

## 2.4. ANALISIS Y PREDICCIÓN DE LA POBLACION ESPAÑOLA 1910-2000 \*

D. ANTONIO GARCÍA FERRER  
D. JUAN DEL HOYO BERNAT  
*Catedráticos de Análisis Económico*  
Universidad Autónoma de Madrid

Nuestra preocupación fundamental en este trabajo era analizar el problema de lo que consideramos el input fundamental de todo análisis del sector de empleo, desempleo, temas de Seguridad Social, temas de pensiones; como es la evolución previsible de la población española.

A partir de ahí los estudios sectoriales, regionales y provinciales son absolutamente necesarios si bien, tras plantear algunos problemas con respecto a la calidad de los datos existentes cualquier investigador en esta dirección tiene un campo muy interesante de actuación y no exento de muchas dificultades, tanto teóricas como empíricas.

En su concepción general el esquema del libro trataba de analizar la evolución de la población española y las causas que la determinan. En este sentido el libro se divide en varios apartados. En uno de ellos se hace una aproximación económica al problema de cómo a lo largo del tiempo diferentes economistas han visto y explicado la situación coyuntural de su momento y nos encontramos con que existe un número importante de nombres que han dedicado tiempo y esfuerzo al estudio de este problema. Pero dejaremos al margen estas cuestiones porque una revisión mínima de las mismas excedería a los límites establecidos para esta exposición.

Nuestro interés fundamental para esta ponencia se centra en un análisis de la evolución demográfica española desde 1900 a 1980 haciendo una pequeña crítica a través del análisis de los datos de los problemas que hemos encontrado, de cómo han evolucionado los principales agregados demográficos, fundamentalmente: población, tasa de fecundidad, tasa de nupcialidad y tasa de mortalidad. Más tarde hablaremos de un primer análisis estadístico univariante basado en la metodología de series tempo-

\* El texto recoge las intervenciones de los profesores GARCÍA FERRER y DEL HOYO BERNAT en el I Congreso de Economía Regional de Castilla y León. La exposición está basada en el libro del mismo título recientemente publicado por estos autores y financiado por FEDEA.

rales y, por último, los efectos y problemas de trabajar con predicciones de tipo extrapolativo.

Posteriormente, vamos a analizar los problemas que nos hemos encontrado al tratar de reconstruir las tablas demográficas españolas por grupos de edad y sexo de año en año y el análisis de los resultados. Aunque en una primera etapa pensábamos que era relativamente sencillo resolver este problema utilizando los datos del movimiento natural de la población y los censos españoles, posteriormente descubrimos que no era así. Para terminar hablaremos de las predicciones poblacionales, de los problemas que tiene la estimación de distintos modelos econométricos donde aparecen relacionadas variables demográficas y socioeconómicas y los efectos que la nueva pirámide demográfica puede ocasionar para la sociedad española.

En nuestra opinión, los problemas metodológicos más importantes que un economista se encuentra cuando intenta efectuar predicciones poblacionales son diversos. Nuestra primera hipótesis de partida es la importancia fundamental de las variables económicas a la hora de explicar determinados cambios en la población. No creemos que sean únicamente los factores económicos los que determinan los cambios fundamentales en los agregados demográficos, pero si se ignora el valor de estas variables se cometerán errores de predicción muy importantes.

La característica fundamental que diferencia en los estudios demográficos a las variables demográficas de las variables económicas es la gran rapidez de los cambios económicos frente a otros posibles. Esas relaciones de «causalidad» no van exclusivamente en la dirección de variables económicas hacia variables demográficas sino que creemos que se pueden detectar relaciones en ambas direcciones e incluso simultáneas. Resulta muy difícil, por no decir imposible, efectuar predicciones demográficas y poblacionales si no somos capaces de efectuar al tiempo predicciones de las variables económicas relevantes con un grado de exactitud similar.

Las principales características de la evolución demográfica española son bien conocidas; los datos, como ya dijimos antes, no son excesivamente fiables. España ha seguido una transición demográfica similar aunque más lenta a la de sus países vecinos europeos. Resulta especialmente curioso el comportamiento de esta evolución en el período 1950-80 que presenta las cuotas máximas y mínimas del crecimiento demográfico.

El comportamiento contradictorio que en ciertos momentos del tiempo aparece entre el ciclo demográfico y el ciclo económico ha llevado a algunos autores a pensar que esa posible relación entre ambos ciclos no es tal.

En la evolución demográfica de este siglo hay tres catástrofes fundamentales (aunque la primera de ellas corresponde al siglo XIX va afectar a los nacimientos del siglo XX por el efecto de la translación):

- 1.<sup>a</sup> La epidemia de cólera de 1885, que vino a causar unos 120.000 muertos.
- 2.<sup>a</sup> La epidemia de gripe 1918-1919, que vino a causar 300.000 muertos.
- 3.<sup>a</sup> La Guerra Civil, sobre la cual hay multitud de cifras sobre el número de víctimas y que en base a diferentes fuentes nosotros consideramos que está en torno a 500.000 personas.

El análisis estadístico no puede ignorar estas influencias cualitativas porque de otra forma se estaría distorsionando. Si se analiza la evolución de la tasa de variación de la población total en el período considerado se detectan los problemas causados por las tres catástrofes comentadas.

Me gustaría terminar planteando algunas cuestiones respecto a la capacidad predictiva de los modelos que hemos utilizado. Una característica fundamental de la modelización ARIMA es que incluso cuando se plantean modelos con un gran número de parámetros resultan ser de memoria muy corta, lo que significa que se hacen muy rápidamente obsoletos con el paso del tiempo. Son modelos que funcionan bien para hacer predicciones a corto plazo pero que pueden producir predicciones a largo plazo bastante erróneas.

El objetivo de este trabajo ha consistido en conseguir predicciones de la población potencialmente activa para el año 2.000 a fin de elaborar políticas de empleo adecuadas. En este sentido, era fundamental, no sólo tener previsiones de la población total sino también por grupos de edad y sexo. Por ello, pensamos que si queríamos llevar a cabo previsiones desde el año 2000, era necesario recoger series lo más largas y minuciosas posibles, para estimar modelos predictivos fiables.

Para elaborar las series de población contábamos con dos tipos de información: una primera derivada de los censos de población y otra, del movimiento natural de población. Esta última permitía establecer la transición de un censo a otro, completando así, las series en los años intercensales.

El primer problema que se nos presentó fue que a la hora de analizar los censos se observaban unos «picos» elevadísimos que representaban grupos de edad con excesos poblacionales que no tenían, en principio, explicación. Estas anomalías, verosímilmente, son debidas a las técnicas de elaboración de censos de la época ya que, al no existir documentación adecuada, los individuos respondían a la pregunta de su edad, agrupándose en torno a las cifras terminadas en cero. De esta forma, si a estos grupos de edad los trasladamos en el tiempo aplicándoles sus tasas de supervivencia, el censo siguiente no concuerda en absoluto con el anterior. En todo caso, estos problemas estadísticos han ido atenuándose a lo largo del tiempo aunque, en los censos recientes aún aparecen vestigios de este problema.

A consecuencia de lo anterior, nos planteamos la forma de homogeneizar los censos y poder tener las cifras de población de cada grupo de edad, desde 1910 hasta 1980. A continuación, lo que hicimos fue buscar un modelo explicativo que nos permitiese extrapolar o predecir las series de población por cada grupo de edad hasta el año 2000.

El modelo que deseábamos debería relacionar, fundamentalmente, variables económicas y demográficas. Por lo que se refiere a las primeras, estarían presente variables de renta, de producción industrial, de cosechas, etc., que representarían índices de actividad económica a lo largo del período que estábamos analizando.

Una vez analizadas las distintas series obtenidas se llegó a la conclusión de que, en principio, no existía ningún modelo posible que explicara de forma sencilla la relación entre variables económicas y demográficas. No obstante, estudiando las series de una forma más exhaustiva, vimos que separando las tendencias de esas series podíamos elaborar dos modelos: uno de tendencias y otro de desviaciones de las tendencias. Con estos modelos se llegaba a resultados bastante buenos desde un punto de vista estadístico y, además, lográbamos el objetivo buscado de relacionar variables demográficas y económicas.

Puesto que lo que deseábamos era predecir la evolución de la población, necesitábamos, en principio, también obtener predicciones de las variables económicas involucradas en el modelo. Sin embargo, el modelo que estábamos utilizando era «de memoria reducida» de forma que, las series de renta que se introducían en el mismo figuraban con retardos muy cortos (por ejemplo, para explicar la tasa de natalidad

del año 1981 solamente era relevante la renta de los años 1978, 1979 y 1980). Esta situación no permitía predecir la población con fiabilidad hasta el año 2000 y, por ello, la solución consistió en buscar modelos específicos para predecir la población de menos de un año y, posteriormente, ir propagando por medio de predicciones de las tasas de supervivencia, de acuerdo también con modelos específicos, los distintos grupos de edad a partir del año base.

A lo largo de este trabajo también es evidente que uno de los fenómenos demográficos más importantes de la última década es la disminución de la natalidad. En nuestra opinión, esta disminución se debe específicamente al incremento de la tasa de actividad de las mujeres en edad fértil, edades comprendidas entre 19 y 40 años aproximadamente. La introducción en nuestros modelos de la evolución de la tasa de actividad de las mujeres en edad fértil eleva el grado de significación estadística desde un 50 por ciento hasta casi un 80 por ciento, lo que demuestra la singular importancia de esta variable.

Además, si se observan los datos referentes a la influencia que sobre las tasas de natalidad tiene la mujer trabajadora en edad fértil, se ve claramente como existe una diferencia importante entre el número de hijos habidos de mujeres activas y de mujeres no activas. Lógicamente, este diferente comportamiento va a tener consecuencias importantes no sólo para las instituciones del mercado laboral sino también para la organización de la sociedad española.

Así pues, y con las naturales reservas derivadas de las consideraciones anteriores, procedimos a propagar la estructura de la población española, tomando como base el año 1980, a partir de la predicción del número de nacimientos, junto con las tasas de supervivencia para los distintos grupos de edad. La previsión del número de nacimientos partió de dos hipótesis:

- a) Hipótesis optimista (asegura los niveles de natalidad hasta el año 2000 en niveles similares al final de los años setenta).
- b) Hipótesis tendencial o pesimista (el nivel de natalidad va cayendo hasta el año 2000).

Los resultados obtenidos bajo las dos hipótesis de natalidad permiten decir, como resumen, que la evolución esperada supone el envejecimiento progresivo de la población y, consecuentemente, del mercado laboral. Esta evolución tendrá grandes repercusiones sobre la organización de los sistemas de pensiones y de Seguridad Social.