

Implementation of new methods of measurement and comparison of longevity in Europe

Patricia Carracedo¹, Ana María Debón²

¹patcarga@posgrado.upv.es, Centro de Gestión de Calidad y del Cambio, Universitat Politècnica de València

²andean@eio.upv.es, Centro de Gestión de Calidad y del Cambio, Universitat Politècnica de València

Abstract

The aim of this study is to show a method for detecting clusters of European countries with similar mortality. The method takes into account the geographical location of countries and the evolution of mortality along time. Given the space-time structure of the data has confirmed with a previously exploratory analysis, we have implemented panel data models. The results measure this space-time dependence, finding a interaction between the two components.

Keywords: mortality, clustering, panel data models

1. Introduccion

La mortalidad ha disminuido fuertemente en todos los países de la Unión Europea durante el siglo pasado. A pesar de estas tendencias generales, todavía hay considerable diferencias en los niveles de mortalidad entre los países [12].

Para poder identificar patrones de mortalidad en los países, es necesario el estudio modelos estocásticos, los cuales han sufrido un gran desarrollo en los últimos años [2, 3]. Así pues, la motivación de este trabajo es debido a que el interés de las desigualdades en salud entre los países y regiones de la UE, así como entre los diversos grupos sociales [10] está creciendo. Además, muchos estudios epidemiológicos europeos y nacionales muestran la diferencia entre las regiones norte y sur de la UE, dentro de los países, así como las regiones [6] y entre grupos socioeconómicos [13]. Por ello, el objetivo principal de este trabajo es presentar una metodología espacio-temporal para detección de grupos (clusters) de los países europeos con una mortalidad similar, teniendo en cuenta la situación geográfica de los países y, en consecuencia, las relaciones de vecindad entre ellos. Esta metodología será útil en varios dominios. En los mercados financieros y, especialmente, para los actuarios, para diseñar rentas vitalicias y seguros de vida. En el ámbito de las políticas públicas ayudan a los responsables políticos y los analistas, pues evalúan cuánto se benefician de los programas públicos, tales como la Seguridad Social y la asistencia sanitaria, los diferentes grupos de ciudadanos.

2. Materiales y Métodos

2.1. Datos

Los datos de panel o longitudinales son una combinación de datos de corte transversal ¹ y

¹Consta de una muestra de individuos, hogares, empresas, ciudades, países, etc. tomada en un momento de tiempo.

series temporales², es decir, cuando los mismos individuos son seguidos a través del tiempo. Por lo que los datos de panel tienen dos dimensiones la espacial y la temporal [15].

Normalmente, los datos de panel se distinguen unos de otros según su amplitud espacial y temporal. Así, los datos de panel con más observaciones espaciales que periodos de tiempo se denominan Paneles Micro. En el caso contrario, más periodos de tiempo que información espacial se conoce con el nombre de Paneles Macro. Por último, hablaremos de un Campo aleatorio (RandomField) ante un panel con amplia dimensión tanto temporal como espacial.

Así mismo, si no tenemos datos faltantes en los datos de panel, es decir, si cada unidad estudiada es observada en todos los periodos de tiempo, hablaríamos paneles equilibrados. Por contra, si hay datos faltantes, el periodo de tiempo varía entre individuos, en este caso estamos ante paneles no equilibrados [1].

Este estudio trata con datos de mortalidad de países europeos para el período comprendido entre 1990 y 2009 y para un rango de edad de cero a más de 110. Estos datos se han obtenido de la base de datos “Human Mortality Database” (<http://www.mortality.org/>) para un total de 26 países europeos: Austria, Bielorrusia, Bélgica, República Checa, Dinamarca, Estonia, Finlandia, Francia, Alemania, Hungría, Irlanda, Italia, Letonia, Lituania, Luxemburgo, Países Bajos, Noruega, Polonia, Portugal, Eslovaquia, Eslovenia, España, Suecia, Suiza, Reino Unido y Ucrania. Así pues, la variable dependiente en el modelo será el ratio de mortalidad estandarizado y como variables independientes la “Esperanza de vida al nacer” y “Tasa de actividad”. En concreto, nuestros datos son un panel Micro y equilibrado en donde hay 26 países y 20 años.

El análisis estadístico se realizó utilizando el software estadístico libre R [11], junto con algunos R-paquetes como demography [9] and splm [8].

2.2. Modelos espacio-temporales para datos en panel

Un modelo de datos de panel es aquel modelo de regresión que utiliza, para la estimación de los parámetros de interés, tanto la variabilidad temporal como la espacial de los datos. Éstos, permiten modelar la heterogeneidad no observada producida tanto por las unidades estudiadas como por el tiempo, dado que esta heterogeneidad no se puede detectar ni con estudios de series temporales ni tampoco con los de corte transversal [4].

Elhorst [7] proporciona una revisión detallada de las cuestiones que se plantean en la estimación de los modelos espaciales de datos de panel más comúnmente utilizados incluyendo la autocorrelación del error y/o el retardo de la variable dependiente. Estos cuatro modelos son:

- Modelo efectos fijos error espacial autocorrelacionado
- Modelo efectos fijos retardo espacial de la variable dependiente
- Modelo efectos aleatorios error espacial autocorrelacionado
- Modelo efectos fijos retardo espacial de la variable dependiente

²Consta de observaciones, de una o más variables a lo largo del tiempo.

Según la tipología de nuestros datos, el modelo que hemos seleccionado para éstos es un modelo de efectos fijos, ya que según [7] éste es más apropiado que el de efectos aleatorios cuando se trabaja con datos espacio-temporales de unidades espaciales adyacentes ubicadas en áreas de estudio ininterrumpidas. Además, consideramos el modelo de efectos fijos con retardo espacial en la variable dependiente y errores espaciales autorregresivos. El primero porque en un análisis exploratorio previo de los datos, se demostró que el valor del ratio de mortalidad estandarizado en un país depende del valor del ratio de mortalidad estandarizado en otro país adyacente y el segundo porque se desea estudiar si existe algún patrón espacial que se refleja en el término del error, es decir, las observaciones están relacionadas debido a factores no medidos, los cuales se correlacionan a través de las distancias entre las observaciones [14].

3. Resultados

Para el cálculo del modelo de efectos fijos con retardo espacial en la variable dependiente y errores espaciales autorregresivos se ha utilizado la librería `splm` de R. En la función `splm` especificamos la fórmula del modelo, los datos, matriz de pesos espaciales y la tipología de modelo. Los principales resultados del modelo han sido los siguientes:

- Tanto el parámetro espacial asociado a la variable dependiente (λ) como el parámetro espacial autorregresivo asociado al error del modelo (ρ) son significativos, es decir, existe una dependencia espacial entre las observaciones debido tanto a la variable dependiente como a los factores no medidos. Pero el ρ es más significativo que el λ , por lo que dependencia espacial entre las observaciones, en mayor medida, es debido al ρ .
- Ambas covariables son significativas, siendo el p-valor de la esperanza de vida al nacer menor que el de la tasa de actividad.
- Los efectos tanto espaciales como temporales son no significativos. Pero al representarlos gráficamente se observa que siguen una tendencia aproximadamente lineal y creciente. Se realiza la prueba F restrictiva para ver si individualmente o en conjunto estos efectos son significativos y por tanto, los debemos incluir en el modelo. La prueba concluye no podemos excluirlos del modelo.

4. Conclusión

Una vez confirmada la estructura espacio-tiempo en un análisis exploratorio previo de los datos [5], hemos implementado un modelo de efectos fijos con retardo espacial en la variable dependiente y errores espaciales autorregresivos. Los resultados del modelo confirman también esta dependencia espacial y temporal en nuestros datos.

5. Bibliography

- [1] Andreß, H., Golsch, K., and Schmidt, A. (2013). *Applied Panel Data Analysis for Economic and Social Surveys*. Springer.

- [2] Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22(3):547-581.
- [3] Booth, H. and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 3(1-2):3-43.
- [4] Burdisso, T. (1997). *Estimación de una función de costos para los bancos privados argentinos utilizando datos en panel*. Banco Central de la República Argentina, Área de Economía y Finanzas.
- [5] Carracedo, P., Debón, A., Iftimi, A., Montes, F., A cluster method for Dynamic Life Tables (2014). Poster in *Satellite workshop of the IBC2014 on Spatio-Temporal Statistics*.
- [6] Debón, A., Montes, F., and Martínez-Ruiz, F. (2011). Statistical methods to compare mortality for a group with non-divergent populations: an application to spanish regions. *European Actuarial Journal*, 1(2):291-308.
- [7] Elhorst, J. P. (2014). Spatial panel data models. In *Spatial Econometrics*, pages 37-93. Springer.
- [8] Giovanni Millo, Gianfranco Piras (2012). splm: Spatial Panel Data Models in R. *Journal of Statistical Software*, 47(1), 1-38. URL <http://www.jstatsoft.org/v47/i>
- [9] Hyndman, R. J., B. H. T. L. and Maindonald, J. . (2014). demography: Forecasting mortality, fertility, migration and population data. *R package version 1.17*. <http://CRAN.R-project.org/package=demography>.
- [10] Spinakis, A., Anastasiou, G., Panousis, V., Spiliopoulos, K., Palaiologou, S., and Yfantopoulos, J. (2011). *Expert Review and Proposals for Measurement of Health Inequalities in the European Union: Full Report*. European Commission.
- [11] Team, R. D. C. (2009). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*, ISBN 3-900051-07-0.
- [12] Vaupel, J. W., Zhang, Z., and van Raalte, A. A. (2011). *Life expectancy and disparity: an international comparison of life table data*. *BMJ open*, 1(1).
- [13] Villegas, A. M. and Haberman, S. (2014). On the modeling and forecasting of socioeconomic mortality differentials: An application to deprivation and mortality in england. *North American Actuarial Journal*, 18(1):168-193.
- [14] Ward, M. and Gleditsch, K. (2008). *Spatial Regression Models*. Number n.º 155 in Quantitative Applications in the Social Sciences. SAGE Publications.
- [15] Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.