



Universidad de Valladolid

FACULTAD DE CIENCIAS SOCIALES, JURÍDICAS Y DE LA
COMUNICACIÓN

Grado en Administración y Dirección de Empresas

TRABAJO DE FIN DE GRADO

Análisis econométrico del gasto de los hogares en educación

Presentado por Noemí Arévalo García

Tutelado por Helena Corrales Herrero

Segovia, 24 de Junio 2022

ÍNDICE

INTRODUCCIÓN.....	7
--------------------------	----------

CAPÍTULO 1

REVISIÓN DE LA LITERATURA

1.1 Revisión de literatura.....	11
---------------------------------	----

CAPÍTULO 2

ANÁLISIS EMPÍRICO

2.1 Encuesta de Gasto de los Hogares en Educación 2019/2020.....	17
2.2 Selección de la muestra y descripción de las variables.....	19
2.3 Metodología.....	21
2.4 Estimación del modelo.....	23
2.5 Interpretación del signo de los coeficientes.....	27

CAPÍTULO 3

CONCLUSIONES

3.1 Conclusiones	31
------------------------	----

REFERENCIAS BIBLIOGRÁFICAS	33
---	-----------

ANEXO 1

Listado de variables.....	35
----------------------------------	-----------

ÍNDICE DE TABLAS

Tabla 2.1 Gastos en Servicios Educativos.....	18
Tabla 2.2 Gastos en Bienes Educativos.....	18
Tabla 2.3 Becas y ayudas.....	19
Tabla 2.4 Análisis descriptivo de las variables.....	20
Tabla 2.5 Estimación por MCO del modelo que explica el gasto de las familias en educación a partir de todos los regresores.....	23
Tabla 2.6 Factor de inflación de la varianza.....	25
Tabla 2.7 Test Reset de Ramsey.....	25
Tabla 2.8 Test de Jarque Bera.....	26

Introducción

El gasto en educación representa para las familias uno de los desembolsos que más crece dentro del hogar, tan solo por detrás de la vivienda, los suministros y la salud. Este crecimiento se explica en parte por la caída de la financiación pública de la Educación tras la crisis económica de 2008. En cierto modo son las familias las que se ven obligadas a compensar el gasto educativo público, cuyos niveles son comparativamente reducidos. Uno de cada tres niños y niñas está en riesgo de pobreza y exclusión social ya que las familias no pueden dedicar ese dinero en gastos educativos, común en familias de renta baja. Esto tiene efectos sobre cuestiones como el aumento de la segregación o el abandono escolar. Las partidas que componen el grueso del gasto son las actividades complementarias, los servicios como el comedor, libros o academias y otros tipos de formación (El diario.es, 2021).

Según datos ofrecidos por el Sistema Estatal de Indicadores de la Educación (SEIE), las familias españolas gastaron la cifra de 11.417 millones de euros en 2019 (El diario, 2021). En comparación con nuestro entorno, según datos de la Oficina Estadística de la Unión Europea (EUROSTAT) de ese mismo año, España se situaba entre los países de Europa que menos gasto público ha destinado a la educación un 4,0% del PIB, mientras que otros países, entre ellos los nórdicos, muestran volúmenes de gasto público superiores, con Suecia a la cabeza con un 6,9%, Dinamarca con un 6,3%, o Bélgica con un 6,2% (Servimedia, 2021).

El presente trabajo tiene como objetivo principal estimar el gasto privado de los hogares españoles en educación mediante la realización de un modelo econométrico. De tal manera que como fin último podamos interpretar el signo de los coeficientes e identificar y analizar aquellas variables que provocan el aumento de dicho gasto. El método que usaré para la estimación será el de mínimos cuadrados ordinarios (MCO) y la base de datos que he utilizado para dicha estimación es la Encuesta de Gasto de los Hogares en Educación para el año 2019/2020, más adelante (EGHE 2019/2020), disponibles en la página web del Instituto Nacional de Estadística (INE).

El trabajo se divide en tres capítulos. En el primer capítulo encontramos el contexto histórico, en el segundo capítulo, todo lo relacionado con el análisis empírico del trabajo, la base de datos a usar, metodología y la estimación del modelo. En el tercer y último capítulo se exponen los efectos reflejados en el modelo y las conclusiones finales de este análisis.

CAPÍTULO 1

REVISIÓN DE LA LITERATURA

1.1 Revisión de la literatura

Es durante el siglo XIX cuando se definen los principios básicos de la educación en España, con la implantación de la Ley de Instrucción Pública en 1857, conocida también como "Ley Moyano", primera ley educativa en España que pretendía acabar con las altas tasas de analfabetismo y mejorar la situación de la educación. En ella se consagró la coexistencia de la enseñanza pública y de la enseñanza privada, siendo la mayoría de carácter religioso, así mismo se estableció una educación organizada en tres niveles: la enseñanza primaria, siendo esta de carácter obligatorio y gratuito, en contraposición de la secundaria y la universidad, de carácter voluntario y oneroso.

Desde entonces, los hogares comenzaron a participar en la inversión en actividades educativas para poder alcanzar el éxito a través del acceso a la educación. Por consiguiente, han sido varios autores los que han tratado de describir y analizar el gasto privado educativo de los hogares en España, mediante la estimación de modelos econométricos. A continuación, vamos a destacar algunos de estos trabajos, los cuales han resultado importantes a la hora de la elección de variables relevantes para llevar a cabo nuestro propio estudio econométrico.

Principalmente hemos utilizado dos trabajos para el desarrollo del presente trabajo. El primero de ellos, es el realizado por Beneito, Ferri, Moltó y Uriel (1995), el objetivo de este trabajo es estimar una ecuación de demanda de educación de los hogares, tanto de educación secundaria como universitaria, utilizando una estimación del coste de oportunidad, con datos procedentes de la Encuesta de Presupuestos Familiares 1990-1991 del INE. Para la realización del modelo econométrico tienen en cuenta una gran variedad de factores, entre ellos los relacionados con el estudiante como pueden ser la edad, el sexo o el nivel de estudios, también factores financieros como el nivel de ingreso familiar, o incluso factores de tipo socioeconómico de la familia como la localización geográfica del lugar de residencia, el tamaño del municipio, la composición familiar, la ocupación del sustentador principal o el estatus social al que pertenece la familia. No obstante, para nuestro trabajo destacamos, cuatro de estos factores, ya que son los que tenemos en común con nuestra base de datos. El primero es la edad, los autores exponen en su trabajo que la edad es significativa y positiva, a mayor edad mayor probabilidad de mayores ingresos, así como de permanecer empleado. El factor, sexo, en sus resultados se obtiene que hay una clara discriminación hacia las mujeres de edades comprendidas entre los 16 y los 26 años, donde estas obtendrían ingresos inferiores y tienen menos probabilidad de encontrar un trabajo que un varón. En cuanto al nivel de estudios, también resulta significativo ya que concluyen que, a mayor nivel de estudios alcanzados, más probabilidades de empleo. Por último, el tamaño del municipio, exponen que es una variable muy enriquecedora ya que supone que un joven encontrara más dificultades para cursar sus estudios si vive en un pueblo pequeño que si reside en una gran ciudad, ya que en los pueblos pequeños puede haber una menor disponibilidad de centros educativos, lo que repercute en el desplazamiento a municipios mayores. Los resultados de este trabajo señalan que el estatus socioeconómico de la familia tiene un impacto mayor en el gasto de los hogares en la educación secundaria que en la educación universitaria de sus miembros.

CAPÍTULO 1

El segundo trabajo considerado es el de Cueto (2017), donde se lleva a cabo un análisis y modelización del gasto de los hogares españoles a través del método de mínimos cuadrados ordinarios, utilizando también la base de datos de la Encuesta de Presupuestos Familiares del año 2012, que recopila información anual sobre la naturaleza y destino de los gastos de consumo. El fin de este trabajo es analizar los efectos de distintas variables sobre el gasto anual total de las familias. Este trabajo, aunque no trate en particular el gasto privado que las familias destinan a la educación, no lo descartamos como un estudio primordial de referencia. Para la realización del modelo econométrico tiene en cuenta una gran variedad de factores como los relacionados con el hogar como es el número de miembros del hogar, el número de estudiantes en el hogar o el número de miembros ocupados del hogar, en cuanto a factores relacionados con el sustentador principal, tiene en cuenta los factores como los estudios completados, o el tipo de contrato o por último factores más generales como el tamaño del municipio. Por lo que al igual que con el trabajo de Beneito et al. (1995) optamos por aquellas variables que tenemos en común con nuestra base de datos, como pueden ser el número de personas en el hogar, el número de estudiantes del hogar, en este último considera que el gasto total de la familia aumenta un 7,31% por cada estudiante que haya en el hogar. Por último, destacamos el tamaño del municipio, ya que señala que según se va reduciendo el tamaño del municipio se reduce el gasto total de los hogares.

Los estudios siguientes son más generales y menos exhaustivos, y tienen en cuenta menos variables. Uno de ellos es el estudio llevado a cabo por García, Lechuga, y Carmona (1998), donde tratan de estimar el gasto educativo de los hogares de la Región de Murcia en el curso 1995/96, diferenciando gastos totales y por alumno, por niveles de enseñanza, según la titularidad del centro. Para ello, estiman un modelo econométrico a través del método de regresión lineal, con el objetivo de identificar únicamente aquellos componentes del gasto que contribuyen en mayor medida a explicar la variabilidad del gasto en educación de los hogares de la Región de Murcia. Las variables a utilizar han sido algunos de los componentes del gasto en educación, entre ellos los bienes como el material escolar y los libros de textos y en cuanto los relativos a los servicios educativos sitúan los de inscripción, el comedor, otras actividades educativas, la asociación de padres, los regulares, los del seguro escolar, los del transporte escolar y por último los gastos varios. Algunas de las principales conclusiones del estudio son que más de la mitad de los gastos educativos no universitarios de los hogares murcianos se destina a los niveles de Preescolar y EGB. Y que las principales partidas que causan este aumento son los gastos regulares y de inscripción, para los niveles de Preescolar o COU, selectividad y pruebas de acceso. Además, también concluye que un hogar con hijos matriculados en un centro privado incurre en un gasto 2,5 veces más elevado que si estuvieran en un centro público, siendo destacable la diferencia de gasto en preescolar (6 veces más) y en COU, selectividad y pruebas de acceso (7,8 veces más).

El otro trabajo es el de Rogero-García y Andrés-Candelas (2014) que se centran en estimar, por un lado, el gasto público a través del coste de la plaza escolar en la educación no universitaria según el tipo de centro, teniendo en cuenta las becas y ayudas. Y, por otro lado, el gasto privado de las familias por estudiante, para así poder conocer el gasto total por estudiante según tipo de centro. Para ello estiman el gasto público por alumno en los niveles de enseñanza no universitaria, lo realizan dividiendo el gasto público destinado a centros públicos o concertados entre el número de alumnos matriculados en cada tipo de centro. En segundo lugar, recogen los datos sobre el gasto privado de las familias, para que, en conjunto con el gasto público, obtengan un gasto global de la inversión por estudiante. El trabajo concluye que el gasto público en centros públicos es mayor que en

REVISIÓN DE LA LITERATURA

centros concertados o lo que es lo mismo la escolarización en centros concertados supone una reducción del gasto público, ya que hay un trasvase de parte de los costes educativos a las familias.

CAPÍTULO 2

ANÁLISIS EMPÍRICO

2.1 Encuesta de Gasto de los Hogares en Educación 2019/2020

Los datos utilizados en este trabajo son los correspondientes a la Encuesta de Gasto de los Hogares en Educación para el conjunto del territorio español durante el curso académico 2019/2020 (EGHE 2019/2020). Esta encuesta, es una investigación estadística de carácter plurianual llevada a cabo por el INE, que atiende a la demanda estadística de información sobre el gasto de los hogares en educación. Se llevó a cabo por primera en el año 2007 como un módulo integrado dentro de la Encuesta de Presupuestos Familiares (EPF). En el curso académico 2011/2012, esta investigación se realizó por segunda vez ya como encuesta independiente. La tercera edición y más reciente es la del curso académico 2019/2020, la elegida para la elaboración del presente trabajo.

El periodo de recogida de los datos comprende de enero hasta mediados de mayo de 2020, coincidiendo parcialmente con la suspensión de las actividades no esenciales, motivada por el estado de alarma como consecuencia de la pandemia del COVID-19. Debido a esta situación excepcional el número de unidades previstas para la investigación se ha visto reducido de forma significativa.

En cuanto a la muestra, está formada por los hogares que han finalizado su periodo de colaboración en la EPF, y que declararon gastos de enseñanza de alguno de sus miembros. No obstante, para garantizar un nivel adecuado de cobertura de la población investigada, también se han tenido en cuenta a los hogares que han demostrado tener algún miembro del hogar menor de 30 años y que no tenga gastos en el estudio.

El tamaño de la muestra comprende un total de 6370 hogares (cada hogar, considerado como unidad de gasto). Sin embargo, los hogares que han declarado que tienen algún miembro de la familia estudiando son un total de 2812 hogares. Y de estos hogares sacamos un total de 4903 individuos que si realizan algún tipo de estudio ya sea de tipo reglado o no reglado.

El contenido de esta encuesta se encuentra resumida en las tablas 2.1 Gastos en bienes educativos, 2.2 Gastos en servicios educativos y 2.3 Becas y ayudas a estudio. Así mismo aparece el desglose de los gastos, tanto para la educación reglada como la no reglada.

Tabla 2.1 Gastos en bienes educativos

- | |
|---|
| <ul style="list-style-type: none">- Libros de texto, manuales, textos universitarios, libros de apuntes y libros complementarios de apoyo al estudio (diccionarios, lecturas recomendadas).- Fotocopias e impresión de documentos.- Uniformes, calzado o ropa deportiva escolar, mochilas.- Artículos de papelería.- Mobiliario y complementos.- Herramientas y productos informáticos.
- Instrumentos musicales, herramientas o materiales de dibujo, instrumentos y material especializado (laboratorio, óptico, etc.). |
|---|

Fuente: Elaboración propia a partir de los datos de la EGHE 2019/2020 (INE)

Tabla 2.2 Gastos en servicios educativos

- | |
|---|
| <ul style="list-style-type: none">- Matrícula y clases lectivas: se considera la actividad docente referida a las disciplinas contenidas en los programas oficiales de enseñanza.- Actividades extraescolares: organizadas por el centro educativo, realizadas de manera voluntaria por los estudiantes fuera del horario lectivo (ballet, deportivas, música, idiomas no obligatorios, etc.).- Clases de apoyo: organizadas en el propio centro educativo (matemáticas, inglés, informática).- Servicios complementarios: transporte escolar, comedor escolar y universitario, residencia escolar y universitaria, servicio de ampliación de horario.- Actividades complementarias: excursiones, visitas culturales, actividades deportivas, etc. realizadas a través del centro y dentro del horario escolar.- Cuota AMPA: correspondiente a la Asociación de Madres y Padres de los Alumnos.- Otras actividades realizadas o servicios proporcionados por el centro educativo: gabinete psicológico, orientador, servicio médico, horas complementarias voluntarias, otras donaciones o gastos (servicios administrativos, material didáctico, etc.).- Gasto en servicios educativos correspondientes a enseñanzas no regladas (pagos realizados a academias, profesores particulares, cursos on-line, idiomas, música, informática, contabilidad, preparación de oposiciones y pruebas de acceso). |
|---|

Fuente: Elaboración propia a partir de los datos de la EGHE 2019/2020 (INE)

Tabla 2.3 Becas y ayudas al estudio

- | |
|--|
| <ul style="list-style-type: none"> - Beca básica o de carácter general: ayuda económica que se concede a determinados estudiantes en función fundamentalmente de tener una renta familiar baja. Puede ser de cuantía fija o variable. - Beca o ayuda de matrícula: reducción del coste de la matrícula o gratuidad de la misma bajo determinados supuestos (condiciones socioeconómicas, familias numerosas, expediente académico, etc.). - Beca de necesidades educativas específicas: dirigidas a cubrir determinadas necesidades específicas de apoyo educativo de los estudiantes, así como orientadas a los alumnos de educación especial. - Beca de movilidad: ayuda económica a los estudios dirigida a los estudiantes que cursan todo o parte del año académico bien en el extranjero (ej. beca Erasmus), bien en otra Comunidad Autónoma a la de residencia (ej. beca Séneca). - Beca de excelencia: aportación económica que contribuye a financiar los estudios de los estudiantes que han demostrado un aprovechamiento académico excelente. - Beca o ayuda de comedor: destinadas a financiar la cuantía total o parcial del gasto en comedor (escolar o universitario). - Beca o ayuda de transporte o desplazamiento: destinadas a financiar la cuantía total o parcial del gasto en transporte. - Beca o ayuda para libros de texto o material escolar - Otros tipos de becas. |
|--|

Fuente: Elaboración propia a partir de los datos de la EGHE 2019/2020 (INE)

2.2 Selección de la muestra y descripción de las variables

Debido a la gran cantidad de variables que tiene el fichero de la encuesta, se ha procedido a realizar una selección previa de aquellas variables más relevantes basándonos en la literatura. Para construir nuestra propia base de datos se ha hecho uso del programa estadístico SPSS 26, que además usaremos para realizar el análisis descriptivo de las variables. De cara a la construcción del modelo econométrico, así como a la realización de contrastes de hipótesis, o a la búsqueda de problemas de especificación en el modelo se ha trabajado con el programa econométrico EViews7.

El tipo de datos que disponemos para realizar la estimación son de corte transversal, es decir, tenemos datos recopilados mediante la observación de muchos sujetos al mismo tiempo, sin tener en cuenta las diferencias en el tiempo.

Las variables que se pueden encontrar en un modelo econométrico por un lado pueden ser variables dependientes, explicadas o endógenas, que son aquellas cuyo comportamiento se pretende explicar con el modelo, es decir son aquellas que se van a ver afectadas por el comportamiento de las variables independientes, en nuestro caso es una sola variable dependiente. Por otro lado, tenemos las variables independientes, explicativas, exógenas o regresores, que son aquellas cuya variación causa un cambio en la variable dependiente y van a determinar su valor. A la hora de realizar el análisis empírico, primero debemos establecer cuál es nuestra variable dependiente y qué variables independientes serán las que ayuden a analizar el comportamiento de ésta.

CAPÍTULO 2

Se parte de un total de diez variables, una variable dependiente y nueve variables exógenas. En el caso de las variables exógenas, por un lado, tenemos tres variables cuantitativas y, por otro lado, seis cualitativas. Para estas últimas se han generado variables ficticias. A continuación, en la tabla 2.4 se presentan en primer lugar los estadísticos descriptivos de las variables originales para posteriormente presentar los resultados de la totalidad de las variables ficticias que han sido generadas a partir de dicha variable original. Las variables para las cuales se han generado ficticias, son tamaño del municipio, sexo, nacionalidad, nivel de estudios, tipo de enseñanza y localización del centro educativo. Antes de estimar el modelo, primero realizaremos un análisis de cada una de las variables en la tabla 2.4.

Esta tabla se ha distribuido de la siguiente manera, en primer lugar, encontramos una columna que engloba la totalidad de las variables con las que vamos a trabajar referentes al total de estudiantes. Y a continuación se recoge los distintos datos descriptivos de cada una de las variables como el número de observaciones, la media, la desviación típica, el mínimo y el máximo. Por último, aparece una columna donde viene expresado el porcentaje para cada variable.

Tabla 2.4. Análisis descriptivo de las variables

TOTAL DE ESTUDIANTES						
Variable	Nº obs.	Media	Des. típica	Mínimo	Máximo	%
Gasto total en educación	4903	1570,3	2327,560	0	46625	
Nº de personas en el hogar	4903	3,8105	1,0248	1	11	
Nº de estudiantes en el hogar	4903	2,1393	0,95519	1	6	
Edad	4903	18,7334	13,6737	0	82	
Tamaño del municipio de residencia	4903					100
tamamu1	966	0,1970	0,39779	0	1	20
tamamu2	1372	0,2790	0,44896	0	1	30
tamamu3	670	0,1367	0,34351	0	1	14
tamamu4	1274	0,2598	0,43859	0	1	26
tamamu5	621	0,1267	0,33262	0	1	13
Sexo	4903					100
mujer	2560	0,5221	0,49956	0	1	52
hombre	2343	0,4779	0,49956	0	1	48
Nacionalidad	4903					100
española	4564	0,9309	0,23372	0	1	93
doble nación.	137	0,0279	0,16482	0	1	3
extranjera	202	0,0412	0,19877	0	1	4
Nivel de estudios	4903					100
infantil	591	0,1205	0,32562	0	1	12
primaria	1125	0,2295	0,42052	0	1	23
secundaria	1468	0,2994	0,45805	0	1	30
superior	1001	0,2042	0,40313	0	1	20
otros estudios	718	0,1464	0,35358	0	1	15

ANÁLISIS EMPÍRICO

Tipo de enseñanza	4903					100
publica	3019	0,6157	0,48647	0	1	62
privada	358	0,0730	0,26019	0	1	7
concertada	808	0,1648	0,37104	0	1	16
Localidad el municipio de residencia	4903					100
localce1	2997	0,6113	0,48751	0	1	61
localce2	817	0,1666	0,37269	0	1	17
localce3	113	0,0230	0,15007	0	1	2
localce4	236	0,0481	0,21407	0	1	5
localce5	22	0,0045	0,06684	0	1	0,45

Nota: En el anexo 1, se muestra el nombre real de cada una de las variables.

Fuente: Elaboración propia a partir de los datos de la encuesta EGHE2019/2020 (INE)

Las principales conclusiones de este análisis descriptivo es que de media los hogares gastan un total de 1570,3 euros en educación por estudiante, no obstante, la desviación típica nos ayuda a medir el grado de dispersión de los datos, es decir, es la variación de los datos respecto a la media, en este caso es muy dispar, esto es debido a que las cifras de gasto comprenden como mínimo 0 euros y 46625 euros lo máximo. Nuestra muestra está compuesta mayoritariamente por estudiantes jóvenes con una media de edad de 18,73 años, en cuanto a su desviación típica, la variación de los datos con respecto a la media también es muy dispar, es decir, tenemos edades comprendidas de 0 a 82 años. De media aproximadamente la mitad de personas que forman un hogar son estudiantes, mayoritariamente de nacionalidad española, un poco más de la mitad son mujeres. Donde más estudiantes hay son en los municipios de entre 10.000 a 49.000 habitantes (tamamu2), hasta un 30% de la muestra y donde menos en las grandes ciudades de más de 500.000 habitantes (tamamu5), tan solo un 13%. En los niveles donde se concentran más estudiantes son en la primaria y en la secundaria, un 23% y un 30% respectivamente. Predomina muy por encima los estudios en un centro público sobre uno concertado, hasta un 62% de la muestra. Por último, cabe resaltar que la mayoría opta por estudiar en el mismo municipio de residencia o en otro municipio de la misma provincia, apenas casi nadie opta por estudiar en otra provincia, otra comunidad autónoma o mucho menos en el extranjero.

2.3 Metodología

Una vez establecidas y definidas las variables principales con las que vamos a trabajar, vamos a proceder a estudiar su comportamiento y como es la relación que las une. Para ello se emplea como herramienta estadística, el análisis de regresión, que intenta encontrar una relación matemática que explique de la mejor manera posible el comportamiento de la variable dependiente a partir de otras variables (exógenas) y además permita una posterior predicción. Esta relación matemática puede lograrse a través de la construcción de modelos matemáticos.

Si introducimos una variable explicativa, el modelo se denomina de regresión simple, pero en este trabajo vamos a introducir más de una variable y pasará a denominarse modelo de regresión múltiple o general, lo que nos va a permitir analizar cómo afecta de forma conjunta un grupo de variables (variables exógenas) al comportamiento de otra

CAPÍTULO 2

variable (variable dependiente) y estudiar de forma pormenorizada la influencia de cada una de las variables explicativas en la variable objeto de estudio. Como anteriormente mencionábamos en este trabajo también vamos a trabajar con variables de tipo cualitativo, las cuales también pueden ser relevantes para explicar el comportamiento de la variable endógena, para incorporarlas en el modelo econométrico vamos a construir variables ficticias que toman dos valores arbitrarios, normalmente 1 y 0.

Sin embargo, para una variable independiente con m atributos categóricos, primero se seleccionará una categoría como referencia, de modo que se puedan generar $m-1$ variables ficticias, ya que si no caeríamos en la trampa de las ficticias. Por ejemplo, para nuestra variable independiente de nacionalidad, se generan tres variables ficticias (española, extranjera y doble nacionalidad), para introducirlas en el modelo, eliminaremos la ficticia española, siendo esta la de referencia. Para realizar un análisis de estas ficticias, se introducirán en el modelo de forma aditiva y se analizarán de tal manera que solo afecte al término constante de la ecuación, es decir a la variable dependiente.

Este modelo de regresión lineal múltiple en cuanto a su especificación se puede expresar en notación escalar:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i \text{ para } i = 1, \dots, N$$

Por un lado, Y , se trata de la variable dependiente, endógena o explicada mientras que X representa las variables explicativas, que tratarán de explicar el comportamiento de Y . β son los coeficientes, se desconocen, y establecen la relación que hay entre Y e X , se pueden estimar. ε es la perturbación o error, es una variable aleatoria desconocida que explica la parte de Y que no es explicada por los regresores del modelo, recogiendo la variabilidad de la relación. Al introducir este componente al modelo lo dotamos de aleatoriedad y de un mayor realismo ya que estamos aceptando incertidumbre en cualquier comportamiento social. N es el número de observaciones o datos.

Para saber si un modelo tiene buenas propiedades, se requiere la incorporación de una serie de hipótesis sobre las variables de dicho modelo, esta incorporación de hipótesis da lugar al modelo de regresión lineal clásico (MRLC). Son varias las hipótesis. En primer lugar, el modelo ha de estar correctamente especificado, esto quiere decir que la relación entre las variables sea lineal y que los coeficientes sean constantes. En segundo lugar, el valor esperado del término de perturbación tiene media cero. En tercer lugar, las varianzas de las perturbaciones han de ser constantes (homoscedasticidad) y que además tengan covarianza nula (incorrelación), en cuarto lugar, tiene que haber normalidad en la distribución de ε , en quinto lugar, no debe haber relación lineal entre los regresores ya que si no estaríamos ante un caso de multicolinealidad y por último los regresores deben de ser no aleatorios o no estocásticos, deben ser fijos en muestras repetidas.

El método de estimación que vamos a usar en nuestro modelo econométrico va a ser el de mínimos cuadrados ordinarios (MCO). Su objetivo es obtener estimadores de los parámetros β y σ^2 y consiste en buscar aquella recta que minimice $\sum e^2$. Si el modelo cumple todas las hipótesis clásicas, la estimación mediante MCO proporciona estimadores con buenas propiedades, o lo que es lo mismo tendríamos un estimador ELIO, un estimador lineal, insesgado, óptimo, eficiente y de distribución finita, por lo que el modelo lo daremos por bueno.

2.4 Estimación del modelo

En este apartado se pretende estimar un modelo definitivo en el que primeramente introduciremos todas las variables (Tabla 2.5), para posteriormente someterlo a una serie de test (las hipótesis clásicas) para ver si tiene algún tipo de problema.

La variable dependiente (logaritmo del gasto total en educación), es el logaritmo de la variable gasto total en educación, tomándose como logaritmo para suavizar los efectos extremos de la muestra.

Tabla 2.5 Estimación por MCO del modelo

Dependent Variable: LOGARITMO DEL GASTO EN EDUCACION

Method: Least Squares

Sample: 1 4903

Included observations: 4903

Variable	Coefficient	Std. Error	t-Statistic	Prob.
EDAD	-0.031177	0.002024	-15.40349	0.0000
EHOGAR	0.095858	0.023545	4.071239	0.0000
NHOGAR	-0.118650	0.022639	-5.240968	0.0000
HOMBRE	-0.052477	0.033456	-1.568543	0.1168
TAMAMU1	-0.147293	0.049503	-2.975437	0.0029
TAMAMU3	0.002338	0.055259	0.042313	0.9663
TAMAMU4	0.100189	0.046400	2.159260	0.0309
TAMAMU5	0.307448	0.058443	5.260653	0.0000
DOBLENA	-0.029856	0.101993	-0.292726	0.7697
EXTRA	-0.390680	0.085179	-4.586594	0.0000
INFANTIL	-0.521379	0.063308	-8.235606	0.0000
PRIMARIA	-0.146077	0.049221	-2.967773	0.0030
SUPERIOR	0.645670	0.058659	11.00716	0.0000
OTROSESTU	0.518394	0.068679	7.548083	0.0000
PRIVADA	1.496962	0.067517	22.17147	0.0000
CONCERTADA	0.793500	0.048590	16.33058	0.0000
LOCALCE2	0.213559	0.052830	4.042391	0.0001
LOCALCE3	0.314035	0.120428	2.607669	0.0091
LOCALCE4	0.652859	0.091086	7.167513	0.0000
LOCALCE5	0.487394	0.252628	1.929294	0.0538
C	7.068003	0.092354	76.53167	0.0000
R-squared	0.246527	Mean dependent var	6.659048	
Adjusted R-squared	0.243440	S.D. dependent var	1.337853	
S.E. of regression	1.163670	Akaike info criterion	3.145309	
Sum squared resid	6610.852	Schwarz criterion	3.173138	
Log likelihood	-7689.724	Hannan-Quinn criter.	3.155072	
F-statistic	79.86652	Durbin-Watson stat	1.641529	
Prob(F-statistic)	0.000000			

Fuente: Elaboración propia

En primer lugar, hemos estudiado la multicolinealidad, una cuestión importante que debe analizarse al estudiar los resultados de un modelo de regresión como éste, esto es comprobar si existe una alta correlación entre más de dos variables explicativas. Esto supone que las variables explican parte de otra por lo que un cambio en una afecta al resto que estén correlacionadas y por tanto los estimadores obtenidos pierden eficiencia y tendremos problemas de multicolinealidad. Para ello un método es estudiar la

CAPÍTULO 2

significación conjunta del modelo, a través del valor del estadístico $F=79,86$ y su p-valor, que es menor que 0,05, podemos afirmar que el modelo es significativo conjuntamente y, por lo tanto, los regresores incluidos dentro del modelo inicial son significativos conjuntamente a la hora de explicar nuestra variable dependiente. Por otro lado, en cuanto a la significación individual de las variables explicativas, nos hemos fijado en sus p-valores. Observando dichos valores observamos que algunas variables tienen un p-valor mayor a 0,05, pareciendo primeramente no significativas, esto se puede deber a que haya errores de especificación consistentes en inclusión de variables irrelevantes o a la existencia de una fuerte multicolinealidad entre las variables explicativas.

Sin embargo, lo que ocurre en este caso, es que la significación individual no se da en algunas de las variables ficticias, las cuales no podemos eliminar ya que queremos asegurar la exactitud del significado de todas las variables ficticias, son variables que nos ayudan a medir el efecto de pertenecer a una categoría determinada frente a la categoría de referencia, si las eliminamos, el resultado se alteraría y no tendría un significado exacto. Además, si eliminásemos alguna de estas variables ficticias no significativas, el R^2 disminuiría. El coeficiente de determinación o R^2 mide qué parte de la variabilidad de la variable dependiente viene explicada por las variables independientes. Tiende a crecer cuando la variable incluida es relevante para explicar el comportamiento de la variable dependiente; por el contrario, se reduce en el caso de que la variable sea no significativa. Cuanto más próximo a uno sea el coeficiente, tendremos una mejor estimación del modelo. Esto significaría que toda la variabilidad de la variable dependiente es explicada por los regresores y, por lo tanto, existe un margen muy pequeño de error. Para nuestro caso tenemos un R^2 de 0,2465 lo que quiere decir que los regresores explican el 24,65% de la variabilidad de la variable dependiente.

Existen otros métodos para detectar multicolinealidad, como es el análisis de los coeficientes de correlación a través de la herramienta matriz de correlaciones, en ella se puede observar qué variables están más o menos relacionadas con la variable dependiente, y qué variables independientes están relacionadas entre sí. Puede tomar valores dentro del intervalo (-1,1), un coeficiente de correlación muy cercano a uno (tanto en positivo como en negativo), indica que la relación entre las variables es muy fuerte, esto supone una alta probabilidad de acierto en la inclusión de la variable en el modelo. Por el contrario, un coeficiente próximo a cero, indica que las variables apenas tienen relación entre sí, y por lo tanto esa variable es posible que no sea importante ni tenga relevancia en el modelo.

Otro de los métodos para detectar multicolinealidad sería comprobar si el Factor de Inflación de la Varianza (FIV) toma valores mayores que 10. Aun sabiendo que no podemos eliminar ninguna variable utilizaremos este último método para demostrar que no existe este problema. A continuación, en la tabla 2.6 podemos ver la información que arroja este contraste.

Tabla 2.6: Factor de inflación de la varianza

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
EDAD	4.10E-06	7.978231	2.772753
EHOGAR	0.000554	11.01736	1.830777
NHOGAR	0.000513	28.89432	1.948837
HOMBRE	0.001119	1.936667	1.011190
TAMAMU1	0.002451	1.748153	1.403728
TAMAMU3	0.003054	1.510871	1.304409
TAMAMU4	0.002153	2.025552	1.499231
TAMAMU5	0.003416	1.566381	1.367987
DOBLENA	0.010403	1.052459	1.023051
EXTRA	0.007255	1.082312	1.037722
INFANTIL	0.004008	1.749215	1.538367
PRIMARIA	0.002423	2.012775	1.550941
SUPERIOR	0.003441	2.543580	2.024281
OTROSESTU	0.004717	2.500986	2.134739
PRIVADA	0.004559	1.205191	1.117192
CONCERTADA	0.002361	1.408779	1.176617
LOCALCE2	0.002791	1.683926	1.403329
LOCALCE3	0.014503	1.210238	1.182345
LOCALCE4	0.008297	1.445951	1.376352
LOCALCE5	0.063821	1.036877	1.032224
C	0.008529	30.88256	NA

Fuente: Elaboración propia

Los factores de inflación de la varianza no están lo suficientemente inflados como para considerar que haya dicho problema, por lo que queda totalmente descartado de que existan problemas de multicolinealidad.

Por lo que se refiere a la especificación correcta del modelo, analizaremos la linealidad, en la que la forma funcional del modelo debe ser lineal. Para analizar si el lineal o no, se especificará un modelo alternativo al inicial, añadiendo una nueva variable para ver si sigue una forma lineal. En este caso utilizaremos el contraste Reset de Ramsey.

Hipótesis:

H_0 : Existe linealidad.

H_1 : No existe linealidad.

Tabla 2.7. Test Reset de Ramsey

	Value	df	Probability
t-statistic	0.528932	4881	0.5969
F-statistic	0.279769	(1, 4881)	0.5969
Likelihood ratio	0.281022	1	0.5960

Fuente: Elaboración propia

El modelo se considera válido, el valor del estadístico F= 0,27 con un p-valor = 0,596 > 0,05, por lo que aceptamos la hipótesis nula. Es decir, nuestro modelo sí tiene una forma funcional lineal. Al ser lineal el modelo, los estimadores β tienen buenas propiedades. El modelo es lineal, insesgado, óptimo.

CAPÍTULO 2

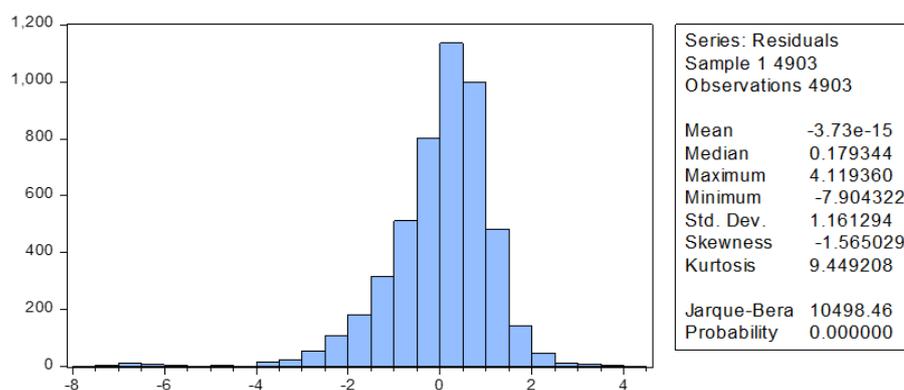
A la hora de analizar la normalidad, en el cual las perturbaciones del modelo deben seguir una distribución normal. Para ello, hemos sometido nuestro modelo final al test de Jarque-Bera.

Hipótesis:

H_0 : si ε sigue una distribución $N(0, \sigma^2)$.

H_1 : si ε no sigue una distribución $N(0, \sigma^2)$.

Tabla 2.8. Test de Jarque-Bera



Fuente: Elaboración propia

En dicha estimación se obtiene un p-valor de $0,00 < 0,05$ por lo que rechazaríamos la hipótesis nula de este contraste, por lo tanto, aceptamos que las perturbaciones de nuestro modelo no siguen una distribución normal, el modelo tiene problemas de normalidad de sus perturbaciones. El hecho de que el modelo no cumpla con esta hipótesis hace que el estimador de MCO no siga distribuciones conocidas. Además, los estadísticos t-student y F-Snedecor para realizar contrastes, no son válidos. Este problema no lo podemos solucionar ya que nuestra variable dependiente ya está en logaritmos.

También analizaremos la hipótesis de la heteroscedasticidad, para que se dé, las varianzas de las perturbaciones no deben ser constantes. El método que utilizaremos para detectarla es a través del contraste asintótico de White, el método más general y más utilizado. Analiza la heteroscedasticidad de forma general, sin saber qué variable es la causante. Este test parte de un modelo auxiliar en la que trata de explicar los residuos al cuadrado en función de sus regresores, sus cuadrados y los productos cruzados, y determina si hay problemas de heteroscedasticidad o no. Debido al número de observaciones que tenemos en el modelo no es posible realizar dicho test, por lo que no podemos comprobar si existe heteroscedasticidad. Por lo que los β siguen siendo lineales, insesgados y consistentes, pero no son de mínima varianza. Además, los estadísticos t-student y F-Snedecor no son válidos. En el caso de que existiera heteroscedasticidad, un método para corregirlo sería a través de la construcción de la matriz de White cuya finalidad es calcular correctamente la desviación típica de los estimadores MCO en presencia de heteroscedasticidad y, por consiguiente, poder seguir utilizando los contrastes de hipótesis de la t y la F.

2.5 Interpretación de los signos de los coeficientes

A continuación, analizaremos los signos de los coeficientes de las variables explicativas sobre la variable dependiente y sacaremos conclusiones sobre el modelo.

El valor de la constante β_0 es 7,06 e indica que, en ausencia de variables explicativas, el gasto total en educación de las familias tomaría ese valor.

En cuanto a las variables cuantitativas, por lo que se refiere a la edad, a medida que se incrementa la edad de un miembro que estudie, el gasto disminuye en 3,11%. Por cada miembro del hogar que estudie, el gasto aumenta un 9,58%. En cambio, por cada miembro en el hogar, el gasto total disminuye en un 11,86%.

Ahora en cuanto a las variables cualitativas, cuando se trata del sexo, si es hombre el gasto disminuye un 5,24% respecto a la referencia si es mujer. Por lo que se refiere al signo del tamaño del municipio en el que se reside, se observa que, si se reside en un municipio con más de 500.000 habitantes (tamamu5) el gasto total aumenta en un 30,74% con respecto a la referencia si un municipio cuenta de ente 10.000 y 49.000 habitantes (tamamu2). Si el municipio tiene entre 100.000 y 499.999 habitantes (tamamu4), el gasto aumenta en un 10%, con respecto a la referencia. Si el municipio tiene entre 50.000 y 99.000 habitantes (tamamu3), el gasto aumenta un 0,23 %. En cambio, si el municipio tiene menos de 10.000 habitantes (tamamu1), el gasto disminuye en un 14,72% con respecto a la referencia. Dependiendo de la nacionalidad, si se trata de un estudiante extranjero el gasto disminuye un 39% y si es de doble nacionalidad, tan solo disminuye el gasto un 2,9% con respecto si se trata de un estudiante español para ambos casos. Con lo que respecta al nivel de estudios en curso, por cada estudiante que realice estudios superiores el gasto aumenta un 64,56%, por debajo se sitúa un estudiante que realice otros estudios cuyo gasto se incrementa en un 51,83%, ambos con respecto a la referencia que es la educación secundaria. Tan solo los niveles situados por debajo de la educación secundaria, es decir, un estudiante de infantil, el gasto disminuiría un 52,1%, mientras que para un estudiante de primaria disminuiría un 14,60%, también con respecto a la referencia que es secundaria. En cuanto al tipo de enseñanza, por cada estudiante que estudie una concertada el gasto aumenta en un 79,35% y en la privada hasta un 149% con respecto a un estudiante que estudie en una pública. Por último, en cuanto a la localidad del centro educativo, por cada estudiante que estudie en un centro educativo fuera de su comunidad autónoma (localce4), el gasto se incrementa hasta en un 65,28%, un 48,73% si estudia en el extranjero (localce5), un 31% si estudia en otra provincia de la misma comunidad (localce3) y un 21% si estudia en otro municipio de la misma provincia (localce2), todos ellos con respecto a la referencia en el caso de que estudiase en su mismo municipio de residencia (localce1).

CAPÍTULO 3

CONCLUSIONES

3.1 Conclusiones

Los resultados del estudio nos han ayudado a identificar aquellos factores que hacen que el gasto de las familias aumente o disminuya, ya que año tras año este gasto por parte de las familias cada vez es mayor, poniendo en evidencia las consecuencias que puede conllevar si este gasto sigue aumentando, como puede ser un aumento de la tasa de abandono escolar o un aumento de la segregación escolar.

Los signos de los coeficientes son especialmente relevantes a la hora de identificar aquellos factores que aumentan dicho gasto, dichos coeficientes coinciden con los resultados esperados ya que además coincide con los resultados de otros estudios anteriores y que hemos querido analizar en el presente trabajo. Como es por ejemplo en el trabajo de Beneito et al. (1995) que indica que, a mayor edad, mayor probabilidad de ingresos y de permanecer en un trabajo, al igual que en nuestro trabajo a mayor edad, menor gasto les supone a las familias ya que alcanzando una edad determinada los estudiantes pueden trabajar y por lo tanto colaborar con los gastos educativos. Otro dato que también llama la atención es el sexo del estudiante, donde si eres varón el gasto disminuye con respecto a si eres mujer, esto puede deberse a como explica Beneito et al. (1995) en su trabajo, las mujeres tienen más probabilidad de recibir ingresos inferiores y de menor probabilidad de encontrar un trabajo, y es por eso que las mujeres deban de gastar más en sus estudios, para encontrar más oportunidades, esto también se plasma en nuestro trabajo ya que el signo del coeficiente del sexo hombre es negativo con respecto al de la mujer. Con lo que respecta al nivel de estudios, según Beneito et al. (1995), un mayor nivel de estudios genera un aumento cada vez mayor de los gastos, esto también se verifica en nuestro trabajo, es decir, por ejemplo, en los estudios superiores. Este nivel educativo es la partida que presenta más gasto, los altos costes de la matrícula y de las clases lectivas de la universidad son el principal causante. No muy atrás se queda la parte de los gastos de dedicada a otros estudios, es decir con lo que respecta a los gastos de academias, profesores particulares, cursos on-line, idiomas, preparación de oposiciones o por ejemplo las pruebas de acceso, todo ello son elementos muy costosos que evidencian un mayor desembolso. Otro de los elementos a destacar es el tamaño del municipio, que al igual que como se expone en el trabajo realizado por Beneito et al. (1995), cuanto más pequeño sea el tamaño del municipio, menos disponibilidad de disponer de un centro educativo, lo que se traduce en desplazamientos a municipios mayores, por lo que ese traslado provoca un incremento de los gastos, en nuestro trabajo esto también se verifica, además de que hay una relación positiva ya que a mayor tamaño de municipio más gastos tendrá un estudiante. Esto es algo que también coincide con el trabajo de Cueto (2017). Otro dato a destacar en comparación con el trabajo de Cueto (2017) es que cuanto más estudiantes haya en un hogar el gasto aumenta, efecto que también se da en nuestro trabajo. Por otro lado, es evidente que a pesar de que la mayor parte de estudiantes estudien en una pública, la educación privada no deja de ser la más cara, hasta un 149% con respecto a la pública, los altos precios de las matrículas, las clases lectivas o los servicios de alojamiento son los principales causantes.

Por otro lado, en cuanto a la construcción del modelo, hemos conseguido un modelo que consiga explicar el 24,65 % de la variabilidad del gasto de las familias en educación a través de las variables independientes. Un porcentaje no muy alto que aparentemente puede indicar que no es del todo un buen modelo, esto puede deberse a que la gran mayoría de las variables que conforma el modelo son variables ficticias y solo pueden

CAPÍTULO 3

tomar valores de 1 y 0. No obstante independientemente del aparente bajo R^2 , hemos comprobado que la gran mayoría de variables son significativas, así como el modelo en su conjunto, por lo que a pesar de todo consideramos que se han sacado conclusiones importantes, además de obtener unos resultados bastante coherentes en comparación con otros estudios.

Referencias bibliográficas

Páginas web

Instituto Nacional de Estadística (INE)

<https://www.ine.es/> “Encuesta sobre el gasto de los hogares en educación”.

Noticias de prensa

- El Diario (agosto, 2021): *Las familias invierten más cada año en Educación y el gasto privado ya supone uno de cada cinco euros del total.*

https://www.eldiario.es/sociedad/familias-invierten-ano-educacion-gasto-privado-supone-cinco-euros-total_1_8173212.html#:~:text=La%20educaci%C3%B3n%2C%20tambi%C3%A9n%20la%20gratuita,Estatal%20de%20Indicadores%20de%20Educaci%C3%B3n.

- Servimedia (marzo, 2021): *España es el quinto país de la UE que menos dinero público destina a educación.*

<https://www.servimedia.es/noticias/1673489>

Literatura

- Beneito, P., Ferri, J., Moltó, M. L., y Uriel, E. (1995). Determinantes de la demanda de educación en España. *Documentos de trabajo: Serie EC-(Instituto Valenciano de Investigaciones Económicas)*, pp 1-47.

<http://www.ivie.es/downloads/docs/wpasec/wpasec-1995-12.pdf>

- Cueto Vicens, F. J. (2017). Análisis y modelización del gasto de los hogares a partir de microdatos de la ECPF.

https://dspace.uib.es/xmlui/bitstream/handle/11201/3427/TFG_GECO_FJCuetoVicens.pdf?sequence=1

- García, J. G., Lechuga, M. L., & Carmona, A. L. (1998). Propuesta metodológica para la evaluación del gasto educativo de los hogares en educación reglada no universitaria. Aplicación a la Comunidad Autónoma de la Región de Murcia. *Revista de Investigación Educativa*, 16(1), 137-153.

<https://revistas.um.es/rie/article/view/122461/115081>

- Rogero-García, J., y Andrés-Candelas, M. (2014). Gasto público y de las familias en educación en España: diferencias entre centros públicos y concertados. *Revista Española de Investigaciones Sociológicas*.

https://reis.cis.es/REIS/PDF/REIS_147_071404906873547.pdf

Material teórico

Corrales Herrero, H.: Material Docente de Econometría I.

Corrales Herrero, H.: Material Docente de Econometría II.

ANEXO 1. Listado de variables

ETIQUETA	DESCRIPCIÓN	AGRUPACIÓN
GTE	Gasto total en educación.	Variable cuantitativa
EDAD	Edad del estudiante.	Variable cuantitativa
EHOGAR	Número de estudiantes en el hogar.	Variable cuantitativa
NHOGAR	Número de personas del hogar.	Variable cuantitativa
HOMBRE	Estudiante del sexo masculino.	Variable ficticia que toma el valor 1 si el individuo es hombre y 0 en el caso contrario.
MUJER	Estudiante del sexo femenino.	Variable ficticia que toma el valor 1 si el individuo es mujer y 0 en el caso contrario.
TAMAMU1	Tamaño del municipio de residencia (menos de 10.000 habitantes).	Variable ficticia que toma el valor 1 si el individuo pertenece a un municipio de menos de 10.000 habitantes y 0 en el caso contrario.
TAMAMU2	Tamaño del municipio de residencia (de 10.000 a 49.000 habitantes).	Variable ficticia que toma el valor 1 si el individuo pertenece a un municipio de 10.000 a 49.000 habitantes y 0 en el caso contrario.
TAMAMU3	Tamaño del municipio de residencia (de 50.000 a 99.000 habitantes).	Variable ficticia que toma el valor 1 si el individuo pertenece a un municipio de 50.000 a 99.000 habitantes y 0 en el caso contrario.
TAMAMU4	Tamaño del municipio de residencia (de 100.000 a 499.000 habitantes).	Variable ficticia que toma el valor 1 si el individuo pertenece a un municipio de 100.000 a 499.000 habitantes y 0 en el caso contrario.
TAMAMU5	Tamaño del municipio de residencia (más de 500.000 habitantes).	Variable ficticia que toma el valor 1 si el individuo pertenece a un municipio de más de 500.000 habitantes y 0 en el caso contrario.
ESPAÑOLA	Nacionalidad española a la que pertenece el estudiante.	Variable ficticia que toma el valor 1 si el individuo es de nacionalidad española y 0 en el caso contrario.
DOBLE NACIONALIDAD	Nacionalidad de doble nacionalidad a la que pertenece el estudiante.	Variable ficticia que toma el valor 1 si el individuo es de nacionalidad doble nacionalidad y 0 en el caso contrario.

EXTRANJERA	Nacionalidad extranjera a la que pertenece el estudiante.	Variable ficticia que toma el valor 1 si el individuo es de nacionalidad extranjera y 0 en el caso contrario
INFANTIL	Nivel de estudios en curso para el caso de estudiantes de infantil.	Variable ficticia que toma el valor 1 si el individuo estudia infantil y 0 en el caso contrario.
PRIMARIA	Nivel de estudios en curso para el caso de estudiantes de primaria.	Variable ficticia que toma el valor 1 si el individuo estudia primaria y 0 en el caso contrario.
SECUNDARIA	Nivel de estudios en curso para el caso de estudiantes de secundaria.	Variable ficticia que toma el valor 1 si el individuo estudia secundaria y 0 en el caso contrario.
SUPERIOR	Nivel de estudios en curso para el caso de estudiantes de superior.	Variable ficticia que toma el valor 1 si el individuo estudia estudios superiores y 0 en el caso contrario.
OTROSESTUDIOS	Nivel de estudios en curso para el caso de estudiantes de otros estudios	Variable ficticia que toma el valor 1 si el individuo estudia otros estudios y 0 en el caso contrario.
PUBLICA	Tipo de enseñanza pública.	Variable ficticia que toma el valor 1 si el individuo estudia en un centro público y 0 en el caso contrario.
PRIVADA	Tipo de enseñanza privada.	Variable ficticia que toma el valor 1 si el individuo estudia en un centro privado y 0 en el caso contrario.
CONCERTADA	Tipo de enseñanza concertada.	Variable ficticia que toma el valor 1 si el individuo estudia en un centro concertado y 0 en el caso contrario.
LOCALCE1	Localidad del centro educativo si se encuentra en el mismo municipio de residencia.	Variable ficticia que toma el valor 1 si el individuo pertenece al mismo municipio de residencia y 0 en el caso contrario.
LOCALCE2	Localidad del centro educativo si se encuentra en otro municipio de la misma provincia.	Variable ficticia que toma el valor 1 si el individuo estudia en otro municipio de la misma provincia y 0 en el caso contrario.
LOCALCE3	Localidad del centro educativo si se encuentra en otra provincia de la misma comunidad autónoma	Variable ficticia que toma el valor 1 si el individuo estudia en otra provincia de la misma comunidad autónoma y 0 en el caso contrario.

LOCALCE4	Localidad del centro educativo si se encuentra en otra comunidad autónoma.	Variable ficticia que toma el valor 1 si el individuo estudia en otra provincia de la misma comunidad autónoma y 0 en el caso contrario.
LOCALCE5	Localidad del centro educativo si se encuentra en el extranjero.	Variable ficticia que toma el valor 1 si el individuo estudia en el extranjero y 0 en el caso contrario.
