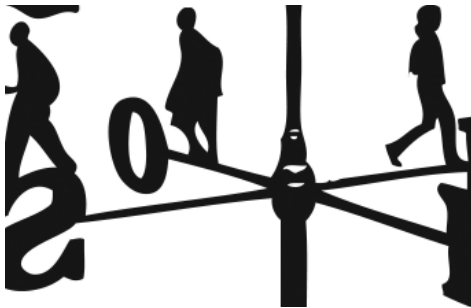


# Modelos de elección discreta

## Una aplicación a la demanda por cupos universitarios en ciudades intermedias de Colombia\*



**Andrés Ramírez Hassan**

Docente e investigador, Departamento de Economía,  
Universidad EAFIT.  
aramir21@eafit.edu.co.

Recepción: 17 de febrero de 2008 | Aceptación: 27 de agosto de 2008

### Resumen

Con esta exposición se presenta una caracterización cuantitativa de la demanda de cupos universitarios por parte de bachilleres egresados de ciudades intermedias colombianas. Se lograron resultados a través de modelos de elección discreta. Específicamente se encuentra que el número de programas ofrecidos aumenta la probabilidad de selección de las universidades por parte de los aspirantes, mientras que el ranking universitario está inversamente relacionado con dicha probabilidad. A otro nivel, se puede observar que la posición socioeconómica, medida a través del estrato, es una variable bastante explicativa de la elección universitaria. Así, los bachilleres que provienen de estratos bajos y medios tienden a optar por universidades públicas, mientras que los bachilleres de estratos altos seleccionan universidades privadas.

### Palabras Clave

Modelos de elección discreta  
Utilidad aleatoria  
Servicios educativos

\* Este artículo es resultado del trabajo investigativo "Caracterización cuantitativa de la demanda de la Universidad EAFIT", que fuera financiado con fondos de la Dirección de Investigación y Docencia de la Institución. El autor agradece enormemente la colaboración prestada por el asistente de investigación Leonel Arango Vásquez, por la jefe de mercadeo institucional de la Universidad, Beatriz Mora y su asistente Hernán Caro.

## Discrete selection models. An application to the demand for university places in intermediate cities of Colombia\*

### Abstract

A quantitative characterization of the demand for university places on the part of high school graduates from Colombian intermediate cities, is presented. Results were achieved through discrete selection models. Specifically, it was found that the number of programs offered to students, increases the possibilities of choosing a given university, whereas the university ranking is inversely related to such possibility. On the other hand, it can be observed that the socio-economic position, measured by the stratum, is a highly explicative variable in the university selection process. Therefore, students from low and middle strata tend to choose public universities, while those from high strata, choose private universities.

### Key words

Discrete selection models  
Random utility  
Educational Services

### Introducción



Una de las herramientas básicas para el proceso de planeación de las instituciones es la caracterización cuantitativa de la demanda, puesto que dicho análisis genera un marco de referencia fundamental para la toma de decisiones estratégicas, como por ejemplo lanzar un nuevo producto al mercado. En la medida que se conozca al potencial demandante se pueden implementar campañas institucionales que estén orientadas a influir sobre los principales inductores de los consumidores. Las entidades universitarias no son una excepción a la regla, especialmente las de propiedad privada, puesto que en general dependen de la evolución de las matrículas para garantizar su supervivencia financiera. En este orden de ideas, identificar las variables relevantes que inciden sobre la probabilidad de elección de las universidades ofrece pistas interesantes para orientar campañas tendientes a maximizar dicha probabilidad.

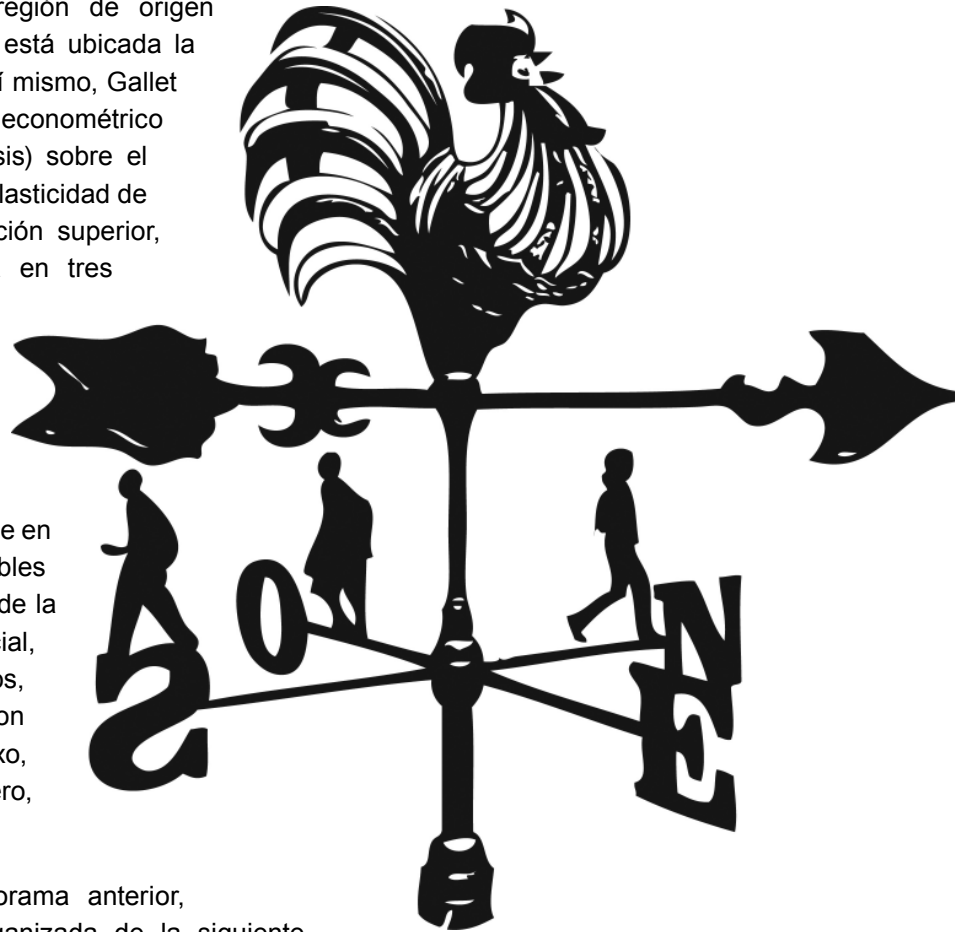
El objetivo del presente artículo es determinar cuáles son las variables más importantes que inciden en la probabilidad de elección de las principales universidades colombianas por parte de bachilleres provenientes de ciudades intermedias con una posición socioeconómica que en general es media – alta, salvo contadas excepciones. El estudio se realizó a través de modelos de elección discreta, a partir de información recolectada por vía de encuestas aplicadas en las ciudades objeto de estudio.

Al momento de llevar a cabo la presente investigación no se encontraron aplicaciones de este tipo en Colombia. A nivel internacional, la literatura al respecto no es muy extensa, pero se puede citar a Modrego Rico (1988), quien plantea un modelo de elección discreta en dos fases. La primera considera la elección de continuar estudiando, y la segunda, optar por formación técnica o universitaria. El objetivo es discriminar las tendencias por sexo considerando variables socioeconómicas. Salas Velasco y Martín-Cobos

\* This article is the result of the research Project “Quantitative characterization of places demand at Eafit University”, financed with funds belonging to the Direction of Research and Teaching of this institution. The author is very grateful to Leonel Arango Vásquez, research assistant, Beatriz Mora, institutional marketing chief at the university and to Hernán Caro, her assistant, for their great help.

Puebla (2006) trabajan con un modelo Logit binomial donde la variable endógena toma valores dependiendo de si los agentes seleccionan formación técnica o universitaria; este modelo se establece en función del nivel educativo de los padres, ingresos familiares y costos de la educación.

Otro estudio, desarrollado por Carla Sa, Raymond Florax y Piet Riedvelt (2004), pone especial énfasis en una variable poco explorada en los trabajos tradicionales: la dimensión espacial de los potenciales estudiantes. El trabajo analiza la demanda por educación superior de estudiantes graduados y tiene en cuenta tanto su región de origen como la región donde está ubicada la universidad elegida. Así mismo, Gallet (2007) hace un análisis econométrico exhaustivo (meta-análisis) sobre el comportamiento de la elasticidad de la demanda de educación superior, para lo cual se basa en tres variables principales: el nivel de ingreso del estudiante, el costo promedio de la educación superior y la oferta educativa. Además, tiene en cuenta algunas variables características propias de la institución (privada, oficial, duración de los cursos, etc.) y otras que lo son de los estudiantes (sexo, raza, nacional, extranjero, entre otras).



Con base en el panorama anterior, la exposición está organizada de la siguiente forma: la primera sección enseña el marco teórico microeconómico; la segunda muestra los resultados obtenidos a partir de la implementación de los diferentes modelos de elección discreta, y finalmente se proponen algunas conclusiones.

## 1. Marco teórico microeconómico

La interpretación económica de los modelos de elección discreta parte de la utilidad que le generan a los individuos los productos o servicios que ellos demandan. Bajo este contexto, la racionalidad de

los agentes económicos hace que se comporten de forma que maximicen la utilidad esperada, proporcionada por cada una de las alternativas disponibles dentro de un conjunto exhaustivo y mutuamente excluyente.

Obviamente, la satisfacción de cada uno de los individuos dentro de la muestra no es una variable observable, y lo que hace el investigador es descomponer dicha variable aleatoria en un componente determinístico y uno aleatorio (no observable). Al respecto del primero, se asume cierta forma funcional que, dependiendo de una serie de variables, recopila atributos propios de las alternativas y características inherentes al individuo. Con referencia al componente aleatorio, se asumen funciones de densidad específicas que generan diversos modelos. Las funciones de densidad más utilizadas son la normal, que da origen al modelo Probit, y la función valor extremo tipo uno, que genera el modelo Logit. De esta forma, la satisfacción que percibe un individuo por elegir la alternativa  $j$ , está dada por:

$$U_j = V_j + e_j \quad [1]$$

$$j = 1, 2, \dots, m$$

Donde  $V$  denota el componente determinístico de la satisfacción del individuo, y  $e$  es el componente aleatorio de la alternativa en cuestión. Generalmente, se encuentra que la especificación para el componente determinístico de la utilidad del individuo  $i$ -ésimo es:

$$V_{ij} = x'_{ij}\beta \quad \text{o} \quad V_{ij} = x'_i\beta_j \quad [2]$$

Aquí,  $x$  representa un vector de variables que explica el componente determinístico de la satisfacción.<sup>1</sup> En la primera especificación, las variables son inherentes a la alternativa en cuestión, y en la segunda especificación, las variables dependen del individuo en consideración.

<sup>1</sup> Train (2003) argumenta que cualquier función puede ser aproximada mediante una función lineal en los parámetros.

La alternativa que se selecciona es aquella que brinda un mayor grado de satisfacción ( $U$ ), es decir, la probabilidad de seleccionar la  $j$ -ésima alternativa es:

$$\begin{aligned} P = [Y = j] &= P[U_j \geq U_k, k \neq j] \\ &= P[U_k - U_j \leq 0, k \neq j] \\ &= P[e_k - e_j \leq V_j - V_k, k \neq j] \\ &= [\tilde{e}_{kj} \leq -\tilde{V}_{kj}, k \neq j] \end{aligned} \quad [3]$$

De la anterior expresión se encuentra que al asumir que el componente aleatorio de la utilidad sigue una distribución de densidad normal, entonces la probabilidad de elección de la  $j$ -ésima alternativa será una variable aleatoria normal, la cual da origen a los modelos Probit. Si se asume que el componente estocástico de la utilidad es una variable con función de densidad valor extremo tipo uno, entonces la probabilidad de elegir la alternativa en cuestión será una variable aleatoria logística, la cual genera los modelos Logit.<sup>2</sup>

En general, dado que hay  $m$  alternativas, la variable dependiente toma el valor  $j$ , si la  $j$ -ésima alternativa es seleccionada. Luego, la probabilidad que la  $j$ -ésima alternativa sea seleccionada será:

$$p_j = P[Y = j]$$

$$j = 1, 2, \dots, \bar{m} \quad [4]$$

Introduciendo  $m$  variables binarias para cada observación  $Y$ ,

$$y_j = \begin{cases} 1 & \text{si } Y = j \\ 0 & \text{si } Y \neq j \end{cases} \quad [5]$$

Esto implica que para cada observación sobre  $Y$ , exactamente solo una de las  $y_j$  tomará un valor diferente de cero. En consecuencia, la función de densidad implícita en el análisis es multinomial, la cual para una de las observaciones se puede escribir como:

<sup>2</sup> McFadden (1978) propuso una clase general de modelos bajo el supuesto de que la función de distribución conjunta de los componentes aleatorios de la utilidad es valor extremo generalizada, este supuesto genera los modelos Logit Anidados.

$$f(y) = \prod_{j=1}^m p_j^{y_j} \quad [6]$$

Introduciendo los regresores  $x$ , la probabilidad que el individuo  $i$  seleccione la alternativa  $j$ , está dada por:

$$\begin{aligned} p_{ij} &= P[Y_i = j] = F_j(x_i, \beta) \\ j &= 1, 2, \dots, m \\ i &= 1, 2, \dots, N \end{aligned} \quad [7]$$

La función  $F$  debe ser tal que las probabilidades pertenezcan al intervalo  $(0 - 1)$ , y sumen uno sobre  $j$ .

Básicamente, el proceso de estimación se fundamenta en Máxima Verosimilitud, donde se asume independencia sobre los individuos y las alternativas. Esto implica que el logaritmo de la función de Verosimilitud está dado por:

$$LL = \sum_{i=1}^N \sum_{j=1}^m y_{ij} \ln p_{ij} \quad [8]$$

De esta forma, la condición de primer orden para el estimador Máximo Verosímil será:

$$\frac{\partial LL}{\partial \beta} = \sum_{i=1}^N \sum_{j=1}^m \frac{y_{ij}}{p_{ij}} \frac{\partial p_{ij}}{\partial \beta} = 0 \quad [9]$$

En este contexto, se aplica la teoría asintótica usual. Si se asume una correcta especificación para el proceso generador de datos de la probabilidad de elección, se tiene  $E[y_{ij}] = p_{ij}$ , dado que la distribución para  $y_j$  es necesariamente multinomial. Luego,

$$\hat{\beta}_{MV} \stackrel{a}{\sim} N \left( \beta_0, \left( \sum_{i=1}^N \sum_{j=1}^m \frac{1}{p_{ij}} \frac{\partial p_{ij}}{\partial \beta} \frac{\partial p_{ij}}{\partial \beta'} - \frac{\partial^2 p_{ij}}{\partial \beta \partial \beta'} \Big|_{\beta_0} \right)^{-1} \right) \quad [10]$$

Se debe tener presente que el modelo Logit implícitamente asume la hipótesis de Independencia de Alternativas Irrelevantes (IIA), lo cual establece que la ratio de probabilidades entre un par de alternativas pertenecientes al

conjunto de elección permanece inalterada ante modificaciones en dicho conjunto. Este supuesto debe ser contrastado empíricamente. Este supuesto no está implícito en los modelos Probit, pero el limitante de este último modelo es el desgaste computacional por la evaluación de integrales en  $(m - 1)$  variables. En los modelos señalados también están presentes los supuestos de homocedasticidad y no autocorrelación, lo cual implica que se deben utilizar estimaciones robustas de la matriz de varianzas y covarianzas.

## 2. Ejercicio econométrico

Dado el objetivo de caracterizar cuantitativamente la demanda potencial de cupos universitarios por parte de bachilleres egresados de colegios ubicados en las ciudades de Manizales, Pereira, Armenia, Barranquilla, Cartagena, Santa Marta, Montería, Sincelejo, Apartadó, Cúcuta y Bucaramanga, se realizó una encuesta que recopila información socioeconómica y perfil de preferencias en cuanto a universidades y pregrados. Se entrevistaron 459 estudiantes que cursaban el grado 11 durante el año 2007, matriculados en colegios que en general corresponden a un nivel socioeconómico alto y con elevada calificación académica por parte del ICFES.

Dada una población de 2731 individuos, pertenecientes a 65 colegios ubicados en las ciudades objeto de estudio, la muestra recolectada implica un nivel de confianza del 95% y un error máximo del 0,035.

A partir de la información recopilada se procedió a estimar una serie de modelos de elección discreta en los cuales la variable dependiente toma el valor de 1, 2, ...,  $m$ , según la universidad que el estudiante haya seleccionado. Esto, con el objetivo de analizar la probabilidad de elección de las universidades seleccionadas por parte de los estudiantes encuestados en función de características propias de ellos, tales como: estrato socioeconómico y disposición a pagar, y de atributos inherentes a las universidades: número de programas de pregrado que ofrece, ranking nacional y carácter

(pública o privada). Todos los ejercicios realizados se fundamentan en una especificación robusta de la matriz de varianzas y covarianzas tomando en consideración la correlación al interior de las regiones.<sup>3</sup>

En la Tabla 1 se muestran los valores medios de las variables que se utilizaron para el análisis. En primera instancia, se debe aclarar que en el rótulo de *otras* se tomaron en consideración todas aquellas universidades que fueron seleccionadas por los bachilleres pero que solo eran nombradas cuatro veces o menos.

En dicha tabla se evidencia que la Universidad EAFIT es la universidad con mayor probabilidad muestral; le siguen, en su orden, la Universidad Javeriana, Universidad Nacional, Universidad Los Andes y Universidad UPB. En general, las universidades públicas son seleccionadas en promedio por los bachilleres de menor estrato socioeconómico y cuya disposición a pagar es inferior comparada con la que exhiben los aspirantes que escogen las universidades privadas.

**Tabla 1.** Variables utilizadas en los modelos de elección discreta

Encuesta —ciudades colombianas intermedias— 2007

UNIVERSIDAD	Ubicación sede principal	Probabilidad muestral	Estrato medio	Rango disponibilidad media a pagar (Pesos colombianos)	Pregrados ofrecidos	Ranking nacional universitario	Carácter (pública = 0 / privada = 1)
EAFIT	Medellín	11,11%	4	[3'500,001-4'500,000]	17	13	1
ICESI	Cali	0,87%	5	[3'500,001-4'500,000]	16	5	1
UDEA	Medellín	4,79%	3	[2'500,001-3'500,000]	64	116	0
SABANA	Bogotá	1,96%	5	[4'500,001-5'500,000]	16	7	1
ANDES	Bogotá	5,66%	5	[5'500,001-6'500,000]	29	156	1
NORTE	Barranquilla	4,57%	4	[3'500,001-4'500,000]	19	14	1
ROSARIO	Bogotá	1,08%	5	[3'500,001-4'500,000]	23	17	1
EXTERNADO	Bogotá	2,39%	5	[4'500,001-5'500,000]	10	15	1
JAVERIANA	Bogotá	9,36%	4	[3'500,001-4'500,000]	43	89	1
TADEO	Bogotá	1,08%	4	[2'500,001-3'500,000]	19	7	1
NACIONAL	Bogotá	7,62%	3	[2'500,001-3'500,000]	81	176	0
USB	Bogotá	1,30%	3	[1'500,001-2'500,000]	18	-	1
UPB	Medellín	5,66%	4	[2'500,001-3'500,000]	44	11	1
UIS	Bucaramanga	3,05%	3	[1'500,001-2'500,000]	30	28	0
OTRAS		39,43%	4	[3'500,001-4'500,000]			1
TOTAL		100%	4	[3'500,001-4'500,000]	17	13	1

Fuentes: Encuesta desarrollada por el Departamento de Mercadeo Institucional de la Universidad EAFIT en el año 2007 (columnas 3, 4 y 5), páginas de internet de las universidades (columna 6) y La Nota Económica guía de las mejores universidades del año 2006 (columna 7).

<sup>3</sup> Se especificaron cuatro regiones: región 1 (Armenia, Manizales y Pereira), región 2 (Barranquilla, Cartagena, Santa Marta, Sincelejo y Montería), región 3 (Apartadó) y región 4 (Cúcuta y Bucaramanga).

En primera instancia se estimó un modelo Logit Condicional del cual se destaca que las universidades que poseen un mayor número de pregrados tienen una mayor probabilidad de ser seleccionadas por los bachilleres. Se encuentra, además, que la variable ranking es estadísticamente significativa, pero posee signo negativo; es decir, según la información muestral y el modelo propuesto, las universidades que poseen mejor escalafón presentan una menor probabilidad de ser seleccionadas “*ceteris páribus*”. La explicación a este hecho se fundamenta en la alta probabilidad muestral que posee el rótulo de *otras*, el cual está conformado por una serie de pequeñas universidades con un ranking nacional reducido pero que ofrecen programas que no son encontrados en las universidades más reconocidas del país.

El modelo Logit Condicional también muestra que las universidades privadas presentan una mayor probabilidad de ser seleccionadas, lo cual básicamente está asociado a que los colegios objeto de análisis pertenecen a estratos altos (Tabla 2).

En la Tabla 2 también se presentan los resultados de un modelo Logit Multinomial. A partir de este, se encontró que la variable *disposición a pagar* no era estadísticamente significativa en la mayoría de los casos, mientras que el intercepto y el estrato socioeconómico si son estadísticamente significativos. Como era de esperarse, los estudiantes de menor estrato tienden a seleccionar las universidades públicas, en tanto que las universidades privadas presentan una mayor probabilidad de ser seleccionadas por los estudiantes de mayor perfil socioeconómico.

**Tabla 2.** Modelos de elección discreta Logit: probabilidad de elección<sup>4</sup>

Colombia —universidades seleccionadas— 2007

Regresor	Carácterística del regresor	Tipo de modelo		
		$p_{ij} = \frac{\text{Exp}(x'_{ij}\beta)}{\sum_{l=1}^m \text{Exp}(x'_{il}\beta)}$	$p_{ij} = \frac{\text{Exp}(x'_i\beta_j)}{\sum_{l=1}^m \text{Exp}(x'_i\beta_l)}$	$p_{ij} = \frac{\text{Exp}(x'_{ij}\beta + w'_i\gamma_j)}{\sum_{l=1}^m \text{Exp}(x'_{il}\beta + w'_i\gamma_l)}$
		Condicional	Multinomial	Mixto
Número de Pregrados ofrecidos	Específica	0,05 (0,004)	-	0,033 (0,017)
Ranking universitario	Específica	-0,002 (0,001)	-	-0,001 (0,007)
Carácter (pública / privada)	Específica	1,640 (0,161)	-	-1,691 (0,884)
Intercepto	Invariante			
	EAFIT	-	0,000	0,000
	ICESI	-	-6,819 (2,121)	-6,861 (2,095)

<sup>4</sup> EAFIT es el caso base en los modelos multinomial y mixto. En el modelo mixto no se introducen las constantes asociadas a la Universidad Nacional, la Universidad Industrial de Santander y *Otras* debido a la existencia de multicolinealidad. \* No estadísticamente significativos al 0,1 y las desviaciones estándar se encuentran entre paréntesis.

Regresor	Característica del regresor	Tipo de modelo		
		$p_{ij} = \frac{\text{Exp}(x'_{ij}\beta)}{\sum_{l=1}^m \text{Exp}(x'_{il}\beta)}$	$p_{ij} = \frac{\text{Exp}(x'_i\beta_j)}{\sum_{l=1}^m \text{Exp}(x'_i\beta_l)}$	$p_{ij} = \frac{\text{Exp}(x'_{ij}\beta + w'_i\gamma_j)}{\sum_{l=1}^m \text{Exp}(x'_{il}\beta + w'_i\gamma_l)}$
		Condicional	Multinomial	Mixto
Intercepto	UDEA	-	1,609* (1,083)	-0,220* (1,242)
	SABANA	-	-6,502 (1,437)	-6,517 (1,422)
	ANDES	-	-5,775 (1,161)	-4,193 (1,843)
	NORTE	-	0,582* (0,409)	0,562* (1,376)
	ROSARIO	-	-3,826 (1,769)	-3,940 (1,711)
	EXTERNADO	-	-2,851 (0,607)	-2,555 (0,652)
	JAVERIANA	-	-0,764* (1,242)	-0,567* (1,155)
	TADEO	-	-1,272* (2,712)	-1,388* (2,708)
	NACIONAL	-	1,581 (1,157)	-
	USB	-	0,410* (1,279)	0,232* (1,191)
	UPB	-	-0,605* (1,357)	-1,506* (0,945)
	UIS	-	1,889 (1,006)	-
	Otras	-	1,659 (0,852)	-
Estrato aspirante	<b>Invariante</b>			
	EAFIT	-	0,000	0,000
	ICESI	-	0,904 (0,271)	0,896 (0,268)
	UDEA	-	-0,656 (0,278)	-0,664 (0,270)
	SABANA	-	0,997 (0,263)	0,989 (0,263)
	ANDES	-	1,060 (0,211)	1,052 (0,211)
	NORTE	-	-0,377 (0,255)	-0,385 (0,251)



Regresor	Característica del regresor	Tipo de modelo		
		$p_{ij} = \frac{\text{Exp}(x'_{ij}\beta)}{\sum_{l=1}^m \text{Exp}(x'_{il}\beta)}$	$p_{ij} = \frac{\text{Exp}(x'_i\beta_j)}{\sum_{l=1}^m \text{Exp}(x'_i\beta_l)}$	$p_{ij} = \frac{\text{Exp}(x'_{ij}\beta + w'_i\gamma_j)}{\sum_{l=1}^m \text{Exp}(x'_{il}\beta + w'_i\gamma_l)}$
		Condicional	Multinomial	Mixto
Estrato aspirante	ROSARIO	-	0,344* (0,455)	0,336* (0,449)
	EXTERNADO	-	0,303 (0,128)	0,295 (0,122)
	JAVERIANA	-	0,140* (0,271)	0,132* (0,266)
	TADEO	-	-0,264 (0,607)	-0,272 (0,604)
	NACIONAL	-	-0,513 (0,272)	-0,521 (0,267)
	USB	-	-0,686 (0,186)	-0,694 (0,185)
	UPB	-	-0,016* (0,242)	-0,024* (0,248)
	UIS	-	-0,881 (0,188)	-0,889 (0,185)
	Otras	-	-0,096* (0,172)	-0,106* (0,175)
<b>Log L</b>		-1040,040	-925,320	-926,212

Fuente. Encuesta Departamento de Planeación Institucional Universidad EAFIT. Cálculos del autor

El último modelo que fue realizado es el Logit Mixto, que combina atributos inherentes a las universidades y características propias de los individuos. Bajo este modelo se observó que la variable número de pregrados tiene un efecto positivo en la probabilidad de selección, en tanto que las universidades con mayor ranking y de carácter privado tienen una menor probabilidad. Al respecto del perfil socioeconómico, de nuevo se encuentra que los bachilleres con mejor posición económica tienden a seleccionar las universidades privadas.

Al comparar el logaritmo de la función de verosimilitud se encuentra que el modelo a seleccionar debe ser el Logit Multinomial. Además,

se debe tener presente que a los tres modelos propuestos se les realizó la prueba de Hausman (*Hausman test*) para corroborar la hipótesis de Independencia de Alternativas Irrelevantes (IIA); en todos los casos se encontró que los modelos cumplen dicha propiedad.

A partir de los modelos propuestos se realizaron una serie de ejercicios de estática comparativa con el ánimo de ilustrar mejor ciertos rasgos característicos de la población encuestada.

En primera instancia se calcularon las sensibilidades de las probabilidades estimadas a partir del modelo Logit Condicional cuando varían marginalmente los betas de los diversos regresores (Tabla 3).

**Tabla 3.** Cambio marginal en los betas asociados a *programas, ranking y carácter*

Modelo logit condicional —universidades seleccionadas— 2007

UNIVERSIDAD	Sensibilidad Número de Programas	Sensibilidad Ranking	Sensibilidad Carácter
EAFIT	-0,95	-0,9	0,00
ICESI	-0,95	-1,12	0,00
UDEA	0,78	3,55	-0,04
SABANA	-0,95	-1,05	0,00
ANDES	-0,71	4,34	0,01
NORTE	-0,98	-0,96	0,01
ROSARIO	-1,02	-1,04	0,01
EXTERNADO	-0,82	-0,59	0,00
JAVERIANA	-0,44	4,35	0,01
TADEO	-1	-1,22	0,01
NACIONAL	3,18	12,81	-0,08
USB	-1	-1,42	0,01
UPB	-0,45	-3,76	0,02
UIS	-0,2	-0,16	-0,01
OTRAS	5,49	-12,84	0,05

Cálculo del autor

Se observa que cuando los bachilleres encuestados se vuelven más sensibles al número de programas de las universidades, entonces aumenta la probabilidad de seleccionar la Universidad Nacional, la Universidad de Antioquía y las que se hallan bajo el rótulo de *Otras*, en detrimento de las demás. Por otra parte, cuando se vuelve más importante el ranking aumenta la probabilidad de elegir a la Universidad Nacional, la Universidad de Antioquía, la Universidad Los Andes y la Universidad Javeriana. Finalmente, se encuentra que el carácter tiene un efecto bastante reducido sobre los cambios en las probabilidades de selección.

En la Tabla 4 se ilustran los cambios en la probabilidad de elección de las universidades objeto de estudio ante la creación de un programa adicional por parte de alguna de aquellas.

Las universidades que incrementan de forma más significativa la probabilidad de ser seleccionadas son: la Universidad Javeriana, la Universidad Pontificia Bolivariana (UPB) y la Universidad

Nacional. De otro lado, las universidades que exhiben un incremento menor en la probabilidad de ser seleccionadas son la Universidad ICESI y la Universidad Externado.

En la Tabla 5 se aprecian los resultados que se obtienen a partir del modelo Logit Multinomial. Específicamente, los cambios marginales en la probabilidad, asociados a la variable estrato, y la elasticidad probabilística estrato de la demanda, evaluados en el estrato medio. Como se evidencia, el mayor incremento en la probabilidad de ser seleccionada ante un mejoramiento en la posición socioeconómica de los bachilleres, se encuentra para la Universidad de Los Andes; esto se logra disminuyendo la probabilidad de seleccionar a las universidades públicas. También se encuentra que la mayor elasticidad probabilística se encuentra también para la Universidad de Los Andes, específicamente un incremento del 1% en el estrato socioeconómico implica un incremento del 4,66% en la probabilidad de que esta sea la *seleccionada*.

Tabla 4. Cambio en la probabilidad de elección ante la creación de un programa adicional

Modelo logit condicional —universidades seleccionadas— 2007

Universidad	EAFIT	Icesi	UdeA	Sabana	Andes	Norte	Rosario	Externado	Javeriana	Tadeo	Nacional	USB	UPB	UIS	Otras
EAFIT	0,149	-0,004	-0,007	-0,004	-0,008	-0,005	-0,006	-0,003	-0,017	-0,005	-0,016	-0,005	-0,018	-0,001	-0,049
ICESI	-0,004	0,142	-0,006	-0,004	-0,008	-0,005	-0,006	-0,003	-0,016	-0,005	-0,015	-0,005	-0,017	-0,001	-0,046
UDEA	-0,007	0,006	0,216	-0,006	-0,012	-0,007	-0,009	-0,005	-0,025	-0,007	-0,023	-0,007	-0,026	-0,002	-0,072
SABANA	-0,004	-0,004	-0,006	0,142	-0,008	-0,005	-0,006	-0,003	-0,016	-0,005	-0,015	-0,005	-0,017	-0,001	-0,046
ANDES	-0,008	-0,008	-0,012	-0,008	0,267	-0,009	-0,011	-0,006	-0,031	-0,009	-0,029	-0,009	-0,009	-0,002	-0,089
NORTE	-0,005	-0,005	-0,007	-0,005	-0,009	0,164	-0,007	-0,004	-0,019	-0,006	-0,017	-0,005	-0,02	-0,001	-0,054
ROSARIO	-0,006	-0,006	-0,009	-0,006	-0,011	-0,007	0,2	-0,004	-0,023	-0,007	-0,021	-0,006	-0,024	-0,002	-0,066
EXTERNADO	-0,003	-0,003	-0,005	-0,003	-0,006	-0,004	-0,004	0,106	-0,012	-0,004	-0,011	-0,003	-0,012	-0,001	-0,034
JAVERIANA	-0,017	-0,016	-0,025	-0,016	-0,031	-0,019	-0,023	-0,012	0,509	-0,019	-0,059	-0,018	-0,067	-0,004	-0,181
TADEO	-0,005	-0,005	-0,007	-0,005	-0,009	-0,006	-0,007	-0,004	-0,019	0,164	-0,017	-0,005	-0,02	-0,001	-0,054
NACIONAL	-0,016	-0,015	-0,023	-0,015	-0,015	-0,018	-0,022	-0,011	-0,059	-0,018	0,479	-0,017	-0,062	-0,004	-0,169
USB	-0,005	-0,005	-0,007	-0,005	-0,009	-0,005	-0,006	-0,003	-0,018	-0,005	-0,017	0,157	-0,019	-0,001	-0,051
UPB	-0,018	-0,017	-0,026	-0,017	-0,033	-0,02	-0,024	-0,013	-0,067	-0,02	-0,062	-0,062	0,531	-0,005	-0,005
UIS	-0,001	-0,001	-0,002	-0,001	-0,002	-0,001	-0,002	-0,001	-0,004	-0,001	-0,004	-0,001	-0,005	0,04	-0,013
Otras	-0,049	-0,047	-0,047	-0,047	-0,09	-0,055	-0,067	-0,035	-0,183	-0,183	-0,171	-0,052	-0,193	-0,013	1,114

Cálculo del autor

**Tabla 5.** Cambio marginal y elasticidad estrato

Modelo logit multinomial —universidades seleccionadas— 2007

UNIVERSIDAD	Cambio marginal	Elasticidad estrato
EAFIT	1,03%* (0,017)	0,347* (0,624)
ICESI	0,52%* (0,006)	4,03 (0,537)
UDEA	-2,39% (0,014)	-2,326 (1,076)
SABANA	1,15% (0,002)	4,41 (0,553)
ANDES	3,23% (0,004)	4,664 (0,316)
NORTE	-1,37%* (0,010)	-1,189 (0,477)
ROSARIO	0,46%* (0,005)	1,749* (1,558)
EXTERNADO	0,94% (0,001)	1,582 (0,371)
JAVERIANA	2,26%* (0,014)	0,919* (0,593)
TADEO	-0,21%* (0,007)	-0,731* (2,389)
NACIONAL	-3,13% (0,008)	-1,743 (0,693)
USB	-0,67%* (0,005)	-2,447 (0,331)
UPB	0,43%* (0,010)	0,279* (0,747)
UIS	-1,77% (0,003)	-3,243 (0,218)
Otras	-0,48%* (0,033)	-0,046* (0,320)

\*No estadísticamente significativos al 0,1 y las desviaciones estándar se encuentran entre paréntesis. Cálculo del autor

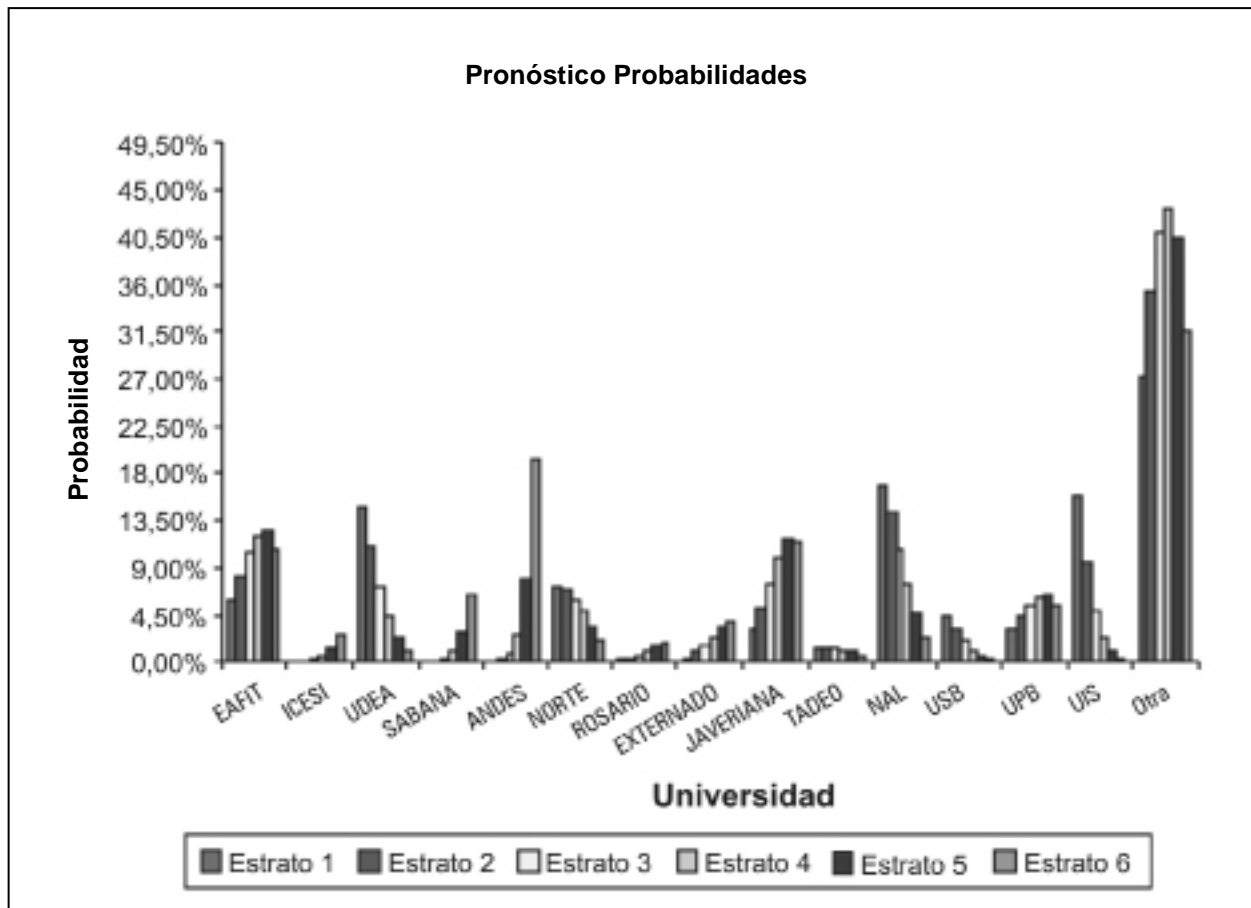
Finalmente, en la Figura 1 está ilustrada la probabilidad de selección para una universidad, obtenida a partir del modelo Logit Multinomial. De estos resultados se infiere que la probabilidad de selección de las universidades públicas es función

decreciente del estrato, mientras que en general, la probabilidad de elección de las universidades privadas es función creciente con relación al estrato social.



**Figura 1.** Probabilidad estimada en función del estrato

Modelo Logit Multinomial —universidades seleccionadas— 2007



Fuente: Elaboración propia.

## Conclusiones

A partir de los ejercicios realizados se encuentra que el número de pregrados es una variable estadísticamente significativa y con un impacto positivo para determinar la probabilidad de elección de las universidades objeto de estudio. Entre tanto, el ranking de las universidades colombianas ejerce un efecto negativo sobre la probabilidad de elección. Al respecto de la discriminación entre pública y privada, los resultados son ambiguos y dependen de la especificación del modelo, aunque en general el cambio marginal asociado a esta variable es bastante reducido.

Las universidades que ganan mayor participación ante la generación de un programa adicional son la Universidad Nacional, la Universidad Javeriana y la Universidad Pontificia Bolivariana (UPB), mientras que la sensibilidad de los encuestados al ranking universitario los inclina, preferentemente, hacia la Universidad de Los Andes, la Universidad Nacional y la Universidad de Antioquia.

Uno de los resultados más contundentes de la investigación se fundamenta en el hecho de que, pese a la corta edad de los bachilleres encuestados, en promedio 16 años, son bastante conscientes de la situación socioeconómica que evidencian en sus hogares. Se manifiesta en los modelos, al observar que los estudiantes de menores estratos tienden a seleccionar las universidades públicas, las cuales, por obvias razones, presentan matrículas a precios más reducidos. De igual forma, los bachilleres de posición socioeconómica más elevada tienden a seleccionar las universidades privadas.

Finalmente, se hace necesario considerar las discrepancias en las probabilidades muestrales de la presente investigación con lo que evidentemente se corrobora cuando se observan las inscripciones en las diferentes universidades del país. Esta diferencia se explica teóricamente por los conceptos de preferencia declarada y preferencia revelada.

En primera instancia, a la hora de realizar análisis de mercado a partir de encuestas, los agentes entrevistados declaran ciertas preferencias, lo cual no implica que lo señalado sea evidentemente lo revelado, puesto que al responder las indagaciones hay una serie de factores que afectan la decisión de los agentes y que no son tomados en consideración, además de ciertos aspectos del entorno específico que pueden sesgar la respuesta de los interrogados. Este tipo de fenómenos ocasionan mayores discrepancias cuando la decisión objeto de análisis involucra una cuantía significativa de recursos, específicamente tiempo y dinero. En este orden de ideas, una posible extensión a la presente investigación es realizar el análisis de selección de universidades a partir de las preferencias reveladas por los bachilleres.

## Bibliografía

Carla Sa, Raymond Florax y Piet Riedvelt. (2004). *Determinants of the regional demand for higher education: a gravity model approach*. Tinbergen Institute Discussion paper. [En línea] Disponible en: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=384861](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=384861).

Gallet, Craig. (2007). "A comparative analysis of the demand for higher education: results from a meta-analysis of elasticities", *Economics Bulletin*, 7 (9). Nashville, pp. 1-14.

McFadden, Daniel L. (1978). "Modelling the choice of residential location", A. Karlquist et al. (eds.): *Spatial interaction theory and planning models*. Amsterdam/New York/North – Holland, pp. 75–96.

Modrego Rico, Aurelia. (1988). "Demanda de educación: resultados de la estimación de un modelo de demanda de educación superior para la provincia de Vizcaya", *Revista vasca de economía*, 12. Biscaya, pp. 87–96.

Salas Velasco, Manuel y Martín-Cobos Puebla, Manuel. (2006). "La demanda de educación superior: un análisis microeconómico con datos de corte transversal", *Revista de Educación*, 339. Madrid, pp. 637–660.

Train, Kenneth E. (2003). *Discrete choice methods with simulations*. Cambridge: University Press, pp. 395–402.