



Departamento de Economía

**Facultad de Ciencias
Administrativas y Económicas**



Icesi Economics Working Papers

**Corrección del sesgo de selección muestral en la
probabilidad de demandar educación universitaria en
Colombia**

Carlos Giovanni González Espitia
Jhon James Mora Rodríguez
Cecilia Albert Verdú

Icesi EWP No. 5
2015

Corrección del sesgo de selección muestral en la probabilidad de demandar educación universitaria en Colombia

**Carlos Giovanni González Espitia
Jhon James Mora Rodríguez
Cecilia Albert Verdú**

Icesi EWP No. 5
2015

Universidad Icesi

Editor:

Carlos Giovanni Gonzalez Espitia
Profesor tiempo completo, Universidad Icesi
cggonzalez@icesi.edu.co

Asistente editorial:

Andres Felipe Hoyos Martin
andres.hoyos1@correo.icesi.edu.co

Gestión Editorial

Departamento de Economía - Universidad Icesi

© Universidad Icesi. Todos los derechos reservados. El contenido de esta publicación es responsabilidad absoluta de los autores y no necesariamente reflejan los puntos de vista y opiniones de la Universidad Icesi.

www.icesi.edu.co

Tel: 5552334. Fax: 5551441

Calle 18 # 122-135 Cali, Valle del Cauca, Colombia

Corrección del sesgo de selección muestral en la probabilidad de demandar educación universitaria en Colombia

Carlos Giovanni González Espitia, Ph.D.*

Jhon James Mora Rodríguez, Ph.D.*

Cecilia Albert Verdú, Ph.D.†

Resumen

El objetivo de este documento es analizar la demanda de educación universitaria y sus determinantes en Colombia. Para ello, se utilizan modelos de elección discreta robustos que tienen en cuenta el sesgo de selección muestral, debido a que la heterogeneidad no observada de los estudiantes que terminan el bachillerato, que está correlacionada con la decisión de demandar educación universitaria. Los resultados muestran la corrección del sesgo de selección y que la probabilidad de demandar educación universitaria en Colombia está condicionada por características propias de algunos grupos sociales.

Palabras clave: Demanda de educación, modelos de elección discreta, sesgo de selección.

Clasificación JEL: I20, J12, C35

Abstract

The objective of this document is to analyze the demand for higher education and its determinants in Colombia. For this purpose, robust discrete choice models are used to correct the selection sample bias; this is the unobserved heterogeneity of the students that graduate from high school, which is also correlated with the decision of demanding higher education. The results show the correction of the selection sample bias and that the probability of demanding higher education in Colombia is conditioned by the characteristics of certain social groups.

Key words: Demand for education, discrete choice models, selection sample bias.

* Departamento de Economía, Universidad Icesi (Colombia), Email: cggonzalez@icesi.edu.co y jjmora@icesi.edu.co, respectivamente.

† Departamento de Economía, Universidad de Alcalá (España), E-mail: cecilia.albert@uah.es.

1. Introducción

En Colombia la forma tradicional de analizar la demanda de educación empíricamente ha sido a través del método de Mínimos Cuadrados Ordinarios (MCO), con la idea de calcular la tasa de rendimiento de los años y de los títulos de educación. Esto ha sido posible gracias a la existencia de datos adecuados, como bien resalta Londoño (1995) al citar a Psacharopoulos (1981), quien anota que Colombia dispone de las series más largas de corte transversal aplicadas a la ecuación minceriana para países en desarrollo.

Los primeros trabajos sobre el tema se pueden clasificar en cuatro generaciones. La primera surge entre los años sesenta y setenta, periodo en el que se pueden ubicar los trabajos de Schultz (1961) y Selowsky (1968); el de aquellos aplicados a la Encuesta Nacional de Hogares (ENH) como Bourguignon (1979, 1983) y Mohan (1978). Una segunda generación que aparece en los años ochenta y principios de los noventa la conforman Ayala (1987), Carrizosa (1987), Sarmiento *et al.* (1987) y Muñoz (1990), los cuales fueron impulsados por las dos misiones de la década de los ochenta, la Misión de Empleo de 1986 y la Misión del Ministerio de agricultura y Departamento Nacional de Planeación de 1990. La tercera generación es impulsada por los trabajos de Tenjo (1993a y 1993b), quien utilizando la ENH realiza varios estudios en los que aplica la ecuación de ingresos para Bogotá y Colombia; otros trabajos de esta generación son los de Londoño (1995) y Núñez y Sánchez (1998). Finalmente, en una cuarta generación se pueden ubicar los primeros trabajos del siglo XXI, Castellar y Uribe (2001), Tenjo (2001), Tenjo *et al.* (2002), Chávez y Arias (2002), Girón *et al.* (2003) y Hernández (2010).

A la par de estos últimos trabajos, un grupo de autores se enfocó en la estimación de la rentabilidad de los títulos y no de los años por educación. De estos trabajos se destacan algunos como los de Mora (2003), Mora *et al.* (2004), Mora (2006), Mora y Muro (2008), Castellar y Uribe (2008) y Cano, Muñoz y Mesa (2009).

Con respecto a la demanda de educación universitaria, en la literatura económica se pueden encontrar varias referencias de estudios internacionales que consideran la influencia de diversos factores sobre la probabilidad de demandar educación universitaria y que discuten la corrección del sesgo de selección. Estos trabajos son Murnane *et al.* (1981), Rice (1987), Kodde y Ritzen (1988), Manski y Wise (1983), Grubb (1988), Betts y McFarland (1995), Modrego (1986), Mora (1990), Albert (1998 y 2000), Martínez (1999), Marcerano y Navarro (2001), Salas (2003), García (1998), Valiente (2003), Marcerano, Galindo y Vignoles (2007) y Albert y Garcia (2010).

El objetivo de este artículo consiste en corregir el sesgo de selección muestral que se presenta en las elecciones de educación universitaria para el caso de Colombia. La corrección se realiza a través de un modelo *probit* estimado por el método de máxima verosimilitud que tiene en cuenta la corrección del sesgo de selección muestral para

datos de corte transversal en la línea de los trabajos de Heckman (1979), Van de Ven y Van Praag (1981) y Miranda y Rabe-Hesketh (2006).

Para realizar las estimaciones en este estudio se seleccionó como población a los jóvenes entre 17 y 31 años de edad. Este rango de edad permitirá tener en cuenta la definición de demanda por años de educación universitaria para los años comprendidos entre 1990 y 2005.

Este documento se estructura de la siguiente forma: además de esta introducción, en la sección 2 se presenta la discusión del sesgo de selección en las elecciones por educación universitaria en Colombia. En la sección 3 se describe la metodología econométrica y la estrategia de estimación. También, se detallan los datos y la selección de la muestra. En la sección 4 se muestran los principales resultados obtenidos. Finalmente, se presentan las conclusiones y las referencias bibliográficas.

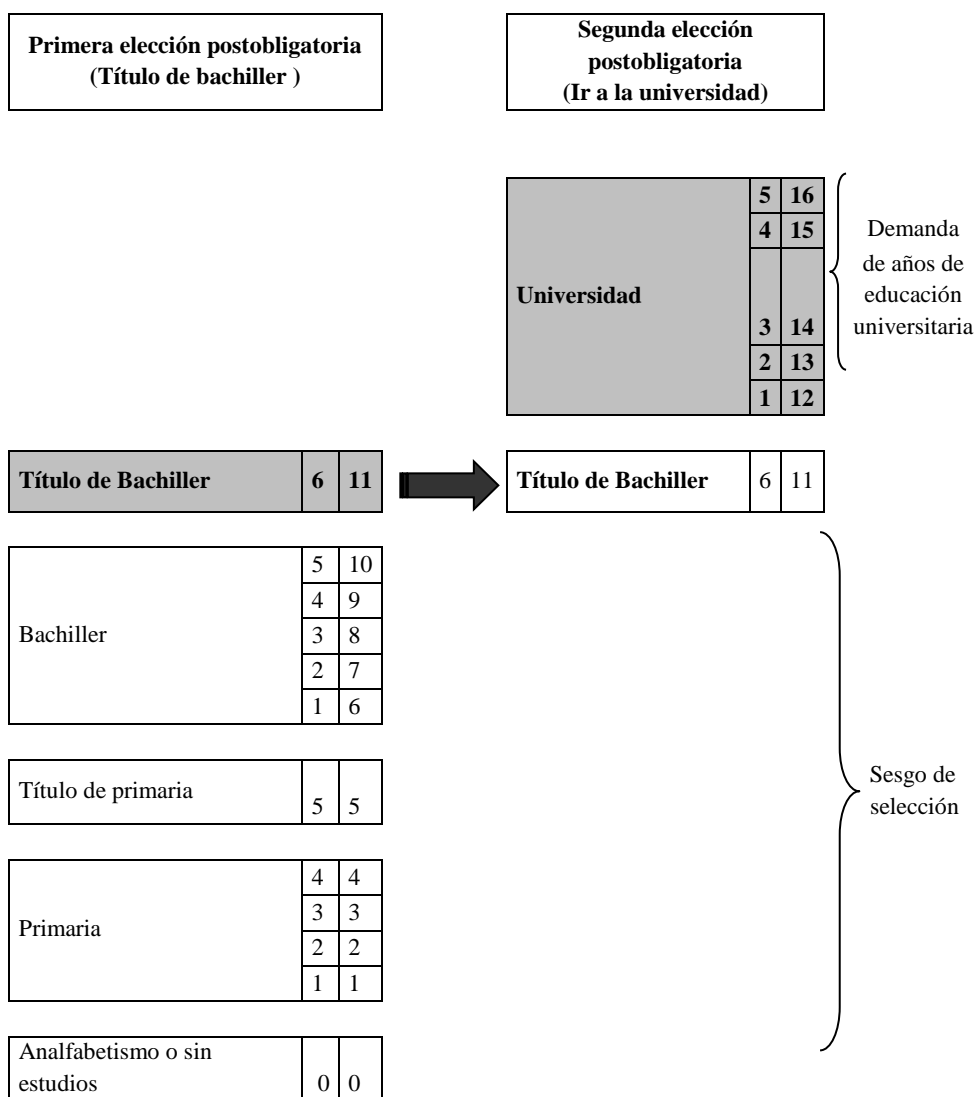
2. El sesgo de selección muestral en la demanda de educación universitaria

El concepto de sesgo de selección fue introducido por James Heckman en un artículo publicado en 1979. Este problema se puede contextualizar en algo más general, conocido como el problema de la endogeneidad. Para Greene (2011) una regresión puede presentar un problema de endogeneidad debido a cuatro causas fundamentales: i. Variables relevantes omitidas que estén correlacionadas con las variables explicativas incluidas en el modelo, ii. Simultaneidad, iii. Error de medición en las variables explicativas, y iv. El sesgo de selección. Este último problema aparece cuando los datos no son aleatorios. Según Heckman (1979), el problema del sesgo de selección se puede dar por dos motivos diferentes. El primero de ellos es la existencia de un proceso de autoselección por parte de los individuos de la muestra. Mientras que el segundo es consecuencia de la decisión del propio investigador de seleccionar únicamente una parte concreta de toda la muestra. En cualquiera de estos dos casos, a la hora de tratar de estimar una ecuación de comportamiento, habría que tener en cuenta este aspecto para evitar incurrir en estimadores sesgados e inconsistentes.

En el caso de la demanda de educación universitaria para Colombia, al haber optado por analizar los jóvenes con edades entre 17 y 31 años, el problema del sesgo de selección, en caso de existir, podría ser debido a un problema de truncamiento incidental, ya que al estimar la ecuación sobre la probabilidad de ir a la universidad, únicamente disponemos de observaciones de la elección para aquellos individuos que habían obtenido el título de bachillerato, mientras que no se observan los que no han terminado este nivel o de los que podrían haber demandado educación secundaria pero que no lo hicieron por diferentes razones (habilidad, financiamiento, entre otras).

En la figura 1 se pueden observar las elecciones educativas postobligatorias que pueden realizar los jóvenes colombianos divididas en dos etapas. En el lado izquierdo aparece la primera elección, que consiste en obtener el título de bachiller, lo cual implica como mínimo once años de educación formal y haber superado los seis años de educación secundaria. Por otra parte, en el lado derecho se observa la elección de la educación universitaria, que tiene como requisito previo obtener el título de bachiller. En este cuadro se observa cómo para tomar la decisión de asistir o no a la universidad es necesario haber alcanzado la primera etapa de educación postobligatoria y estar en posesión de un título previo.

Figura 1 Del título de bachiller a la demanda de educación universitaria



Fuente: Elaboración propia.

Dado que no es aleatorio para un joven el obtener el título de bachiller y que existe un truncamiento incidental de la muestra, entonces se evidencia teóricamente un sesgo de selección muestral, al no observar de forma aleatoria a toda la población. También, si

observamos los datos de la Encuesta de hogares para el año 2005, podremos intuir si efectivamente estamos ante este problema econométrico. En la tabla 1 se presenta la frecuencia y las estadísticas descriptivas de la variable años de educación así como la forma en que se convierte en una variable discreta (o dummy) que toma el valor de 1 cuando el individuo tiene más de 11 años de educación y 0 cuando alcanza el título de bachiller. En la misma tabla se observa que la frecuencia de mayor peso porcentual es la de 11 años de educación con un 35,01 por ciento, que son los jóvenes que han obtenido el título de bachiller. Para el nivel universitario corresponden los años de educación del 12 al 16. También se puede resaltar que la media de años de educación alcanzados por los jóvenes en este rango de edad es 10,3 años, con una desviación estándar de 3,3.

Tabla 1 Frecuencia de la variable continua años de educación y generación de la variable discreta demanda de años de educación universitaria, año 2005

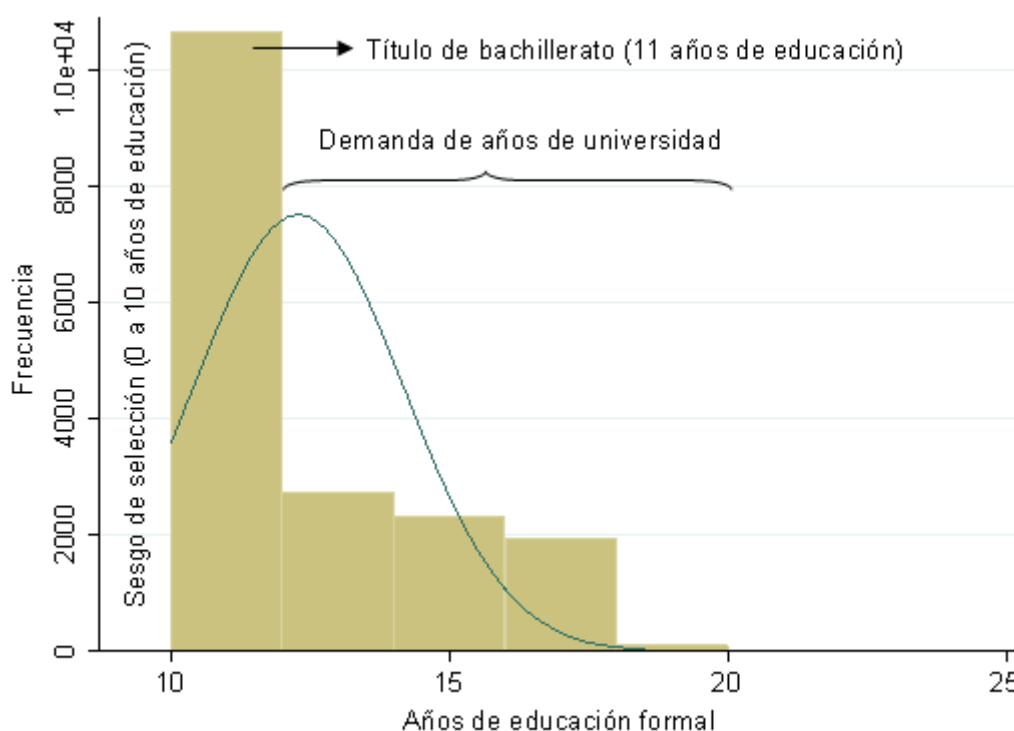
Años de educación	Porcentaje	Acumulado	Demanda por años de universidad
0	0.94	0.94	NA
1	0.29	1.23	NA
2	0.65	1.88	NA
3	1.22	3.10	NA
4	1.27	4.37	NA
5	6.43	10.80	NA
6	3.72	14.52	NA
7	4.95	19.47	NA
8	6.37	25.84	NA
9	8.24	34.08	NA
10	7.47	41.55	NA
11	35.01	76.55	0
12	4.46	81.02	1
13	4.55	85.57	1
14	4.55	90.12	1
15	3.05	93.17	1
16	5.87	99.04	1
17	0.54	99.59	1
18	0.22	99.81	1
19	0.10	99.91	1
20	0.04	99.95	1
21	0.03	99.98	1
22	0.02	100	1

Fuente: Cálculos del autor con Encuesta de hogares de Colombia.

Ahora bien, en la figura 2 se presenta la distribución de la parte de la muestra que supera la regla del truncamiento incidental, que en este caso sería el obtener el título de

bachiller o terminar 11 años de educación. Así, a partir de la primera elección de educación postobligatoria (bachillerato), se pueden observar los jóvenes que deciden ir o no a la universidad. Por lo tanto, el sesgo de selección se evidencia al no poder observar las características de todos aquellos jóvenes que no cumplieron con la regla de selección por motivos que no son aleatorios, por ejemplo: la habilidad, la deserción estudiantil, el abandono escolar temprano, el difícil acceso al mercado de capitales y un mercado de trabajo altamente informal con altas tasas de trabajo infantil, entre otros factores.

Figura 2 Distribución del sesgo de selección muestral en las elecciones de demanda de educación universitaria en Colombia, año 2005



Fuente: Cálculos del autor con Encuesta de hogares de Colombia.

Para obtener estimadores insesgados y consistentes se plantea un modelo en dos etapas que corrige el problema del sesgo de selección, obteniendo en una primera etapa una variable que corrige el problema de selección, para posteriormente incluirla como regresor en la estimación de la probabilidad de ir o no a la universidad en Colombia.

3. Metodología econométrica

El proceso de adquirir educación universitaria exige un requisito previo que consiste en aprobar primero los estudios de bachillerato. Además, las decisiones sobre inversión en educación, si el individuo es racional y tiene perfecta información, son tomadas en un único momento del tiempo. Según Valiente (2003), tras finalizar la educación

obligatoria, el individuo o su familia, decide qué nivel educativo desea obtener: educación media, bachiller o universitaria. Pero la situación familiar o social del momento en que se realiza esta decisión puede verse alterada por cambios en la capacidad de financiación, en las características familiares, en las condiciones del mercado de trabajo, entre otras razones, y en consecuencia se podría reconsiderar la elección previa. Por lo tanto, para tener en cuenta la incertidumbre que rodea a las decisiones que abarcan más de un nivel educativo, se debe suponer que la elección de invertir en educación se realiza por etapas o completando niveles educativos durante todo el proceso y no como una única decisión (Willis y Rosen, 1979 y Marcerano y Navarro, 2001).

La estimación de los modelos *probit* univariantes robustos pueden presentar una limitación importante para que tengan validez, esto es la necesidad de que exista independencia entre los factores no observados, es decir, que la covarianza entre los términos de error de cada uno de los dos modelos de elección (educación bachiller y educación universitaria) sea cero. Si esto no se cumpliera, el modelo *probit* univariante no podría ser empleado ya que los estimadores serían sesgados e inconsistentes. Por lo tanto, si los términos de error asociados a cada una de las dos elecciones no son independientes entre sí, se debe recurrir a una especificación más sofisticada del modelo *probit*.

En este caso, la omisión de la condición necesaria de poseer un título de bachiller para acceder a la universidad condiciona la muestra ya que esta deja de ser aleatoria. Este problema podría ser corregido por medio de un modelo *probit* con una estimación en dos etapas, para la cual el modelo apropiado es un Heckprob, donde se incluye un término de corrección del sesgo de selección muestral en la misma línea de los trabajos de Heckman (1979), Van de Ven y Van Praag (1981) y Miranda y Rabe-Hesketh (2006).

El sesgo de selección muestral en la elección de educación universitaria en Colombia se da por el problema de endogeneidad derivado de la omisión de una variable relevante que sería la corrección del sesgo de selección muestral. De ahí que se deban estimar dos etapas. En la primera etapa se estima la probabilidad de tener un título de bachiller, con la siguiente ecuación de selección:

$$Y_{li} = Y_{li}^* = W_{li} \hat{\beta}_1 + \varepsilon_{li} \quad (1)$$

Donde Y_{li}^* es una variable aleatoria no observable que representa la utilidad que proporciona al individuo invertir en educación, específicamente en alcanzar el título de bachiller. Por lo que se puede definir una variable aleatoria discreta Y_{li} que refleja la elección:

$$Y_{1i} = \begin{cases} 0 = \text{si el joven est\u00e1 cursando el bachillerato, un nivel inferior o sin estudios.} \\ 1 = \text{si tiene acabados los estudios de bachiller pero no va a la universidad.} \end{cases}$$

Adem\u00e1s, W_{1i} es el vector de caracter\u00edsticas individuales, β_1 es el vector de coeficientes asociados al modelo y ε_{1i} es el t\u00e9rmino de error aleatorio que se distribuye seg\u00fan una normal $[0, \sigma_1^2]$.

En una segunda etapa se estima la ecuaci\u00f3n de inter\u00e9s, que en este caso es la elecci\u00f3n por adquirir educaci\u00f3n universitaria.

$$Y_{2i} = Y_{2i}^* = W_{2i}'\beta_2 + \varepsilon_{2i} \quad (2)$$

Donde Y_{2i}^* es una variable aleatoria no observable, que representa la utilidad que proporciona al individuo invertir en educaci\u00f3n universitaria. Se puede definir una variable aleatoria discreta Y_{2i} que refleja la elecci\u00f3n:

$$Y_{2i} = \begin{cases} 0 = \text{si ha finalizado los estudios de bachillerato pero no va a la universidad.} \\ 1 = \text{si ha finalizado los estudios de bachillerato y tiene alg\u00fan a\u00f1o de educaci\u00f3n universitaria o ya termin\u00f3 sus estudios universitarios.} \end{cases}$$

Igualmente, W_{2i} es el vector de caracter\u00edsticas individuales, β_2 es el vector de coeficientes asociados al modelo y ε_{2i} es el t\u00e9rmino de error aleatorio que se distribuye seg\u00fan una normal $[0, \sigma_2^2]$. El vector de variables explicativas estar\u00e1 determinado por caracter\u00edsticas propias del individuo como el g\u00e9nero, caracter\u00edsticas de los padres como su educaci\u00f3n y situaci\u00f3n laboral, caracter\u00edsticas del hogar como el n\u00famero de hermanos menores de 16 a\u00f1os y el tama\u00f1o de la familia, el ingreso del hogar y la regi\u00f3n, entre otros controles incluidos en la estimaci\u00f3n.

Para analizar las correlaciones entre las variables explicativas incluidas en el modelo con aquellas excluidas, se toma la esperanza matem\u00e1tica condicional de la ecuaci\u00f3n (2), que representa la ecuaci\u00f3n de inter\u00e9s:

$$E(Y_{2i} | Y_{2i}^*, W_{1i}'\beta_1 + \varepsilon_{1i} > 0) = W_{2i}'\beta_2 + E(\varepsilon_{2i} | W_{2i}, W_{1i}'\beta_1 + \varepsilon_{1i} > 0)$$

Asumiendo que ε_{1i} y ε_{2i} siguen una distribuci\u00f3n normal bivalente $[0, 0, 1, 1, \rho]$, representando ρ el coeficiente de correlaci\u00f3n entre ambos t\u00e9rminos de error aleatorio, puede escribirse:

$$E(\varepsilon_{2i} | W_{2i}, W_{1i}'\beta_1 + \varepsilon_{1i} > 0) = \sigma_2 \rho \lambda_2$$

Si se realiza la normalización $\sigma_2 = 1$, resulta:

$$E(\varepsilon_{2i} | W_{2i}, Y_{1i}^* > 0) = \rho \lambda_1$$

Así, el parámetro λ_1 se interpreta como el término de corrección del sesgo de selección muestral y se puede escribir de la siguiente forma:

$$\lambda_1 = \frac{\phi(W_i'\beta)}{1 - \Phi(W_i'\beta)}, \quad \text{si } Y_1 > 0, \quad \text{y } \lambda_1 = \frac{-\phi(W_i'\beta)}{\Phi(W_i'\beta)}, \quad \text{si } Y_1 \leq 0$$

Donde ϕ representa la función de distribución normal acumulativa y Φ la pdf. De esta forma, dado que $E(\varepsilon_{2i} | W_{2i}, Y_{1i}^* > 0) = \rho \lambda_1$, el modelo por estimar tendría la forma de un modelo Heckprob:

$$Y_{2i} = W_{2i}'\beta_2 + \rho\lambda_i + \zeta_{2i} > 0 \quad (3)$$

En la ecuación de interés aparece el término de corrección del sesgo de selección muestral λ_1 . Por lo tanto, la ecuación de selección vendría definida por:

$$W_{1i}'\beta_1 + \zeta_{1i} > 0 \quad (4)$$

Además se deben cumplir los siguientes supuestos sobre el comportamiento del término de error aleatorio asociado a la ecuación de interés (3), así como de la ecuación de selección (4).

$$\begin{aligned} \zeta_{1i} &\approx N(0,1) \\ \zeta_{2i} &\approx N(0,1) \quad , \\ \text{corr}(\zeta_{1i}, \zeta_{2i}) &= \rho \end{aligned}$$

Finalmente, cuando existe correlación $\rho \neq 0$ los estimadores de los modelos *probit* univariantes son sesgados e inconsistentes, mientras que la estimación del modelo Heckprob provee estimadores consistentes y asintóticamente eficientes para la ecuación de interés. La estimación del modelo se puede realizar de forma robusta con la corrección a la matriz de varianzas y covarianzas propuesta por White (1980 y 1982), que permite obtener errores estándar robustos y resolver el problema de heteroscedasticidad presente por lo general en datos de corte transversal. Por otro lado, la interpretación de los coeficientes no se puede hacer de la forma convencional, por lo que es necesario calcular el efecto marginal evaluado en la media y la probabilidad condicional de que ocurra la elección.

3.1 Los datos

Los datos utilizados en este artículo provienen de varias etapas de la Encuesta de hogares de Colombia entre los años 1990 al 2005. Esta encuesta es realizada por el Departamento Administrativo Nacional de Estadísticas (DANE). Se debe tener en cuenta que para la realización de las estadísticas descriptivas y las estimaciones econométricas se hizo previamente una depuración de la base de datos ya que existen registros para los cuales no se tiene información para la elaboración de las estadísticas o las regresiones. En la tabla 3 aparece un resumen de las variables utilizadas a la luz de la teoría y los hechos estilizados.

Tabla 3 Descripción de variables

Tipo de variable	Nombre de la variable	Modalidad de variable	Categorías	Descripción
Dependiente	UNIV	Dicotómica	1: ≥ 11 años educa 0: En otro caso	Demanda de años de educación universitaria
Selección	BACHI	Dicotómica	1: $= 11$ años educa 0: En otro caso	Demanda del título de bachillerato
Personales	Género	Dicotómica	1: Mujer 0: Varón	Sexo (o género) del individuo.
Relación familiar	No es hijo	Dicotómica	1= No es hijo 0= En otro caso	Relación con respecto a los padres
Características de los padres	Educación de los padres	Dicotómicas	1: Padre sin estudios 1: Padre bachiller 1: Padre universitario 1: Madre sin estudios 1: Madre bachiller 1: Madre universitaria	Nivel educativo máximo alcanzado por los padres.
	Actividad laboral de los padres	Dicotómicas	1: Padre ocupado 1: Padre inactivo 1 Padre desempleado 1: Madre ocupada 1: Madre inactiva 1: Madre desempleada	Condición laboral de los padres en el mercado laboral
Características del hogar	Tamaño del hogar	Dicotómicas	1: ≤ 3 personas 1: 4 personas 1: 5 personas 1: ≥ 6 personas	Número de personas que son miembros del hogar
	Número de hermanos < 16 años	Dicotómicas	1: Ninguno 1: 1 hermano 1: 2 hermanos 1: Mas de 3 hermanos	Número de hermanos menores de 16 años
Ingresos del hogar	Ingresos	Dicotómicas	1: < 1 SMLV 1: 1 hasta 2 SMLV 1: 2 hasta 3 SMLV 1: > 3 SMLV	Total ingresos del hogar ajustados por inflación: Salario mínimo legal vigente (SMLV)
Región	Bogotá	Dicotómica	1: Bogotá 0: En otro caso	Región donde reside el individuo. En este caso Bogotá Distrito Capital

Fuente: Elaboración propia con Encuesta de hogares. **Individuo de referencia:** Hombre con padre y madre sin estudios (analfabetos, con primaria, sin bachillerato terminado) estando ambos ocupados, en una familia de tres o menos personas, sin hermanos menores de 16 años, con unos ingresos familiares inferiores a 1 SMLV y que además reside en la ciudad de Bogotá. Las categorías padres sin estudios e inactivos recogen a los no hijos.

3.2 Estadísticas descriptivas

En la tabla 4 se presentan las estadísticas descriptivas de la muestra y las variables utilizadas para el análisis. Se puede observar que la característica propia del individuo, que en este caso es el género, muestra que las mujeres son mayoría en el rango de edad de 17 a 31 años, llegando a ser el 54,14 por ciento en el año 2005. Igualmente, los individuos que son hijos superan el 70 por ciento.

Por otra parte, dentro de las características del entorno familiar del individuo están el nivel de educación de los padres y la actividad laboral de los mismos. De la educación de los padres se puede resaltar que la gran mayoría de estos se encuentran en la categoría sin estudios 61,53 y 62,43 por ciento, para padres y madres respectivamente, en el año 2005. Mientras que en la categoría de padres y madres con educación superior el porcentaje es muy cercano al 10 por ciento. De otro lado, la actividad laboral de los padres muestra una estructura del mercado laboral donde hay un gran porcentaje de individuos inactivos y otros ocupados, mientras que los desempleados que son padres llegan al 5 por ciento y los que son madres al 10 por ciento.

En cuanto a las variables que están midiendo la composición de la familia se encuentran el tamaño del hogar, donde estos están conformados por menos de 3 y hasta más de 6 miembros; y el número de hermanos menores de 16 años, donde el que mayor peso tiene es el de 1 hermano menor de 16 años con un 33 por ciento en 2005. Por otra parte, los ingresos de la familia se calcularon sumando los ingresos nominales de todos los miembros del hogar y se llevaron a pesos constantes del año 2000. Así, se decidió agruparlos en cuatro categorías que dependen del salario mínimo mensual legal vigente en Colombia (SMMLV). Esta es la variable que más ha cambiado en el tiempo, ya que en 1990 los hogares con 1 SMMLV eran del 94 por ciento y en 2005 son cerca del 20 por ciento. Finalmente, la variable región discrimina entre los individuos que residen en la capital de Colombia, Bogotá D.C., que son aproximadamente el 20 por ciento de la población para 2005.

Tabla 4 Estadísticas descriptivas

	1990	1995	2000	2005
	(%)	(%)	l(%)	(%)
Género				
Hombre	45.74	44.70	46.90	45.86
Mujer	54.26	55.30	53.10	54.14
Relación familiar				
Hijo	88.56	82.39	77.11	74.35
No es hijo	11.44	17.61	22.89	25.65
Educación de los padres				
Padre sin estudios	76.82	71.92	78.09	61.53
Padre bachiller	17.43	20.33	16.08	29.06
Padre universitario	5.75	7.76	5.83	9.41
Madre sin estudios				
Madre sin estudios	81.37	73.81	74.52	62.43
Madre bachiller	16.22	21.87	20.93	30.64
Madre universitaria	2.41	4.32	4.56	6.93
Actividad de los padres				
Padre empleado	23.33	20.02	43.35	19.10
Padre inactivo	73.79	77.45	51.83	76.13
Padre desempleado	2.89	2.53	4.82	4.77
Madre empleada				
Madre empleada	71.60	64.00	52.62	49.90
Madre inactiva	24.82	32.32	37.52	42.01
Madre desempleada	3.58	3.67	9.86	8.09
Tamaño del hogar				
<=3 personas	21.40	19.69	24.43	21.97
4 personas	18.97	21.20	21.01	24.53
5 personas	17.62	20.09	18.15	21.21
>6 personas	42.00	39.01	36.40	32.28
Hermanos <16 años				
0 hermanos < 16	22.46	19.96	25.37	31.05
1 hermano < 16	28.27	30.47	29.91	33.30
2 hermanos < 16	24.95	26.85	23.38	24.37
>3 hermanos < 16	24.31	22.71	21.35	11.28
Ingresos del hogar				
< 1 SMLV	94.41	59.30	41.66	18.55
De 1 hasta 2 SMLV	4.44	26.28	25.76	29.11
De 2 hasta 3 SMLV	0.60	7.46	12.83	19.40
Más de 3 SMLV	0.55	6.96	19.75	32.93
Región				
Resto	75.55	76.17	87.80	80.52
Bogotá D.C.	24.45	23.83	12.20	19.48
Observaciones	14828	15382	17905	18003

Fuente: Cálculos del autor con Encuesta de hogares de Colombia (DANE). (SMLV): Salario mínimo legal vigente de Colombia deflactado por inflación. **Individuo de referencia:** Hombre con padre y madre sin estudios (analfabetos, con primaria, sin bachillerato terminado) estando ambos ocupados, en una familia de tres o menos personas, sin hermanos menores de 16 años, con unos ingresos familiares inferiores a 1 SMLV y que además reside en la ciudad de Bogotá. Las categorías padres sin estudios e inactivos recogen a los no hijos.

4. Resultados

Los resultados de la tabla 5 muestran como la influencia de las variables socioeconómicas que fueron especificadas en el modelo sobre la probabilidad de demandar años de educación universitaria es la esperada según las hipótesis planteadas con la teoría económica y las regularidades empíricas (González, 2011). Así, el género resulta ser un factor determinante ya que el coeficiente asociado a ser mujer resultó ser positivo y estadísticamente significativo. Este efecto parece ser reciente y haberse acentuado después de la Ley 30 de 1992, pues el signo era negativo para el año 1990. Lo anterior implica que en esa época ser mujer reducía la probabilidad de ir a la universidad. No obstante, según los resultados, actualmente la situación es diferente, y esta diferencia entre hombres y mujeres refleja el aumento extraordinario que han tenido estas últimas en su participación en la educación superior. Según las estadísticas del Ministerio de educación, en la actualidad las mujeres son la mayoría en la matrícula de educación superior en Colombia. Otra posible explicación a este fenómeno, que no es exclusivo para el país, sería que por lo general las mujeres tienen unas mayores tasas de desempleo que los hombres, lo que las incentiva a aumentar su nivel de educación formal (Rice, 1987; Kodde y Ritzen, 1988; Kodde, 1988; Micklewright *et al.*, 1990; Calero, 1996; Mora, 1996; Albert, 1998 y 2000).

El nivel educativo de los padres también resultó ser un factor determinante de la demanda por años de educación universitaria. Los coeficientes asociados son positivos y estadísticamente significativos. Así, padres y madres con bachillerato o educación superior afectan positivamente la probabilidad de demandar educación universitaria, comparado con padres analfabetas/sin estudios o que no obtuvieron un grado de educación postobligatoria. Lo anterior indica que los hijos de padres con un alto nivel educativo poseen un entorno familiar propicio para ascender en la pirámide educativa, crear los incentivos para este tipo de inversiones y heredar las preferencias por el capital humano. Este resultado es acorde con la literatura internacional que asocia el nivel de educación superior de los padres con el estatus social y económico del hogar (Leibowitz, 1974 y Bowles y Nelson, 1974)

El efecto de la actividad laboral de los padres cuando estos están desempleados es negativo y estadísticamente significativo comparado con padres que se encuentran ocupados. Este efecto podría ser derivado de las restricciones presupuestales y de acceso a créditos cuando los padres pasan por episodios de desempleo, que efectivamente limitan las posibilidades de inversión en educación, más aun en países con mercados de capitales y de créditos educativos que funcionan de forma imperfecta. Igualmente, la situación de madre inactiva tiene un signo positivo y es estadísticamente significativo, lo cual podría estar mostrando que en Colombia todavía existen hogares donde las madres se dedican al cuidado de los hijos, lo que supone un aumento en el

tiempo de permanencia de uno de los padres en el hogar y es posible que esto repercuta en la atención y el cuidado de los hijos.

Las restricciones económicas asociadas al tamaño del hogar y al número de hermanos menores de 16 años resultan estadísticamente significativas. Al parecer existe un fuerte efecto negativo de pertenecer a un hogar numeroso sobre la probabilidad de demandar educación universitaria y la existencia de hermanos menores de 16 años en el hogar, ya que estos pueden ser una carga económica que afronta la familia al no ser generadores de ingresos, pero sí contribuir al gasto familiar.

Muchos autores registran la variable del ingreso del hogar como uno de los factores de mayor peso entre los determinantes de la demanda de educación. En el caso de Colombia esto parece acentuarse, ya que la educación superior no es gratuita y la oferta está compuesta por los sectores oficial y privado. Los coeficientes asociados a la variable ingresos resultan estadísticamente significativos y muestran la importancia de tener un ingreso familiar alto para mejorar las probabilidades de demandar educación universitaria.

La variable región, que mide la diferencia de la capital del país con el resto de áreas metropolitanas, resulta estadísticamente significativa y con un signo positivo. Lo anterior confirma el peso que tiene Bogotá en materia económica, social, política y demográfica sobre el país. Por lo tanto, un joven que resida en Bogotá tiene mayores probabilidades de acceder a una universidad que si vive en otra región.

Finalmente, la especificación y estimación de un modelo *probit* con corrección del sesgo de selección muestral resulta adecuada para el análisis de la probabilidad de demanda por años de educación universitaria en Colombia, dado que el término de corrección del sesgo de selección muestral (λ_1) resulta ser positivo y estadísticamente significativo, probando la existencia del sesgo de selección y corrigiendo los resultados (Cameron y Trivedi, 2005; Wooldridge, 2010 y Greene, 2011). Así, con el modelo Heckprob se puede realizar inferencia estadística sobre la población objeto de estudio. También es importante resaltar que las regresiones se realizan con errores estándar robustos por medio del método de White (1980 y 1982), lo que corrige el posible problema de heteroscedasticidad y garantiza, además de la corrección del sesgo de selección, que los estimadores sean consistentes y asintóticamente eficientes. Estos modelos, Heckprob, presentan un buen ajuste y al realizar una prueba de significancia conjunta (wald test), esta resulta significativa con un valor-p asociado de 0.00. Lo que indica que todos los coeficientes en su conjunto son válidos para el modelo en términos estadísticos, dando aún más validez a las estimaciones realizadas.

Tabla 5 Resultados de la estimación robusta de los modelos *probit* con corrección del sesgo de selección

	1990		1995		2000		2005	
	Universidad	Bachillerato	Universidad	Bachillerato	Universidad	Bachillerato	Universidad	Bachillerato
Constante	-0.01	-4.05***	-0.03***	0.18***	-0.06*	0.34***	-0.12***	0.16***
Género								
Mujer	-0.05*	0.08***	0.10***	0.20***	0.17***	0.23***	0.03***	0.09***
Relación familiar								
No es hijo	-0.02	-0.02	-0.01	-0.02	-0.01*	-0.02*	-0.02*	-0.02*
Educación de los padres								
Padre bachiller	0.08	0.72***	0.47***	0.59***	0.14***	0.39***	0.28***	0.59***
Padre universitario	0.86***	0.54***	0.99***	0.57***	0.67***	0.34***	0.79***	0.42***
Madre bachiller	0.13	1.19***	0.25***	0.78***	0.49***	0.83***	0.32***	0.78***
Madre universitaria	1.02***	1.18***	1.10***	0.74***	1.14***	0.72***	0.85***	0.61***
Actividad laboral de los padres								
Padre desempleado	-0.20*	0.06	-0.18*	0.29	-0.04***	-0.12***	-0.04**	0.10
Padre inactivo	-0.02	0.27	-0.03	0.17	-0.21***	0.24***	-0.42**	0.46
Madre desempleada	-0.32*	-0.04*	-0.02*	-0.11***	-0.16***	0.16***	-0.15**	0.20***
Madre inactiva	-0.08	0.02*	0.03	0.12***	0.02***	0.00	0.01	0-06***
Tamaño del hogar								
4 personas	0.24***	0.03	0.15*	-0.10	0.01***	0.02	0.06*	-0.04
5 personas	0.34***	0.23**	0.17**	-0.04	0.14***	0.24***	0.18**	0.04
>6 personas	0.51***	0.46**	-0.22**	0.07	-0.08***	0.32***	-0.08**	-0.15***
Número de hermanos < 16 años								
1 hermano < 16	-0.31***	-0.62***	-0.47***	-0.49***	-0.57***	-0.60***	-0.41***	-0.17***
2 hermanos < 16	-0.59***	-0.89***	-0.81***	-0.85***	-0.89***	-0.97***	-0.71***	-0.45***
>3 hermanos < 16	-0.59***	-1.33***	-1.08***	-1.23***	-1.29***	-1.40***	-0.86***	-0.65***
Ingresos del hogar (corregidos por inflación)								
De 1 hasta 2 SMLV	-0.46*		-0.61***		-0.63***		-0.66***	
De 2 hasta 3 SMLV	-0.93*		-0.27***		-0.29***		-0.45***	
Más de 3 SMLV	0.02*		0.15***		0.17***		0.22***	
Región								
Capital (Bogotá D.C)	-0.13***	0.04	0.03***	-0.02	0.02***	0.03	0.13***	0.109
/athrho	0.21*		0.93***		0.53***		0.69***	
Obs.	14828		15382		17905		18003	
% Obs. censuradas	46.97		40.12		35.60		30.84	
Wald-Prob>chi2	0.00		0.00		0.00		0.00	

Fuente: Cálculos del autor con Encuesta de hogares de Colombia. **Nota:** Errores estándar robustos. *** p<0.01, ** p<0.05, * p<0. **Individuo de referencia:** Hombre con padre y madre sin estudios (analfabetos, con primaria, sin bachillerato terminado) estando ambos ocupados, en una familia de tres o menos personas, sin hermanos menores de 16 años, con unos ingresos familiares inferiores a 1 SMLV y que además reside en la ciudad de Bogotá. Las categorías padres sin estudios e inactivos recogen a los no hijos.

Los coeficientes del modelo *probit* con corrección del sesgo de selección carecen de una interpretación convencional. Esta se hace a través de los efectos marginales. En la tabla 6 se presentan los efectos marginales y las probabilidades condicionales a la ocurrencia de la elección. Así, para el año 2005, ser mujer aumenta la probabilidad de demandar años de educación universitaria en 14 puntos porcentuales con respecto a los hombres, manteniendo todos los demás factores constantes. Las variables socioeconómicas que mayor efecto marginal tienen sobre la probabilidad de elección son que el padre y la madre tengan estudios universitarios con 31.4 y 35.2 puntos porcentuales, respectivamente; y, el tener un ingreso alto, que aumenta la probabilidad de demandar años de educación en 14.5 puntos porcentuales.

Algunas de las variables que tienen un efecto marginal negativo sobre la probabilidad son: la actividad laboral del padre cuando este es desempleado ($-0.012*100$) o inactivo ($-0.163*100$) comparado con un padre ocupado, la condición de madre desempleada (5.1 puntos porcentuales), el número de hermanos (20 puntos porcentuales) y la probabilidad de pertenecer a un hogar con ingresos familiares bajos. Sin embargo, para la madre inactiva el efecto es positivo.

Por otro lado, las probabilidades de demandar años de educación universitaria condicionadas a las características del individuo promedio continúan siendo muy bajas, para el año 2005 era de aproximadamente 32.35 por ciento. Esta probabilidad ha ido aumentando con el tiempo, pasando de 18.53 por ciento en el año 1990 a 22.69 por ciento en el año 1995 y 29.06 por ciento en el año 2000. El aumento de 13.82 puntos porcentuales que se registró de 1990 a 2005 corresponde con el crecimiento del sector educativo durante este periodo. Sin embargo, la probabilidad continúa siendo muy baja, teniendo en cuenta que la tasa de cobertura para la educación superior es de alrededor del 30 por ciento para el año 2005 y existe una deserción de alrededor del 40 por ciento de los alumnos que ingresan a la educación superior según cifras del Ministerio de educación.

Tabla 6 Efectos marginales y las probabilidades condicionadas en media

	1990	1995	2000	2005
Género				
Mujer	-0.023*	0.036***	0.056***	0.140***
Relación familiar				
No es hijo	-0.01	-0.01	-0.02*	-0.02*
Educación de los padres				
Padre bachiller	0.031	0.173***	0.047***	0.114***
Padre universitario	0.314***	0.378***	0.256***	0.301***
Madre bachiller	0.052	0.096***	0.177***	0.115***
Madre universitaria	0.352***	0.417***	0.433***	0.329***
Actividad de los padres				
Padre desempleado	-0.082*	-0.068*	-0.015***	-0.012**
Padre inactivo	-0.006	-0.010	-0.074***	-0.163**
Madre desempleada	-0.126*	-0.007*	-0.056***	-0.051**
Madre inactiva	-0.029	0.007	0.003***	0.001
Tamaño del hogar				
4 personas	0.097***	0.053*	0.003***	0.019*
5 personas	0.136***	0.058**	0.048***	0.069**
>6 personas	0.199***	-0.081**	-0.026***	-0.028***
Hermanos <16 años				
1 hermano < 16	-0.125***	-0.164***	-0.184***	-0.141***
2 hermanos < 16	-0.226***	-0.254***	-0.256***	-0.230***
>3 hermanos < 16	-0.229***	-0.320***	-0.336***	-0.247***
Ingresos del hogar				
De 1 hasta 2 SMLV	-0.176*	-0.024***	-0.212***	-0.227***
De 2 hasta 3 SMLV	-0.331*	-0.091***	-0.093***	-0.149***
Más de 3 SMLV	0.123*	0.058***	0.065***	0.145***
Región				
Bogotá D.C.	-0.041***	0.014***	0.007***	0.045***
<i>Prob (Y=1 X=media)</i>	<i>18.71%</i>	<i>22.83%</i>	<i>29.44%</i>	<i>33.14%</i>
<i>% Predicciones correctas</i>	<i>64.21%</i>	<i>65.93%</i>	<i>67.04%</i>	<i>70.25%</i>

Fuente: Cálculos del autor. **Individuo de referencia:** Hombre con padre y madre sin estudios (analfabetos, con primaria, sin bachillerato terminado) estando ambos ocupados, en una familia de tres o menos personas, sin hermanos menores de 16 años, con unos ingresos familiares inferiores a 1 SMLV y que además reside en la ciudad de Bogotá. Las categorías padres sin estudios e inactivos recogen a los no hijos.

A continuación se comparan los resultados de las estimaciones de los modelos *probit* robustos con corrección del sesgo de selección con los modelos *probit* que no tienen en cuenta la corrección. En los modelos *probit* univariantes se utilizó el supuesto de que no existía correlación entre las perturbaciones de las dos elecciones de educación postobligatoria (bachillerato y universidad) y que por lo tanto no existía un sesgo de selección. Sin embargo, los resultados muestran que la estimación del modelo *probit* robusto con corrección del sesgo de selección permitió corregir el problema de endogeneidad derivado del mismo. Al mismo tiempo se ganó consistencia y eficiencia en los estimadores comparados con los del modelo *probit* robusto.

Para realizar la comparación de las dos metodologías de estimación se seleccionaron los efectos marginales en la media de los años 1990 y 2005. En la tabla 7 se observa que el efecto marginal del modelo *probit* univariante subestima los incrementos en la probabilidad comparado con los efectos marginales del modelo *probit* robusto con corrección del sesgo de selección muestral. Por ejemplo, el efecto marginal de ser mujer en el modelo con corrección del sesgo de selección es de 14 puntos porcentuales, mientras que para el modelo sin corrección del sesgo es de 1.6. La diferencia, entre uno y otro, es de más de 10 puntos porcentuales.

La diferencia en el resultado del efecto marginal para la educación de los padres no es muy grande. En el modelo *probit* con corrección del sesgo de selección cuando el padre y la madre tienen educación universitaria es de 30.1 y 32.9 puntos porcentuales, respectivamente. Mientras que para el modelo *probit* univariante es tan sólo de 7.3 y 12.7 puntos porcentuales, en los mismos casos. Otros ejemplos de la corrección que realiza la estimación del modelo Heckprob frente al modelo *probit*, se observan en el número de hermanos menores de 16 años, el nivel de ingresos y la región.

Del mismo modo, la probabilidad del modelo *probit* univariante también está subestimada si se compara con la probabilidad condicional del modelo *probit* robusto con corrección del sesgo de selección. Para el año 2005, la probabilidad del primero es de 25.63 por ciento, mientras que para el segundo la probabilidad es de 33.14 por ciento. Lo mismo ocurre en el año 1990, donde la probabilidad del modelo *probit* con corrección del sesgo de selección muestral es mayor en 9.36 puntos porcentuales.

En suma, la estimación de la probabilidad de demanda por años de educación universitaria debe plantearse como un modelo *probit* robusto con corrección del sesgo de selección, ya que hay evidencia estadística de que existe un problema derivado del sesgo de selección muestral. Por lo tanto, las elecciones sobre inversión en educación postobligatoria en Colombia están interrelacionadas según los resultados obtenidos.

Tabla 7 Efecto marginal *probit* vs *probit* con corrección del sesgo de selección

	1990		2005	
	EM- <i>Probit</i>	EM-Heckprob	EM- <i>Probit</i>	EM-Heckprob
Género				
Mujer	0.001	-0.023*	0.016***	0.140***
Relación familiar				
No es hijo	-0.018	-0.01	-0.078*	-0.02*
Educación de los padres				
Padre bachiller	0.022***	0.031	0.018***	0.114***
Padre universitario	0.113***	0.314***	0.073***	0.301***
Madre bachiller	0.012***	0.052	0.019***	0.115***
Madre universitaria	0.199***	0.352***	0.127***	0.329***
Actividad de los padres				
Padre desempleado	-0.022*	-0.082*	-0.053***	-0.012**
Padre inactivo	-0.015	-0.006	-0.024***	-0.163**
Madre desempleada	-0.003	-0.126*	-0.016***	-0.051**
Madre inactiva	-0.014	-0.029	-0.020***	0.001
Tamaño del hogar				
4 personas	0.001	0.097***	-0.010	0.019*
5 personas	-0.001	0.136***	-0.005**	0.069**
>6 personas	0.009	0.199***	-0.019***	-0.028**
Hermanos <16 años				
1 hermano < 16	-0.018***	-0.125***	-0.019***	-0.141***
2 hermanos < 16	-0.023***	-0.226***	-0.036***	-0.230***
>3 hermanos < 16	-0.032***	-0.229***	-0.032***	-0.247***
Ingresos del hogar				
< 1 SMLV	0.029***	-0.176*	-0.006	-0.227***
De 1 hasta 2 SMLV	0.039***	-0.331*	0.022***	-0.149***
De 2 hasta 3 SMLV	0.007***	0.123*	0.065***	0.145***
Más de 3 SMLV				
Región				
Bogotá D.C.	0.001	-0.041***	0.004	0.045***
<i>Prob (Y=1 X=media)</i>	16.27%	18.71%	25.63%	33.14%

Fuente: Cálculos del autor. **Individuo de referencia:** Hombre con padre y madre sin estudios (analfabetos, con primaria, sin bachillerato terminado) estando ambos ocupados, en una familia de tres o menos personas, sin hermanos menores de 16 años, con unos ingresos familiares inferiores a 1 SMLV y que además reside en la ciudad de Bogotá. Las categorías padres sin estudios e inactivos recogen a los no hijos.

5. Conclusiones

En este estudio se han analizado los determinantes de la demanda por años de educación universitaria de los jóvenes colombianos, teniendo en cuenta el sesgo de selección muestral que se presenta en este tipo de elecciones educativas. La evidencia estadística muestra la presencia del sesgo de selección y de ahí la necesidad de realizar la estimación a través de un modelo *probit* robusto con corrección del sesgo de selección (Heckprob), del cual se obtuvieron estimadores que son consistentes y asintóticamente eficientes. La estimación de este modelo permitió obtener mejores resultados comparados con la estimación del modelo *probit* robusto. De la estimación del modelo que tiene en cuenta el sesgo de selección se pueden derivar las siguientes conclusiones:

En primer lugar, dentro del conjunto de variables que han sido tenidas en cuenta para estimar la probabilidad de demandar años de educación universitaria, algunas de las que presentan una influencia positiva y estadísticamente significativa son: ser mujer, la educación de los padres (en especial la educación universitaria) y pertenecer a un hogar pequeño de hasta 5 personas. También resultó positivo el pertenecer a un hogar con ingresos familiares altos y residir en la ciudad de Bogotá. Por otra parte, algunas variables que tienen una influencia negativa fueron el tener hermanos menores de 16 años y un ingreso familiar bajo. Estos resultados son coherentes con las hipótesis planteadas por la teoría económica y las características del sistema educativo en Colombia.

En segundo lugar, la probabilidad condicionada a las características medias de un joven entre los 17 y 31 años de edad es muy baja. Para el año 2005 la probabilidad era de 33.14 por ciento. En cuarto lugar, los resultados se ajustan a un sistema de educación de un país en vía de desarrollo como es el caso de Colombia, donde existen factores asociados a las características personales, del entorno familiar y del entorno más próximo al individuo que afectan sus elecciones educativas y por lo tanto generan un alto grado de desigualdad en las oportunidades educativas postobligatorias. Entonces es innegable la necesidad de mejorar las políticas educativas en materia de cobertura, calidad, eficiencia y equidad para la educación universitaria, además de un papel todavía más activo y propositivo del sector privado.

Finalmente, es imperativo que existan más iniciativas público-privadas que permitan aumentar la oferta educativa en el sector de la educación superior, como por ejemplo una política de financiación (becas) mucho más agresiva por parte del Estado en las universidades públicas y privadas.

Referencias

- Albert, C. (1998). La evolución de la demanda de enseñanza superior en España. *Hacienda Pública Española*. Monografía educación y economía, p. 119-137.
- Albert, C. (2000). Higher education demand in Spain; the influence of labour market signals and family background. *Higher Education*, 40(2), p. 147-162.
- Albert C. y García-Serrano C. (2010). Cleaning the slate? School choice and educational outcomes in Spain. *Higher Education*, 60(6), p. 559–582.
- Ayala, U. (1987). Hogares, Participación Laboral e Ingresos. En: J. Ocampo y M, Ramirez (Eds. 1987). *El Problema Laboral Colombiano*. Bogota: Contraloría General de la República.
- Barr, N. (1993). Alternative Funding Resources for Higher Education. *The Economic Journal*. 103(418), p. 718–728.
- Betts, J. y McFarland L. (1995). Safe port in a storm. The impact of labor market condition on community college enrollments. *The Journal of Human Resources*, 30(4), p. 741-765.
- Bourguignon, F. (1979). Pobreza y Dualismo en el Sector Urbano de las Economías en Desarrollo: El caso de Colombia. *Desarrollo y Sociedad*. (1). P.39-72
- Bourguignon, F. (1983). El Papel de la Educación en el Mercado de Trabajo Urbano en el Proceso de Desarrollo: el Caso de Colombia. En V, Urkidy y Trejos (Eds. 1983). México: Recursos Humanos, Empleo y Desarrollo en América Latina, Fondo de Cultura Económica.
- Bowles, S. y Nelson, V. (1974). “The inheritance of IQ’ and the Intergenerational Reproduction of Economic Inequality”, *Review of Economics and Statistics* 56,1.
- Calero, J. (1996). La financiación de la educación superior en España. Sus implicaciones en el terreno de la equidad. *Fundación BBVA*
- Cameron, A. C., y P. K. Trivedi. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Cano, C., Muñoz, J. y Mesa, M. (2009). *El mercado laboral colombiano. Análisis desde la teoría de la señalización*. Bogotá: Investigaciones Universidad EAFIT.
- Carrizosa, M. (1987). Evolución y Determinantes de la Pobreza en Colombia. En J, Ocampo. y M, Ramírez (Eds.1987). Bogotá: El Problema Laboral Colombiano, SENA, DNP, Contraloría General de la República.
- Castellar, C. y Uribe, J. (2001). Una aproximación Econométrica a la Tasa de Retorno Social de la Educación. *Revista sociedad y economía*, 3(1). P. 24-40.
- Castellar, C. y Uribe, J. (2008). Capital humano y señalización: Evidencia para el área metropolitana de Cali. *Revista sociedad y economía*. 14(2).P.26-41.
- Chávez, A. y Arias, H. (2002). *Cálculo de la tasa interna de retorno de la educación en Colombia*. Universidad Externado de Colombia, (2).P.1-23.

- García, M. (1998). *Recursos formativos e inserción laboral de los jóvenes* (ed. 158 pp. 1-281). Siglo XXI.
- Girón, L., González, C., y Sánchez, H. (2003). Algunos cálculos de la tasa de retorno de la educación para Cali. *Revista Asuntos Económicos*. 5(1).P.7-25.
- González, C. G. (2011). *Análisis económico de la demanda de educación universitaria en Colombia: Un análisis cuantitativo*. Tesis Doctoral. Universidad de Alcalá, España.
- Greene, W. (2011). *Econometrics Analysis*. 7th. Prentice Hall,
- Grubb, W. (1988). Vocationalizing higher education. The causes of enrollment and completion in public Two-Year Colleges, 1970-1980. *Economic of Education Review*, 7(3), p. 301-319.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometría*. 7(1). P. 42-68.
- Hernández, G. (2010). ¿Cuán rentable es la educación superior en Colombia?. *Revista Lecturas de economía*. 73(1), P. 181-214.
- Kodde, D. (1988). "Unemployment expectation and human capital formation". *European Economic Review*, 32(8), p. 1645-1660.
- Kodde, D. y Ritzen, J. (1988). Direct and indirect effects of parental education level on the demand for higher education. *The Journal of Human Resources*, 23(3), p. 356-371.
- Leibowitz, A. (1974). "Home Investment in Children". *Journal of Political Economy*, 82(2), p. 111-131.
- Londoño, J. (1995). *Distribución del Ingreso y Desarrollo Económico; Colombia en el Siglo XX* (1ra ed., pp. 1-261). Bogotá: Tercer Mundo Editores.
- Manski, C. y Wise, D. (1983). *College Choice in America*. Harvard University Press. Cambridge, Massachusetts, an London England, P. 1-221.
- Marcerano, O. y Navarro, M. (2001). Un análisis microeconómico de la demanda de educación superior en España. *Estudios de Economía Aplicada*. (19). P. 69-86.
- Marcerano, O. y Navarro, M. (2007). El éxito en la Universidad: Una aproximación cuantílica. *Revista de Economía Aplicada*, 44 (15). P. 5-39.
- Marcerano, O., Galindo, F., Vignoles, A. (2007). "Who actually goes to university?" *Empirical Economics*, Springer, vol. 32(2), P. 333-357.
- Martínez, J. (1999). *La demanda de educación universitaria en un contexto de alto desempleo*. Tesina CEMFI. (9907).P. 1-46
- Micklewright, J., Pearson, M. y Smith, S. (1990). "Unemployment and early school leaving". *The Economic Journal*, 100(1), p. 163-169.
- Miranda, A. y Rabe-Hesketh, S. (2006). Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *Stata Journal*. 6(1). P. 285-308.
- Modrego, A. M. (1986). *Determinantes de la demanda de educación superior. Estimación de un modelo de educación superior para la provincia de Vizcaya*. Universidad del País Vasco, 329.

- Mohan, R. (1978). *Workers of Bogotá: Who They Are, What they do and Where they Live. The World Bank, Mimeo.*
- Mora, J. (1990). *La demanda de educación superior*, Consejo de Universidades.
- Mora, J. (1996). “Influencia del origen familiar en el acceso a la educación, en la obtención de empleo y en los salarios”. *Economía de la Educación*. Temas de estudio e investigación. Servicio Central de Publicaciones del Gobierno Vasco. 22. P. 195-211
- Mora, J., González, N., Gómez, J. y Zuluaga, B. (2004). Las Ganancias de Señalizar en el Mercado Laboral en Cali. *Revista Estudios Gerenciales*. 92(1).P. 105-128.
- Mora, J. (2006). El efecto de los títulos, la segmentación y el funcionamiento del mercado de trabajo: un análisis cuantitativo. Tesis Doctoral Universidad de Alcalá.
- Mora, J. y Muro, J. (2008). “Sheepskin effects by cohorts in Colombia”. *International Journal of Manpower*. 29 (2). P. 111-121.
- Muñoz, M. (1990). Algunos aspectos de la distribución del ingreso en Colombia, *Cuadernos de Economía*, 14(3). P. 35-70.
- Murnane, R.; Maynard, R. y Ohls, J. (1981). Home resources and Children`s Achievement. *The Review of Economics and Statistics*, 63(3), p. 369-377.
- Núñez, J. y Sánchez, F. (1998). Educación y Salarios Relativos en Colombia 1976-1995. Determinantes, Evolución e Implicaciones para la Distribución del Ingreso. *Archivos de Macroeconomía*. 74 (3).P. 1-51.
- Psacharopoulos, G., (1981). Returns to education: an update international comparison. *Comparative Education*, 17(3), p. 14-57.
- Rice, P. (1987). The demand for post-compulsory Education in the UK and the effects of educational maintenance allowances. *Económica*, 54(216), p. 465-475.
- Salas, M. (2003). *Educación superior y Mercado de trabajo*. Grupo editorial universitario.
- Sarmiento, A., Ramírez, C., Baldion, E. y De Gómez, M. (1987). *El Problema Laboral Colombiano*, En J, Ocampo. y Ramírez, M. (Eds.1987). Bogotá: Contraloría General de la República,.
- Schultz, T. (1961). Investment in human capital. En Blaug, M. (ed) *Economía de la educación*. Textos escogidos. 1972. Madrid editorial Tecnos.
- Selowsky, M. (1968). The Effect of Unemployment Labour-Market Segmentation and the Returns to Educated: the Case of Colombia. *Center for International Affairs, Harvard University y Revista de Planeación y Desarrollo*. Bogotá.
- Tenjo, J. (1993). Evolución de los Retornos de la Inversión en Educación 1976-1989. *Planeación y Desarrollo*. Bogotá.
- Tenjo, J. (1993b). Educación, habilidad, conocimientos e Ingreso. *Revista Planeación y Desarrollo*. 24 (Diciembre).
- Tenjo, J. (2001). *Pobreza y Mercados Rurales*, Documentos CEDE, (10). P. 1-35.
- Tenjo, J., Ribero, R. y Bernat, L. (2002). Evolución de las diferenciales Salariales por Sexo en seis Países de Latinoamérica. *Documentos de Economía*. 1(2) .P. 1-59

- Valiente, A. (2003). La demanda de educación universitaria y el rendimiento privado de la educación en España. Secretariado de Publicaciones e Intercambio Editorial. Universidad de Valladolid.
- Van de Ven, W. y Van Pragg B. (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics*. 17(1). P. 229–252.
- White, H. (1980). Maximum likelihood estimation of misspecified models. *Econometrica* 50. P. 1–25.
- White, H. (1982). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48.P 817–838.
- Willis, R. Y Rosen, S. (1979). Education and Self -Selection, *Journal of Political Economy*. 87(5), p. 7-36.
- Wooldridge, J. M. (2010). *Introductory Econometrics: A Modern Approach*. 4th ed. Cincinnati, OH: South-Western.