

Análisis de la demanda de másteres y programas de postgrado en España

por

María Martínez Aguas

Una tesis presentada conforme a los requisitos del Máster en Economía, Finanzas y Computación

Universidad de Huelva y Universidad Internacional de Andalucía

uhu.es

un
i Universidad
Internacional
de Andalucía
A

Noviembre 2018

Análisis de la demanda de másteres y programas de postgrado en España

María Martínez Aguas

Máster en Economía, Finanzas y Computación

Mónica Carmona
Universidad de Huelva

2018

Abstract

The goal of this work is performing an analysis from the results that the Family Budget Survey eases us with about the group of teaching expenses of the postgraduate. Where information related to the expenditure destined from Spanish households to higher education is facilitated, apart from the level of education that the household members have.

The data is obtained from the National Institute of Statistics from the Family Budget Survey.

We have obtained the database from micro data to carry out the study. Next, the Heckman model composed of two stages to know if an individual decides or not conducting postgraduate studies will be estimated, and conditioning to which they decide to continue their studies, they will decide how much they want to spend.

Finally, there will be conclusions elaborated from the obtained results.

Key words: higher education, master, Heckman, expenditure, econometric model, income, EPF

Resumen

Este trabajo tiene como objetivo realizar un análisis de los resultados que nos facilita la Encuesta de Presupuestos Familiares (EPF) sobre el grupo de gastos de enseñanza de postgrado. Donde se facilita información relacionada con el gasto destinado de los hogares españoles a la enseñanza superior además de, el nivel de educación que tienen los miembros del hogar.

Los datos se obtienen del Instituto Nacional de Estadística, a partir de la Encuesta de Presupuestos Familiares.

Hemos obtenido la base de datos a partir de microdatos para llevar a cabo el estudio. A continuación, se estimará un modelo Heckman compuesto por dos etapas para saber si un individuo decide o no realizar estudios de postgrado, y condicionado a que decide continuar con sus estudios, decidirá cuánto desea gastar

Finalmente, se elaboran conclusiones de los resultados obtenidos.

Palabras clave: educación superior, máster, Heckman, gasto, modelo econométrico, ingresos, EPF.

Tabla de contenidos

1.- Introducción	p. 1-3
2.- Revisión de la literatura	p. 4-5
2.1.- Determinantes en la demanda de Educación Superior	p. 5-6
3. Metodología	p. 7
3.1.- Datos	p. 7-9
3.2.- Descripción de las variables	p. 9-11
3.3.- Modelo teórico	p. 11-13
4.- Resultados	p. 14-22
5.- Conclusiones	p. 23-24
Referencias	p. 25-26
Apéndice	p. 27-29

Lista de Tablas

Tabla 1. Clasificación grupos de gasto de la EPF	p. 8
Tabla 2. Descripción de las variables	p. 10-11
Tabla 3. Estadísticos descriptivos	p. 14
Tabla 4. Estadísticos descriptivos de variables discretas	p. 15-16
Tabla 5. Correlaciones entre variables	p. 17
Tabla 6. Coeficientes y errores estándar del modelo de selección muestral	p. 18-20

Lista de Figuras

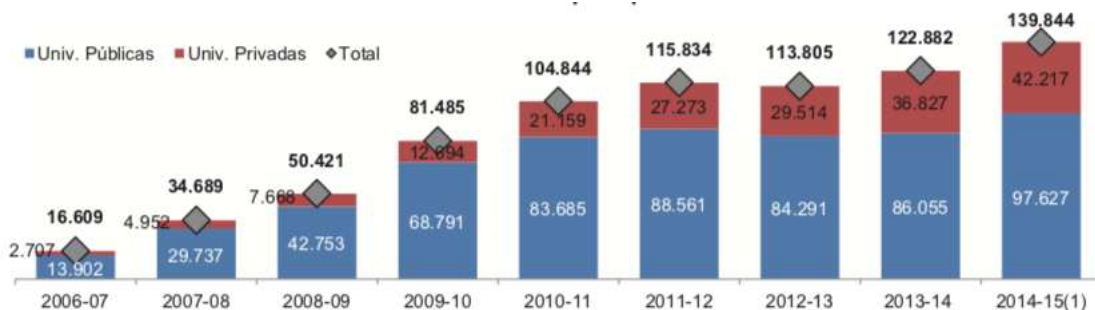
Gráfico 1. Evolución de estudiantes matriculados en máster por tipo de Universidad	p. 1
Gráfico 2. Distribución de la oferta de másteres por rama de enseñanza	p. 2
Imagen 1. Fichero de usuario de hogar 2017 formato plano	p. 27
Imagen 2. Diccionario en STATA	P. 28
Imagen 3. Base de Datos	p. 29

1 Introducción

Actualmente, nos encontramos ante un entorno laboral que exigirá máxima cualificación y una de las salidas más acertadas para continuar con la formación es la de hacer un máster. Durante el año 2014/2015 el número total de estudiantes matriculados en Másteres ascendió a 139.844, según datos oficiales del Ministerio de Educación, Cultura y Deporte. De esta cifra las mujeres representaban el 53,4% de la demanda total. El máster aportará la diferenciación que demanda el mercado laboral, especialización y potenciación de ciertas habilidades, la posibilidad de optar a un trabajo mejor remunerado percibiendo un mayor salario anual y otorga mayor probabilidad de acceder a puestos de trabajo de más responsabilidad. En la actualidad, la cuarta parte de los trabajadores en activo tiene formación media y un 32% formación alta, se espera que en 2026 ambas cualificaciones asciendan al 65% de la oferta total frente al 57% actual (“El País”, 2018).

A continuación, se muestra una tabla con la evolución que ha experimentado la demanda de másteres en España atendiendo al tipo de universidad.

Gráfico 1. Evolución de los estudiantes matriculados en máster por tipo de Universidad.

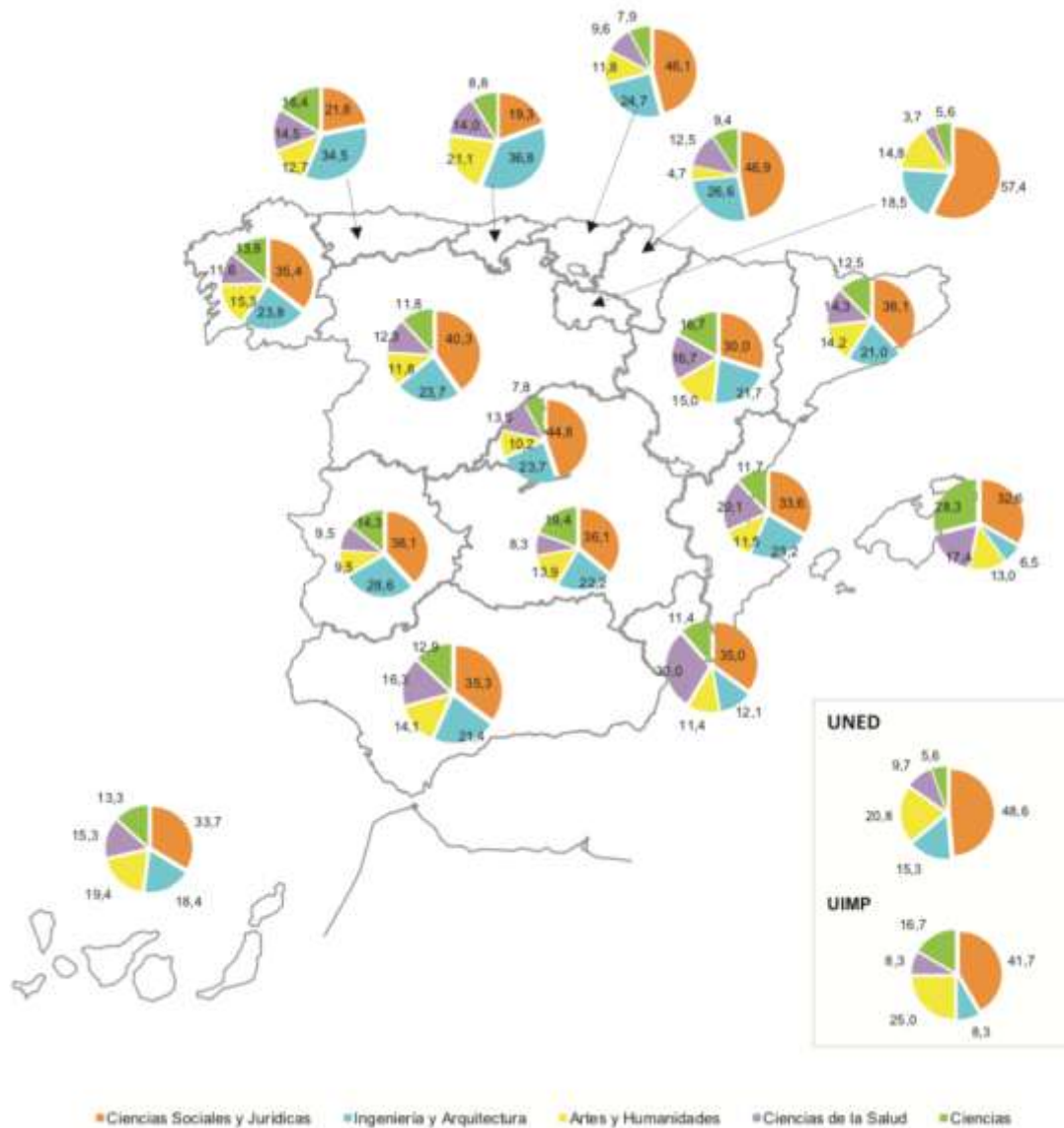


Fuente: Ministerio de Educación, Cultura y Deporte.

Con la crisis económica se incrementó la demanda en educación superior, debido a que se reducía el coste de oportunidad de continuar en el sistema educativo y no acceder al mercado laboral, según datos del Ministerio de Educación, la demanda de másteres online se incrementó en más de un 300% tras la crisis. Por este motivo aumentó el nivel de exigencia de las empresas, el 90% de las ofertas de trabajo que se publican cada día exigen tener un máster y conocimientos de un segundo idioma como requisitos mínimos (“El Economista”, 2018).

En el siguiente gráfico, se muestra la distribución de la oferta de másteres por rama de enseñanza en el territorio español para el año 2015/2016.

Gráfico 2. Distribución de la oferta de másteres por rama de enseñanza. Curso 2015/2016



Fuente: Ministerio de Educación, Cultura y Deporte.

Es posible demandar dos tipos de másteres:

- ✓ Títulos oficiales dirigidos a la investigación, profesional y académico. El tipo de máster que más ha crecido, sin tener en cuenta los títulos propios, ha sido el máster profesionalizante.

- ✓ Máster propio de una universidad, programa diseñado según las necesidades que demanda el mercado laboral.

Las empresas están demandando profesionales capaces de adaptarse a la continua transformación digital que están experimentando. Por este motivo cada vez son más los profesionales con experiencia que deciden realizar másteres para renovar su perfil, de esta manera tienen la posibilidad de promocionar accediendo a puestos de trabajo de mayor categoría o liderar proyectos estratégicos dentro de una compañía.

Sin embargo, existe una gran incertidumbre en torno a los másteres o programas de postgrado, ¿cuál es el mejor momento para matricularse en un máster? en algunos casos es imprescindible matricularse en el máster tras finalizar los estudios universitarios, ya que resulta obligatorio para ejercer una determinada profesión o acceder a determinados puestos de trabajo. No obstante, existe una amplia gama de másteres, cursos de postgrado o cursos de especialización para profesionales con cierta experiencia, de manera que consiguen ser más competitivos y aportan un valor añadido a la empresa. En los últimos 15 años, las universidades online están siendo más competitivas frente a las presenciales para un perfil de estudiante más adulto que compagina sus estudios con su trabajo u otras obligaciones (“El Economista”, 2018).

Este trabajo realiza un análisis sobre la decisión de los individuos de demandar educación superior. Condicionado a la decisión de continuar con la formación académica realizaremos un estudio de cuánto estarían dispuestos a invertir estos hogares, basándonos en diferentes teorías económicas de la educación desarrolladas a lo largo del tiempo, que serán explicadas a continuación, y utilizando diferentes variables explicativas que influyen en la toma de esta decisión.

2 Revisión de la literatura

Para realizar este estudio sobre la decisión de invertir en educación superior, nos hemos basado en numerosas teorías sobre la demanda de educación que analizan diferentes factores y variables socioeconómicas que influyen en la toma de esta decisión.

En el contexto de la teoría del capital humano (Becker 1975 y Mincer 1974) se consideran las inversiones en educación superior como una inversión en capital humano, de manera que el individuo resultará más cualificado de cara al mercado laboral y esto se traduce en mayores ganancias futuras. Por tanto, esta teoría entiende la educación como un bien de inversión y no como un bien de consumo.

Sin embargo, algunos economistas consideran esta teoría insuficiente por no tener en cuenta el factor consumo en la demanda de educación. Schaafsma (1976), Gullason (1989), Kodde y Ritzen (1994) proponen la teoría de consumo, según ésta, la demanda de educación podría estar motivada por el deseo de adquirir conocimientos, y dependerá del nivel de renta, de los precios y de otros factores de consumo.

Para otros autores (MacConnell 2003) no existe algún método para saber qué parte de los gastos en la demanda de educación superior es inversión y cuál es consumo. Por tanto, también consideran la teoría de consumo incompleta ya que no tendría en cuenta los aspectos relacionados con la inversión.

El supuesto de señalización es otra perspectiva en el análisis de acceso a la educación de Arrow (1973), Spence (1973) y Psacharopoulos (1979), de esta manera, se rechaza la idea de la educación como un medio para aumentar la productividad y los conocimientos. Esta hipótesis se basa en el supuesto de que la educación se utiliza como una señal de distinción entre los trabajadores (Stiglitz 1975).

La teoría institucionalista desarrollada por Doeringer y Piore (1971), Piore (1974) y Thurrow (1975) sostiene que la productividad está vinculada al puesto de trabajo, de esta manera la empresa sería un factor más activo dentro del proceso de contratación, en contraposición a lo que consideran otras teorías, por ejemplo, la teoría del capital humano.

La teoría de la socialización Bowles y Gintis (1975) resalta que la educación más que proporcionar igualdad de oportunidades, ejerce de nexo de transmisión de la clase social entre generaciones. Por tanto, se extrae como conclusión que no existe una relación de causalidad entre la inversión en educación y las rentas futuras percibidas, si no que la renta percibida por un individuo se debe a la herencia familiar.

La diferencia principal con la teoría del capital humano es la influencia del factor socioeconómico familiar.

Atendiendo a las diferentes teorías que consideran la educación como un bien de consumo, vamos a analizar diferentes factores socioeconómicos relacionados con el hogar, la localización, características personales, etc... que, según las teorías antes mencionadas, podrían influir en el gasto en másteres de los hogares.

2.1. Determinantes en la demanda de educación superior

En el acceso a la educación influyen numerosos factores como las características personales del individuo, el entorno familiar, el entorno socioeconómico, algún factor derivado del mercado laboral, las instituciones y el desarrollo del sistema de educación del país. También cabe destacar el género del individuo, debido a que en los últimos años, según numerosos estudios, la demanda de educación postobligatoria se ha incrementado por parte de las mujeres (Calero 1996, Rahona 2006, Jerrim y Vignoles 2015).

Otras características estudiadas han sido, el coeficiente intelectual del individuo, el talento Glomm y Ravikumar (1992) y la edad (López, S. F., Gómez, I. N. y Sanmartín, E. R. 2001). También influyen las motivaciones personales, el tiempo dedicado a estudiar, las expectativas Plug (2002).

Por otra parte, también se han analizado las características familiares y el hogar donde reside el individuo, la educación de los padres y el nivel de ingresos familiar (Valiente 2003), que es un indicador de la posibilidad de financiar los estudios postobligatorios, considerando el nivel educativo de los padres como un indicador de la renta familiar.

En esta línea, podemos considerar que otro factor socioeconómico que influye en la decisión de invertir en educación, es la tasa de desempleo, ya que si el sustentador principal ha pasado por etapas de desempleo puede disminuir su capacidad de financiación.

Además de los determinantes mencionados anteriormente, podemos indicar algunos factores macroeconómicos, demográficos, institucionales y de carácter político que tienen influencia en la demanda de educación, por ejemplo, la incertidumbre relativa a las rentas futuras, los impuestos, la oferta de trabajo, etc... (Levhari y Weiss 1974, Eaton y Rosen 1980, Kodde 1986 y Lundborg y Nilsson 2014).

3 Metodología

Como hemos mencionado anteriormente, el objetivo de este trabajo es estimar un modelo econométrico que determine la inversión de los hogares en educación superior. El modelo está dividido en dos procesos, la primera etapa formada por la ecuación de participación que determina si en el hogar se decide gastar en estudios de postgrado. La segunda etapa, formada por la ecuación de interés y condicionada a consumir estudios de postgrado, determina cuánto están dispuestos a invertir los individuos en su formación.

3.1. Datos

Para realizar este análisis vamos a utilizar los datos recogidos a través de la Encuesta de Presupuestos Familiares, realizada por el INE para el año 2017, que ofrece información sobre la naturaleza y destino de los gastos de consumo, así como sobre diversas características referente a las condiciones de vida de los hogares.

A lo largo del tiempo se han realizado dos tipos de EPF, estructurales o básicas cada ocho o diez años y coyunturales o trimestrales. A partir de 1997 se implantó una encuesta que tratase de agrupar los aspectos más importantes de los dos tipos de encuestas que se estaban realizando hasta el momento.

Entre los objetivos principales de la EPF se encuentran:

- Obtención de estimaciones del agregado gasto de consumo anual de los hogares para el conjunto nacional y para las comunidades autónomas, así como su clasificación según diversas variables del hogar.
- Estimación del cambio interanual del agregado gasto de consumo para el conjunto nacional y para las comunidades autónomas.
- Estimación del consumo en cantidades físicas de determinados bienes alimenticios para el conjunto nacional.
- Estimación del gasto como instrumento para la obtención del consumo privado en la Contabilidad Nacional.
- Estimación de la estructura de ponderaciones a partir del gasto necesario para el cálculo del IPC.

Los gastos de consumo que se registran en la EPF además de referirse al flujo monetario que destina el hogar al pago de bienes y servicios de consumo final se refiere también al valor de los consumos efectuados por los hogares en concepto de autoconsumo, autosuministro, salario en especie, comidas gratuitas o bonificadas y alquiler imputado a la vivienda en la que reside el hogar.

La clasificación que realiza la EPF está basada en la nueva clasificación europea de consumo (ECOICOP) y ofrece un mayor desglose de algunas partidas de gasto, además permite la comparación con otras estadísticas como el Índice de Precios de Consumo (IPC).

Esta clasificación divide las partidas de gasto realizadas en el hogar en 12 grupos:

Tabla 1. Clasificación grupos de gasto de la EPF

1. Alimentos y bebidas no alcohólicas

2. Bebidas alcohólicas y tabaco

3. Vestido y calzado

4. Vivienda, agua, electricidad, gas y otros combustibles

5. Muebles, artículos del hogar y artículos para el mantenimiento corriente del hogar

6. Sanidad

7. Transporte

8. Comunicaciones

9. Ocio y cultura

10. Enseñanza

11. Restaurantes y hoteles

12. Otros bienes y servicios

Fuente: Elaboración propia a partir de la Encuesta de Presupuestos Familiares recogida en el INE

De esta tabla he utilizado los datos del grupo 10 correspondiente a “Enseñanza”, y dentro de este, el subgrupo 10.3.0.3 que hace referencia a los másteres.

3.2. Descripción de las variables

Para la estimación del modelo nos hemos basado en diferentes teorías económicas sobre educación, introduciendo una serie de variables contenidas en la EPF y que a priori pensamos que pueden influir en la toma de la decisión de invertir en educación.

En primer lugar, la ecuación de gasto contará con la variable que identifica la cantidad destinada al gasto en nuestra partida clave, esta variable dependiente toma un valor distinto de cero si realiza gasto en másteres y en caso contrario no tendrá ningún valor (missing values), debido a que Heckman no trabaja con ceros sino con “missing values”.

Los ingresos mensuales netos totales del hogar será una variable explicativa del modelo, ya que numerosos autores (Petrongolo y San Segundo 1999 y Modrego 1987) consideran que existe una relación positiva entre los ingresos del hogar y la demanda en educación. Por ello consideramos que a mayor renta, existe mayor posibilidad de acceder a niveles de educación superior. Esta variable ha sido modificada con logaritmos para poder interpretar los resultados en términos de elasticidades.

Son varios los trabajos (Rahona 2005, Calero 1996, González y Dávila 1998, Albert 1998, Marcenaro y Navarro 2001, San Segundo 2003) que destacan la importancia del nivel educativo de los padres o sustentador principal debido a la transmisión generacional. Podemos suponer que a mayor nivel de formación de los padres mayor es la propensión a invertir parte de los recursos en formación de los hijos, con lo cual se elevará la partida relacionada con los estudios superiores.

Con el objetivo de controlar la situación laboral del sustentador principal hemos utilizado la variable OCUPACION, que es una variable dicotómica que toma el valor 1 si el sustentador principal estaba ocupado una semana antes de la realización de la muestra y el valor 0 en caso contrario, influye positivamente en la demanda de educación superior una mejor posición en el mercado laboral o mayores ingresos percibidos (Cea y Mora 1992, Calero 1996, González y Dávila 1998, Albert 2000, Rahona 2005).

Sexo, variable dicotómica que toma el valor 1 si el individuo es mujer y el valor 0 si el individuo es hombre. Son varios los trabajos que afirman que se ha incrementado la demanda de másteres por parte de las mujeres, esta situación se explica mediante diferentes razones: Las mujeres

necesitan un mayor nivel educativo para competir con los hombres (Mora 1997 “Teoría del filtro” y Albert 2000), señalan la discriminación que pueden sufrir las mujeres en el mercado laboral. Marcenaro y Navarro (2001) ponen de manifiesto las mayores tasas de desempleo y el mayor coste de oportunidad que sufren las mujeres. Además, otros trabajos exponen que existen variables que afectan de forma diferente en la demanda según si el individuo es hombre o mujer (Modrego 1987, Cañada 1999, Cea y Mora 1992, Marcenaro y Navarro 2001, Albert 2000).

El estado civil es una variable que afecta a la demanda de determinados bienes, de hecho en algunos casos suele ser un buen criterio de segmentación, para el caso de la demanda de estudios de postgrado existen estudios que avalan esta relación entre demanda y género (Marcenaro y Navarro 2001), y que se espera que los casados tengan menos probabilidad de demandar educación superior debido a que tienen cargas familiares que atender. Pero en nuestro análisis vamos a considerar la demanda en el hogar, independientemente de quién consuma este tipo de bienes, a pesar de ello vamos a incluirla entre las variables retenidas en el modelo.

Finalmente, se han utilizado otras variables relacionadas con las características del hogar como, la situación geográfica y las características del individuo (número de miembros del hogar de 18 años o más), presencia de hijos dependientes, localización geográfica y edad del sustentador principal. Estas variables mantienen los valores originales que presentan en la Encuesta de Presupuestos Familiares y son indicadores que afectan a la decisión de los individuos de destinar parte de los ingresos a la partida de estudios de postgrado.

Con la ayuda del programa estadístico STATA (versión 14), se ha procedido a la modificación de alguna de las variables, que se exponen en la tabla a continuación, que serán las variables utilizadas para estimar el modelo econométrico.

Tabla 2. Descripción de las variables

Nombre de la variable	Naturaleza de la variable	Descripción variable	Valores (Base de Datos)
GASTO_HEC*	Dependiente	Hogares que realizan gasto en educación	
SEXO*	Independiente	Sexo del sustentador principal	<ul style="list-style-type: none"> • 1 Mujer • 0 Hombre
NUTS1	Independiente	Regiones	<ul style="list-style-type: none"> • 1 Noroeste • 2 Noreste • 3 Comunidad de Madrid • 4 Central

			<ul style="list-style-type: none"> • 5 Este • 6 Sur • 7 Canarias
ESTUDREDSP	Independiente	Nivel de estudios del sustentador principal reducidos	<ul style="list-style-type: none"> • 1 Inferior primera etapa de Educación Secundaria • 2 Primera etapa de Educación Secundaria • 3 Segunda etapa de Educación Secundaria • 4 Educación Superior
INGRESOS*	Independiente	Ingresos mensuales netos	
OCUPACION*	Independiente	¿Estaba el sustentador principal ocupado una semana antes de la realización de la encuesta?	<ul style="list-style-type: none"> • 1 Ocupado • 0 No ocupado
NMIEMB6	Independiente	Número de miembros del hogar de 18 años o más	0-20
NHIJOSD	Independiente	Número de hijos dependientes	0-19
EDADSP	Independiente	Edad del sustentador principal	85 personas de 85 o más años 16-84 resto de personas
ECIVILLEGALSP	Independiente	Estado civil del sustentador principal	<ul style="list-style-type: none"> • 1 Soltero • 2 Casado • 3 Viudo • 4 Separado • 5 Divorciado

*Variables modificadas o creadas a partir de las existentes en la Encuesta de Presupuestos Familiares

Fuente: Elaboración propia a partir de la Encuesta de Presupuestos Familiares

3.3. Modelo teórico

Para que un individuo pueda acceder a programas de Máster o postgrado, es necesario que haya completado sus estudios universitarios. Por tanto, hay que tener en cuenta que la decisión de continuar con la educación superior se basa en una decisión anterior, es decir, una vez que el individuo haya finalizado sus estudios universitarios decidirá si demanda o no educación de postgrado. Puesto que en la muestra aparecerán muchas familias que no registran este tipo de

partidas, sólo podríamos estudiar los factores que determinan la cuantía del gasto. La aplicación del modelo Heckman nos permitirá estimar por un lado la decisión de consumir este tipo de productos y por otro la cuantía del gasto. Por tanto, la ecuación de decisión o de participación consiste en la decisión de acceder o no a Másteres o programas de postgrado. La segunda ecuación o ecuación de interés corresponde al gasto en educación superior condicionada a que decida continuar con su formación.

En primer lugar, un individuo decide si quiere continuar o no con su formación académica cursando un máster o programa de postgrado. En segundo lugar, condicionado que decida continuar con su formación, decidirá cuánto desea gastar en dicha formación.

De manera que:

$$y_i = x_i' \beta + \varepsilon_i$$

$$d_i = 1\{w_i' \gamma + u_i > 0\}$$

La primera ecuación es la ecuación de interés y la segunda es la ecuación de participación.

La ecuación de cuánto decide gastar es observada sólo si la variable indicador toma el valor 1 ($d_i = 1$) y 0 en caso contrario.

Para estimar la ecuación de interés es necesario que:

$$E(\varepsilon_i | x_i) = 0$$

Sin embargo, esto no garantiza que:

$$E(\varepsilon_i | x_i, d_i = 1) = 0$$

Si el término error en la ecuación de participación está correlacionado con el término de error de la ecuación de interés, entonces:

$$E(\varepsilon_i | x_i, d_i = 1) \neq 0$$

Corregir directamente por el sesgo de selección calculando el valor de $E(\varepsilon_i|x_i, d_i = 1)$, e incluyéndolo en la ecuación de interés como una nueva variable explicativa en la regresión para las observaciones para las que observamos cuánto gasta en educación superior.

Suponiendo que (ε_i, U_i) se distribuye como una Normal bivalente $(0,0,1,\rho)$, donde ρ recoge la correlación entre las dos decisiones.

$$\begin{aligned} E(\varepsilon_i|x_i, d_i = 1) &= E(\varepsilon_i|x_i, w_i'\gamma + u_i > 0) = \sigma_\varepsilon E\left[\frac{\varepsilon_i}{\sigma_\varepsilon} \middle| x_i, w_i'\gamma + u_i > 0\right] \\ &= \rho \sigma_\varepsilon E[u_i|x_i, w_i'\gamma + u_i > 0] \end{aligned}$$

$$E[u_i|x_i, w_i'\gamma + u_i > 0]$$

Este término no es más que la media de una variable aleatoria normal estándar truncada por debajo de $-w_i'\gamma$

$$E[u_i|x_i, w_i'\gamma + u_i > 0] = \frac{\phi(-w_i'\gamma)}{1 - \Phi(-w_i'\gamma)} = \frac{\phi(w_i'\gamma)}{1 - \Phi(w_i'\gamma)} = \lambda(w_i'\gamma)$$

$\lambda(w_i'\gamma)$ Este término se denomina lambda de Heckman o inversa del ratio de Mills.

La estimación del modelo de selección de muestra se lleva a cabo en dos pasos:

1. Se estima la ecuación de selección por medio de un modelo probit para construir la lambda de Heckman, utilizando los parámetros γ del probit.
2. Se lleva a cabo la regresión de la ecuación de interés (con la muestra de observaciones positivas), incluyendo como variable explicativa la nueva variable construida en el primer paso.

$$y_i = x_i'\beta + \rho\sigma_\varepsilon\lambda(w_i'\hat{\gamma}) + \varepsilon_i - \rho\sigma_\varepsilon\lambda(w_i'\hat{\gamma}) = x_i'\beta + \rho\sigma_\varepsilon\lambda(w_i'\hat{\gamma}) + e_i$$

Los errores de esta ecuación (e_i) tienen media cero por construcción.

4 Resultados

Procedemos a estimar el modelo con ayuda de STATA sobre los datos (más o menos transformados) procedentes de la Encuesta de Presupuestos Familiares del año 2017. El INE publica los microdatos (fichero de hogares y fichero de gasto) correspondientes a esta encuesta en ficheros planos, por ello, para poder trabajar con ellos se elaboraron dos diccionarios para convertirlo a formato de datos de STATA. En el apéndice se muestra un ejemplo de la transformación que ha experimentado la base de datos. Se ha procedido a la fusión de las bases de datos obtenidas para tener en la misma los datos correspondientes a las características de los hogares y las partidas de gasto que realizan cada una de las familias. De esta base eliminamos todas las partidas de gasto menos la relevante para nuestro estudio, el grupo correspondiente al código “10303” estudios de másteres y programas de postgrado. Finalmente, para estimar el modelo se han tenido en cuenta un conjunto de variables que influyen en la decisión de demandar educación superior, que han sido seleccionadas en base a la revisión de la literatura previa.

A continuación, mostramos una tabla (tabla 3) en la que aparecen los estadísticos descriptivos de las variables que hemos utilizado en la estimación del modelo econométrico.

Tabla 3. Estadísticos descriptivos

Variable	Obs.	Mean	Std. Dev.	Min	Max
INGRESOS	22.028	7,406234	0,8884414	0	10,00256
NMIEMB6	22.028	2,134601	0,8832493	1	8
NHIJOSD	22.028	0,6083621	0,9006674	0	8
EDADSP	22.028	56,16261	14,97034	18	85

Fuente: Elaboración propia a partir de los datos de la EPF utilizando el programa STATA

En la siguiente tabla se muestran los estadísticos descriptivos de las variables discretas consideradas.

Tabla 4. Estadísticos descriptivos de variables discretas

Variable	Freq.	Percent	Cum.
Sexo			
• 0 Hombre	14,893	67,61	67,61
• 1 Mujer	7,135	32,39	100,00
Estudios Completados			
• 1 Inferior a la primera etapa de Educación Secundaria	4,452	20,21	20,21
• 2 Primera etapa Educación Secundaria	6,669	30,28	50,49
• 3 Segunda etapa Educación Secundaria	4,016	18,23	68,72
• 4 Educación Superior	6,891	31,28	100,00
Ocupación			
• 0 No ocupado	9,460	42,95	42,95
• 1 Ocupado	12,568	57,05	100,00
Regiones			
• 1 Noroeste	3,027	13,74	13,74
• 2 Noreste	4,618	20,96	34,71
• 3 Comunidad de Madrid	1,663	7,55	42,26
• 4 Central	3,627	16,47	58,72
• 5 Este	4,500	20,43	79,15
• 6 Sur	3,593	16,31	95,46
• 7 Canarias	1,000	4,54	100,00

Estado Civil Sustentador Principal			
• 1 Soltero			
• 2 Casado	4,153	18,85	18,85
• 3 Viudo	13,083	59,39	78,25
• 4 Separado	2,595	11,78	90,03
• 5 Divorciado			
	632	2,87	92,90
	1,565	7,10	100,00

Fuente: Elaboración propia a partir de los datos de la EPF utilizando el programa STATA

De esta última tabla podemos destacar respecto a la variable sexo que aproximadamente el 68% de los hogares dependen del hombre como sustentador principal, frente al 32% en el caso de las mujeres.

Referido a la variable estudios del sustentador principal, podemos observar que en un 20% de los hogares el sustentador principal solo tiene estudios inferiores a la primera etapa de Educación Secundaria, en el 30% de los hogares estudios correspondientes a la primera etapa de Educación Secundaria, un 18% estudios correspondientes a la segunda etapa de Educación Secundaria y en el 31% de los hogares el sustentador principal posee estudios superiores.

La variable ocupación representa que en el 57% de los hogares analizados, el sustentador principal estaba ocupado frente al 43% de hogares con sustentadores principales desocupados.

Finalmente, en cuanto al estado civil del sustentador principal podemos afirmar que en el 59% de los hogares el sustentador principal está casado, como ya mencioné anteriormente se espera que si el sustentador principal está casado disminuya la probabilidad de demandar másteres y programas de postgrado, debido a que puede tener más responsabilidades y cargas familiares.

En la siguiente tabla, se muestran las correlaciones existentes entre las variables seleccionadas para la estimación del modelo.

Tabla 5. Correlaciones entre variables

	INGRESOS	ESTUDRED	OCUP	SEXO	NMIEMB6	NHIJOSD	NUTS1
INGRESOS	1,0000						
ESTUDRED	0,3079	1,0000					
OCUP	0,3045	0,4163	1,0000				
SEXO	-0,1388	0,0295	-0,0546	1,0000			
NMIEMB6	0,3039	-0,0134	0,1487	-0,2489	1,0000		
NHIJOSD	0,1243	0,1964	0,4095	-0,0845	0,2362	1,0000	
NUTS1	-0,0808	-0,0642	0,0292	-0,0065	0,0327	0,0633	1,0000
EDADSP	-0,0729	-0,4173	-0,6974	0,0351	-0,0320	-0,4566	-0,0581
ECIVIL	-0,0939	-0,1003	-0,1453	0,2514	-0,1372	-0,0392	0,0156

	EDADSP	ECIVILLEGALSP
EDADSP	1,0000	
ECIVILLEGALSP	0,2284	1,0000

Fuente: Elaboración propia a partir de los datos de la EPF utilizando el programa STATA

Consideramos que un coeficiente cercano a -1, 1 es correlación elevada. Podemos observar que entre las variables seleccionadas no existe una correlación elevada ya que ninguno de los coeficientes se aproxima a los valores antes mencionados. Debido a la ausencia de correlación, evitamos que los resultados de la estimación puedan estar distorsionados.

Podemos destacar que entre la variable ingresos y estudios existe una correlación positiva, es decir, a medida que aumentan los ingresos se incrementa la inversión en estudios. También podemos mencionar que entre la edad y la ocupación existe una correlación negativa, es decir, a

medida que aumenta la edad disminuye la probabilidad de estar ocupado y esto es explicado por diferentes motivos, por ejemplo, la jubilación.

A continuación, se muestra la tabla correspondiente al modelo Heckman estimado y descrito anteriormente para capturar los dos procesos de demandar o no educación superior y condicionada a demandar educación superior se muestra el gasto de los hogares destinado a continuar con la formación académica.

Tabla 6. Coeficientes y errores estándar del modelo de selección muestral.

Gasto en Máster	Coef.	Std. Err.
Ingresos	299,7901***	64,18037
Estudios Sustentador principal reducidos		
1. Inferior primera etapa Educación Secundaria	-	-
2. Primera etapa Educación Secundaria	137,9577	131,3292
3. Segunda etapa Educación Secundaria	207,8667	142,0483
4. Educación Superior	749,4163***	132,3191
Constante	-1141,449**	575,7362
Decisión de participación.	Coef.	Std. Err.
Sexo	0,0702001**	0,0310729
Número de miembros de 18 años o más	0,3877346***	0,014603
Número de hijos dependientes	0,109619***	0,0148809
Estudios Sustentador Principal reducidos		
1. Inferior primera etapa Educación Secundaria	-	-
2. Primera etapa Educación Secundaria	0,04096	0,041244
3. Segunda etapa Educación Secundaria	0,1205066***	0,0460124

4. Educación Superior	0,1835992***	0,0443645
Regiones		
1. Noroeste	-	-
2. Noreste	0,1890629***	0,0457706
3. Comunidad de Madrid	0,0914144	0,0586016
4. Central	0,2213526***	0,0475491
5. Este	0,0223006	0,0473226
6. Sur	0,4243007***	0,0458455
7. Canarias	0,4399718***	0,0630628
Ingresos	0,076701***	0,0200057
Edad Sustentador Principal	-0,0015966	0,001448
Ocupación	0,1505666***	0,0377139
Estado Civil Sustentador Principal		
1. Soltero	-	-
2. Casado	0,0238734	0,0408135
3. Viudo	0,2259069***	0,0630325
4. Separado	0,2485473***	0,0818462
5. Divorciado	0,2551343***	0,0585845
Constante	-3,234162***	0,1596633
Mills		
Lambda	-543,3705***	117,66

	Rho	-0,27991
	Sigma	1941,2504
<hr/>		
Heckman Selection model –two-step estimates		Number of obs = 22.028
(regression model with sample Selection)		Censored obs = 19.845
		Uncensored obs = 2.183
		Wald chi2 (5) = 104,99
		Prob > chi2 = 0,0000
<hr/>		
*Significativo al 90%		
**Significativo al 95%		
***Significativo al 99%		

Fuente: Elaboración propia a partir de los datos de la EPF utilizando el programa STATA

Observamos que lambda es significativo, por tanto, podemos decir que en nuestro modelo existen dos etapas. Por un lado, la decisión de acceder o no a másteres o programas de postgrado (ecuación de selección) y, por otro lado, condicionado a que decida continuar con su formación académica, el gasto que se destina a la misma (ecuación de interés). También cabe destacar que las variables son conjuntamente significativas al 99%.

De la ecuación de participación podemos comentar los siguientes resultados:

Los ingresos, como se esperaba, influyen positivamente en la probabilidad de demandar educación superior, es decir, a mayores ingresos mayor probabilidad de acceder a programas de másteres y postgrado. Por tanto, existe una gran influencia de la capacidad económica del sustentador principal para demandar educación superior.

Un segundo factor socioeconómico que incide en el acceso a másteres es el nivel de estudios del sustentador principal. En este sentido, la variable estudios la hemos dividido en cuatro categorías de manera que, las dos primeras correspondientes a un nivel de educación inferior presentan menos probabilidad de continuar con la formación académica, es decir, no son significativas en

el modelo. Mientras que, las dos últimas etapas correspondientes a niveles de educación más elevados presentan una elevada probabilidad de acceder a másteres y programas de postgrado.

En cuanto a la localización podemos decir que las zonas de Noreste, Central, Sur y Canarias presentan mayor probabilidad de demandar educación superior debido a que resultan significativas en la muestra. Algunos trabajos resaltan la significatividad de la localización en la demanda de educación debido a la proximidad o no con una universidad, sin embargo, en nuestro estudio sólo podemos determinar las regiones en las que existe mayor probabilidad de realizar estudios de másteres y postgrado, ya que en nuestra muestra no existe información relativa a las universidades.

La variable edad guarda una relación negativa con la decisión de realizar estudios de postgrado, de modo que cuanto mayor sea la edad del sustentador principal menor será la probabilidad de ampliar estudios mediante la realización de cursos de postgrado. Como he mencionado anteriormente, la variable edad y la variable ocupación presentan una relación negativa, es decir, si a mayor edad del sustentador principal se reduce la probabilidad de estar ocupado, el hogar percibe unos ingresos menores de manera que se reduce la probabilidad de demandar educación superior.

En cuanto al estado civil del sustentador principal, no es una variable muy significativa en la decisión de demandar educación superior. Sin embargo, como mencionamos anteriormente si el sustentador principal está casado la probabilidad de demandar másteres disminuye.

Las variables número de miembros de 18 o más años y número de hijos dependientes son estadísticamente significativas al 99%. Por una parte, si aumenta el número de hijos dependientes aumentará la probabilidad de que se demande educación superior y de esta manera se realice gasto en la misma. Por otra parte, si aumenta el número de miembros 18 o más años en el hogar también aumentará la probabilidad de demandar estudios de másteres o postgrado ya que lo normal en España es que los estudios de postgrado se empiecen a cursar a partir de 18 años y no en niveles de edad inferiores.

En cuanto a la situación laboral comprobamos, como parecería lógico a priori que, existe mayor probabilidad de continuar con la formación académica si el sustentador principal se encuentra ocupado, ya que sus expectativas de obtener ingresos son mayores.

Por último, podemos observar que la variable sexo es estadísticamente significativa en el modelo. Hay mayor preocupación en las mujeres por demandar educación superior que por parte de los hombres. Como he mencionado anteriormente, varios trabajos señalan que las mujeres demandan más educación superior debido a que necesitan mayor formación para competir con los hombres debido a las condiciones de desigualdad del mercado laboral. Las mujeres experimentan mayores tasas de desempleo y mayores costes de oportunidad de continuar con su formación y no incorporarse al mercado laboral.

De la ecuación de interés podemos extraer las siguientes conclusiones:

Respecto a la variable estudios del sustentador principal podemos afirmar que, a medida que el sustentador principal presenta un nivel de formación más elevado aumenta el gasto en másteres. Por tanto, es un indicador de que el nivel educativo del sustentador principal influye positivamente en la decisión de realizar gasto en másteres.

Los ingresos es otro factor económico que influye positivamente en el gasto realizado en educación, es decir, si aumenta la cantidad de ingresos aumentará el gasto en másteres y programas de postgrado.

Por tanto, los ingresos y los estudios del sustentador principal son dos variables fuertemente relacionadas con el gasto en másteres.

5 Conclusiones

A través del análisis realizado se pone de manifiesto qué factores socioeconómicos se asocian con la demanda de estudios de postgrado en España y con la decisión de participar en el gasto por parte de los hogares españoles. De hecho, podemos concluir que la decisión de demandar estudios de másteres o programas de postgrado es un proceso formado por dos etapas afectadas por diferentes factores.

En una primera etapa se decide si participar destinando parte de los recursos del hogar a este tipo de bienes (servicios) y en la segunda la decisión gira en torno a la cuantía de los recursos que destinan los hogares a másteres o programas de postgrado condicionado a que decida demandar educación superior.

Además, con la muestra analizada hemos confirmado cómo las variables señaladas por la teoría del capital humano de Becker (1975) y Mincer (1974) son válidas para explicar la decisión de cursar estudios de postgrado, es decir, la educación es considerada como una inversión, debido a que cuando aumentan los ingresos también aumenta el gasto en educación. Considerando la educación como una inversión que se espera recuperar a lo largo del tiempo.

Sin embargo, también podríamos considerar como válidos algunos aspectos destacados en la teoría de consumo Kodde y Ritzen (1994), por ejemplo, el nivel de renta ya que hemos podido observar que cuando los ingresos se incrementan aumenta la probabilidad de acceder a másteres o cursos de postgrado.

Según la teoría institucionalista desarrollado por Doeringer y Piore (1971) y Thurrow (1975) podríamos vincular el acceso a la educación superior a la productividad y el puesto de trabajo, y en este caso hemos comprobado cómo el género del sustentador principal afecta a la decisión de participación, en el sentido de que las mujeres (cuando juegan el rol de sustentador principal en el hogar) parecen presentar una mayor preocupación y, por tanto, probabilidad de destinar parte de los recursos del hogar al consumo de cursos de postgrado, posiblemente, debido a una mayor conciencia de la competitividad en el mercado laboral como señalaban los trabajos de Mora (1997) y Albert (2000).

En cuanto a la teoría de la socialización Bowles y Gintis (1975) que afirma que la educación es un nexo de transmisión de la clase social entre generaciones, podríamos destacar la significatividad dentro del modelo del nivel de estudios del sustentador principal y de cómo el grupo de individuos con mayor nivel educativo es el que más gasto realiza en másteres.

De este mismo modo, hemos podido comprobar que también influyen sobre el gasto en educación superior otros factores socioeconómicos relacionados con el hogar, por ejemplo, el número de miembros de 18 años o más o el número de hijos dependientes, factores relacionados con la situación laboral como es el caso de que si el sustentador principal se encuentra ocupado existe una mayor probabilidad de demandar educación superior, factores demográficos relacionados con la localización, factores relacionados con las características personales del sustentador principal como, por ejemplo, la edad que influye de manera negativa y el estado civil.

Referencias

Velasco, S. A. L. A. S., & Puebla, M. M. C. (2006). La demanda de educación superior: un análisis microeconómico con datos de corte transversal. *Revista de educación*, 339, 637-660.

Gómez, M. L. N., & Gutiérrez, O. D. M. (2001). Un análisis microeconómico de la demanda de educación superior en España. *Estudios de economía aplicada*, (19), 69-86.

López, S. F., Gómez, I. N., & Sanmartín, E. R. (2001). La demanda de másteres y cursos de especialización: Análisis económico a través de un modelo Logit Binomial. In *X Jornadas de la Asociación de la Economía de la Educación* (pp. 169-185). Servicio de Publicaciones.

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The internet and higher education*, 7(2), 95-105.

Verdú, C. A., Espitia, C. G. G., & Rodríguez, J. J. M. (2016). La demanda de educación superior: breve revisión de la literatura. *Ensayos de Economía*, 26(48), 209-228.

Izquierdo, M. G., & de Pablos Escobar, L. (2009). Los determinantes socioeconómicos de la demanda de educación superior en España y la movilidad educativa intergeneracional. In *XVI Encuentro de Economía Pública: 5 y 6 de febrero de 2009: Palacio de Congresos de Granada* (p. 76).

Figuroa, C. S., Vázquez, P. C., Martín, I. T., & del Rey, S. (2012). James Heckman, el sesgo de selección muestral.

Material electrónico:

- Para qué sirve un máster. Disponible en:
https://elpais.com/economia/2018/09/06/actualidad/1536233264_893344.html
- Haga un máster para impulsar su carrera. Disponible en:
<http://www.expansion.com/emprendedores-empleo/desarrollo-carrera/2017/09/15/59bc14e6ca474146678b458d.html>
- La cantidad de títulos que hay en el mercado sobrepasa la demanda. Disponible en:
<https://www.eleconomista.es/ecoaula/noticias/9104320/04/18/la-crisis-disparo-un-900-la-demanda-de-los-masteres.html>
- Datos y cifras del Sistema Universitario Español, curso 2015/2016 Ministerio de Educación, Cultura y Deporte. Disponible en:
<https://www.mecd.gob.es/dms/mecd/servicios-al-ciudadano-mecd/estadisticas/educacion/universitaria/datos-cifras/datos-y-cifras-SUE-2015-16-web.pdf>

Apéndice

Para llevar a cabo este análisis hemos utilizado los datos de la Encuesta de Presupuestos Familiares elaborada por el INE, que publica los microdatos en ficheros planos, por tanto, es necesario elaborar diccionarios para poder trabajar con ellos en STATA (versión 14).

Imagen 1. Fichero de usuario de hogar 2017 formato plano

```

20170000110563212012577902040330000000201010201020100000002000002010201000301010220180101 7 7 6 9 9 6 4 3 5 4 4 2 20238 6 1 1
2017000021211122906525713210220000000200000200020000000001000100020002000200000170150101101010 5 5 3 7 4 2 6 6 8 80185 6 1 1
201700003133111120142051933902200000002000002000200000000010100020002000200000170150100101010 5 5 4 7 4 2 6 6 8 80265 6 1 1
2017000040746531100571238405011000000010000010001000000000100001000010000010000000 2 2 2 1 1 1 1 2 1 1 1-9-90147 1 1 1
20170000512164222005268378390220000002000002000200000000200002000101000200000170150101 6 6 5 6 5 3 2 2 3 3 5 20155 1 1 1
201700006162642220059331538980330000000201010201020100000101000002010201000301010220180101 7 7 6 9 9 6 4 3 5 4 4 2 20138 1 1 1
2017000070956532201571909178011000000010000010001000000001000001000001000001000000 2 2 2 1 1 1 1 2 1 1 1-9-90147 1 2 2 1
2017000080616421100446154148044000000400000400000200020000030102010301010310250202111101212 8 5 4 6 5 5 2 10150 1 1 1
20170000910564311009069107120220000002000002000200000001010000020002000200000170150101 5 5 4 5 5 3 1 2 3 3 8 80166 1 1 1
201700010121121120037603783904400000003010202020002000002000002020202000402020290230202 8 8 71010 7 5 3 5 4 4 2 20150 1 1 1
20170001105762111006637315170440000000400010301030001010002000003010202000401010310250303111101212 7 5 4 6 5 5 2 10145 6 1 1
2017000120746532200454367109033000000030000030003000000001010101020102000300000240200202101010 7 7 3 7 4 3 6 6-9-90185 6 1 1
201700013074121220041739537002200000020000020002000000002000001010101000200000170150101 6 6 5 6 5 3 2 2 3 3 5 50158 1 1 1
201700014016621220091776072005500000004010104010400010100010200010400050104020203602003021212101212 9 8 4 6 5 5 8 60174 1 1 1
201700015152653110016142229304400000002020202020200000101000002020202000402020270210101 8 8 71010 6 5 3 5 4 4 2 20132 1 1 1
20170001609565322009376025350440000000202020202020101000002000002020103000402020270210101 8 8 71010 7 5 3 5 4 4 5 20139 1 1 1
2017000170226532200503069976011000000010000010001000000000100000100010001000000 1 1 1 2 2 1 2 1 1 1 1 8 80167 1 1 1
2017000180166531201020162359022000000020000020002000000002000002000200000170150101 6 6 5 6 6 5 3 2 2 3 3 8 20162 1 1 1
2017000190741111100652486970022000000020000020002000000002000002000200000170150101 6 6 5 6 5 3 2 2 3 3 2 20160 1 1 1
20170002014664312005568368010330000000201010201020001000002000001020102000301010220180101 7 7 6 9 9 7 4 3 5 4 4 5 50239 6 1 1
201700021146633220053420916502200000002000002000200000000020000020002000200000170150101 5 5 4 5 5 3 1 2 3 3 8 80171 1 1 1
201700022074111220143838534001100000001000001000100000000010000010001000001000000 2 2 2 1 1 1 1 2 1 1 1-9-90149 1 1 1
20170002307412222004367575560440000000301020202020002000002020103000402020290230200 8 8 71010 7 5 3 5 4 4 5 20245 6 1 1
20170002409565322015995490810330000000300000300030000010200000003000300000300000240200201121210 7 710 8 4 3 6 6 1 10132 1 1 1
2017000250456422200910034831033000000030000030003000000101010001020003000300000240200200111110 7 7 9 4 4 3 6 6 8 60165 1 1 1
2017000261521112200134495302022000000020000020002000000001000101010002000200000170150101121210 5 5 4 8 4 2 6 6 8-90185 6 1 1

```

Esta imagen representa un fichero plano correspondiente al “fichero de usuario de hogar 2017”.

A partir de este archivo se elaboraron los diccionarios para poder extraer toda la información contenida en estos ficheros de la EPF.

En el diccionario hay que indicar en qué posición empieza cada variable, su longitud, su nombre, su formato en STATA (float, byte, int) y una breve descripción de esta. Para poder realizar este proceso hemos contado con un informe elaborado por el INE que recoge toda la información necesaria para elaborar los diccionarios en STATA.

Imagen 2. Diccionario en Stata

The screenshot shows a Stata Do-file Editor window titled "Do-file Editor - diccionario hogar tfm". The window contains a dictionary file named "diccionario hogar tfm" with the following content:

```

1 dictionary using "C:\Users\Maria\Desktop\Trabajo Fin de Máster\datos_2017\datos_epf2017\F
2 _column(1) float AÑOENC %4.0f "año encuesta"
3 _column(5) float NUMERO %5.0f "orden del hogar"
4 _column(10) byte CCAA %2.0f "comunidad autonoma"
5 _column(12) byte NUTS1 %1.0f "region"
6 _column(13) byte CAPROV %1.0f "capital provincia 1=si 6=no"
7 _column(14) byte TAMAMU %1.0f "tamaño municipio"
8 _column(15) byte DENSI %1.0f "densidad poblacion"
9 _column(16) byte CLAVE %1.0f "clave de colaboracion efectiva del hogar"
10 _column(17) byte CLATEO %1.0f "clave de colaboracion teorica del hogar"
11 _column(18) float FACTOR %11.6f "factor poblacional"
12 _column(29) byte NMIEMB %2.0f "numero de miembros del hogar"
13 _column(31) byte TAMAÑO %1.0f "tamaño del hogar"
14 _column(32) byte NMIEMBSD %2.0f "numero de miembros que pertenecen al servicio domestico"
15 _column(34) byte NMIEMBHU %2.0f "numero de miembros del hogar que son huéspedes"
16 _column(36) byte NMIEMBIN %2.0f "numero de miembros del hogar que son invitados"
17 _column(38) byte NMIEMB1 %2.0f "numero de miembros del hogar de 14 o mas años"
18 _column(40) byte NMIEMB2 %2.0f "numero de miembros del hogar menores de 14 años"
19 _column(42) byte NMIEMB3 %2.0f "numero de miembros del hogar menores de 16 años"
20 _column(44) byte NMIEMB4 %2.0f "numero de miembros del hogar de 16 años o mas"
21 _column(46) byte NMIEMB5 %2.0f "numero de miembros del hogar menores de 18 años"
22 _column(48) byte NMIEMB6 %2.0f "numero de miembros del hogar de 18 o mas años"
23 _column(50) byte NMIEMB7 %2.0f "numero de miembros del hogar de 0 a 4 años"
24 _column(52) byte NMIEMB8 %2.0f "numero de miembros del hogar de 5 a 15 años"

```

The status bar at the bottom indicates "Línea: 1, Col: 1 CAP NUM OVR".

Podemos observar el diccionario elaborado en STATA correspondiente al “fichero de usuario de hogar 2017” que se mostraba en la imagen anterior.

Para realizar este trabajo hemos utilizado el fichero de hogar y el fichero de gastos. En primer lugar, se elaboraron dos diccionarios (uno para cada fichero) y poder así extraer los datos. En segundo lugar, fusionamos ambas bases de datos, utilizando como nexo de unión el número de hogar, con el fin de obtener la información relevante para nuestro estudio.

Seguidamente, se llevó a cabo un proceso de depuración de datos, es decir, solo mantuvimos los gastos que eran relevantes en nuestro análisis, asignando “missing values” a aquellos hogares que no realizaban gasto en másteres. Para realizar la selección muestral se ha utilizado la variable “CÓDIGO” ya que cada partida de gasto tiene asignado un código de gasto dentro de la EPF, el valor de este código varía dependiendo del tipo de gasto que realice el hogar.

Finalmente, se ha procedido a la creación y modificación de determinadas variables para poder quedarnos con una muestra formada, únicamente, por las variables relevantes en nuestro estudio.

Imagen 3. Base de datos

Editor de Datos (Navegación) - [DEFINITIVA - copia]

Archivo Edición Ver Datos Herramientas

NUTS1[1] 5

	NUTS1	SEMIEMB	SHIJOED	EDAD6V	ECIVILLEGA-P	ESTUDOS6V	SEXO	INGRESOS	GASTO_REC	OCUPACION
1	5	2	1	38	2	4	1	7.716506	-	1
2	1	2	0	88	3	4	1	6.488205	-	0
3	3	2	3	65	1	4	1	9.131531	-	9
4	4	1	0	47	1	2	0	7.055064	-	1
5	1	2	0	55	2	4	0	7.318321	-	1
6	2	2	1	38	2	4	0	7.888913	-	1
7	5	1	0	47	1	3	0	6.907755	-	1
8	1	4	1	50	2	4	0	7.645638	-	1
9	5	2	0	66	2	4	0	7.124891	-	0
10	1	2	2	50	2	3	0	7.491645	-	1
11	7	3	1	45	5	1	1	7.410952	-	1
12	4	3	0	86	3	1	1	7.574045	-	0
13	4	2	0	59	2	3	0	7.46049	-	1
14	6	4	2	74	2	1	0	7.088893	-	0
15	2	2	2	32	2	4	0	7.638483	-	1
16	5	2	2	39	2	2	0	7.149917	-	1
17	2	1	0	67	5	3	0	5.906342	-	0
18	6	2	0	62	2	2	0	6.747527	-	0
19	4	2	0	60	2	3	0	7.524796	-	1
20	6	2	1	29	2	2	1	7.549609	-	1
21	6	2	0	71	2	1	0	8.248732	-	0
22	4	1	0	49	1	4	0	7.495542	48.1708	1
23	4	2	2	45	2	4	1	7.789868	.0173054	1
24	5	3	0	32	1	3	0	7.938902	-	1
25	3	2	3	65	2	1	0	7.782224	-	8
26	2	2	0	85	3	1	1	6.954413	-	0

Valor Vers: 10 Orden: Dataset Obs: 22.028

Esta imagen representa la muestra final que hemos utilizado para estimar el modelo econométrico.