



**GRADO EN ECONOMIA
CURSO ACADEMICO 2019/2020**

TRABAJO FIN DE GRADO

**DETERMINANTES SOCIODEMOGRAFICOS Y
DEMANDA DE SEGUROS SANITARIOS.**

**SOCIODEMOGRAPHIC DETERMINANTS AND DEMAND
FOR HEALTH INSURANCE**

**AUTOR: ESPERANZA-AGUEDA MBASOGO NGUEMA
EYANG**

TUTORA: SOBERON VELEZ ALEXANDRA PILAR

Septiembre de 2020

ÍNDICE

RESUMEN	3
1. INTRODUCCION.	5
1.1. Justificación.....	7
1.2. Objetivos	8
2. Revisión de Literatura	9
3.1 Estado de la cuestión	9
3.2. Modelos Empíricos: Demanda y utilización de Servicios Sanitarios....	13
3.2.1. Modelos de elección discreta	13
3.2.2. Modelos censurados	14
3.2.3. Modelos Count Data.....	15
3.2.4. Modelos Hurdle	15
3. Modelo y Procedimiento de estimación.....	16
3.1. Método de investigación.....	16
3.2. Población y Muestra.....	16
3.3. Estimación del modelo	17
3.4. Modelo econométrico	18
3.5. Análisis de la muestra	20
3.6. Estimación lineal (mínimos cuadrados).....	21

4. Resultados	22
5. Conclusiones.....	25
6. Bibliografía	28
7. Anexos	33
7.1. Gráficos de cajas para cada variable	33
7.2. Scripts en R.....	36
Nota: No computa para el número de páginas exigidas	36

RESUMEN

El presente trabajo tiene como objetivo determinar los factores sociodemográficos que intervienen en la demanda de seguros sanitarios en España.

Para este análisis se trabajará con dos muestras o cohortes. El primero será el correspondiente a los datos recogidos en la Encuesta Nacional de Salud antes del 2008 y el segundo serán los correspondientes después del antes mencionado año. Se ha utilizado un modelo econométrico dicotómico con variables discretas que se acopla al comportamiento de la demanda sanitaria (pública y privada) con respecto a la demanda de cobertura privada en funcionarios públicos. Para el análisis se utilizó la regresión logística binaria (logit) y el método probit de variables dependientes cualitativas relativas a sexo, edad, clase social, nivel de ingresos y estudios.

Se encontró que todas las variables sociodemográficas estudiadas influyen en la demanda de servicios sanitarios privados en la población española, tendencia que se mantuvo en los años 2011 y 2017 y que se había visto en decrecimiento entre 2006 y 2011. El sexo fue uno de los que más relevancia estadística ha mostrado en este aspecto, siendo que las mujeres son más proclives en adquirir este tipo de servicios que los varones. La clase social y el nivel de estudios también impactan en el criterio de decisión, siendo positivo para las clases sociales altas y en personas con estudios superiores culminados.

Palabras clave: modelo econométrico, seguros sanitarios, factores demográficos, logit, probit.

ABSTRACT

The present work aims to determine the sociodemographic factors that intervene in the demand for health insurance in Spain.

For this analysis, two samples or cohorts will be used. The first will be that corresponding to the data collected in the National Health Survey before 2008 and the second will be those corresponding after the aforementioned year.

A dichotomous econometric model with discrete variables has been used that is coupled to the behavior of health demand (public and private) with respect to the demand for private coverage in public officials. For the analysis, the binary logistic regression (logit) and the probit method of qualitative dependent variables related to sex, age, social class, income level and studies were used.

It was found that all the sociodemographic variables studied influence the demand for private health services in the Spanish population, a trend that was maintained in 2011 and 2017 and that had decreased between 2006 and 2011. Sex was one of those that More statistical relevance has been shown in this regard, with women being more likely to acquire this type of service than men. Social class and level of studies also impact the decision criteria, being positive for upper social classes and for people with completed higher education.

1. INTRODUCCION.

El análisis de indicadores relacionados al contexto sociodemográfico se fundamenta en la integración de conceptos relacionados con desigualdad y las dificultades en el acceso a derechos fundamentales por parte de una población en particular (Reiners et al., 2019). De esto se desprende la importancia en abordar la existencia de desigualdades significativas en salud en Europa, el cual, de acuerdo con Hernández Quevedo & Jiménez Rubio (2009), suele beneficiar más aquella población con mayor renta. El grado de desigualdad viene relacionado esencialmente a factores mesurables como el nivel educativo, la ocupación actual del individuo y el ingreso económico mensual por familia (Costa et al., 2014). Más concretamente en España, este ocupa el quinto lugar entre los países con mayores tasas de pobreza de la Unión Europea¹

Al respecto vale destacar algunos estudios realizados en España que han explorado la prevalencia y correlación entre la existencia de la desigualdad en salud y el nivel de ingresos, encontrándose como resultado una inequidad en el acceso a la atención sanitaria para la población española en su conjunto, lo que incluye a nativos como población migrante (Hernández Quevedo & Jiménez Rubio, 2009, 2009; Jiménez-Martín & Jorgensen, 2009).

En los tiempos que corren se puede afirmar sin mucha dilación que los cambios sociodemográficos causados por los movimientos migratorios a escala internacional representan un elemento central del debate público y político en Europa (Bas-Sarmiento et al., 2015; Rechel et al., 2013). Los cambios de este tipo han supuesto una profunda transformación de los diversos sistemas de salud encada país miembro en aras de satisfacer las necesidades de los nuevos ciudadanos y residentes y, que además deberán mantener la coexistencia diferentes enfoques de salud, enfermedad y atención (Henderson et al., 2018). Como consecuencia, se requiere adoptar una perspectiva focalizada en la diversidad de la gestión de los servicios de salud en el marco de la inclusividad y conocimiento de las necesidades sociales de este tipo de población.

¹De acuerdo al informe Eurostat (2017): https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Income_poverty_statistics/es

Ante tal hecho, el fenómeno de la inmigración creciente toma importancia para las entidades gubernamentales para lo relacionado a los servicios de salud a través de la aprobación de la Ley 4/2000 del 11 de enero sobre los derechos y libertades de los extranjeros en España, mediante la cual se materializa que todos los individuos, sin importar la nacionalidad de procedencia, tendrán acceso a los servicios sanitarios bajo las mismas condiciones que los ciudadanos españoles, y cuyo único requisito para dicho acceso sería estar registrado en el padrón de su localidad habitual.

Vale destacar además el Plan Estratégico de Ciudadanía e Integración 2007–2010, que pretendía promover la cohesión social a través de políticas asentadas en la igualdad de oportunidades y derechos y deberes. Como consecuencia, también se han desarrollado planes de inserción e inclusión a nivel regional en las comunidades autónomas, que incluyen de manera prioritaria la disminución de tasas de desigualdades en salud y en el acceso a los servicios sanitarios para este colectivo.

En tal sentido, la literatura científica se mantiene consistente en cuanto a mostrar que la salud de los inmigrantes puede verse amenazada debido a la prevalencia de barreras que les dificulta el acceso al sistema de salud, bien sea por una falta de sensibilidad y adaptación a una población diversa y multicultural, y por el «*status quo*» de los migrantes mismos (Gea-Sánchez et al., 2017; Porthé et al., 2018; Terraza et al., 2010). De esto se desprende el hecho que las experiencias con otros sistemas de salud pueden influir en sus expectativas y el concepto de salud-enfermedad y atención, y podrían entrar en conflicto con nuestras prácticas de atención médica, siguiendo un esquema más occidental y más ortodoxo científicamente hablando (Ruiz et al., 2020; Urrutia, 2018; Yagüe Ruipérez, 2019). Asimismo, influyen las circunstancias del proceso de migración y la ignorancia de los servicios disponibles o cómo abordarlos.

Por consiguiente, en algunos estudios se ha determinado que la población migrante podría encontrar obstáculos que dificultan el acceso, el uso y la navegación en los servicios de salud, debido, entre otras cosas, a requisitos legales, burocracia, dificultades de lenguaje y baja alfabetización en salud (Albar et al., 2015; González, 2013). Estas barreras influyen en el deterioro de su bienestar, son una fuente de desigualdades porque son factores determinantes de su salud y crean situaciones de desigualdad que podrían evitarse.

Ante tales circunstancias el presente trabajo tiene como propósito cuantificar y validar el modelo econométrico propuesto por Jiménez-Martín & Jorgensen (2009) que relaciona el estudio de las desigualdades socioeconómicas y demográficas frente al acceso que tienen tanto españoles nativos como población migrante a los servicios sanitarios. Para ello, se tomará en consideración la relación entre el nivel de estudios, ocupación actual del individuo e ingresos económicos con respecto a la demanda por seguros sanitarios, sean estos de carácter privado, público o doble cobertura.

Por consiguiente, se realizará una revisión exhaustiva de la bibliografía relacionada con este tópico siguiendo el paradigma documental de cribado y selección de referencias (Meth & Williams, 2006). Con ello se pretende construir un estado del arte lo suficientemente riguroso en aras de comprender mejor los posibles resultados estadísticos obtenidos del modelo econométrico.

Seguidamente, se señala el uso de la ecuación original del modelo presentado por Jiménez-Martín & Jorgensen (2009) y de esta se tomará como puntos a considerar el los factores sociodemográficos de toda la población española (incluida la población inmigrante) para determinar la correlación de estos estados y la demanda de servicios sanitarios.

1.1. Justificación

Cuando se hace referencia al Sistema Nacional de Salud (SNS) es necesario enfatizar crecimiento poblacional que ha tenido que afrontar en los últimos 20 años. Actualmente atiende a más de 8 millones de personas donde tres cuartas partes son pertenecientes a la población migrante (Picatoste et al., 2018).

A este crecimiento poblacional se debe sumar la excepcional heterogeneidad del colectivo inmigrante y el movimiento interno de nativos en toda la península que ha llevado a concentración demográfica en determinadas comunidades autónomas. Como consecuencia, el SNS ha tenido que hacer frente a una presión extra llevándola a adaptar su oferta y evitar posible colapso y/o congestión (Fayos Jordán, 2017).

En este mismo orden de ideas, el carácter universal que esgrime el SNS en España se encamina en favor del acceso a las personas con situación de vulnerabilidad

(inmigrantes y nativos de clase baja) a los servicios y, desde luego, esto se convierte un instrumento eficaz hacia las políticas de integración (Peiró & Ridaó, 2018). Sin embargo, también se ha visto que el SNS ha tenido que redirigir esfuerzos en aras de ofrecer una respuesta adaptada a las demandas surgidas de los nuevos movimientos socioculturales y demográficos de la actualidad.

Estudios de la misma índole también han encontrado que la demanda de los servicios SNS por la población inmigrante, en líneas generales, suele ser similar a la de los autóctonos que comparten un mismo rango de edad (Canals Barceló, 2018; López Nicolás & Ramos Parreño, 2009). Tomando en cuenta además que estas poblaciones cuentan con mayor número de jóvenes y en líneas generales con un mejor estado de salud, el impacto sobre la actividad del SNS puede llegar a resultar menor en proporción a la fracción que representa la población migrante frente la población española autóctona. Pese a ello, el sistema de financiamiento no ha sido del todo eficaz en cuanto a evitar colapsos del sistema sanitario a mediano y largo plazo.

Así pues, en vista de la delicada relación entre el número de población vulnerable, tanto españoles como inmigrantes, cabe preguntarse si factores sociodemográficos como el nivel educativo, ocupación actual y nivel de estudios pueden influenciar de manera proporcional la demanda sanitaria en general. Esto incluye no solo a los servicios sanitarios estatales, sino además de la demanda de servicios de carácter privado, siendo esta una alternativa que surge como respuesta a la falta de confianza en la solidez del esquema estatal (cada vez más pensionados y menos cotizantes) y a la eficiencia en cuanto a la atención del paciente (tiempos de espera menor).

1.2. Objetivos

En vista de todo lo antes mencionado, surge este trabajo cuyo objetivo principal es identificar los factores sociodemográficos en la población nativa e inmigrante que determinan el uso y demanda de seguros sanitarios y servicios afines en España.

Por consiguiente, será necesario analizar los datos de las ENS (Encuesta Nacional de Salud) dividiendo en dos grupos, el primero conformado por aquellos antes del 2008 (antes de la crisis) y el segundo por aquellos de la era postcrisis.

Seguidamente se tendrá que validar el modelo dicotómico de elección discreta modelizando la demanda sanitaria de población total (inmigrante y autóctona) y determinar la probabilidad de que los factores sociodemográficos antes mencionados sean determinantes en cuanto al acceso y calidad de servicio sanitario.

2. Revisión de Literatura

3.1 Estado de la cuestión

Mesurar la eficiencia de los servicios sanitarios en función de su coste, cantidad y calidad de bienes y servicios que produce, es un punto fundamental que concierne tener con claridad al momento de formalizar y seleccionar el modelo econométrico más adecuado y con menor error relativo. Tal como señalaba González Álvarez & Clavero Barranquero (2005), es razonable pensar que un servicio sanitario puede producir mayores niveles de salud al tiempo que contribuyen a la disminución de los casos de enfermedad y evitar muertes prematuras, siendo este un factor estrechamente relacionado con factores económicos como la redistribución de la renta pública. De ello se desprende que el análisis de demanda y uso de los servicios de salud representa un aspecto clave de la investigación en economía de la salud, ya que, en ambas, además, influyen también muchos otros aspectos del sistema de salud, como la calidad de la oferta, la planificación estratégica y su configuración en cada país o territorio (Kabongo & Mbonigaba, 2016).

La relevancia en estudiar la demanda y uso de los servicios de salud desde un punto de vista cuantitativo estriba en la identificación de los factores que tienen mayor importancia en el incremento observado las últimas décadas, especialmente por los movimientos económicos y sociales de gran impacto que han tomado lugar (Catalán et al., 2012). Este tipo de análisis también se derivan en diferencias en el uso tanto a nivel individual como en función de las áreas geográficas donde se aplican.

Conocer a fondo estas variables y su influencia determinista en la oferta y demanda de servicios sanitarios se esgrime como un instrumento avocado en diseñar políticas diferenciadas relativas al individuo, el profesional de la salud o las autoridades sanitarias pertinentes.

Tomando en cuenta estos aspectos se han llevado a cabo diversos estudios en esta materia con el propósito de establecer un modelo matemático que permita simular la relación entre factores sociodemográficos de una determinada población y la demanda de los servicios sanitarios; siendo este último un indicador de calidad de vida de los habitantes de una región concreta. En tal sentido se podría comenzar por destacar el estudio publicado por Jiménez-Martín & Jorgensen (2009) quienes plantearon un modelo dicotómico de elección discreta para modelar la demanda de doble cobertura (pública y privada) en funcionarios públicos en España. Estos investigadores partieron de la premisa de que aumento del flujo migratorio en los pasados 10 años había repercutido en la provisión y calidad de los servicios médicos clave. Por consiguiente, su modelo plantea una herramienta de análisis del efecto del aumento migratorio con una congestión sobre la demanda de seguros privados.

Como resultado, estos investigadores encontraron que el aumento de la migración (y población en general en España) ha repercutido de manera proporcional al incremento de la demanda de seguros médicos privados como consecuencia del congestionamiento del servicio estatal. Esta tendencia se evidenciaba con mayor fuerza en aquellas personas cuya clase social estaba entre la media y media-alta y con hijos. La razón de este incremento se debe a la necesidad de cubrir la asistencia primaria expedita donde el servicio estatal ha sido deficiente.

De manera concluyente, estos investigadores encontraron que, si bien la población inmigrante en España gozaba de buena salud con respecto a los españoles nativos, han supuesto un incremento importante de la demanda de atención sanitaria, especialmente en seguros privados.

Siguiendo con los determinantes relacionados al gasto sanitario en materia económica se tiene el estudio llevado a cabo por Martín et al. (2011) quienes llevaron a cabo un cribado profundo sobre esta tema y encontraron que el ingreso es el principal de cada individuo era un factor determinante hacia gasto en salud; a esto también se le debe añadir la relación con el envejecimiento de la población y la proximidad con la muerte.

Del mismo modo, estos investigadores encontraron que once de los artículos totales se enfocaban en calcular la elasticidad de la demanda, y solo dos se llegaba a obtener un valor superior a 1, entrando en la categoría de atención médica como un bien de lujo.

Para ellos no existe unanimidad en las variables econométricas y regresiones del gasto en salud en los países de la Organización para la Cooperación y el Desarrollo Económicos (OCDE). Por tanto, llegaron a la conclusión de que no existía evidencia empírica que factores como el envejecimiento de la población sea el principal determinante del gasto público en salud, mientras que factores como el progreso tecnológico, la proximidad a la muerte y la descentralización territorial de la atención de la salud se consideran cada vez más importantes en el desarrollo de modelos explicativos del gasto en salud.

Siguiendo con esta misma línea se destaca el trabajo de Cantarero et al., (2016) quienes se dedicaron a probar los determinantes del uso de los servicios sanitarios en España. Los datos utilizados proceden de la “*European Health Interview Survey*” (EHIS) para España (2014). Estos investigadores aplicaron un modelo de elección discreta y datos de enumeración para cada uno de los servicios de salud pública analizados (médico general, consultas de especialistas y urgencias, y hospitalizaciones).

Para el modelo planteado tomaron las variables de salud y otras características socioeconómicas y demográficas que eran controladas a nivel individual. Estos encontraron que las variables mencionadas están principalmente relacionadas a características sociales como la edad, el sexo, el nivel educativo o el estado de salud autoevaluado. A partir de sus resultados han destacado la relevancia que estos tienen para los debates en la literatura sobre economía de la salud. Finalmente encontraron que, desde una perspectiva de política económica, esta contribución podría ser invaluable para los tomadores de decisiones al planificar la mejora de la gestión de los servicios de salud de cara al futuro.

Dos años más tarde, Cantarero et al. (2018) publicaron un nuevo estudio donde examinaron uso de la atención médica entre las personas mayores en dieciséis países europeos utilizando la última publicación de la Encuesta sobre salud, envejecimiento y jubilación en Europa (SHARE, por sus siglas en inglés). Para ello aplicaron una regresión binomial negativa enfocada en el estudio de los principales factores que motivan el uso de la asistencia sanitaria (visitas a médicos generales y estancias hospitalarias).

Sus resultados empíricos han sugerido que la edad, el sexo, el nivel educativo, el estado de salud autoevaluado, las limitaciones y el estado de salud, y otras variables socioeconómicas son los principales factores determinantes. Además, estos investigadores demostraron que las variables socioeconómicas no juegan el mismo papel en todos los países, este último punto es relevante, tomando en cuenta los movimientos demográficos y económicos que han suscitado en España en la última década.

Desde un enfoque de política económica, Cantarero et al. (2018) presentan información relevante para los debates actuales tanto en la literatura sobre economía de la salud como sobre bienestar social; siendo además relevantes por las implicaciones en los propósitos de las políticas de mejorar la eficiencia, equidad y calidad de la atención médica que se puede brindar a la población europea en general.

Ya en tiempos más recientes, Cantarero et al. (2020) publican un nuevo artículo titulado: “Utilización de la atención médica y estilos de vida saludables entre las personas mayores que viven en el sur de Europa: evidencia reciente de SHARE²”. El objetivo de este trabajo ha sido examinar el efecto de los estilos de vida saludables entre los adultos mayores en sur de Europa en función a la utilización de la atención médica. Para ello, centraron su estudio en el número de consultas médicas ambulatorias y noches de hospitalización. Al igual que en el estudio de Cantarero et al. (2018) se realizaron regresiones binominales negativas utilizando un conjunto de datos de panel que estaba constituido por cinco publicaciones realizadas entre 2004 y 2015 de la Encuesta sobre salud, envejecimiento y jubilación en Europa (SHARE, siglas en inglés).

Como resultado encontraron nueva evidencia empírica sobre el efecto de los estilos de vida de los adultos mayores del sur de Europa en la utilización y demanda de la atención médica. La participación en actividades físicas vigorosas reduce el número de consultas y noches de hospitalización en un 7,8% y un 28,25%, respectivamente. De manera adicional, hábitos como el tabaquismo aumentaba las noches de hospitalización en un 14,22%. De estos resultados se desprende la importancia de los estados miembros

² Siglas en inglés para: Survey of Health, Ageing and Retirement in Europe

destinen esfuerzos en promover políticas focalizadas en la actividad física vigorosa hacia los individuos.

3.2. Modelos Empíricos: Demanda y utilización de Servicios Sanitarios.

Gracias a estudios como los anteriormente citados se puede contar con un marco referencias robusto que permitan ver los modelos teóricos que participan dentro de la demanda de los servicios sanitarios del SNS español. Para la formalización y validación de un modelo de aplicación empírica se debe estar condicionadas por las fuentes de información que se dispongan; de tal manera que se pueda explicar las diferencias encontradas con la utilización de los SS con respecto a las variables como edad, clase social, ubicación geográfica, etc. (González Álvarez & Clavero Barranquero, 2005).

Huelga decir que, el uso de este tipo de información por métodos estadísticos convencionales trae una limitación que consiste en obviar algunas características individuales que suelen ser determinantes en el estudio de la demanda de estos servicios (Browning & Zupan, 2020). Es por esta razón que, con el pasar del tiempo, los modelos econométricos se han ido mejorando y ajustando de tal manera que sus resultados puedan ser lo más cercanos a la realidad estudiada y posteriormente posibilite la vía del debate en aras de establecer soluciones adecuadas a mediano y largo plazo (Baumol & Blinder, 2015).

Así es como, dentro del campo de la estadística probabilística cualitativa surgen clasificaciones de modelos basados en información individual, entre los que vale mencionar los modelos de elección discreta, modelos censurados, modelos “*count data*” y modelos “*Hurdle*”. A continuación, se describirán los fundamentos de estos modelos.

3.2.1. Modelos de elección discreta

Estos modelos pueden ser binarios, múltiples y discretos e incluyen el proceso de toma de decisión como una comparativa de funciones de utilización indirecta (Tamayo, 2014). Dentro de este grupo los más empleados en el análisis de la demanda sanitaria han sido los modelos *logit* y *probit* donde se toma como supuesto a los individuos como el factor maximizador de su utilidad.

Estos modelos parten de las premisas que se ejemplifican a continuación. Sea U_0 la utilidad que reporta al individuo la decisión de no solicitar asistencia sanitaria, y U_1 , la utilidad en caso de que el individuo sí la solicite; entonces la probabilidad de que el individuo decida demandar la asistencia $Y = 1$, viene dada por la expresión:

$$\Pr(Y = 1|x) = \Pr(U_{1i} > U_{0i})$$

Así pues, un modelo de utilidad aleatoria lineal podría explicar las utilidades basadas en especificaciones como las siguientes:

$$U_{1i} = x_i' \beta_1 + \varepsilon_{1i}$$

$$U_{0i} = x_i' \beta_0 + \varepsilon_{0i}$$

Ambas especificaciones analizan de manera simultánea la probabilidad de que un individuo demande o no un servicio sanitario al tiempo que se tiene en cuenta las variables que el paciente asume cuando toma esta decisión (Edad, Ubicación, Clase social y ocupación), e incluso se puede estudiar por qué el paciente elige la sanidad privada frente a la pública (Rodríguez & Stoyanova, 2004) o si la inmigración influye en este mismo sentido (Jiménez-Martín & Jorgensen, 2009).

3.2.2. Modelos censurados

Cuando lo que se busca es analizar el comportamiento del gasto sanitario, especialmente el privado, se suelen utilizar este tipo de modelos. Dentro de esta categoría se encuentran los modelos *Tobit*, de selección muestral y TPM³ (González Álvarez & Clavero Barranquero, 2005). Es muy importante señalar que, para llevar a cabo estudios de este tipo en España, sería necesario realizar encuestas especializadas en el gasto sanitario, utilización de SS, y salud que permitan además estudiar en profundidad la relevancia del sector privado dentro de este contexto. En tal caso, suele proponerse unificar los datos ofrecidos por la ENS (información de salud) y EPF (información de gasto sanitario y privado) (Baumol & Blinder, 2015).

³ Two-part-models

3.2.3. Modelos Count Data

Se trata de un modelo de regresión basada en la estimación mínima cuadrática donde se ha usado generalmente para medir la probabilidad de utilización de un SS o el número de visitas a una consulta. Dentro de estos modelos encontramos el de Pole & Grossman (1974) que considera la asistencia sanitaria como valor de entrada para la conservación de la salud al tiempo que plantean modelos de múltiples ecuaciones que se resuelven con mínimos cuadrados en dos etapas (Hidayat & Pokhrel, 2010).

Los modelos *Count Data* fueron utilizados por primera vez en el análisis de la utilización de los servicios sanitarios por Cameron & Trivedi (1986) donde tomaron solo valores positivos debido al comportamiento no paramétrico de los datos analizados. Este criterio de valores aunado al recurso de aleatoriedad deja ver que los modelos *count data* en su sentido más básico se fundamentan en la distribución de Poisson y que, por lo tanto, la probabilidad de que un valor sea observado “y” viene regido bajo la hipótesis de Poisson a través de la siguiente expresión:

$$\Pr(Y = y|\lambda) = \frac{e^{-\lambda} \lambda^y}{y!}, \quad y = 0, 1, \dots, \infty.$$

La limitación de este modelo estriba en las restricciones puestas por un único parámetro “λ”, ya que se asume que la media y las varianzas son iguales, siendo esta una hipótesis muy poco frecuente en el comportamiento económico dentro del ámbito sanitario (Browning & Zupan, 2020).

3.2.4. Modelos Hurdle

Este modelo se basa en el alto porcentaje de no utilización de un SS. Este a su vez está relacionado con los modelos “principal-agente” donde el agente es el personal médico y el principal es el paciente una vez se ha hecho la primera visita (Deb & Trivedi, 2002; Zweifel, 1981). La modelización empírica de este modelo viene dada por la aportación de Pohlmeier & Ulrich (1995) quienes consideraron la toma de decisiones como un proceso de dos etapas; en la primera el paciente (principal) es quien decide ir al médico (agente), pero una vez este último toma acción, entonces es quien toca determinar la intensidad y frecuencia del tratamiento sanitario, siendo esta la segunda etapa.

3. Modelo y Procedimiento de estimación

3.1. Método de investigación

Pasando al esquema metodológico del proyecto se ha utilizado el método deductivo de análisis puesto que se parten de datos generales validados (ENSE 2003, 2006, 2011, 2017) y desde los cuales se plantean las hipótesis que han sido contrastadas y corroboradas dentro de los resultados obtenidos. En cuando al diseño de la investigación el presente proyecto se presenta como no experimental puesto que no se contempla manipulación deliberada de las variables de estudio y solo se concentra en el análisis de la causa-efecto de dichas variables. Por consiguiente, la investigación que caracteriza este trabajo es de tipo explicativo, ya que estudia las variables de una situación en particular y como estas influyen en otras. En el marco del análisis causa-efecto, se aíslan y descartan aquellos factores que pudieran ocasionar resultados muy particulares.

3.2. Población y Muestra

En el presente trabajo se han utilizados los datos recolectados en la Encuesta Nacional de Salud de España (ENSE) en sus versiones 2003, 2006, 2011 y 2017 que se encuentran publicadas en la página del Ministerio de Sanidad, Consumo y Bienestar Social. Esta población es dividida en dos muestras que serán llamadas “grupo 1” y “grupo 2”. El grupo 1 estará constituido por los datos publicados entre 2003 y 2006 (antes de la crisis económica) y el grupo 2 está conformado por los resultados de las encuestas de 2011 y 2017 (después de la crisis).

El presente estudio estará focalizado en la población adulta (edad > 16 años), considerando el sexo, el nivel educativo, la actividad económica actual, el nivel de ingresos percibidos (clase social) y la ciudad de residencia. Cabe señalar que la nacionalidad de los encuestados no será relevante en el presente estudio.

La muestra total estará conformada por 95.224 adultos desglosados en 21.650 en 2003, 29.478 en 2006, 21.007 en 2011 y 23.089 en 2017. Para el grupo 1 se tendrá una muestra de 51.128 y para el grupo 2 será de 44.096.

3.3. Estimación del modelo

La formulación del modelo econométrico adecuado se fundamenta en la definición de las variables dependientes como binarias o dicotómicas. Es decir, la variable dependiente solo podrá tener dos valores únicos (0 y 1) y que están en consonancia con las posibles alternativas de decisión del individuo en cuanto a la elección de un seguro sanitario privado. Por consiguiente, siguiendo con el esquema propuesto por Jiménez-Martín & Jorgensen (2009), se establece Y_i como variable que corresponde a la disponibilidad de seguro privado por parte del sujeto i , siendo este un adulto con residencia en España.

No obstante, como también es de interés para el estudio la región de vive el sujeto y la influencia de la crisis sobre su decisión en la demanda de seguros sanitarios privados, se tiene Y_{ijt} , que se define como la disponibilidad de seguro privado que tiene un sujeto i , en una región j , en un año t .

Con base en lo anteriormente mencionado y asumiendo un comportamiento lineal en la valoración de opciones, se tiene:

$$Y_{ijt} = \begin{cases} 1 & \text{si el adulto cuenta con seguro privado} \\ 0 & \text{si el adulto no cuenta con seguro privado} \end{cases}$$

Por tanto, la demanda potencial o latente del seguro privado por parte de la población adulta española que vive en una región j y en un tiempo t vendrá expresada por el modelo dicotómico mostrado a continuación:

$$Y_{ijt} = \alpha X_{ijt} + \alpha Z_{ijt} + \tau p_{ijt} + \varphi S_{jt} + \beta DQ_{jt} + v_{ijt}$$

φS_{jt} : Como variable exógena se ha considerado la actividad económica del encuestado.

Donde: X_{ijt} se refiere a los regresores relacionados con la clase social, el nivel de estudios y actividad económica actual de los participantes. Z_{ijt} Corresponde a los regresores relativos a la edad y el género. τp_{ijt} denota el precio del seguro (prima); DQ_{jt} es el diferencial de calidad del sistema público de sanidad y el privado y v_{ijt} será el error.

En el caso del diferencial de calidad se puede determinar a través de la siguiente ecuación:

$$DQ_{jt} = \overline{Q}_j + \delta \rho_{jt} + \xi_{jt} \quad \delta \geq 0$$

Siendo $\overline{Q}_j$ el diferencial promedio de la calidad por región, ρ_{jt} el cociente entre el número de usuarios totales de dicha región y la capacidad del sector público y ξ_{jt} el error.

3.4. Modelo econométrico

En los estudios relacionados a las variables dependientes cualitativas dicotómicas se encuentran dos tipos de modelos: El modelo Logit, con variable cualitativa dependiente binomial; y el modelo Probit, con variable cualitativa dependiente multinomial. Tomando en cuenta los regresores propuestos en el apartado anterior se tendría la siguiente ecuación general

$$Y_{ijt} = \alpha X_{ijt} + \alpha Z_{ijt} + \tau p_{ijt} + \varphi S_{jt} + \beta(\overline{Q}_j + \delta \rho_{jt} + \xi_{jt}) + v_{ijt}$$

Siendo Y la variable dependiente que puede ser 0= individuo no tiene seguro privado; o 1 = individuo tiene un seguro privado. Por su parte, las variables independientes serán:

$X_{ijt}: \{CS, NE\}$

- CS: Clase social (Nivel de ingresos)
- NE: Nivel educativo
- OC: ocupación

$Z_{ijt}: \{ED, GEN\}$

- ED: Edad
- GEN: Género

p_{ijt} : Para la estimación del precio del seguro se tiene la media en España de primas para 2003 era de 345 € y de 488€ para 2006.

φS_{jt} : Como variable exógena se ha considerado el PIB de España en el año correspondiente. El PIB estará puesto en miles de millones de euros.

El modelo econométrico se usará para los años 2003 y 2006 (pre-crisis) y 2011 y 2017 (post-crisis)

Aplicando el método de la máxima verosimilitud se obtienen los siguientes modelos:

Modelo Probit:

La probabilidad que un individuo adulto demande un servicio sanitario privado antes y después de la crisis del 2008 asumiendo un comportamiento gaussiano viene expresado como:

$$\Pr(y = 1|x) = \alpha X_{ijt} + \alpha Z_{ijt} + \tau p_{ijt} + \varphi S_{jt}$$

Los efectos marginales del modelo probit se obtienen para variables independientes tanto cualitativas como cuantitativas. Para las variables cuantitativas se calculan mediante la siguiente derivada parcial:

$$\frac{\partial \phi(X_{ijt}\alpha)}{\partial x_{kl}} = \phi X_{ijt}\alpha [1 - \phi X_{ijt}\alpha] \alpha_k$$

Para una variable independiente cualitativa se obtiene mediante la diferencia de probabilidades cuando la variable dependiente toma los valores de 1 y 0.

Modelo Logit:

La probabilidad que un individuo adulto demande el uso de un servicio sanitario privado antes y después de la crisis del 2008 viene expresada como:

$$\Pr(y = 1|x) = \frac{\exp(\alpha X_{ijt} + \alpha Z_{ijt} + \tau p_{ijt} + \varphi S_{jt})}{1 + \exp(\alpha X_{ijt} + \alpha Z_{ijt} + \tau p_{ijt} + \varphi S_{jt})}$$

Los efectos marginales de las variables cuantitativas en el modelo logit se obtienen mediante la siguiente derivada parcial:

$$\frac{\partial \Lambda(X_{ijt}\alpha)}{\partial x_{kl}} = \Lambda X_{ijt}\alpha [1 - \Lambda X_{ijt}\alpha] \alpha_k$$

Por otro lado, para variables independientes cualitativas, el efecto marginal se obtiene a través de la diferencia de probabilidades cuando dicha variable toma valor de 0 y 1; o lo que es lo mismo, se evalúa en ambos puntos y se calcula la diferencia.

3.5. Análisis de la muestra

En la presente sección se muestran las variables recolectadas y procesadas que servirán, en primer lugar, para el análisis de regresiones de efectos fijos y aleatorios y posteriormente para la validación del modelo econométrico final (ver tabla 1). La tabla 1 muestra a los individuos adultos varones que vivan en ciudades con más de 100mil habitantes y esté activos en el mercado laboral que poseen un determinado tipo de seguro sanitario; 1 para quienes lo tenga exclusivamente privado y 0 para exclusivamente público.

Se ha visto prudente seleccionar los datos de esta forma puesto que con ello se mantiene el carácter dicotómico del estudio y mutuamente excluyente, donde además de excluyeron los datos referentes a las personas que tienen seguros sanitarios mixtos (doble cobertura).

Tabla 1. Variables de estudio agrupadas por años. Cifras absolutas

Año	2003	2006	2011	2017	2003	2006	2011	2017
Tipo de Cobertura	1	1	1	1	0	0	0	0
Adultos	32,6	107,4	72,60	64,3	5222	5241,9	6178,2	5165,4
Varones	70,0	3474	3088	3732	10425	21319	22371	22407
más 100 mil hab	40,7	342,4	187,4	185	2693,0	17711,	18930,	18318,
Clase Social I	35,8	108,10	67,40	41,40	1125,3	1523,10	1769,10	1657,10
Trabajando	60,0	184,20	103,80	97,10	12569,	8940,6	7557,1	7605,5
Educación Superior	49,4	0,00	46,80	67,6	2217,4	0,00	1760,7	2930,7
PIB	0,91	1,25	1,47	1,30	0,91	1,25	1,47	1,30

Fuente INE.

De la tabla 1 se calculan los estadísticos descriptivos para cada co-variables encontrándose que el PIB mostraba una tendencia alcista entre 2003 y 2006 y posteriormente tendría un decrecimiento a partir del 2011, pasando de 1,47 mil millones de euros a 1,30 (ver tabla 2).

Tabla 2. Resultados estadísticos descriptivos

Variables	Media	Desv Std.	Mín	Pctl(25)	Pctl(75)	Max
Adultos	2760,55	2894,71	32,60	70,52	5226,98	6178,20
Varones	10861,12	9692,45	70	3,378	21582,5	22407
Habitantes	7301,14	9169,98	40,70	186,80	17863,35	18930,20
Clase social I	790,91	799,71	35,80	60,90	1556,60	1769,10
Trabajadores	4639,70	5081,03	60	102,1	7939,3	12,569
Educación Superior	884,08	1216,69	0,00	35,10	1874,88	2930,70
PIB	1,25	0,36	0,90	1,29	1,40	1,47

Con los valores estimados y mostrados en la tabla 1 se han realizado las siguientes estimaciones:

3.6. Estimación lineal (mínimos cuadrados)

La estimación lineal por mínimos cuadrados es el primero que se aplica para las variables de estudio. La tabla 3 recoge el intercepto y los coeficientes correlacionados de sexo, clase social, ocupación, nivel de estudios, habitantes de la ciudad de residencia y PIB.

Tabla 3. Coeficientes de regresiones lineales por co-variables

	Resultado de mínimos cuadrados					
	Intercepto	CS	OC	NE	HB	PIB
Sexo	-4836,10	1,33	0,27	-0,27	0,66	5846,54
Clase Social	92,16	0,01	0,06	0,07	0,05	-62,51
Población	3370,53	6,70	-0,70	-0,34	0,84	-4619,02
Ocupación	1878,77	12,55	0,52	-0,69	-1,06	-2931,56

Nivel de Estudios	-2605,59	7,34	-0,32	-0,25	-0,24	2364,81
-------------------	----------	------	-------	-------	-------	---------

Para el caso del sexo, la clase social y el número de habitantes se encontró un R ajustado de 0,98; mientras que la ocupación se enmarca dentro de un nivel de ajuste del 92% en linealidad. El Nivel de estudios por su parte muestra un ajuste del 82% con respecto a los datos.

Así pues, con la bondad de ajuste de esta estimación podemos inferir que las variables sexo, clase social y el número de habitantes de la zona donde habita el individuo tiene un carácter lineal con respecto al resto. No es el caso del nivel de estudios y la ocupación

4. Resultados

A continuación, se muestran los resultados de los regresores del modelo logit y probit propuestos. La tabla 4 muestra los resultados del modelo econométrico para cada año junto con los intervalos confianza al 95%.

Tabla 4. Resultados de los coeficientes Logit y Probit

Tabla comparativa logit/ probit								
	Elección de seguro privado España							
	2003		2006		2011		2017	
	<i>logistic</i>	<i>probit</i>	<i>logistic</i>	<i>probit</i>	<i>logistic</i>	<i>probit</i>	<i>logistic</i>	<i>probit</i>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sexo	-0.561 (0.787)	-0.337 (0.421)	17.765 (1.734.119)	5.179 (390.617)	-0.398*** (0.085)	-0.191*** (0.042)	-0.156*** (0.045)	-0.086*** (0.024)
Edad	0.056* (0.029)	0.032** (0.015)	0.035 (0.030)	0.020 (0.015)	-0.003 (0.002)	-0.001 (0.001)	-0.004*** (0.001)	-0.002*** (0.001)
Clase Social	0.335 (0.236)	0.178 (0.127)	0.385 (0.290)	0.187 (0.149)	0.008 (0.024)	0.003 (0.012)	-0.307*** (0.015)	-0.167*** (0.008)
Nivel de Estudios	0.042 (0.140)	0.037 (0.075)	0.238 (0.169)	0.116 (0.087)	0.019 (0.016)	0.009 (0.008)	0.203*** (0.013)	0.109*** (0.007)
Actividad económica	0.020 (0.220)	0.020 (0.119)	0.019 (0.233)	0.018 (0.126)	-0.089** (0.043)	-0.043** (0.020)	-0.082*** (0.017)	-0.045*** (0.009)
Habitantes	3.438*** (0.882)	1.896*** (0.463)	-0.069 (0.214)	-0.010 (0.114)	-0.190*** (0.020)	-0.091*** (0.010)	-0.005 (0.010)	-0.001 (0.005)
Constante	-4.077* (0.882)	-2.351* (0.463)	-2.072 (0.214)	-1.104 (0.114)	-1.664*** (0.020)	-1.024*** (0.010)	-1.868*** (0.010)	-1.070*** (0.005)

	(2.208)	(1.221)	(2.489)	(1.337)	(0.191)	(0.094)	(0.166)	(0.088)
Observaciones	99	99	99	99	8,605	8,605	23,089	23,089
Verosimilitud	- 24.657	- 24.472	-21.716	-21.839	- 2,172.007	- 2,172.918	-7,155.050	-7,133.803
Criterio Akaike	63.315	62.943	57.432	57.677	4,358.014	4,359.836	14,324.100	14,281.610
Nota: * p ** p *** p<0.01								

De cada modelo se obtuvo un R McFadden de 0,0729 y 0,0722 respectivamente. El logaritmo de verosimilitud es para cada modelo en cada año ha sido de -24,657, -24,472, -21,716; -21,839; -2172,007; -2172,918; -7155,050; -7133,803;-9659,82. Los coeficientes de cada modelo en cada año se muestran en la siguiente tabla:

Tabla 5. Coeficientes en años. Modelo logit y probit

Logit							
Año	Intcp.	Sexo	Edad	Clase social	Nivel de Estudios	Ocupación	Habitantes
2003	0,017	0,571	1,057	1,398	1,043	1,021	31,109
2006	0,126	0,519	1,035	1,470	1,269	1,019	0,934
2011	0,189	0,672	0,997	1,008	1,019	0,914	0,827
2017	0,154	0,856	0,996	0,736	1,225	0,921	0,995
Probit							
Año	Intcp.	Sexo	Edad	Clase social	Nivel de Estudios	Ocupación	Habitantes
2003	-0,316	-0,045	0,004	0,024	0,005	0,003	0,255
2006	-0,134	0,629	0,002	0,023	0,014	0,002	-0,001
2011	-0,138	-0,026	0,000	0,000	0,001	-0,006	-0,012
2017	-0,179	-0,014	0,000	-0,028	0,018	-0,008	0,000

También se ha estimado el modelo probit/logit total, cuyo resultado se muestra en la siguiente tabla:

Tabla 4. Resultados logit/probit 2003-2017

Tabla comparativa logit/ probit 2003 - 2017		
	Variable dependiente:	
	Elección de seguro privado España <i>logit</i> (1)	<i>probit</i> (2)
Año	-0.036*** (0.007)	-0.022*** (0.004)
Sexo	-0.197*** (0.038)	-0.105*** (0.020)
Edad	-0.008*** (0.001)	-0.004*** (0.001)
Clase Social	-0.235*** (0.012)	-0.125*** (0.006)
Nivel de Estudios	0.176*** (0.009)	0.091*** (0.005)
Actividad económica	-0.0003 (0.013)	-0.003 (0.007)
Habitantes	-0.054*** (0.009)	-0.028*** (0.005)
Constante	71.179*** (14.072)	42.563*** (7.406)
Observaciones	31,892	31,892
Verosimilitud	-9,960.085	-9,950.093
Criterio Akake	19,936.170	19,916.190
Nota:		* ** *** p p p<0.01

Los coeficientes para logit y probit del modelo final que considera los años desde 2003 hasta el 2017 se muestran en la siguiente tabla:

Tabla 5. Coeficientes modelo logit y probit

Modelo	Intercepto	Sexo	Edad	Clase social	Nivel de estudios	Ocupación	Habitantes	Año
Logit	81,72	0,964	0,821	0,993	0,790	1,192	1,000	0,947
Probit	42,56	-0,02	-0,10	-0,004	-0,125	0,091	-0,003	-0,03

En consecuencia, la probabilidad pronosticada de que en 2003 un individuo varón (Sexo = 1) con edad adulta (Edad = 30), actividad económica trabajando (OC = 1), clase social alta (CS = 1) y nivel de estudios superiores (NE=3) cuente con una cobertura médica privada es un 18,57% para el modelo logit; porcentaje que aumenta en el caso de las mujeres, pasando a un 21,72% en las mismas condiciones demográficas del varón.

De manera análoga se percibe algo parecido con la ocupación. En promedio, la probabilidad de que en 2017 un individuo varón, con 30 años trabajando y sea de clase social alta demande este mismo seguro privado disminuyó a 12,08% es de incrementándose a 14,33% en caso de las mujeres con la misma edad.

5. Conclusiones

Como resultado de las pruebas estadísticas en los modelos desarrollados, el uso de servicios sanitarios privados de cada individuo depende en gran medida de las variables socioeconómicas y demográficas en las que se desenvuelva un individuo en España. Los coeficientes de Pseudo R2 en ambos modelos indica que alrededor de al menos un 10% de la variación de la variable dependiente (uso de servicio privado) puede ser explicada por las variaciones de las covariables.

En concordancia con los efectos marginales, la demanda de servicios de salud privado se rige la edad, el sexo, la actividad económica y el nivel educativo con probabilidades que oscilan entre el 10% y 25%. En este mismo orden de ideas, la estadística descriptiva muestra que los descriptores socioeconómicos se relacionan positivamente y con la demanda de la salud de los individuos en el España a lo largo de los años. Por ejemplo, el uso los servicios privados disminuye con la edad y a su vez aumenta con la clase social, el nivel educativo y la actividad económica del individuo.

Con respecto a la eficacia de ambos modelos, los resultados muestran que el método logit se ajusta mejor que probit. Así, con el modelo logit para los años, queda probado en todos los factores que determinan la demanda de servicios sanitarios privados.

Un punto destacable se encuentra en el nivel educativo presentado por los individuos, los cuales se correlacionan positivamente. Como se ha mencionado anteriormente, esto se debe a que los individuos con niveles más altos de educación suelen ser más efectivos en la búsqueda y el uso de la información sobre los planes de salud disponibles, que aquellas personas con poca o nula educación. Por otra parte, el efecto marginal de la actividad económica sobre el uso de cobertura privada es bajo; el impacto del incremento del ingreso en el hogar tiene que ser muy por encima de la media de salario interprofesional español para que se convierta en criterio de decisión más relevante. En cuanto al sexo, se presenta como una variable significativa en términos estadísticos, siendo los varones quienes presentan una menor probabilidad de adquirir un seguro privado en contraste con los varones. La clase social también es significativa en términos estadísticos e impacta de manera positiva en la probabilidad de que los individuos demanden servicios sanitarios privados.

Finalmente, los resultados aquí mostrados indican que el incremento de factores como la población, clase social, ingresos y niveles educativos contribuye

sustancialmente en la demanda de servicios sanitarios privados, tendencia que se mantiene en crecimiento desde el 2011 hasta el 2017.

6. Bibliografía

- Albar, M., García, M., Perez Moreno, P., Luque, V., Garrido, R., & Bocchino, A. (2015). Adaptation to spanish of an ethnocultural empathy scale. *Texto & Contexto-Enfermagem, AHEAD*, 0.
- Bas-Sarmiento, P., Fernández-Gutiérrez, M., Albar-Marín, M. a. J., & García-Ramírez, M. (2015). Percepción y experiencias en el acceso y el uso de los servicios sanitarios en población inmigrante. *Gaceta Sanitaria*, 29(4), 244–251. <https://doi.org/10.1016/j.gaceta.2015.03.008>
- Baumol, W. J., & Blinder, A. S. (2015). *Microeconomics: Principles and policy*. Nelson Education.
- Browning, E., & Zupan, M. (2020). *Microeconomics: Theory and applications*.
- Cameron, A. C., & Trivedi, P. K. (1986). Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*, 1(1), 29–53.
- Canals Barceló, M. P. (2018). *¿Cómo impacta en la salud de los inmigrantes en España las últimas reformas de la política de acceso al Sistema Nacional de Salud?*
- Cantarero, D., Pascual, M., & Blázquez, C. (2016). *Statistical models of economic evaluation in public policies. Analysis of health care resource utilization*. <https://doi.org/http://www.ief.es/docs/destacados/publicacio>
- Cantarero, D., Pascual, M., & Lera, J. (2020). Healthcare Utilization and Healthy Lifestyles among Elderly People Living in Southern Europe: Recent Evidence from the SHARE. *Atlantic Economic Journal*, 48(1), 53–66. <https://doi.org/10.1007/s11293-020-09657-3>
- Cantarero, D., Pascual, M., & Torres, J. L. (2018). Socioeconomic determinants and

health care utilization among elderly people living in Europe : Evidence from the Survey of Health , Ageing and Retirement . Socioeconomic determinants and health care utilization among elderly people living in Europe : Evid. In *GEN Working Paper B: Regional and sectoral economics, 2018-1* (Issue January). Universidad de Vigo. <https://repositorio.unican.es/xmlui/handle/10902/14065>

Catalán, M. A., Moya, T., Angurera, M., de Armas, R., Cevalco, C., Esteve, J., Sató, E., & Fornells, Antonio José Monreal, T. (2012). *El impacto de la crisis económica en la salud mental de la población*. <http://hdl.handle.net/10481/40609>

Costa, J., Hernández-Quevedo, C., & Jiménez, D. (2014). Income inequalities in unhealthy life styles in England and Spain. *Economics and Human Biology*, 13(1), 66–75. <https://doi.org/10.1016/j.ehb.2013.03.003>

Deb, P., & Trivedi, P. K. (2002). The structure of demand for health care: Latent class versus two-part models. *Journal of Health Economics*, 21(4), 601–625. [https://doi.org/10.1016/S0167-6296\(02\)00008-5](https://doi.org/10.1016/S0167-6296(02)00008-5)

Fayos Jordán, P. (2017). *Servicios Sanitarios en España: caracterización y evolución*.

Gea-Sánchez, M., Alconada-Romero, Á., Briones-Vozmediano, E., Pastells, R., Gastaldo, D., & Molina, F. (2017). Undocumented Immigrant Women in Spain: A Scoping Review on Access to and Utilization of Health and Social Services. *Journal of Immigrant and Minority Health*, 19(1), 194–204. <https://doi.org/10.1007/s10903-016-0356-8>

González Álvarez, M., & Clavero Barranquero, A. (2005). Una revisión de modelos econométricos aplicados al análisis de demanda y utilización de servicios sanitarios. *Hacienda Pública Española*, 173, 129–162. http://www.ief.es/docs/destacados/publicaciones/revistas/hpe/173_Art5.pdf

González, J. (2013). La salud en la población inmigrante: la política como determinante

social de la salud. *Fundación de Enfermería de Cantabria*, 2(12), 1–13.
www.enfermeriadecantabria.com/nuberoscientica

Henderson, S., Horne, M., Hills, R., & Kendall, E. (2018). Cultural competence in healthcare in the community: A concept analysis. *Health and Social Care in the Community*, 26(4), 590–603. <https://doi.org/10.1111/hsc.12556>

Hernández Quevedo, C., & Jiménez Rubio, D. (2009). Las diferencias socioeconómicas en salud entre la población española y extranjera en España: evidencia de la Encuesta Nacional de Salud. *Gaceta Sanitaria*, 23(SUPPL. 1), 47–52.
<https://doi.org/10.1016/j.gaceta.2009.07.009>

Hidayat, B., & Pokhrel, S. (2010). The selection of an appropriate count data model for modelling health insurance and health care demand: Case of Indonesia. *International Journal of Environmental Research and Public Health*, 7(1), 9–27.

Jiménez-Martín, S., & Jorgensen, N. (2009). Inmigración y demanda de seguros sanitarios. *Gaceta Sanitaria*, 23(SUPPL. 1), 19–24.
<https://doi.org/10.1016/j.gaceta.2009.09.005>

Kabongo, W., & Mbonigaba, J. (2016). Demand-side determinants of access to healthcare service: Empirical evidence from Africa. *Working Papers*.
https://saef.ukzn.ac.za/wp-content/uploads/2018/05/SAEF_Health-Care-Services_JM-SK.pdf

López Nicolás, Á., & Ramos Parreño, J. M. (2009). Utilización de servicios sanitarios por parte de las poblaciones inmigrante y nativa en la Comunidad Autónoma de la Región de Murcia. *Gaceta Sanitaria*, 23(SUPPL. 1), 12–18.
<https://doi.org/10.1016/j.gaceta.2009.04.007>

Martín, J. J. M., del Amo Gonzalez, M. P. L., & Dolores Cano Garcia, M. (2011). Review of the literature on the determinants of healthcare expenditure. *Applied Economics*,

43(1), 19–46. <https://doi.org/10.1080/00036841003689754>

Meth, P., & Williams, G. (2006). Literature reviews and bibliographic searches. V. *Desai & RB Potter*.

Peiró, S., & Ridao, M. (2018). La atención especializada del Sistema Nacional de Salud durante la Gran Recesión. Gasto, utilización y variabilidad en el periodo 2004-2015. *Cuadernos Económicos de ICE*, 96.

Picatoste, J., Ruesga-Benito, S. M., & González-Laxe, F. (2018). Economic environment and health care coverage: Analysis of social acceptance of access restrictive policies applied in Spain in the context of economic crisis. *Journal of Cleaner Production*, 172, 3600–3608.

Pohlmeier, W., & Ulrich, V. (1995). An econometric model of the two-part decisionmaking process in the demand for healthcare. *Journal of Human Resources*, 30(2), 339–361. <https://doi.org/10.2307/146123>

Pole, J. D., & Grossman, M. (1974). The Demand for Health: A Theoretical and Empirical Investigation. *Journal of the Royal Statistical Society. Series A (General)*, 137(2), 279. <https://doi.org/10.2307/2344574>

Porthé, V., Vargas, I., Ronda, E., Malmusi, D., Bosch, L., & Vázquez, M. L. (2018). Has the quality of health care for the immigrant population changed during the economic crisis in Catalonia (Spain)? Opinions of health professionals and immigrant users. *Gaceta Sanitaria*, 32(5), 425–432. <https://doi.org/10.1016/j.gaceta.2017.03.010>

Rechel, B., Mladovsky, P., Ingleby, D., Mackenbach, J. P., & McKee, M. (2013). Migration and health in an increasingly diverse Europe. In *The Lancet* (Vol. 381, Issue 9873, pp. 1235–1245). Lancet Publishing Group. [https://doi.org/10.1016/S0140-6736\(12\)62086-8](https://doi.org/10.1016/S0140-6736(12)62086-8)

- Reiners, F., Sturm, J., Bouw, L. J. W., & Wouters, E. J. M. (2019). Sociodemographic factors influencing the use of ehealth in people with chronic diseases. In *International Journal of Environmental Research and Public Health* (Vol. 16, Issue 4). <https://doi.org/10.3390/ijerph16040645>
- Rodríguez, M., & Stoyanova, A. (2004). The effect of private insurance access on the choice of GP/specialist and public/private provider in Spain. *Health Economics*, 13(7), 689–703. <https://doi.org/10.1002/hec.832>
- Ruiz, A., Escudero Carretero, M., López, L. A., Gil García, E., March Cerdà, J. C., & López Jaramillo, D. (2020). La perspectiva de personas migrantes sobre el acceso a la atención sanitaria en el contexto de políticas de austeridad en Andalucía. *Gaceta Sanitaria*, 34(3), 261–267. <https://doi.org/10.1016/j.gaceta.2018.09.006>
- Tamayo, W. S. (2014). *Modelos de elección discreta aplicados a la asistencia sanitaria*. Universidad del País Vasco-Euskal Herriko Unibertsitatea.
- Terraza, R., Toledo, D., Vargas, I., & Vázquez, M. L. (2010). Perception of the ecuadorian population living in Barcelona regarding access to health services. *International Journal of Public Health*, 55(5), 381–390. <https://doi.org/10.1007/s00038-010-0180-1>
- Urrutia, R. (2018). Percepciones sobre la educación intercultural y relación médico-paciente inmigrante en médicos de Chile. *Revista Peruana de Medicina Experimental y Salud Pública*, 35(2), 205. <https://doi.org/10.17843/rpmesp.2018.352.3404>
- Yagüe Ruipérez, S. (2019). *Perspectiva enfermera en la Atención Primaria a la población inmigrante*.
- Zweifel, P. (1981). *Supplier-Induced Demand in a Model of Physician Behavior*. J. van der Gaag and Perlman, M.(Eds.) *Health, Economics, and Health Economics*.

7. Anexos

7.1. Gráficos de cajas para cada variable

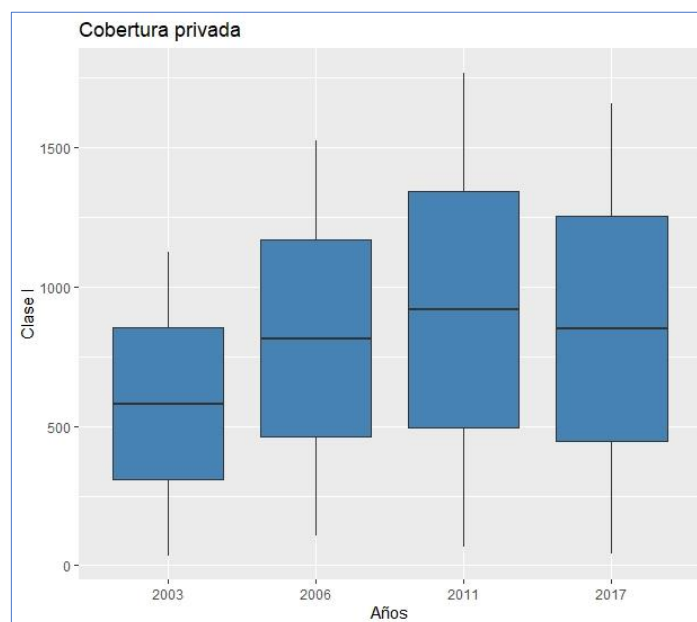


Figura 1. Variación demanda seguro médico privado en función de la clase social

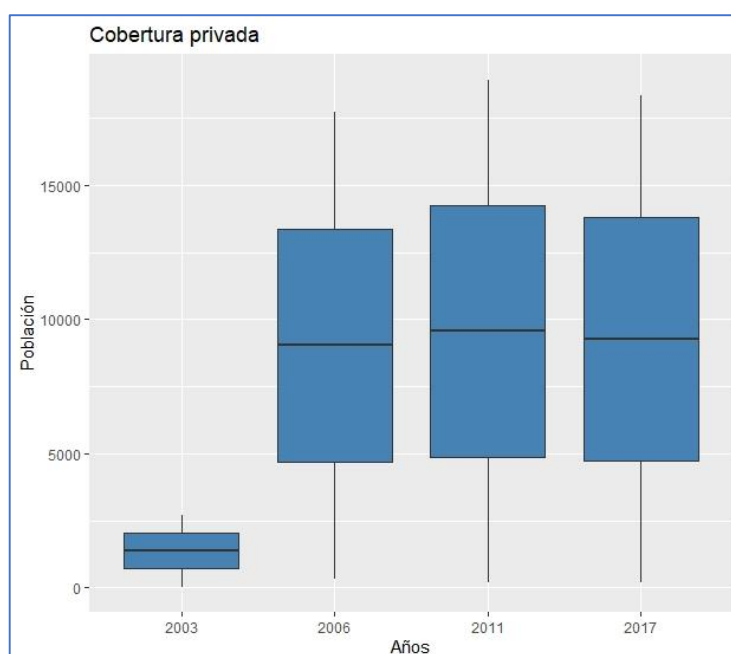


Figura 2. Variación demanda seguro médico privado en función de la región con mayor número de habitantes

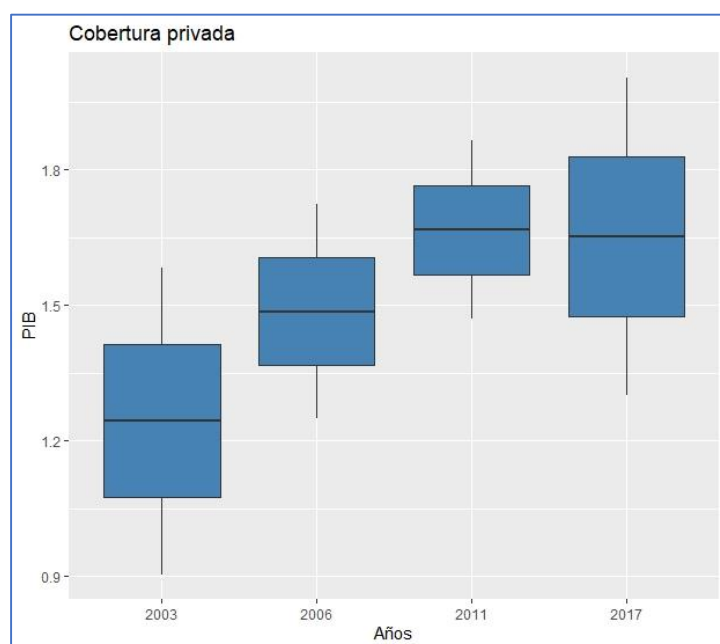


Figura 3. Variación demanda seguro médico privado en función del PIB

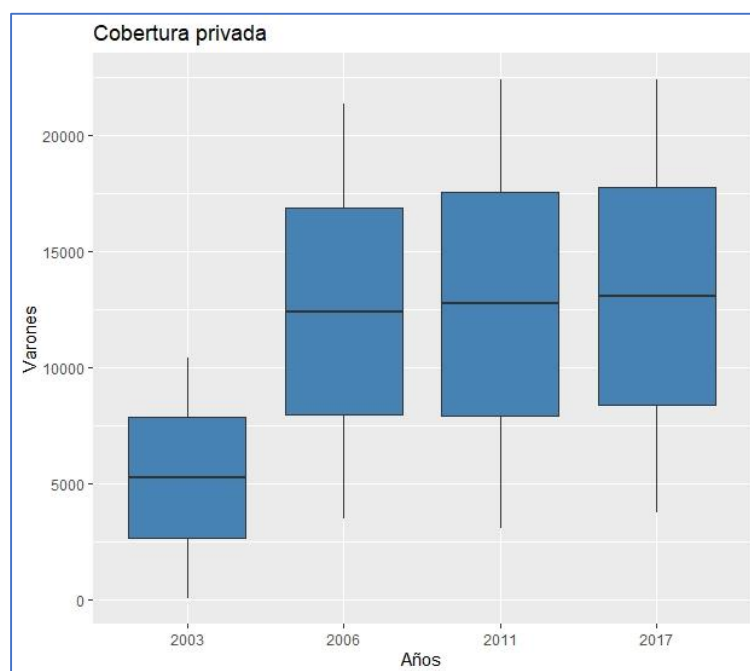


Figura 4. Variación demanda seguro médico privado en función del sexo (varones)

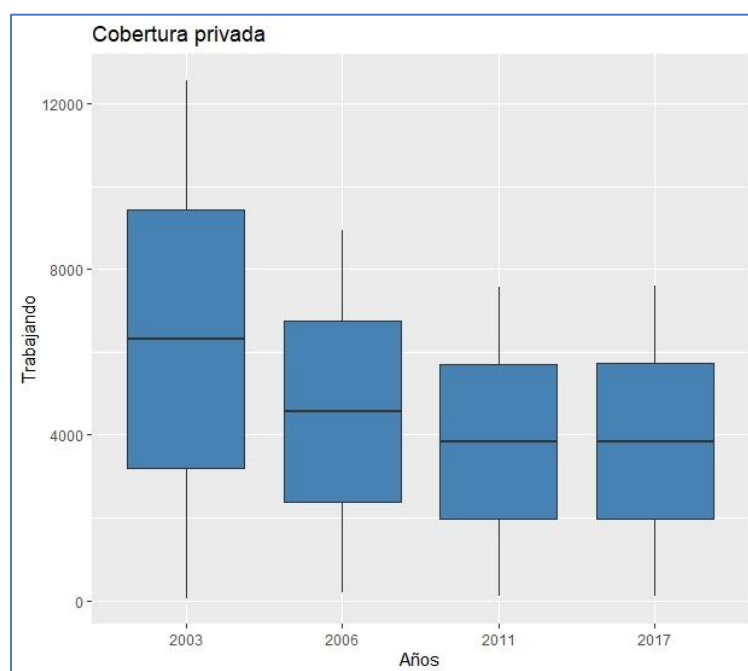


Figura 5. Variación demanda seguro médico privado en función de la ocupación (Trabajadores)

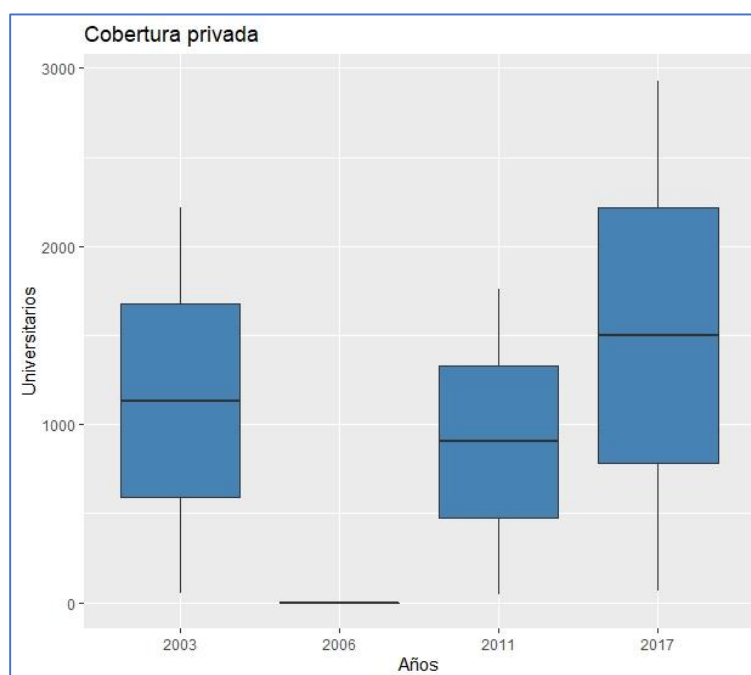


Figura 6. Variación demanda seguro médico privado en función del nivel de estudios (superior universitarios)

7.2. Scripts en R.

Nota: No computa para el número de páginas exigidas

Estimaciones

```
library(stargazer)
library(plm)
library(readxl)
library(dplyr)
library(texreg)
library(gplots)
library(car)
library(ggplot2)

db <- excel_sheets("DB.xlsx")
data <- read_excel("DB.xlsx", sheet= "Data")
head(data)
summary(data)

names(data) = c("A", "Y", "ED", "SX", "HB", "CS", "OC", "NE", "PIB")
str(data)
data<-data.frame(data)
head(data)

descrip<-stargazer(data, type="html",
  digits = 2,
  single.row = FALSE,
  intercept.bottom = TRUE,
  df = TRUE,
  align = TRUE,
  no.space=TRUE,
  ci = TRUE,
  ci.level = 0.95,
  out = "descriptivos.doc")

ggplot(data = data, aes(x=as.character(A), y=SX)) +
  geom_boxplot(fill="steelblue") +
  labs(title="Cobertura privada", x="Años", y="Varones")
```

```
ggplot(data = data, aes(x=as.character(A), y=HB)) +  
  geom_boxplot(fill="steelblue") +  
  labs(title="Cobertura privada", x="Años", y="Población")  
  
ggplot(data = data, aes(x=as.character(A), y=CS)) +  
  geom_boxplot(fill="steelblue") +  
  labs(title="Cobertura privada", x="Años", y="Clase I")  
  
ggplot(data = data, aes(x=as.character(A), y=OC)) +  
  geom_boxplot(fill="steelblue") +  
  labs(title="Cobertura privada", x="Años", y="Trabajando")  
  
ggplot(data = data, aes(x=as.character(A), y=NE)) +  
  geom_boxplot(fill="steelblue") +  
  labs(title="Cobertura privada", x="Años", y="Universitarios")  
  
ggplot(data = data, aes(x=as.character(A), y=PIB)) +  
  geom_boxplot(fill="steelblue") +  
  labs(title="Cobertura privada", x="Años", y="PIB")  
  
mco <- plm(SX~CS+OC+NE+HB+PIB, index=c("Y"), model="pooling", effect = "individual",  
data=data)  
summary(mco)  
  
regFixByInd <- plm(SX~CS+OC+NE+HB+PIB, index = c("Y"), model="within", data=data)  
  
summary(regFixByInd)  
  
#Regresión efectos fijos entre grupos  
regFixByGr <- plm(SX ~ CS+OC+NE+HB+PIB, index=c("Y"), model="between", effect =  
"individual", data=data)  
  
summary(regFixByGr)  
  
#Regresión con efectos aleatorios  
  
regRand<- plm(SX~CS+OC+NE+HB+PIB, index=c("Y"), model="random", random.method =  
"walhus", data=data)  
  
summary(regRand)
```

```
#Constrastes de Hipótesis
```

```
pFtest(regFixByInd,regRand) #Test F para efectos individuales
```

```
phptest(regFixByInd,regRand) #Test Hausman
```

```
RR<-stargazer(mco, regFixByInd, regFixByGr, regRand, type="html",  
              title="Tabla comparativa de regresiones",  
              dep.var.labels = c("Variable dependiente (Sexo)",  
              covariate.labels = c("Clase Social", "Ocupación", "Nivel de estudios",  
              "Población", "PIB"),  
              column.labels = c("Lineal", "Fijo Individual", "Fijo Grupos","Efecto  
Aleatorio"),  
              digits = 2,  
              single.row = FALSE,  
              intercept.bottom = TRUE,  
              df = TRUE,  
              align = TRUE,  
              no.space=TRUE,  
              out = "regresiones_sexo.doc")
```

```
mco <- plm(CS~SX+OC+NE+HB+PIB, index=c("Y"), model="pooling",effect = "individual",  
data=data)  
summary(mco)
```

```
regFixByInd <- plm(CS~SX+OC+NE+HB+PIB, index = c("Y"), model="within", data=data)
```

```
summary(regFixByInd)
```

```
regFixByGr <- plm(CS ~ SX+OC+NE+HB+PIB, index=c("Y"), model="between", effect =  
"individual", data=data)
```

```
summary(regFixByGr)
```

```
regRand<- plm(CS~SX+OC+NE+HB+PIB, index=c("Y"), model="random", random.method =  
"walhus", data=data)
```

```
summary(regRand)
```

```
pFtest(regFixByInd,regRand) #Test F para efectos individuales

phtest(regFixByInd,regRand) #Test Hausman

RR<-stargazer(mco, regFixByInd, regFixByGr, regRand, type="html",
              title="Tabla comparativa de regresiones",
              dep.var.labels = c("Variable dependiente (Clase social)"),
              covariate.labels = c("Clase Social", "Ocupación", "Nivel de estudios",
"Población", "PIB"),
              column.labels = c("Lineal", "Fijo Individual", "Fijo Grupos","Efecto
Aleatorio"),
              digits = 2,
              single.row = FALSE,
              intercept.bottom = TRUE,
              df = TRUE,
              align = TRUE,
              no.space=TRUE,
              out = "regresiones_clase.doc")

#-----Ocupación -----#
mco <- plm(OC~CS+SX+NE+HB+PIB, index=c("Y"), model="pooling",effect = "individual",
data=data)
summary(mco)

regFixByInd <- plm(OC~CS+SX+NE+HB+PIB, index = c("Y"), model="within", data=data)

summary(regFixByInd)

#Regresión efectos fijos entre grupos
regFixByGr <- plm(OC ~ CS+SX+NE+HB+PIB, index=c("Y"), model="between", effect =
"individual", data=data)

summary(regFixByGr)

#Regresión con efectos aleatorios

regRand<- plm(OC~CS+SX+NE+HB+PIB, index=c("Y"), model="random", random.method =
"walhus", data=data)

summary(regRand)
```

```
#Constrastes de Hipótesis
```

```
pFtest(regFixByInd,regRand) #Test F para efectos individuales
```

```
phptest(regFixByInd,regRand) #Test Hausman
```

```
RR<-stargazer(mco, regFixByInd, regFixByGr, regRand, type="html",  
              title="Tabla comparativa de regresiones",  
              dep.var.labels = c("Variable dependiente (Ocupación)"),  
              covariate.labels = c("Clase Social", "Ocupación", "Nivel de estudios",  
"Población", "PIB"),  
              column.labels = c("Lineal", "Fijo Individual", "Fijo Grupos", "Efecto  
Aleatorio"),  
              digits = 2,  
              single.row = FALSE,  
              intercept.bottom = TRUE,  
              df = TRUE,  
              align = TRUE,  
              no.space=TRUE,  
              out = "regresiones_ocupación.doc")
```

```
#----- Nivel Educativo -----#
```

```
mco <- plm(NE~CS+OC+SX+HB+PIB, index=c("Y"), model="pooling",effect = "individual",  
data=data)  
summary(mco)
```

```
#Regresión con efectos fijos cada individuo
```

```
regFixByInd <- plm(NE~CS+OC+SX+HB+PIB, index = c("Y"), model="within", data=data)
```

```
summary(regFixByInd)
```

```
#Regresión efectos fijos entre grupos
```

```
regFixByGr <- plm(NE ~ CS+OC+SX+HB+PIB, index=c("Y"), model="between", effect =  
"individual", data=data)
```

```
summary(regFixByGr)
```

```
#Regresión con efectos aleatorios
```

```
regRand<- plm(NE~CS+OC+SX+HB+PIB, index=c("Y"), model="random", random.method =  
"walhus", data=data)  
  
summary(regRand)  
  
#Constrastes de Hipótesis  
  
pFtest(regFixByInd,regRand) #Test F para efectos individuales  
  
phptest(regFixByInd,regRand) #Test Hausman  
  
#RR: Regresion Results  
  
RR<-stargazer(mco, regFixByInd, regFixByGr, regRand, type="html",  
              title="Tabla comparativa de regresiones",  
              dep.var.labels = c("Variable dependiente (Nivel Estudios)"),  
              covariate.labels = c("Clase Social", "Ocupación", "Nivel de estudios",  
"Población","PIB"),  
              column.labels = c("Lineal", "Fijo Individual", "Fijo Grupos","Efecto  
Aleatorio"),  
              digits = 2,  
              single.row = FALSE,  
              intercept.bottom = TRUE,  
              df = TRUE,  
              align = TRUE,  
              no.space=TRUE,  
              out = "regresiones_NE.doc")  
  
# ----- Población Habitantes -----#  
  
mco <- plm(HB~CS+OC+NE+SX+PIB, index=c("Y"), model="pooling",effect = "individual",  
data=data)  
summary(mco)  
  
#Regresión con efectos fijos cada individuo  
  
regFixByInd <- plm(HB~CS+OC+NE+SX+PIB, index = c("Y"), model="within", data=data)  
  
summary(regFixByInd)
```

```
#Regresión efectos fijos entre grupos
regFixByGr <- plm(HB ~ CS+OC+NE+SX+PIB, index=c("Y"), model="between", effect =
"individual", data=data)

summary(regFixByGr)

#Regresión con efectos aleatorios

regRand<- plm(HB~CS+OC+NE+SX+PIB, index=c("Y"), model="random", random.method =
"walhus", data=data)

summary(regRand)

#Constrastes de Hipótesis

pFtest(regFixByInd,regRand) #Test F para efectos individuales

phptest(regFixByInd,regRand) #Test Hausman

RR<-stargazer(mco, regFixByInd, regFixByGr, regRand, type="html",
              title="Tabla comparativa de regresiones",
              dep.var.labels = c("Variable dependiente (Habitantes)"),
              covariate.labels = c("Clase Social", "Ocupación", "Nivel de estudios",
"Población", "PIB"),
              column.labels = c("Lineal", "Fijo Individual", "Fijo Grupos","Efecto
Aleatorio"),
              digits = 2,
              single.row = FALSE,
              intercept.bottom = TRUE,
              df = TRUE,
              align = TRUE,
              no.space=TRUE,
              out = "regresiones_habitantes.doc")
```

Modelado

```
library(wooldridge)
```

```
library(stargazer)
library(pscl)
library(mfx, betareg)
library(car)
library(readxl)
library(dplyr)
library(texreg)

datos<- excel_sheets("DataEconomica.xlsx")
data_2003 <- read_excel("DataEconomica.xlsx", sheet = "2003")
data_2006 <- read_excel("DataEconomica.xlsx", sheet = "2006")
data_2011 <- read_excel("DataEconomica.xlsx", sheet = "2011")
data_2017 <- read_excel("DataEconomica.xlsx", sheet = "2017")

names(data_2003) = c("Y", "SX", "ED", "CS", "NE", "OC", "HB", "PIB")
str(data_2003)

names(data_2006) = c("Y", "SX", "ED", "CS", "NE", "OC", "HB", "PIB")
str(data_2006)

names(data_2011) = c("Y", "SX", "ED", "CS", "NE", "OC", "HB", "PIB")
str(data_2011)

names(data_2017) = c("Y", "SX", "ED", "CS", "NE", "OC", "HB", "PIB")
str(data_2017)

# ----- Modelo Logit -----
logit_2003 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2003, family =
"binomial")
summary(logit_2003)
pR2(logit_2003)
exp(coefficients(logit_2003))

log.odds_2003 <- predict(logit_2003, data.frame(SX = 0, ED = 35, CS = 6, NE=
7, OC= 4, HB = 1))
log.odds_2003
```

```
lg_2003 <- exp(log.odds_2003)/(1+exp(log.odds_2003))  
lg_2003
```

```
logit_2006 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2006, family =  
"binomial")  
summary(logit_2006)  
pR2(logit_2006)  
exp(coefficients(logit_2006))
```

```
log.odds_2006 <- predict(logit_2006, data.frame(SX = 0, ED = 35, CS = 6, NE=  
7, OC= 4, HB = 1))  
log.odds_2006
```

```
lg_2006 <- exp(log.odds_2006)/(1+exp(log.odds_2006))  
lg_2006
```

```
logit_2011 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2011, family =  
"binomial")  
summary(logit_2011)  
pR2(logit_2011)  
exp(coefficients(logit_2011))
```

```
log.odds_2011 <- predict(logit_2011, data.frame(SX = 0, ED = 35, CS = 6, NE=  
7, OC= 4, HB = 1))  
log.odds_2011
```

```
lg_2011 <- exp(log.odds_2011)/(1+exp(log.odds_2011))  
lg_2011
```

```
logit_2017 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2017, family =  
"binomial")  
summary(logit_2017)  
pR2(logit_2017)  
exp(coefficients(logit_2017))
```

```
log.odds_2017 <- predict(logit_2017, data.frame(SX = 0, ED = 35, CS = 6, NE=
7, OC= 4, HB = 1))
log.odds_2017

lg_2017 <- exp(log.odds_2017)/(1+exp(log.odds_2017))
lg_2017

# ----- Modelo Probit -----

probit_2003 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2003, family =
binomial(link = "probit"))
summary(probit_2003)
pR2(probit_2003)
coefficients(probit_2003)
probitvec <- mean(dnorm(predict(probit_2003, type="link")))
probitvec * coef(probit_2003)

pre_2003 <- predict(probit_2003, data.frame(SX = 0, ED = 35, CS = 6, NE = 7,
OC = 4, HB=1),
                    type = "response")
pre_2003

probit_2006 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2006, family =
binomial(link = "probit"))
summary(probit_2006)
pR2(probit_2006)
coefficients(probit_2006)
probitvec <- mean(dnorm(predict(probit_2006, type="link")))
probitvec * coef(probit_2006)

pre_2006 <- predict(probit_2006, data.frame(SX = 0, ED = 35, CS = 6, NE = 7,
OC = 4, HB=1),
                    type = "response")
pre_2006

probit_2011 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2011, family =
binomial(link = "probit"))
```

```
summary(probit_2011)
pR2(probit_2011)
coefficients(probit_2011)
probitvec <- mean(dnorm(predict(probit_2011, type="link")))
probitvec * coef(probit_2011)

pre_2011 <- predict(probit_2011, data.frame(SX = 0, ED = 35, CS = 6, NE = 7,
OC = 4, HB=1),
                    type = "response")
pre_2011

probit_2017 <- glm(Y ~ SX + ED + CS + NE+ OC + HB, data = data_2017, family =
binomial(link = "probit"))
summary(probit_2017)
pR2(probit_2017)
coefficients(probit_2017)
probitvec <- mean(dnorm(predict(probit_2017, type="link")))
probitvec * coef(probit_2017)

pre_2017 <- predict(probit_2017, data.frame(SX = 0, ED = 35, CS = 6, NE = 7,
OC = 4, HB=1),
                    type = "response")
pre_2017

comp <- stargazer(logit_2003, probit_2003, logit_2006, probit_2006,
logit_2011, probit_2011, logit_2017, probit_2017, type="html",
                  title="Tabla comparativa logit/ probit",
                  dep.var.labels = c("Elección seguro privado España"),
                  covariate.labels = c("Sexo", "Edad", "Clase Social", "Nivel
de Estudios", "Actividad económica", "Habitantes"),
                  digits = 3,
                  single.row = FALSE,
                  intercept.bottom = TRUE,
                  df = TRUE,
                  align = TRUE,
                  no.space=TRUE,
                  out = "probit_logit.doc")
```

```
#-----Evaluación total -----

data_tot <- read_excel("DataEconomica.xlsx", sheet = "Totales")
names(data_tot) = c("A", "Y", "SX", "ED", "CS", "NE", "OC", "HB", "PIB")
str(data_tot)

logit_tot <- glm(Y ~ A + SX + ED + CS + NE + OC + HB, data = data_tot, family =
"binomial")
summary(logit_tot)
pR2(logit_tot)
exp(coefficients(logit_tot))

log.odds_tot <- predict(logit_tot, data.frame(SX = 0, ED = 35, CS = 6, NE = 7,
OC = 4, HB = 1, A = 2003))
log.odds_tot

lg_tot <- exp(log.odds_tot)/(1+exp(log.odds_tot))
lg_tot
probit_tot <- glm(Y ~ A + SX + ED + CS + NE + OC + HB, data = data_tot, family
= binomial(link = "probit"))
summary(probit_tot)
pR2(probit_tot)
coefficients(probit_tot)
probitvec <- mean(dnorm(predict(probit_tot, type="link")))
probitvec * coef(probit_tot)

pre_tot <- predict(probit_tot, data.frame(SX = 0, ED = 35, CS = 6, NE = 7, OC
= 4, HB=1, A = 2003),
                    type = "response")
pre_tot
```

```
comp_tot <- stargazer(logit_tot, probit_tot, type="html",  
                      title="Tabla comparativa logit/ probit",  
                      dep.var.labels = c("Elección seguro privado España"),  
                      covariate.labels = c("Año", "Sexo", "Edad", "Clase Social",  
"Nivel de Estudios", "Actividad económica", "Habitantes"),  
                      digits = 3,  
                      single.row = FALSE,  
                      intercept.bottom = TRUE,  
                      df = TRUE,  
                      align = TRUE,  
                      no.space=TRUE,  
                      out = "probit_logit_tot.doc")
```