso sería inferior a 1, lo que acarrea serios problemas a la hora de tratar con variables endógenas. Por ello, diversos autores aconsejan la práctica de ciertas medidas para corregir este problema, entre la que se encuentra la división del periodo de tiempo de análisis en bienios, trienios, etc. En este caso, se ha optado por la reducción del periodo a bienios, corrigiendo ostensiblemente los citados problemas. Para ello, se ha calculado a partir de los datos anuales presentes en nuestra base de datos, la media para cada uno de los $T = 11$ bienios creados.

Previo ajuste de los distintos modelos de la forma (4.2), debemos seleccionar las variables explicativas que incluiremos y su tratamiento en el mismo. Para la selección de las variables, es recomendable hacer un análisis de correlaciones entre éstas, con el fin de excluir aquellas que tengan un mismo comportamiento. No existe un criterio exacto de a partir de qué correlación debemos valorar la eliminación de una variable, sin embargo se debería esperar una correlación menor o igual a 0.7 entre las variables presentes en el modelo.

	Gasto Sanitario Total _{PC}	PIB _{PC}	Tasa de Paro
Gasto Sanitario Total _{PC}	1.00		
PIB _{PC}	0.5601	1.00	
Tasa de Paro	-0.3941	-0.6999	1.00

Cuadro 4.3: Correlaciones entre las variables explicativas del modelo.

En el Cuadro 4.3, podemos ver que la correlación entre el PIB per cápita y la tasa de paro se encuentra cerca del límite antes mencionado, pero no lo sobrepasa. Ante este hecho, no eliminaremos ninguna de las variables del modelo, pero sí debemos tener en cuenta la posible repercusión de esta alta correlación en los resultados.

Respecto al tratamiento de cada una de las variables explicativas dentro del modelo, no existe una prueba o método que nos permita dilucidar la naturaleza de cada una de estas variables dentro del modelo. Para deducir el mismo, se valoran las diferentes opciones atendiendo al comportamiento de los estadísticos para cada una de ellas. Este proceso concluye, en nuestro caso particular, en la naturaleza endógena del gasto sanitario consolidado total per cápita, y exógena del resto de variables (PIB per cápita y tasa de paro). Además, se valoró la posible inclusión de retardos de las variables explicativas en el modelo, lo que nos llevó a eliminar el gasto sanitario total consolidado per cápita del modelo, y la inclusión de su primer retardo, ante su mejor comportamiento y mayor relevancia en el modelo.

En este punto, únicamente falta abordar el problema del número de instrumentos a emplear. Como primera medida, se llevará a cabo el colapsamiento de los instrumentos en todas y cada una de las estimaciones valoradas. Procede comentar, que es deseable emplear un número de instrumentos menor o igual que el número de unidades de corte transversal N , lo que evitaría resultados sesgados a la baja. Al usar todos los instrumentos disponibles en el presente caso descrito, tras colapsarlos, el número de instrumentos sería igual a 21, es decir, mayor que N . Por ello, se ha optado por combinar la técnica de colapsamiento de los instrumentos con la limitación de retardos considerados como tales.

Modelo dinámico con la tasa de mortalidad evitable estandarizada como variable respuesta

En primer lugar, se llevó a cabo la estimación del modelo descrito, considerando como variable dependiente la tasa de mortalidad evitable estandarizada, mediante el estimador GMM de sistema en dos etapas de Blundell y Bond con las debidas correcciones para muestras pequeñas y heterocedásticas. Teniendo en cuenta las condiciones antes explicadas y llevada a cabo con la función `xtabond2` de Stata, reporta los siguientes resultados:

```
>xtabond2 taxust3_t L.taxust3_t L.Gasto_Total pib_pc tasa_paro, gmm(l.taxust3_t, lag(2 7) collapse)
gmm(l.Gasto_Total, lag(2 7) collapse) iv(pib_pc tasa_paro) small robust twostep ortho
```

Favoring space over speed. To switch, type or click on mata: `mata set matafavor speed, perm.`

Dynamic panel-data estimation, two-step system GMM

```
-----
Group variable: ccaa                Number of obs   =    170
Time variable : bienio              Number of groups =    17
Number of instruments = 17          Obs per group: min =   10
F(4, 16)      =    664.71           avg           =  10.00
Prob > F      =    0.000            max           =   10
-----
```

	Coef.	Corrected Std. Err.	t	P> t	[95% Conf. Interval]	
taxust3_t						
L1.	.7291095	.122594	5.95	0.000	.4692218	.9889972
Gasto_Total						
L1.	-.0222825	.0087255	-2.55	0.021	-.0407797	-.0037852
pib_pc	-.000409	.0004555	-0.90	0.383	-.0013747	.0005567
tasa_paro	.2859079	.1082092	2.64	0.018	.0565147	.5153012
_cons	77.11149	43.86473	1.76	0.098	-15.87759	170.1006

Instruments for orthogonal deviations equation

```
Standard
FOD.(pib_pc tasa_paro)
GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/7).L.Gasto_Total collapsed
L(2/7).L.taxust3_t collapsed
```

Instruments for levels equation

```
Standard
pib_pc tasa_paro
_cons
GMM-type (missing=0, separate instruments for each period unless collapsed)
DL.L.Gasto_Total collapsed
DL.L.taxust3_t collapsed
```

```
-----
Arellano-Bond test for AR(1) in first differences: z = -2.67 Pr > z = 0.008
Arellano-Bond test for AR(2) in first differences: z = 0.64 Pr > z = 0.520
-----
```

```
Sargan test of overid. restrictions: chi2(12) = 15.65 Prob > chi2 = 0.208
(Not robust, but not weakened by many instruments.)
```

```
Hansen test of overid. restrictions: chi2(12) = 14.21 Prob > chi2 = 0.287
(Robust, but weakened by many instruments.)
```

Difference-in-Hansen tests of exogeneity of instrument subsets:

```
GMM instruments for levels
Hansen test excluding group: chi2(10) = 12.35 Prob > chi2 = 0.262
Difference (null H = exogenous): chi2(2) = 1.86 Prob > chi2 = 0.394
gmm(L.taxust3_t, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 9.19 Prob > chi2 = 0.102
Difference (null H = exogenous): chi2(7) = 5.02 Prob > chi2 = 0.658
gmm(L.Gasto_Total, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 9.53 Prob > chi2 = 0.090
Difference (null H = exogenous): chi2(7) = 4.68 Prob > chi2 = 0.699
iv(pib_pc tasa_paro)
Hansen test excluding group: chi2(10) = 13.45 Prob > chi2 = 0.199
Difference (null H = exogenous): chi2(2) = 0.76 Prob > chi2 = 0.684
```

En la anterior salida, podemos ver en la primera línea el comando utilizado para la estimación GMM de sistema, seguido de ciertos datos descriptivos de la estimación del modelo, entre los que destaca el número de instrumentos empleados. A continuación, se puede apreciar la tabla de coeficientes asociados a las variables explicativas y el término constante μ común a todas las comunidades autónomas y sus respectivas significaciones. Por último, se obtiene una descripción de los diferentes instrumentos usados junto con los diferentes test de validación del modelo.

Como resultado se observa una relación directa entre la tasa de mortalidad evitable estandarizada y su primer retardo, siendo el valor del coeficiente asociado estimado de 0.7291 y el correspondiente p-valor menor que los niveles de significación usuales. Por otra parte, los coeficientes asociados al gasto sanitario total per cápita rezagado y el PIB per cápita muestran una relación inversa de éstas respecto de la variable dependiente, siendo solamente la primera de ellas significativa. Así, se tiene que, bajo condiciones fijas del resto de regresores, el incremento de 100 euros en gasto sanitario total per cápita del bienio anterior provocaría una bajada de aproximadamente 2.2 en la tasa de mortalidad evitable estandarizada. Por otro lado, y al contrario de lo que ocurría al tratar este mismo modelo sin la variable respuesta rezagada, el coeficiente asociado a la tasa de paro sí evidencia en este caso una relación directa respecto de la variable dependiente, con un p-valor que pone de manifiesto el carácter significativo de la misma. La interpretación resultante del mismo concluye que una subida de un punto en la tasa de paro provocaría un aumento en la mortalidad evitable entorno a 0.28, fijadas el resto de variables explicativas.

Al igual que ocurría con el estimador intragrupos, Stata únicamente nos devuelve un coeficiente constante común a todas las comunidades autónomas (μ), el cual no resulta significativo al 5% de significación. El efecto individual $\mu_i = \mu + v_i$ se calcula a posteriori empleando los valores de los coeficientes estimados asociados a cada uno de los regresores, como se realizaba en el caso de los modelos de efectos fijos, obteniendo los valores de los v_i que se muestran en el Cuadro 4.4

Comunidad Autónoma	v_i
Andalucía	2.06581
Aragón	2.486589
Asturias (Principado de)	7.93716
Balears (Illes)	1.456362
Canarias	5.192396
Cantabria	-1.051472
Castilla-La Mancha	-6.076802
Castilla y León	-5.663634
Catalunya	0.3601572
Comunidad Valenciana	2.006579
Extremadura	-1.150608
Galicia	-1.294753
Madrid (Comunidad de)	-5.830367
Murcia (Región de)	-0.4206341
Navarra (Comunidad Foral de)	2.171762
País Vasco	3.55948
Rioja (La)	-0.1708148

Cuadro 4.4: Efectos v_i para cada una de las 17 comunidades autónomas.

Visualmente, se pueden representar los mismos mediante el siguiente gráfico de barras, en el cual se puede apreciar el distinto efecto sobre la mortalidad evitable en cada una de ellas.

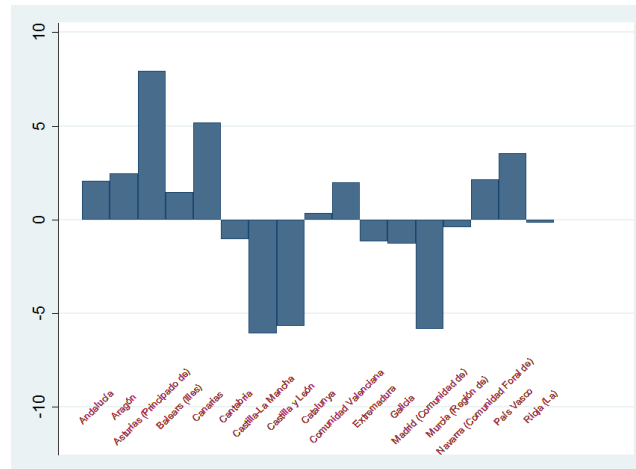


Figura 4.8: Efectos individuales v_i para cada una de las 17 comunidades autónomas.

El Cuadro 4.4 y la Figura 4.8 ilustran la diferente influencia sobre la mortalidad evitable causada por las características sociales, gestión y otros aspectos específicos de cada comunidad autónoma fijadas las otras covariables. De este modo, los coeficientes v_i negativos evidenciarán una menor mortalidad evitable respecto a la media del conjunto del país, interpretándose los coeficientes positivos de manera opuesta. En esta línea, destacan las comunidades de Asturias y Canarias como aquellas donde se produce mayor mortalidad evitable por causas específicas, mientras que por el contrario, en Castilla y León, Castilla-La-Mancha y Madrid se caracterizan por presentar menor mortalidad evitable.

En este punto, procede hacer referencia a la validación del presente modelo, para lo que se aprecia en primer lugar en la salida anterior de Stata el número de instrumentos empleados (17), el cual es igual que el número de unidades de corte transversal N , lo que no nos hace sospechar de problemas de sesgo en los resultados causados por la proliferación de instrumentos. Junto a esto, en la parte inferior aparecen ilustrados los p-valores relativos a la ejecución de los test de autocorrelaciones de primer y segundo orden de Arellano y Bond y los test de validez de los instrumentos de Hansen (tanto el global como los Difference-in-Hansen test), que nos permiten concluir la adecuación del presente modelo. Los test de autocorrelación nos llevan a afirmar la presencia de correlación de primer orden (como era de esperar), y la ausencia de autocorrelación de segundo orden. En cuanto a los p-valores relativos a los diferentes test de Hansen ilustrados, no se obtienen valores cercanos a 1, lo cual descarta la posibilidad de un incumplimiento de las propiedades asintóticas de estos test (lo cual es frecuente en este tipo de modelos). Por otra parte, éstos son mayores que 0.05, y concretamente el test de Hansen relativo a todo el conjunto de instrumentos es mayor que 0.25, de lo que se concluye que los instrumentos utilizados son válidos. En resumen, a la vista de los resultados, podemos afirmar que el modelo es adecuado.

Por último, procedemos a representar, para cada una de las comunidades, los valores ajustados frente a la tasa de mortalidad evitable estandarizada real. Al tratarse de un modelo dinámico, no se obtiene el valor ajustado referente al bienio 1990-1991, como es lógico. En la Figura 4.9 vemos como los valores ajustados describen de manera fiel la curva representada por la tasa de mortalidad evitable estandarizada real, concluyendo en un buen ajuste del presente modelo.

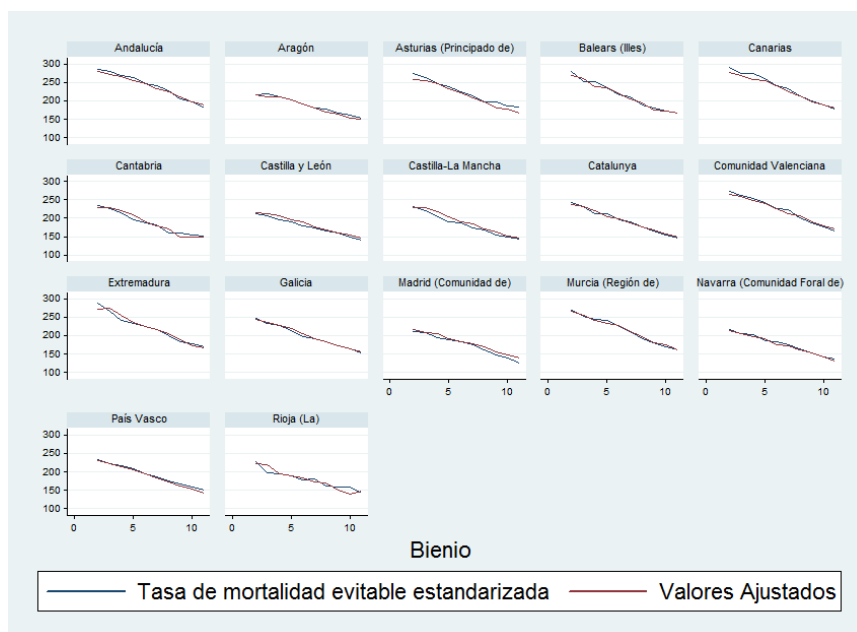


Figura 4.9: Evolución temporal de la tasa de mortalidad evitable estandarizada t_{est} (Azul) junto a los valores ajustados (Rojo) para cada una de las 17 comunidades autónomas.

A pesar del buen ajuste obtenido (Figura 4.9), a continuación se procede a valorar sendos modelos derivados del presente, excluyendo primero el PIB per cápita y luego la tasa de paro, a fin de ver si la alta correlación entre estas dos variables ha causado alguna influencia en los resultados. En primer lugar, consideramos un modelo completamente análogo sin el PIB como regresor, lo que nos reporta los siguientes resultados:

```
>xtabond2 taxust3_t L.taxust3_t L.Gasto_Total tasa_paro, gmm(1.taxust3_t, lag(2 7) collapse)
gmm(1.Gasto_Total, lag(2 7) collapse) iv(tasa_paro) small robust twostep ortho
```

Favoring space over speed. To switch, type or click on mata: mata set matafavor speed, perm.

Dynamic panel-data estimation, two-step system GMM

```
-----
Group variable: ccaa                Number of obs   =    170
Time variable : bienio             Number of groups =    17
Number of instruments = 16          Obs per group:  min =    10
F(3, 16)      = 1051.71              avg   =    10.00
Prob > F      = 0.000                 max   =    10
-----
```

	taxust3_t	Coef.	Corrected Std. Err.	t	P> t	[95% Conf. Interval]	
taxust3_t							
L1.		.7684059	.0928483	8.28	0.000	.5715763	.9652356
Gasto_Total							
L1.		-.0210801	.0079038	-2.67	0.017	-.0378354	-.0043248
tasa_paro							
_cons		.4034121	.1182978	3.41	0.004	.152632	.6541923
		56.75724	27.0858	2.10	0.052	-.6621051	114.1766

Instruments for orthogonal deviations equation

Standard

FOD.tasa_paro

GMM-type (missing=0, separate instruments for each period unless collapsed)

L(2/7).L.Gasto_Total collapsed

L(2/7).L.taxust3_t collapsed

Instruments for levels equation

Standard

tasa_paro

```

_cons
GMM-type (missing=0, separate instruments for each period unless collapsed)
DL.L.Gasto_Total collapsed
DL.L.taxust3_t collapsed
-----
Arellano-Bond test for AR(1) in first differences: z = -2.81 Pr > z = 0.005
Arellano-Bond test for AR(2) in first differences: z = 0.64 Pr > z = 0.523
-----
Sargan test of overid. restrictions: chi2(12) = 17.13 Prob > chi2 = 0.145
(Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(12) = 13.68 Prob > chi2 = 0.321
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:
GMM instruments for levels
Hansen test excluding group: chi2(10) = 12.08 Prob > chi2 = 0.280
Difference (null H = exogenous): chi2(2) = 1.60 Prob > chi2 = 0.449
gmm(L.taxust3_t, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 8.76 Prob > chi2 = 0.119
Difference (null H = exogenous): chi2(7) = 4.92 Prob > chi2 = 0.670
gmm(L.Gasto_Total, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 10.21 Prob > chi2 = 0.070
Difference (null H = exogenous): chi2(7) = 3.47 Prob > chi2 = 0.838
iv(tasa_paro)
Hansen test excluding group: chi2(11) = 13.68 Prob > chi2 = 0.251
Difference (null H = exogenous): chi2(1) = 0.00 Prob > chi2 = 0.990

```

En este caso, se ha reducido el número de instrumentos a 16, lo que junto con los p-valores relativos a los test de validación realizados indican que el modelo sigue siendo adecuado. Respecto a los coeficientes asociados a las variables explicativas, se ha incrementado la influencia de la tasa de mortalidad evitable estandarizada rezagada y disminuido ligeramente el coeficiente asociado al gasto sanitario total del bienio anterior. En relación al coeficiente asociado a la tasa de paro, que es el de mayor interés en este caso, vemos cómo tanto su valor como su significación han aumentado de manera considerable. De hecho, en este caso se tiene que bajo condiciones fijas del resto de regresores, un incremento de 1 punto en la tasa de paro provocaría un aumento de aproximadamente 0.4 en la tasa de mortalidad evitable estandarizada, mientras que al incluir el PIB per cápita en el modelo este aumento era de 0.28.

Por otra parte, al eliminar la tasa de paro del modelo dinámico inicial, sí apreciamos un cambio mayor, como se muestra a continuación.

```

>xtabond2 taxust3_t L.taxust3_t L.Gasto_Total pib_pc, gmm(l.taxust3_t, lag(2 7) collapse)
gmm(l.Gasto_Total, lag(2 7) collapse) iv (pib_pc) small robust twostep ortho

```

Favoring space over speed. To switch, type or click on mata: mata set matafavor speed, perm.

Dynamic panel-data estimation, two-step system GMM

```

-----
Group variable: ccaa                Number of obs   =    170
Time variable : bienio              Number of groups =    17
Number of instruments = 16          Obs per group:  min =    10
F(3, 16) = 841.05                  avg =    10.00
Prob > F = 0.000                   max =    10
-----

```

	Coef.	Corrected Std. Err.	t	P> t	[95% Conf. Interval]	
taxust3_t						
L1	.7486692	.1037084	7.22	0.000	.5288173	.9685211
Gasto_Total						
L1	-.0193917	.0076848	-2.52	0.023	-.0356827	-.0031007
pib_pc	-.0006918	.000332	-2.08	0.054	-.0013956	.0000121
_cons	79.7818	36.95727	2.16	0.046	1.435891	158.1277

Instruments for orthogonal deviations equation

```

Standard
FOD.pib_pc
GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/7).L.Gasto_Total collapsed
L(2/7).L.taxust3_t collapsed

```

```

Instruments for levels equation
Standard
  pib_pc
  _cons
GMM-type (missing=0, separate instruments for each period unless collapsed)
DL.L.Gasto_Total collapsed
DL.L.taxust3_t collapsed
-----
Arellano-Bond test for AR(1) in first differences: z = -2.75 Pr > z = 0.006
Arellano-Bond test for AR(2) in first differences: z = 0.88 Pr > z = 0.376
-----
Sargan test of overid. restrictions: chi2(12) = 18.84 Prob > chi2 = 0.092
(Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(12) = 13.86 Prob > chi2 = 0.310
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:
GMM instruments for levels
Hansen test excluding group: chi2(10) = 13.12 Prob > chi2 = 0.217
Difference (null H = exogenous): chi2(2) = 0.74 Prob > chi2 = 0.692
gmm(L.taxust3_t, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 10.58 Prob > chi2 = 0.060
Difference (null H = exogenous): chi2(7) = 3.28 Prob > chi2 = 0.858
gmm(L.Gasto_Total, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 9.47 Prob > chi2 = 0.092
Difference (null H = exogenous): chi2(7) = 4.39 Prob > chi2 = 0.734
iv(pib_pc)
Hansen test excluding group: chi2(11) = 13.73 Prob > chi2 = 0.248
Difference (null H = exogenous): chi2(1) = 0.13 Prob > chi2 = 0.723
    
```

Los resultados referentes a la validación del modelo no muestran ningún problema. Los coeficientes asociados resultan ser similares al caso inicial, apreciándose una ligera bajada en el coeficiente asociado al gasto sanitario total per cápita rezagado y, sobre todo una mayor significación de los coeficientes en general. Sin embargo, el interés del ajuste del presente modelo reside en el coeficiente asociado al PIB per cápita, el cual, al igual que ocurría en el caso anterior con la tasa de paro, muestra una influencia mayor sobre la tasa de mortalidad evitable. Además, éste se vuelve significativo al 10%, cuando anteriormente no lo era para ningún nivel usual.

Como nota adicional, se presentan en la Figura 4.10 los efectos individuales v_i relativos a cada comunidad autónoma obtenidos en estos dos últimos modelos ajustados. Cabe comentar que los resultados serán ligeramente diferentes al caso anterior, ya que por ejemplo, en el caso de eliminar la tasa de paro, este pasaría a formar parte de las características específicas de cada comunidad.

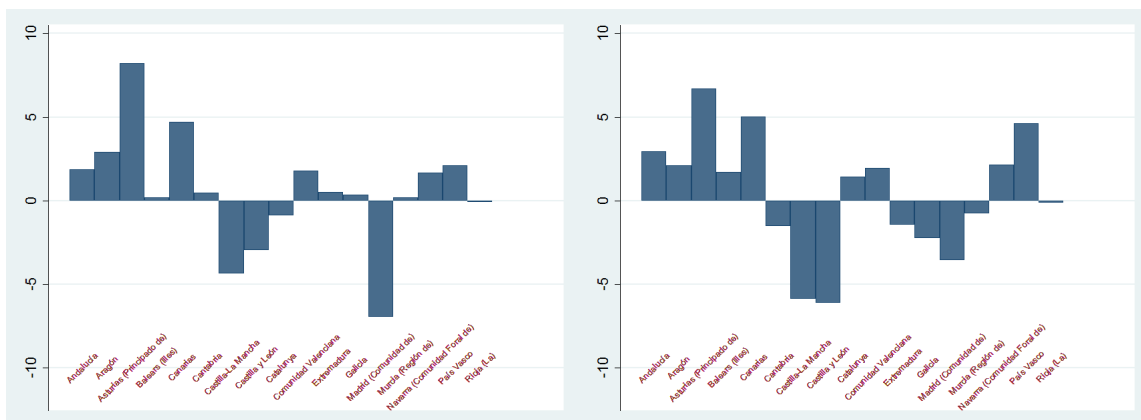


Figura 4.10: Efectos v_i para cada comunidad autónoma para el modelo sin la inclusión de la tasa de paro y sin el PIB per cápita, respectivamente.

Modelo dinámico con los años de esperanza de vida perdidos como variable respuesta

Una vez ajustados los diferentes modelos empleando como variable dependiente la tasa de mortalidad evitable estandarizada, debemos pasar al tratamiento de un modelo dinámico análogo considerando como variable respuesta el otro indicador de mortalidad evitable, los años de esperanza de vida perdidos. De este modo, la estimación GMM de sistema en dos etapas con las debidas correcciones para muestras pequeñas y heterocedásticas, reportan los resultados siguientes:

```
>xtabond2 aevp L.aevp L.Gasto_Total pib_pc tasa_paro, gmm(1.aevp, lag(2 7) collapse)
gmm(1.Gasto_Total, lag(2 7) collapse) iv(pib_pc tasa_paro) small robust twostep ortho
```

Favoring space over speed. To switch, type or click on mata: mata set matafavor speed, perm.

Dynamic panel-data estimation, two-step system GMM

```
-----
Group variable: ccaa                Number of obs   =       170
Time variable : bienio              Number of groups =        17
Number of instruments = 17          Obs per group:  min =         10
F(4, 16) = 369.17                  avg =          10.00
Prob > F = 0.000                   max =           10
-----
```

	aevp	Coef.	Corrected Std. Err.	t	P> t	[95% Conf. Interval]
aevp	L1.	.8105199	.0745407	10.87	0.000	.6525006 .9685391
Gasto_Total	L1.	-.0004442	.0001913	-2.32	0.034	-.0008496 -.0000387
pib_pc		-.0000132	9.17e-06	-1.44	0.170	-.0000326 6.27e-06
tasa_paro		-.0027642	.0038831	-0.71	0.487	-.0109961 .0054676
_cons		1.774101	.8170538	2.17	0.045	.042024 3.506177

Instruments for orthogonal deviations equation

Standard

FOD.(pib_pc tasa_paro)

GMM-type (missing=0, separate instruments for each period unless collapsed)

L(2/7).L.Gasto_Total collapsed

L(2/7).L.aevp collapsed

Instruments for levels equation

Standard

pib_pc tasa_paro

_cons

GMM-type (missing=0, separate instruments for each period unless collapsed)

DL.L.Gasto_Total collapsed

DL.L.aevp collapsed

Arellano-Bond test for AR(1) in first differences: z = -3.24 Pr > z = 0.001

Arellano-Bond test for AR(2) in first differences: z = 2.58 Pr > z = 0.010

Sargan test of overid. restrictions: chi2(12) = 28.75 Prob > chi2 = 0.004
(Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(12) = 13.61 Prob > chi2 = 0.326
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:

GMM instruments for levels

Hansen test excluding group: chi2(10) = 11.36 Prob > chi2 = 0.330

Difference (null H = exogenous): chi2(2) = 2.26 Prob > chi2 = 0.323

gmm(L.aevp, collapse lag(2 7))

Hansen test excluding group: chi2(5) = 7.80 Prob > chi2 = 0.168

Difference (null H = exogenous): chi2(7) = 5.82 Prob > chi2 = 0.561

gmm(L.Gasto_Total, collapse lag(2 7))

Hansen test excluding group: chi2(5) = 10.29 Prob > chi2 = 0.069

Difference (null H = exogenous): chi2(7) = 3.70 Prob > chi2 = 0.825

iv(pib_pc tasa_paro)

Hansen test excluding group: chi2(10) = 13.40 Prob > chi2 = 0.202

Difference (null H = exogenous): chi2(2) = 0.21 Prob > chi2 = 0.898

La estimación ejecutada concluye en una relación directa entre los años de esperanza de vida perdidos y su valor rezagado, con un coeficiente asociado algo mayor que en el caso de la tasa de mortalidad evitable como variable dependiente, y que resultó significativo. De igual modo, el coeficiente asociado

al gasto sanitario total per cápita retardado pone de manifiesto la relación inversa existente entre esta variable y los años de esperanza de vida perdidos, siendo también ésta significativa. Por tanto, se tiene que bajo condiciones fijas del resto de variables explicativas, un incremento de 100 euros en el gasto sanitario total per cápita del bienio anterior se asocia con una reducción de 0.04 en los años de esperanza de vida perdidos. Sin embargo, a diferencia de lo que ocurría en el caso del otro indicador de mortalidad evitable considerado, el coeficiente asociado a la tasa de paro resultó ser negativo, y sin carácter significativo, al igual que el del PIB per cápita. Este coeficiente negativo asociado a la tasa de paro, puede indicar que dado un PIB fijado, una mayor tasa de paro provoca una reducción en la mortalidad evitable. Por otra parte, se debe valorar también que éste derive de la alta correlación entre la tasa de paro y el PIB per cápita y la posible existencia de colinealidad entre las mismas, lo cual valoraremos posteriormente mediante el ajuste de sendos modelos excluyendo cada una de estas variables.

Con respecto a los efectos individuales μ_i , se obtiene una constante de $\mu = 1.77$ común a todas las comunidades autónomas, que resulta significativa, mientras que los v_i calculados a posteriori, muestran resultados similares al caso de la tasa de mortalidad evitable estandarizada en términos de qué comunidades poseen características específicas con influencia positiva o negativa en la mortalidad evitable.

En cuanto a la validación del presente modelo, los test de Hansen (tanto el global como el Difference-in-Hansen tests) reportan p-valores que no nos aportan motivos para sospechar de la no correcta especificación del modelo. El test de autocorrelaciones de segundo orden de Arellano y Bond sí resulta menor que 0.05, lo que indica la presencia de autocorrelación de segundo orden en el error, lo que provoca la inconsistencia del estimador. Sin embargo, este test, ante muestras con un N pequeño no es del todo fiable, lo que nos lleva a no rechazar la especificación de este modelo.

En este punto, es conveniente proceder al ajuste de sendos modelos a partir de éste, excluyendo respectivamente la tasa de paro primero y después el PIB per cápita, a fin de dilucidar si el coeficiente asociado a cada una de estas variables podría ser diferente en ausencia de la otra variable con la que está altamente correlacionada. En primer lugar, procedemos al ajuste del modelo con la exclusión de la tasa de paro como regresora.

```
>xtabond2 aevp L.aevp L.Gasto_Total pib_pc , gmm(1.aevp, lag(2 7) collapse)
gmm(1.Gasto_Total, lag(2 7) collapse) iv(pib_pc) small robust twostep ortho

Favoring space over speed. To switch, type or click on mata: mata set matafavor speed, perm.
```

```
Dynamic panel-data estimation, two-step system GMM
-----
Group variable: ccaa                                Number of obs   =    170
Time variable : bienio                             Number of groups =    17
Number of instruments = 16                          Obs per group:  min =    10
F(3, 16)      = 552.70                               avg =   10.00
Prob > F      = 0.000                                max =    10
-----
              |          Corrected
              |          Coef.   Std. Err.   t   P>|t|   [95% Conf. Interval]
-----+-----+-----
              |
aevp         |
  L1.        |   .8266553   .0825082   10.02  0.000   .6517458   1.001565
              |
Gasto_Total  |
  L1.        |  -.0004266   .0001902   -2.24  0.039  -.0008298  -.0000234
              |
pib_pc       |  -9.35e-06   9.03e-06   -1.04  0.316  -.0000285   9.78e-06
_cons       |   1.522129   .8873371    1.72  0.106  -.3589419   3.403199
-----
Instruments for orthogonal deviations equation
Standard
FDD.pib_pc
GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/7).L.Gasto_Total collapsed
L(2/8).L.aevp collapsed
Instruments for levels equation
```

```

Standard
  pib_pc
  _cons
GMM-type (missing=0, separate instruments for each period unless collapsed)
DL.L.Gasto_Total collapsed
DL.L.aevp collapsed
-----
Arellano-Bond test for AR(1) in first differences: z = -3.22 Pr > z = 0.001
Arellano-Bond test for AR(2) in first differences: z = 2.59 Pr > z = 0.010
-----
Sargan test of overid. restrictions: chi2(13) = 29.58 Prob > chi2 = 0.005
(Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(13) = 16.00 Prob > chi2 = 0.249
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:
GMM instruments for levels
Hansen test excluding group: chi2(11) = 13.13 Prob > chi2 = 0.285
Difference (null H = exogenous): chi2(2) = 2.87 Prob > chi2 = 0.238
gmm(L.aevp, collapse lag(2 8))
Hansen test excluding group: chi2(5) = 8.27 Prob > chi2 = 0.142
Difference (null H = exogenous): chi2(7) = 7.73 Prob > chi2 = 0.460
gmm(L.Gasto_Total, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 10.63 Prob > chi2 = 0.058
Difference (null H = exogenous): chi2(7) = 3.01 Prob > chi2 = 0.874
iv(pib_pc)
Hansen test excluding group: chi2(12) = 15.95 Prob > chi2 = 0.193
Difference (null H = exogenous): chi2(1) = 0.05 Prob > chi2 = 0.826

```

Como resultado obtenemos tanto valores de los coeficientes asociados a los regresores como significaciones de las mismas muy similares a los del modelo completo (considerando también la tasa de paro). Sin embargo, se aprecia la pérdida de significación de la constante μ común a cada una de las comunidades autónomas. La validación de este modelo nos conduce a las mismas conclusiones que en el caso anterior, por lo que no tenemos motivos para sospechar de una mala especificación.

Por otro lado, al excluir del modelo completo el PIB, y ajustar mediante el estimador GMM de sistema de Blundell y Bond en dos etapas de modo idéntico al resto de casos anteriores se obtiene:

```

. xtband2 aevp L.aevp L.Gasto_Total tasa_paro, gmm(l.aevp, lag(2 7) collapse) gmm(l.Gasto_Total, lag(2 7) collapse) iv( tasa_paro)
> small robust twostep ortho
Favoring space over speed. To switch, type or click on mata: mata set matafavor speed, perm.

```

```

Dynamic panel-data estimation, two-step system GMM
-----
Group variable: ccaa                Number of obs   =    170
Time variable : bienio              Number of groups =    17
Number of instruments = 16          Obs per group: min =    10
F(3, 16) = 718.66                  avg = 10.00
Prob > F = 0.000                    max = 10
-----

```

		Coef.	Corrected Std. Err.	t	P> t	[95% Conf. Interval]
aevp	L1.	.8322922	.0670904	12.41	0.000	.6900668 .9745175
Gasto_Total	L1.	-.0004622	.0001891	-2.44	0.026	-.0008631 -.0000613
tasa_paro		.002553	.0031807	0.80	0.434	-.0041897 .0092958
_cons		1.293062	.6134243	2.11	0.051	-.0073397 2.593463

```

-----
Instruments for orthogonal deviations equation
Standard
  FDD.tasa_paro
GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/7).L.Gasto_Total collapsed
L(2/7).L.aevp collapsed
Instruments for levels equation
Standard
  tasa_paro
  _cons

```

```

GMM-type (missing=0, separate instruments for each period unless collapsed)
DL.L.Gasto_Total collapsed
DL.L.aevp collapsed
-----
Arellano-Bond test for AR(1) in first differences: z = -3.25 Pr > z = 0.001
Arellano-Bond test for AR(2) in first differences: z = 2.59 Pr > z = 0.010
-----
Sargan test of overid. restrictions: chi2(12) = 27.66 Prob > chi2 = 0.006
(Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(12) = 13.17 Prob > chi2 = 0.357
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:
GMM instruments for levels
Hansen test excluding group: chi2(10) = 10.97 Prob > chi2 = 0.360
Difference (null H = exogenous): chi2(2) = 2.20 Prob > chi2 = 0.333
gmm(L.aevp, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 8.32 Prob > chi2 = 0.139
Difference (null H = exogenous): chi2(7) = 4.84 Prob > chi2 = 0.679
gmm(L.Gasto_Total, collapse lag(2 7))
Hansen test excluding group: chi2(5) = 11.02 Prob > chi2 = 0.051
Difference (null H = exogenous): chi2(7) = 2.35 Prob > chi2 = 0.892
iv(tasa_paro)
Hansen test excluding group: chi2(11) = 12.98 Prob > chi2 = 0.294
Difference (null H = exogenous): chi2(1) = 0.18 Prob > chi2 = 0.668

```

En la anterior salida de Stata, se puede observar cómo el coeficiente de la tasa de paro ha cambiado de signo respecto al del modelo completo, constatando el hecho de que el signo negativo del coeficiente del modelo inicial está causado por la alta correlación existente entre dicha variable y el PIB per cápita. Esto parece indicar que en el modelo completo se da un problema de colinealidad entre estas dos variables.

En términos de validación del modelo, los p-valores asociados a los diferentes test junto al número de instrumentos empleados nos llevan a conclusiones idénticas al caso previo, por lo que concluiremos en la correcta especificación del mismo.

Por último, en la Figura 4.11 se puede observar cómo los efectos individuales v_i que nos indican la influencia de las características específicas de cada comunidad autónoma sobre la mortalidad evitable, son similares al caso inicial en ambos modelos.

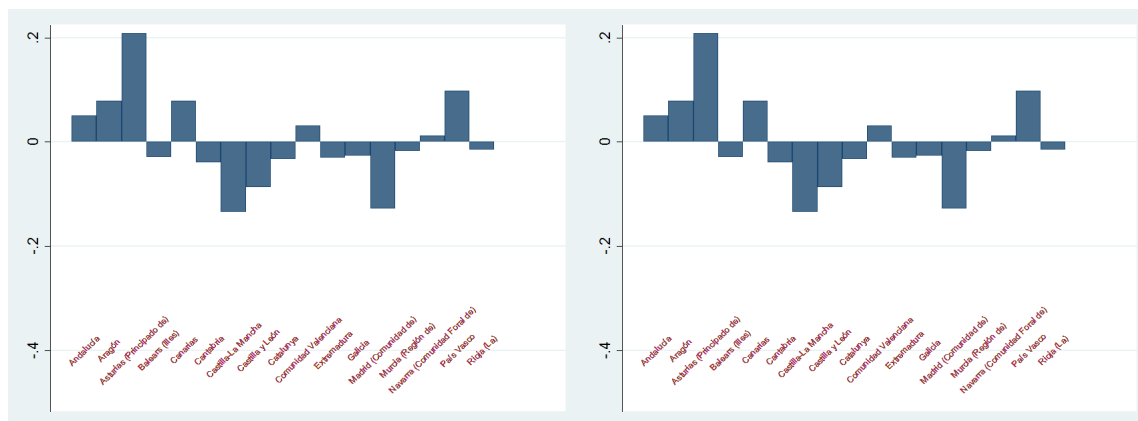


Figura 4.11: Efectos v_i para cada comunidad autónoma para el modelo sin la inclusión de la tasa de paro y sin el PIB per cápita, respectivamente.

En conclusión, el ajuste de modelos dinámicos empleando como indicador los años de esperanza de vida perdidos nos lleva a tener que valorar dos modelos diferenciados, al existir un problema de colinealidad entre la tasa de paro y el PIB per cápita. Ambos concluyen en que, bajo circunstancias

fijas del resto de regresores, el incremento de 100 euros en el gasto sanitario total per cápita se asocia con una reducción de aproximadamente 0.045 en los años de esperanza de vida perdidos. En cuanto a la tasa de paro y el PIB per cápita, no se obtiene una relación significativa respecto de la variable dependiente de los modelos. En cuanto a los efectos individuales, se obtienen resultados similares a los obtenidos para el modelo dinámico completo ajustado con la tasa de mortalidad evitable estandarizada.

Modelos dinámicos con el gasto sanitario per cápita desglosado

Para finalizar, procede comentar, que se ha valorado la consideración de modelos en los cuales la variable de gasto sanitario apareciera desglosada en 3 variables más específicas. Para ello, a partir de las 8 subcategorías en las que se divide el gasto sanitario total consolidado per cápita (Sección 3.2), se crearon las 3 variables de gasto principales como sigue:

- Gasto 1: Servicios hospitalarios y especializados
- Gasto 2: el gasto conjunto de las siguientes subcategorías:
 - Servicios primarios
 - Servicios de salud pública
 - Servicios colectivos de salud
 - Traslado, prótesis y cuidados terapéuticos
 - Transferencias
- Gasto 3: Farmacias

El objetivo de la creación de estas variables era utilizarlos en lugar del gasto sanitario total en los modelos de datos de panel, con el fin de obtener una información más detallada de la influencia que el gasto sanitario ha tenido sobre la mortalidad evitable en nuestro país a lo largo del periodo de estudio. Por desgracia, el carácter endógeno de estas variables de gasto, conllevan en la proliferación de instrumentos necesarios para llevar a cabo la estimación ($> N$), lo que convierte en algo imposible ajustar un modelo sin obtener resultados sesgados, lo que provocaría la no fiabilidad de los mismos.

En caso de ajustar un modelo reduciendo al máximo el número de instrumentos empleados para que este fuese menor o igual que N , estaríamos perdiendo demasiada información al tener que prescindir de la gran mayoría de los instrumentos disponibles, lo cual no evidenciaría resultados realistas. Por estas razones, se ha tenido que desestimar esta opción.

Conclusiones y Limitaciones

Como paso final, se exponen una serie de conclusiones globales derivadas del análisis temporal y geográfico de la mortalidad evitable en España desde 1990 hasta 2011 que se ha desarrollado en el presente texto mediante el ajuste de modelos de datos de panel, así como las limitaciones encontradas en la aplicación de estos métodos.

Conclusiones

- A lo largo del periodo de estudio, se ha observado un gran decrecimiento en la mortalidad evitable en España, y en concreto en todas y cada una de las comunidades autónomas (se excluyen Ceuta y Melilla).
- En cuanto a las listas de causas de mortalidad evitable, al comparar la evolución temporal de la tasa de mortalidad evitable estandarizada y su cambio relativo para cada una de las cuatro listas consideradas, se han apreciado diferencias que indican que la consideración de una u otra lista no resulta indiferente en cuanto a resultados.
- El ajuste de modelos de datos de panel nos condujo a la consideración de modelos dinámicos, donde la variable respuesta retardada es incluida como variable explicativa, apreciándose una relación directa entre ésta y la variable respuesta.
- En cuanto a la influencia de otras variables (gasto sanitario total per cápita del año anterior, PIB per cápita y tasa de paro) el ajuste de este tipo de modelos pone de manifiesto que bajo condiciones fijas de riqueza y desempleo, el incremento del gasto sanitario total per cápita se asocia con una disminución de la mortalidad evitable, tanto al considerar modelos con la tasa de mortalidad evitable estandarizada como con los años de esperanza de vida como variable respuesta.
- Sin embargo, los resultados referentes al PIB per cápita y a la tasa de paro son diferentes al tratar uno u otro indicador.

El empleo de la tasa de mortalidad evitable estandarizada concluye que el incremento de la tasa de paro, bajo condiciones fijas del resto de covariables, se asocia con un aumento de muertes de carácter evitable. Por otra parte, no se aprecia una influencia significativa del PIB per cápita sobre la variable dependiente. Sin embargo, al considerar un modelo sin incluir la tasa de paro, el PIB per cápita sí resulta significativo. En ese caso, se tiene que bajo condiciones fijas del resto de variables, un incremento de éste se asocia con una bajada de la tasa de mortalidad evitable estandarizada.

Al emplear los años de esperanza de vida perdidos como indicador, ni la tasa de paro ni el PIB per cápita resultan significativos.

- Geográficamente, destacar a Asturias y Canarias como aquellas comunidades autónomas donde se produce mayor mortalidad evitable debido a características específicas de la zona una vez

ajustado por las variables explicativas del modelo. Por otro lado, Castilla y León, Castilla-La-Mancha y Madrid son las comunidades donde menor mortalidad evitable se produce por estos motivos.

Limitaciones

- La necesidad de llevar a cabo el ajuste de modelos dinámicos de datos de panel, unido a la pequeña dimensión transversal de nuestra base de datos de panel ($N = 17$) y una dimensión temporal grande ($T = 22$), nos ha obligado a reducir la dimensión temporal, pasando a realizar un análisis por bienios, en lugar de uno anual como inicialmente se pretendía. Esto se debe a que la base de datos inicial, tiene una gran tendencia a la proliferación de instrumentos, lo que provoca resultados sesgados.
- Además, se ha valorado llevar a cabo el ajuste de modelos dinámicos de datos de panel considerando el gasto sanitario desglosado en tres variables, a fin de obtener una mayor información de la influencia de este sobre la mortalidad evitable. Sin embargo, el carácter endógeno de estas variables de gasto provoca la proliferación de instrumentos necesarios para llevar a cabo la estimación ($> N$), lo que convierte en algo imposible ajustar un modelo sin obtener resultados sesgados, con la consiguiente pérdida de fiabilidad de los mismos.

Apéndice A

Listas de causas de mortalidad evitable

En el presente Anexo se presentan de manera detallada cada una de las 4 listas de causas de mortalidad evitable empleadas para la realización de este estudio.

Para cada lista, se exponen la lista de causas de mortalidad evitable y sus respectivos códigos de acuerdo con la novena y décima Clasificación Internacional de Enfermedades (CIE-9 y CIE-10 respectivamente). Además, para las listas de Nolte y McKee(2204) y la de Gispert et al(2006), aparece el intervalo de edad en el cual, de ocurrir una muerte por una causa de la lista, se consideraría evitable.

Lista 1 Nolte y McKee(2004)

Causa	Edad	CIE-9	CIE-10
Infecciones intestinales	0-14	001-009	A00-A09
Tuberculosis y secuelas	0-74	010-018, 137	A15-A19, B90
Otras infecciones (difteria, tétanos, polio)	0-74	032, 037, 045	A36, A35, A80
Tos ferina	0-14	033	A37
Septicemia	0-74	038	A40-A41
Sarampión	1-14	055	B05
Tumor maligno de colon y recto	0-74	153-154	C18-C21
Tumor maligno de la piel	0-74	173	C44
Tumor maligno de mama femenino	0-74	174	C50
Tumor maligno de cuello uterino	0-74	180	C53
Tumor maligno de cuerpo uterino	0-44	179, 182	C54, C55
Tumor maligno de los testículos	0-74	186	C62
Enfermedad de Hodgkin	0-74	201	C81
Leucemia	0-44	204-208	C91-C95
Enfermedades de tiroides	0-74	240-246	E00-E07
Diabetes mellitus	0-49	250	E10-E14
Epilepsia	0-74	345	G40-G41
Cardiopatía reumática crónica	0-74	393-398	I05-I09
Enfermedades hipertensivas	0-74	401-405	I10-I13, I15
Enfermedades cerebrovasculares	0-74	430-438	I60-I69
Enfermedades del sistema respiratorio (excepto neumonía y gripe)	1-14	460-479, 488-519	J00-J09, J20-J99
Gripe	0-74	487	J10-J11
Neumonía	0-74	480-486	J12-J18
Úlceras pépticas	0-74	531-533	K25-K27
Enfermedades del apéndice	0-74	540-543	K35-K38
Hernia abdominal	0-74	550-553	K40-K46
Colelitiasis y colecistitis	0-74	574-575.1	K80-K81
Nefritis y nefrosis	0,74	580-589	N00-N07, N17-N19, N25-N27
Hiperplasia benigna de próstata	0-74	600	N40
Mortalidad materna (complicaciones del embarazo, parto y puerperio)	Todas	630-676	O00-O99
Malformaciones congénitas del sistema circulatorio	0-74	745-747	Q20-Q28
Afecciones originadas en el período perinatal	Todas	760-779	P00-P96, A33
Complicaciones debidas a atención médica o cirugía	Todas	E870-E876, E878-E879	Y60-Y69, Y83-Y84
Cardiopatía isquémica	0-74	410-414	I20-I25

Cuadro A.1: Causas de mortalidad evitable de la lista 1 (Nolte, 2004).

Lista 2: Gispert y cols.(2006)

Causa	Edad	CIE-9	CIE-10
Tuberculosis y secuelas	0-74	010-018, 137	A15-A19, B90
Cáncer de cuello uterino	15-74	180	C53
Cáncer de cuerpo uterino	15-74	182, 179	C54, C55
Enfermedad de Hodgkin	0-74	201	C81
Reumatismo cardíaco	0-74	393-398	I05-I09
Respiratorias aguda	0-74	480-486,460-466, 487	A48.1, J12-J18 (exc. J18.2), J00-J06 (exc. J02.0-J03.0), J20-J22, J10-J11
Asma	5-49	493	J45-J46
Enfermedad del apéndice	0-74	540-543	K35-K38
Hernia abdominal	0-74	550-553	K40-K46
Colicitiasis/Colecistitis	0-74	574-575	K80-K82
Hipertensión	0-74	401-405	I10-I15
Enfermedades cerebrovasculares	0-74	430-438	I60-69, G45, F01.1
Mortalidad materna	Todas	630-676	O00-O09, A34
Causas perinatales	Todas	760-779	P00-P96, A33
Cáncer de mama femenino	0-74	174	C50
Enfermedad isquémica del corazón	35-74	410-414	I20-I25
Úlceras pépticas	0-74	531-534	K25-K28
Enfermedades con vacuna	0-74	032, 037, 033, 055, 056, 072, 045, 070.0, 070.1, 070.2-70.3	A36, A35, A37, A49.2, B05, B06, B26, A80, B15, B16, B17.0, B18.0, B18.1
Anemias carenciais	0-74	280,281	D50-D53
Tumor maligno de la piel	0-74	172-173	C43, C44, C46.0, C46.9
Tumor maligno de testículos	0-74	186	C62
Leucemia	<15	204-208	C91-C95
Enfermedades de las tiroides	0-74	240-246	E00-E07
Diabetes Mellitus	0-49	250	E10-E14
Hiperplasia de próstata	0-74	600	N40
Anomalías congénitas cardiovasculares	0-74	745-747	Q20-Q28, I51.0
Latrogenia	Todas	E870-E879	Y60-Y84
Cáncer de pulmón	0-74	162	C33, C34
Hepatopatía alcohólica	15-74	571.0-571.3	K70.0, K70.1, K70.2-K70.3, K70.4, K70.9
Sida e infección por VIH	Todas	279.5, 042, 279.6, 795.8	B20-24, R75
Accidente de tráfico	Todas	E810-825	V02-V04, V09 (exc. V09.1 y V09.9), V12-V14, V19.0-V19.2, V19.4, V19.6, V20-V79, V80.3-V80.5, V81.0-V81.1, V82.0-V82.1, V83-V88, (exc. V88.9), v89 (exc. V89.1)
Suicidio	Todas	E950-959	X60-X84, Y87.0
Homicidio	Todas	E960-969	X85-Y09, Y87.1
Resto de causas externas	Todas	E800-807, E826-849, E850-858, E860-869, E880-949, E970-999	V01, V05-V06, V09.1, V09.9, V10-V11, V15-V18, V19.3, V19.8-V19.9, V80.0-V80.2, V80.6-V80.9, V81.2-V81.9, V82.2, V82.9, V88.9, V89.1, V90-V99, W00-W99, X00-X49, X50-X59, Y10-Y59, Y85-Y86, Y87.2, Y88-Y89, F10.0

Cuadro A.2: Causas de mortalidad evitable Lista 2 (Gispert, 2006) divididas en causas susceptibles de intervención por los servicios de asistencia sanitaria (primer bloque) y causas susceptibles de intervención por políticas sanitarias intersectoriales (segundo bloque).

Lista 3: AMEIHS(2011)

Causa	CIE-9	CIE-10
Infección por VIH	042-044	B20-B24
Tumor maligno de colon y recto	153-154	C18-C21
Tumor maligno de mama femenino	174	C50
Tumor maligno de cuello uterino	180	C53
Tumor maligno de los testículos	186	C62
Enfermedad de Hodgkin	201	C81
Leucemia linfocítica	204-205	C91
Cardiopatía reumática crónica	390-398	I00-I09
Enfermedades hipertensivas	401-404	I10-I13
Cardiopatía isquémica	410-414	I20-I25
Insuficiencia cardíaca	428-429	I50-I51
Enfermedades cerebrovasculares	430-438	I60-I69
Úlceras pépticas	531-532	K25-K26
Insuficiencia renal	584-586	N17-N19
Malformaciones congénitas del sistema circulatorio	745-746	Q20-Q24
Afecciones originadas en el período perinatal	760-779	P00-P96
Tumor maligno del estómago	151	C16
Tumor maligno de la laringe	161	C32
Tumor maligno de la tráquea, bronquios y pulmón	162	C33-C34
Tumor maligno de los huesos	170	C40-C41
Diabetes mellitus	250	E10-E14
Apendicitis aguda	540	K35
Hernia abdominal	550-553	K40-K46
Suicidio	E950-E958	X60-X84
Tuberculosis respiratoria	010-012	A15-A16, J65
Infección meningocócica	036	A39
Bronquitis y neumonía	466, 480-486	A48.1, J12-J18
Hepatitis	070	B15-B19
Tumor maligno del hígado	155	C22
Tumor maligno de la piel y labios	172-173	C43-C44
Tumor maligno del cuerpo uterino	182	C54
Tumor maligno del riñón	189	C64
Tumor maligno de la vejiga	188	C67
Anemia aplásica	284	D60-D61
Meningitis bacteriana	320	G00
Enfermedad de Parkinson	332.0	G20
Esclerosis múltiple	340	G35
Gripe	487	J10-J11
Bronquitis crónica, enfisema y obstrucción pulmonar crónica	491-492, 496	J41-J44
Asma	493	J45-J46
Obstrucción intestinal distinta de hernia	560	K56
Mortalidad materna (complicaciones del embarazo, parto y puerperio)	630-676	O00-O99
Malformaciones congénitas	740-759	Q00-Q99
Caídas accidentales	E880-E888	W00-W19
Muerte súbita infantil	798	R95

Cuadro A.3: Causas de mortalidad evitable (entre 0 y 74 años) de la lista 3 (AMEIHS, 2011).

Lista 4: Plug in (2012)

Causa	CIE-9	CIE-10
Enfermedades infecciosas		
Tuberculosis y secuelas	010-018, 137	A15-A19, B90
Neumonía y gripe	487, 480-486	J10-J18
Otras enfermedades infecciosas	Resto 001?139	Rest(A00-B99)
Cáncer		
Tumor maligno de cuello uterino	180	C53
Tumor maligno de los testículos	186	C62
Tumor maligno de colon y recto	153-154	C18-C21
Enfermedad de Hodgkin y leucemia	201,204-208	C81, C91-95
Condiciones cardiorrespiratorias		
Cardiopatía isquémica	410-414	I20-I25
Enfermedades cerebrovasculares	430-438	I60-69
Cardiopatía reumática crónica	390-398	I00-09
Enfermedades hipertensivas	401-405	I10-I15
Asma	493	J45-46
Otras enfermedades cardíacas	416,420-429	I26-I52, I98
Condiciones gastrointestinales		
Apendicitis, hernias, úlceras pépticas	531-534, 540?543, 550?553,560	K25-28, K35-38, K40-46, K56
Colelitiasis y colecistitis	574-576	K80-83

Cuadro A.4: Causas de mortalidad evitable (entre 0 e 74 años) de la lista 4 (Plug, 2012).

Apéndice B

Comparación listas de causas de muerte en las diferentes comunidades autónomas

En la sección 3.2, se analiza la evolución temporal de las tasas de mortalidad estandarizadas para el total de causas evitables, así como de su cambio relativo anual, con el objetivo de comparar las 4 listas de causas de mortalidad evitable consideradas en el presente estudio. Para ello se calcularon las tasas de mortalidad estandarizadas para el total de causas de muerte evitables y su cambio relativo anual a lo largo del periodo de estudio (1990-2011) para cada una de las 17 comunidades autónomas.

En la citada sección, únicamente se exponen los resultados correspondientes a la comunidad autónoma de Galicia (Figura 4.7). En este anexo, se ilustran los resultados análogos para el resto de comunidades autónomas.

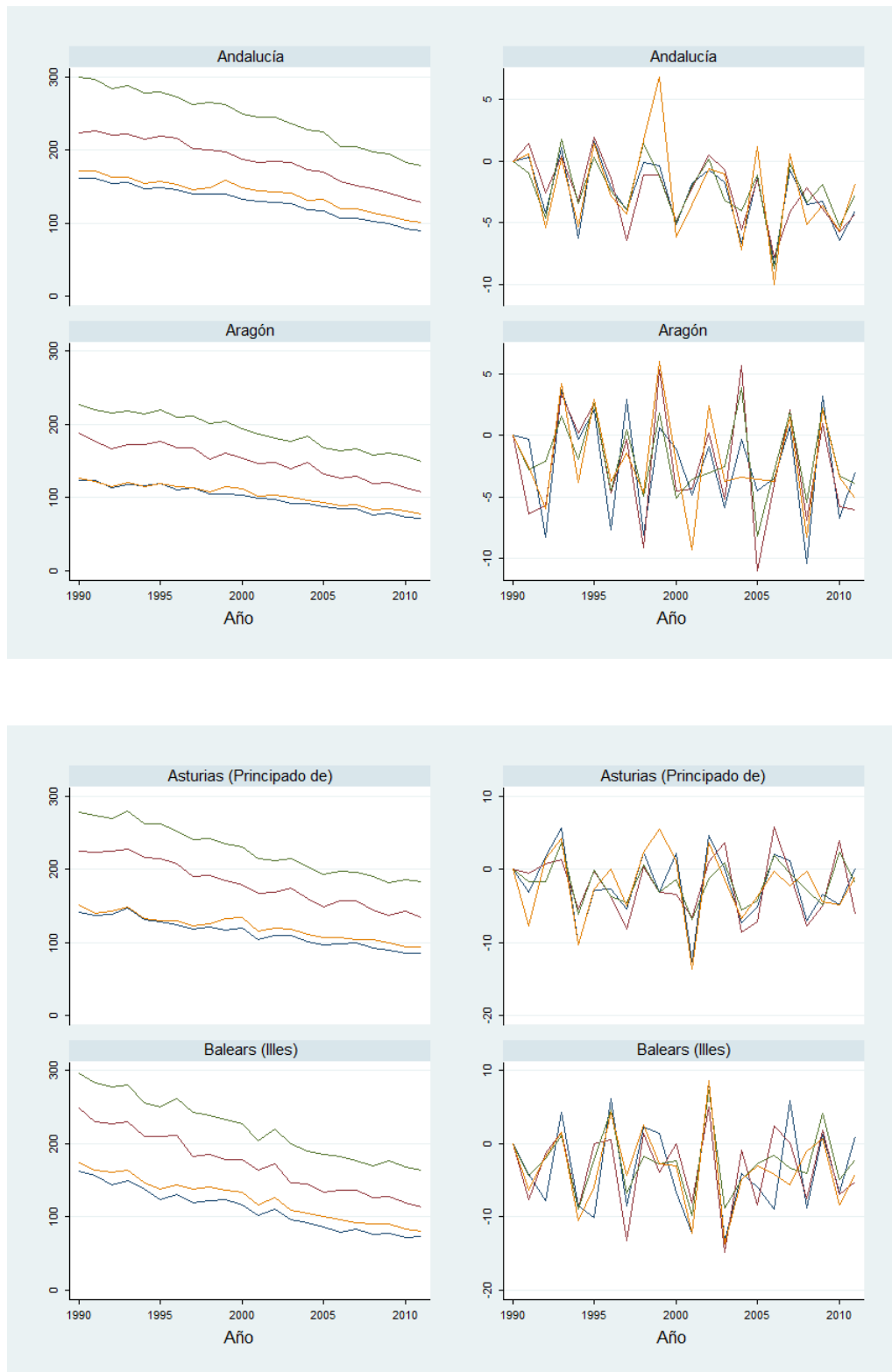


Figura B.1: Tasas de mortalidad estandarizadas para el total de causas de muerte evitables (izquierda) y su cambio relativo anual (derecha) para cada una de las 4 listas en las comunidades autónomas de Andalucía, Aragón, Asturias y Baleares respectivamente entre 1990 y 2011.

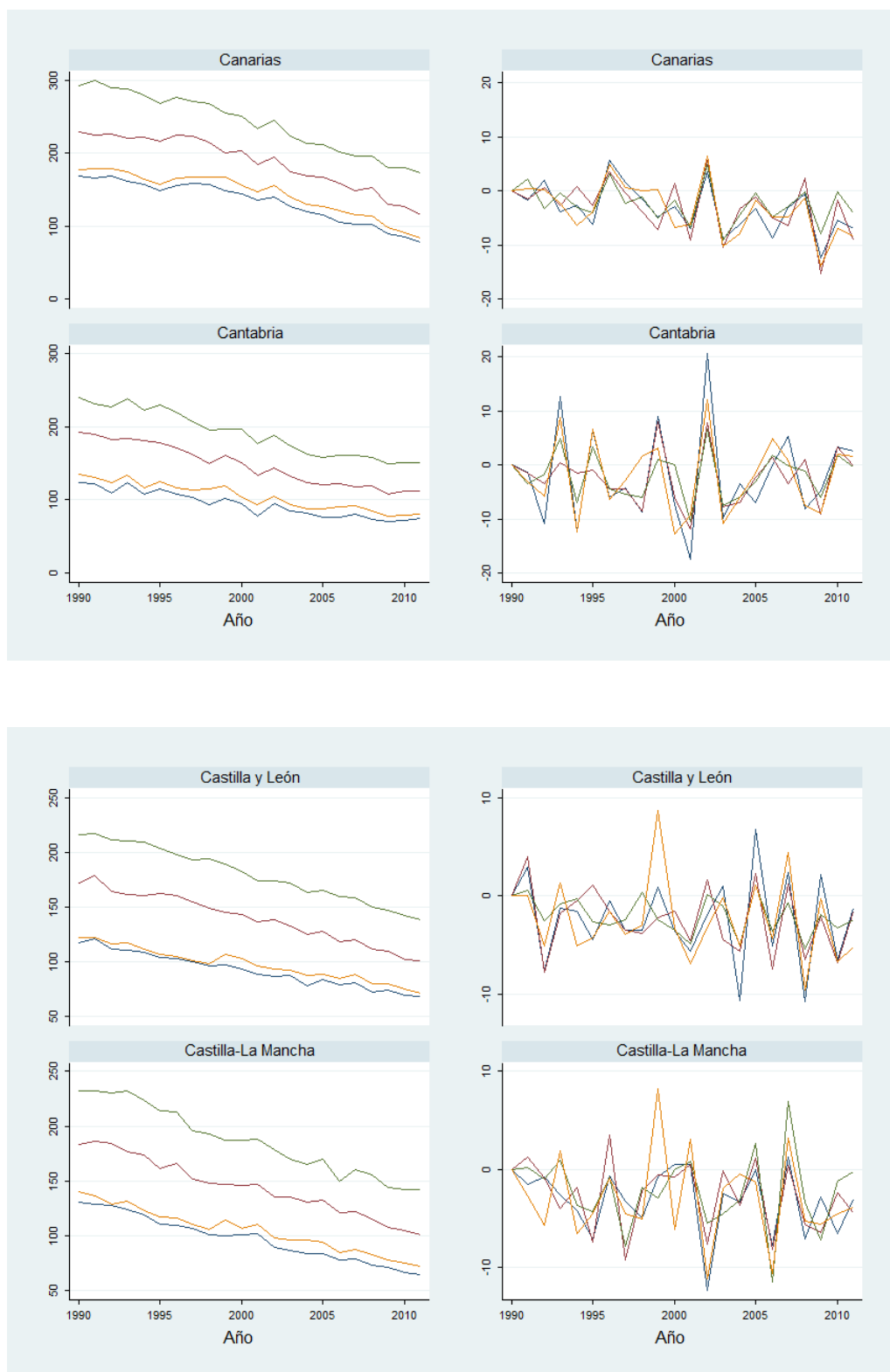


Figura B.2: Tasas de mortalidad estandarizadas para el total de causas de muerte evitables (izquierda) y su cambio relativo anual (derecha) para cada una de las 4 listas en las comunidades autónomas de Canarias, Cantabria, Castilla y León y Castilla La Mancha respectivamente entre 1990 y 2011.

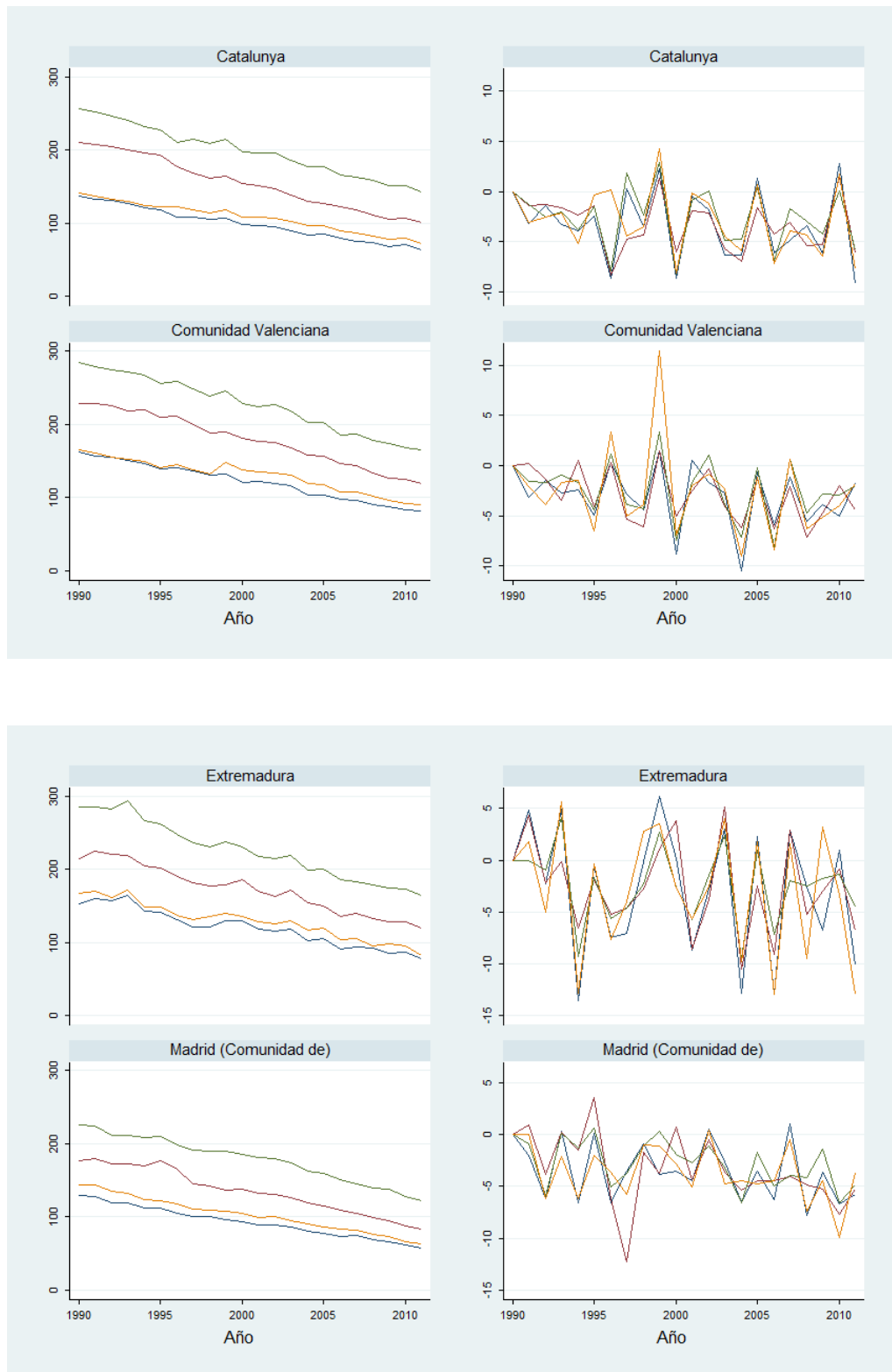


Figura B.3: Tasas de mortalidad estandarizadas para el total de causas de muerte evitables (izquierda) y su cambio relativo anual (derecha) para cada una de las 4 listas en las comunidades autónomas de Cataluña, Comunidad Valenciana, Extremadura y Madrid respectivamente entre 1990 y 2011.



Figura B.4: Tasas de mortalidad estandarizadas para el total de causas de muerte evitables (izquierda) y su cambio relativo anual (derecha) para cada una de las 4 listas en las comunidades autónomas de Murcia, Navarra, País Vasco y La Rioja respectivamente entre 1990 y 2011.

Apéndice C

Evolución temporal de las variables consideradas en las diferentes comunidades autónomas

En el presente apéndice se presentan como complemento al análisis temporal y geográfico de la mortalidad evitable en España distintos gráficos de la evolución temporal de las variables involucradas en los modelos de datos de panel expuestos en el Capítulo 4,

- Tasa de mortalidad evitable estandarizada
- Años de esperanza de vida perdidos
- Gasto sanitario total consolidado per cápita
- Producto Interior Bruto per cápita
- Tasa de paro

El propósito de estas representaciones será puramente descriptivo, y nos servirá de complemento a la hora de analizar los resultados obtenidos en las diferentes estimaciones de los diferentes modelos de datos de panel.

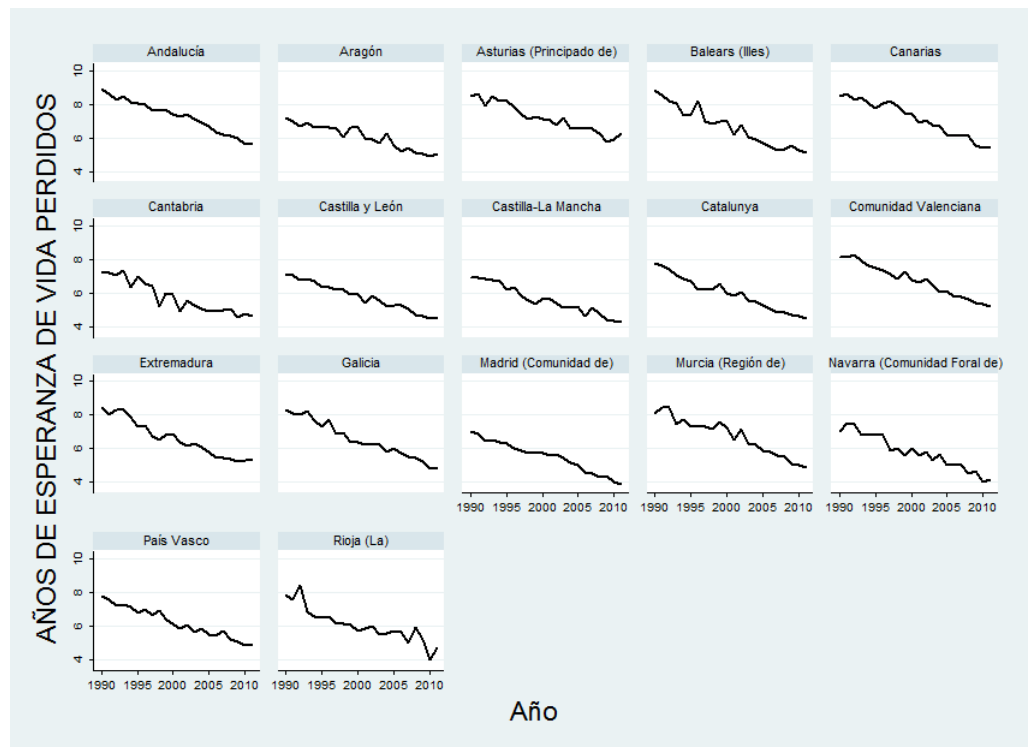
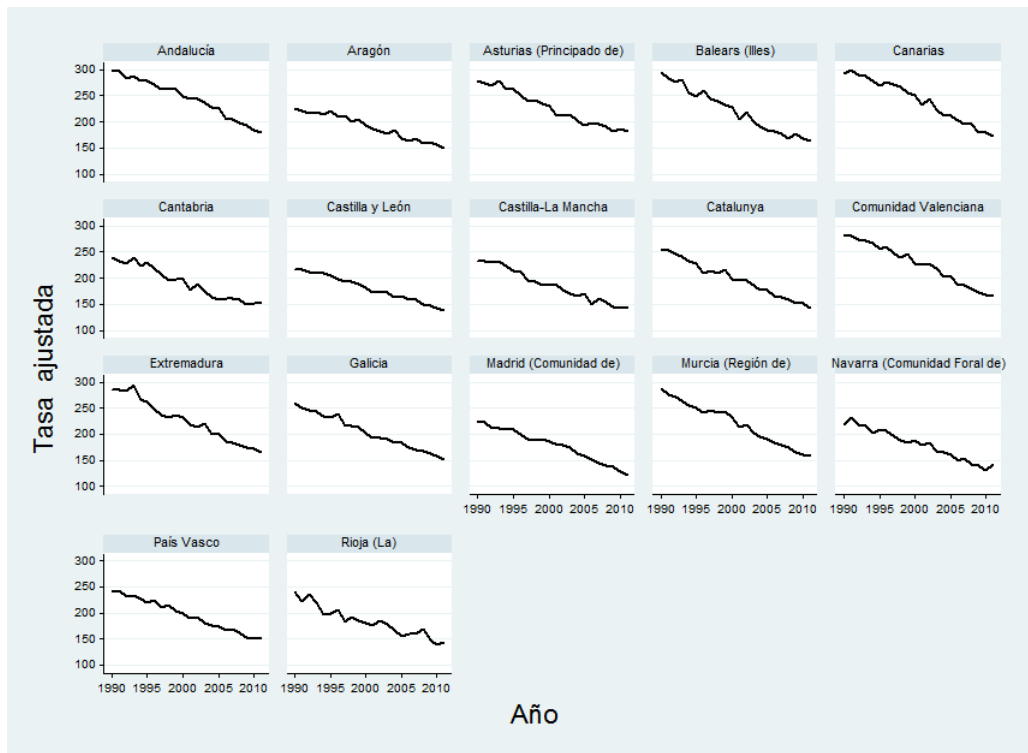


Figura C.1: Evolución temporal de la tasa de mortalidad evitable estandarizada (Arriba) y de los años de esperanza de vida perdidos (Abajo) por comunidad autónoma.

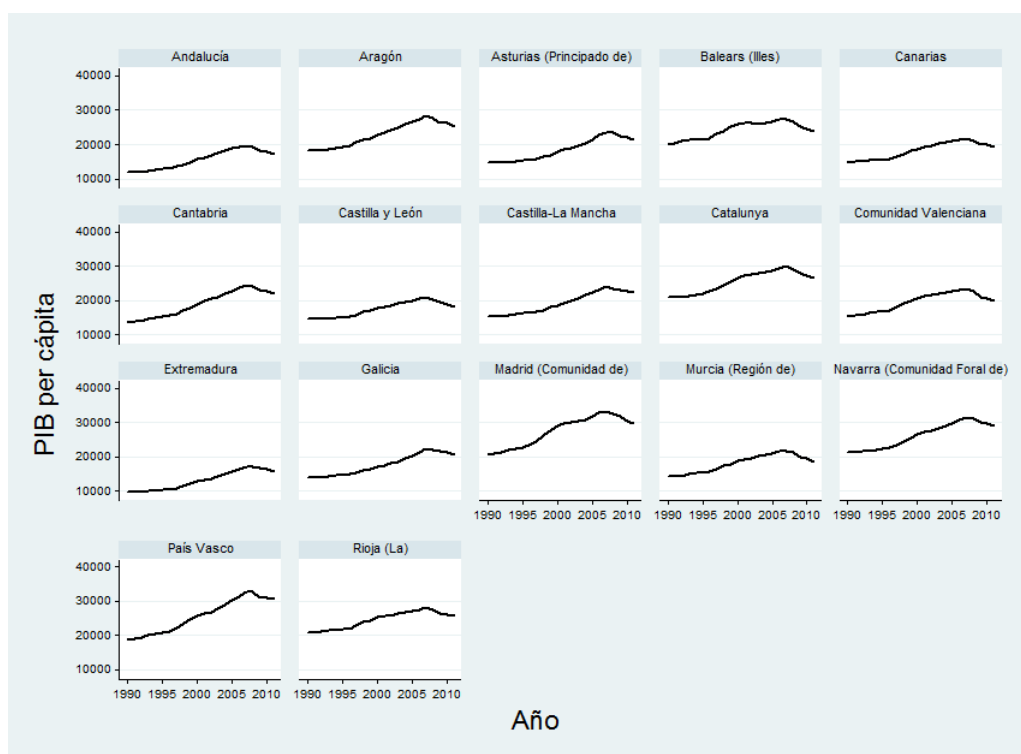
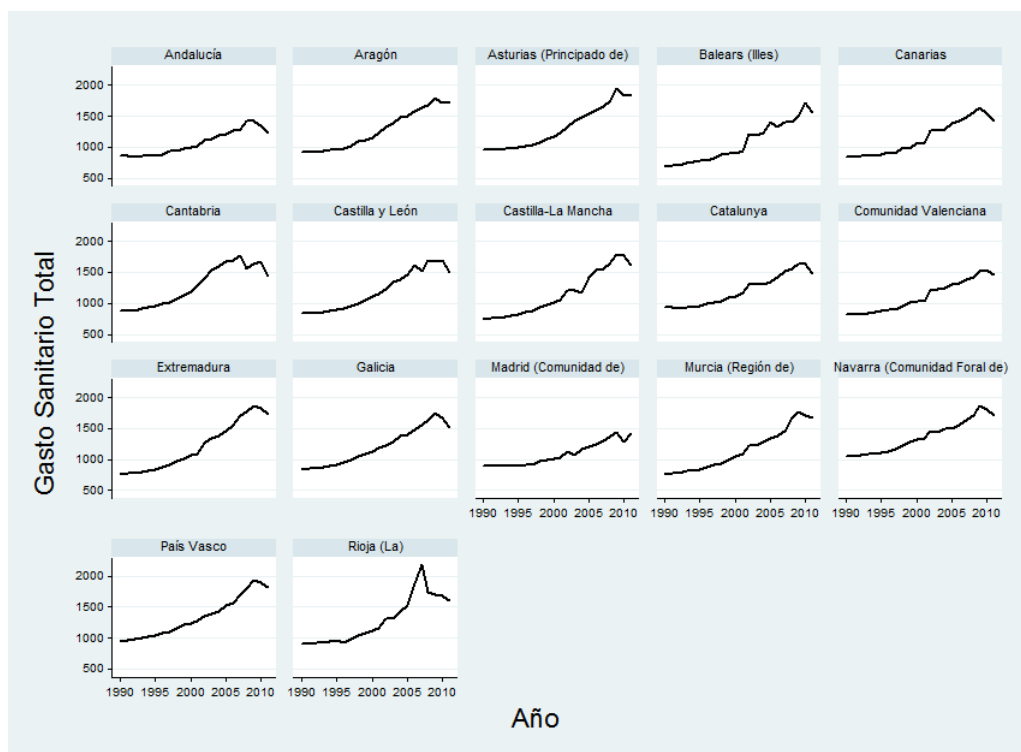


Figura C.2: Evolución temporal del gasto sanitario total consolidado per cápita (Arriba) y el PIB per cápita (Abajo) por comunidad autónoma.

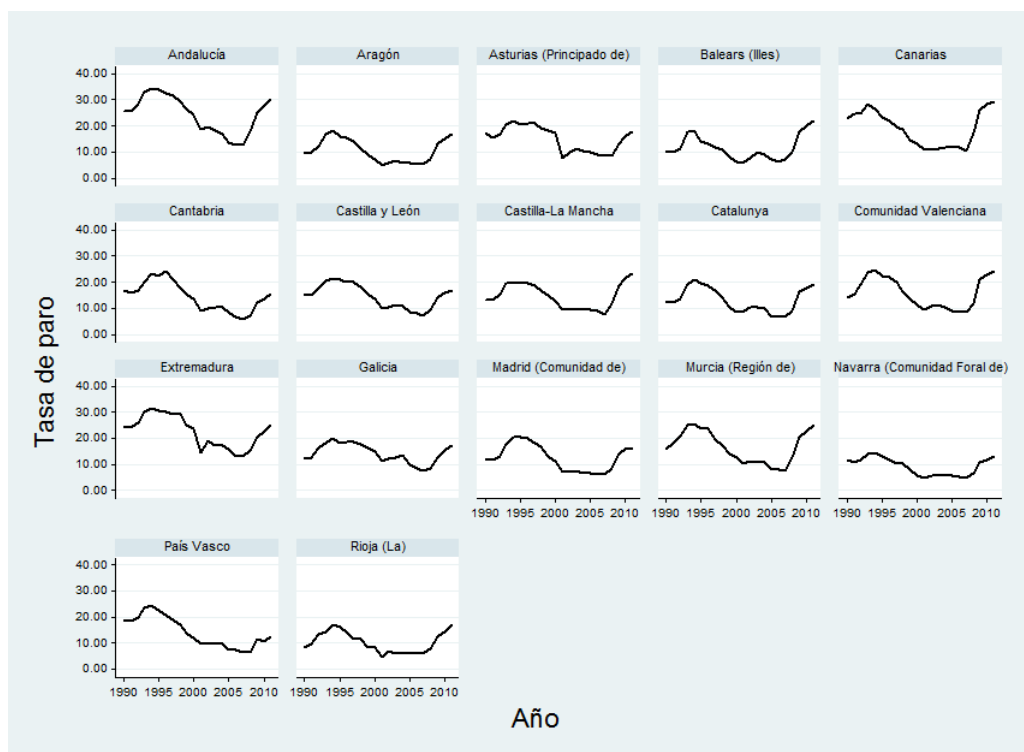


Figura C.3: Evolución temporal de la tasa de paro por comunidad autónoma.

Bibliografía

- [1] Albert X., Bayo A., Alfonso J. L., Cortina P. y Corella D. (1996). *Distribución geográfica de la mortalidad evitable en la Comunidad Valenciana(1975-1990)*. Med Clin (Barc) 1996; **106**, 571-577.
- [2] Andersen T. G. y Sorensen B. E.(1996). *GMM estimation of a sthochastic volatility model: a Monte Carlo study*. Journal of Business and Economic Statistics, **14**, 328-352.
- [3] Snderson T. W. y Hsiao C. (1982). *Formulation and estimation of dynamic models using panel data*. Journal of Econometrics, **18**, 47-82.
- [4] Arellano M., Bond S.(1991). *Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations*. Review of Economic Studies, **58**, 277-297.
- [5] Barés M., Gispert R., Puig X., Puigdefábregas A. y Tresserras R. (2005). *Distribución geográfica y evolución temporal de la mortalidad evitable en Cataluña (1986-2001)*. Gaceta Sanitaria, **19(4)**, 307-315.
- [6] Barrasa J. I., Castán S., Estupiñan F. R., Valderrama M. y Moliner F. J. (2013). *Mortalidad reducible como indicador de efectividad de los sevicios sanitarios en España antes y después de las transferencias*. Rev. Cal. Asist., **28(2)**, 84-95.
- [7] Blundell R. y Bond S. (1998). *Initial conditions and moment restrictions in dynamic panel data models*. Journal of Econometrics, **87(1)**, 115-143.
- [8] Bowsher C. G. (2002). *On testing overidentifying restrictions in dynamic panel data models*. Economic Letters, **77**, 211-220.
- [9] Breusch T. y Pagan A. (1980). *The LM test and its application to model specification in econometrics*. Review of Economic Studies, **47**, 239-254.
- [10] Charlton J. R. H., Hartley R., Silver, R. y Holland W. (1983). *Geographical variation in mortality from conditions amenable to medical intervention in England and Wales*. Lancet, **1**, 858.
- [11] Gispert R., De Arán Barés M., Puigdefábregas A. (2006). *La mortalidad evitable: lista de consenso para la actualización del indicador en España*. Gacet Sanit, **20(3)**, 184-93.
- [12] Gispert R., Barés M. A., Freitas A., Torné M., Puigdefábregas A., Alberquilla A., Albert X., Alfonso J. L., Caminal J., Fernández-Cuenca R., García F., Gervas J., Libroero J., Martos C., Medrano M. J. y Ruiz M. (2006). *Medida del resultado de las intervenciones sanitarias en España: una aproximación mediante el análisis temporal y espacial de la mortalidad evitable entre 1986-2001*. Revista Española de Salud Pública, **80**, 139-155.
- [13] Gispert R., Serra I., Barés M. A., Puig X., Puigdefábregas A. y Freitas A. (2008). *The impact of avoidable mortality at birth in Spain: changes between three periods, from 1987 to 2001*. Journal of Epidemiology Community Health, **62**, 783-789.
- [14] Hausman J. A. (1978). *Specification tests in Econometrics*. Econometrica, **46**, 1251-1272.

- [15] Hansen L. (1982). *Large sample properties of generalized method of moments estimators*. *Econometrica*, **50**(3), 1029-54.
- [16] Holland W. (Ed.). (1988). *European Community atlas of 'avoidable death'*. Commission of the European Communities Health Services Research Series No.3. Oxford: Oxford University Press.
- [17] Holland W. (Ed.). (1991). *European Community atlas of 'avoidable death' (2nd ed., Vol. I)*. Commission of the European Communities Health Services Research Series No.6. Oxford: Oxford University Press.
- [18] Holland W. (Ed.). (1993). *European Community atlas of 'avoidable death' (2nd ed., Vol. II)*. Commission of the European Communities Health Services Research Series No.9. Oxford: Oxford University Press.
- [19] Holland W. (Ed.). (1997). *European Community atlas of 'avoidable death' 1985-89*. Oxford: Oxford University Press.
- [20] Mackenbach J. P., Looman C., Kunst A., Habbema J. y van der Maas P. (1988). *Post-1950 mortality trends and medical care: gains in life expectancy due to declines in mortality from conditions amenable to medical interventions in Netherlands*. *Social Science Medicine*, **27**, 889-894
- [21] Melchor I., Nolasco A., García-Senchermes C., Pereyra-Zamora P., Pina J. A., Moncho J., Martínez P., Valero S. y Zurriaga O. (2008). *La mortalidad evitable. ¿Cambios en el nuevo siglo?*. *Gaceta Sanitaria*, **22**(3), 200-209.
- [22] Nickell S. (1981). *Biases in Dynamic Models with Fixed Effects*. *Econometrica*, **49**, 1417-1426.
- [23] Nolasco A., Melchor M., Pina J. A., Pereyra-Zamora P., Moncho J., Tamayo N., García-Senchermes C., Zurriaga O. y Martínez-Beneito M. A. (2009). *Preventable avoidable mortality: Evolution of socioeconomic inequalities in urban areas in Spain, 1996-2003*. *Health & Place* **15** (2009), 732-741
- [24] Nolte E. y McKee M. (2004). *Does health care save lives?. Avoidable mortality revisited*. London: The Nuffield Trust.
- [25] Ortún V., Gispert R. (1988). *Exploración de la mortalidad prematura como guía política sanitaria e indicador de caidad asistencial*. *Med Clin (Barc)*, **90**, 399-403.
- [26] Pérez G., Rodríguez-Sanz M., Cirera E., Pérez K., Puigpinós R. y Borrell C. (2014). *Approaches, strengths, and limitations of avoidable mortality*. *Journal of Public Health Policy*, Vol 35, 2, 171-184.
- [27] Roodman D. (2007). *How to do xtabond2: an introduction to Difference and System GMM in Stata*. *Stata Journal*, **9**(1), 86-136.
- [28] Roodman D. (2009). *A Note on the Theme of Too Many Instruments*. *Oxford Bulletin of Economics and Statistics*, Departamento de Economía, University of Oxford, **71**(1), 135-158.
- [29] Rutstein D. D., Berenberg W., Chalmers T. C., Child C. G., Fishman A. P. y Perrin E. B. (1976). *Measuring the quality of medical care. A clinical method*. *N Engl J Med*, **294**(11), 582-588.
- [30] Tobias M. y Jackson G. (2001). *Avoidable mortality in New Zealand, 1981-97*. *Aust N Z J Public Health* **25**, 12-20.
- [31] Tauchen G. (1986) *Statistical properties of generalized method-of-moments estimator of structural parameters obtained from financial market data*. *Journal of Business and Economic Statistics*, **4**, 397-416.

- [32] Vergara M., Benach J., Martínez J. M., Buxó M. y Yasui Y. (2009). *La mortalidad evitable y no evitable en áreas pequeñas de España (1990-2001)*. Gaceta Sanitaria, **23(1)**, 16-22.
- [33] Wheller L., Baker A., Griffiths C. y Rooney C. (2007). *Trends in avoidable mortality in England and Wales, 1993-2005*. Health Statistics Quarterly **34**, 6-25.
- [34] Ziliak J. P. (1997) *Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators*. Journal of Business and Economic Statistics, **16**, 419-431.
- [35] Windmeijer F. (2005). *A finite sample correction for the variance of linear efficient two-step GMM estimators*. Journal of Econometrics **126**. 25-51.
- [36] Wooldrige J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.