

**LA DEMANDA POR PROGRAMAS EN LA EDUCACION
SUPERIOR EN COLOMBIA : CASODEL PROGRAMA DE
ECONOMIA EN LA UNIVERSIDAD EAFIT**

SEBASTIAN OSORIO NEGRETE

**Universidad Eafit
Facultad de ciencias económicas
Medellín, Colombia
2008**

**LA DEMANDA POR PROGRAMAS EN LA EDUCACION
SUPERIOR EN COLOMBIA : CASO DEL PROGRAMA DE
ECONOMIA EN LA UNIVERSIDAD EAFIT**

SEBASTIAN OSORIO NEGRETE

**Trabajo realizado para optar por el titulo de Economista de la Universidad
EAFIT**

Febrero 2008

Asesor temático

GUSTAVO LOPEZ

**Universidad Eafit
Facultad de ciencias económicas
Medellín, Colombia**

TABLA DE CONTENIDO

INTRODUCCIÓN	5
II. MARCO TEORICO.	7
1. EL CONSUMIDOR Y SUS PREFERENCIAS	7
2. ANTECEDENTES: EVIDENCIA EMPÍRICA	9
III. MODELO ECONOMETRICO	11
1. COMPORTAMIENTO DE LAS VARIABLES.....	11
2. PRUEBAS SOBRE EL MODELO LOGARÍTMICO	15
IV. DEMANDA DE EDUCACION SUPERIOR SEGÚN LOS CONDICIONAMIENTOS DE LOS ESTUDIANTES EN LA ELECCION DE LA CARRERA DE ECONOMIA.....	18
1. EL «MODELO LOGIT BINOMIAL» APLICADO A LAS ELECCIONES EDUCATIVAS.....	19
2. UN «MODELO LOGIT MULTINOMIAL» DE ELECCIÓN DE TITULACIÓN UNIVERSITARIA ...	21
2.1 <i>Objetivo del modelo logístico</i>	22
2.2 <i>Variables propuestas bajo la metodología Logit</i>	23
V. ANALISIS Y CONCLUSIONES	25
VI. BIBLIOGRAFIA	28
VII. ANEXOS.....	31
ANEXO 1 MODELO LOGARÍTMICO CON UNA SOLA VARIABLE	31
ANEXO 2 MODELO LOGARÍTMICO CON DOS VARIABLES REGRESORAS	31
ANEXO 3 DATOS TOMADOS EN EL MODELO LOGÍSTICO	32
ANEXO 4 TRANSFORMACIÓN DE VARIABLES Y MODELOS LOGARÍTMICOS	32
ANEXO 5 MODELOS DE PROBABILIDAD: MODELO DE PROBABILIDAD LOGÍSTICO (LOGIT)	34

TABLA DE FIGURAS

FIGURA 1. Numero de estudiantes inscritos, admitidos y atriculados.....	6
FIGURA 2. Número de estudiantes matriculados y valor del semestre ..	12
FIGURA 3. Gráfica del modelo logarítmico ajustado	14
FIGURA 4. Autocorrelación de los residuales del modelo logarítmico... 	16
FIGURA 5. Correlación de los residuos del modelo logarítmico.....	16
FIGURA 7. Histograma y prueba de normalidad de los residuos.....	17

INTRODUCCIÓN

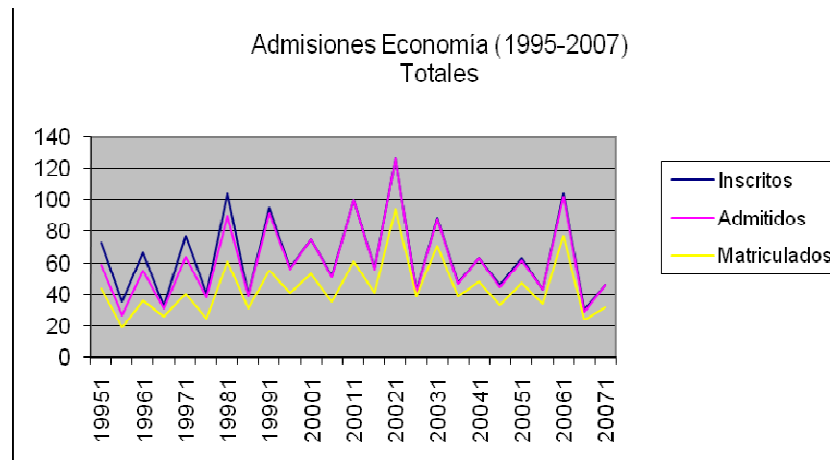
La presente investigación se realiza con el ánimo de identificar la demanda de los estudiantes a la carrera de economía e identificar las posibles causas de la irregularidad de la misma a partir de una estimación econométrica ya que observando la información brindada por el departamento se ve un comportamiento variable sin una tendencia definida tal como lo muestra el gráfico 1. La importancia de este análisis es la contribución a la productividad académica de estudios relacionados con la competitividad educativa y el desarrollo del sistema de educación superior. Como resultado de las cambiantes condiciones en que debe desenvolverse, la educación superior (ES) latinoamericana se halla frente a un conjunto de nuevas demandas¹ Este planteamiento constituye, en verdad, el eje central de los más recientes debates y propuestas sobre la reforma de la ES en el mundo y, también, en nuestra región². Como ha dicho Burton Clark (1998), los sistemas nacionales de ES (SNES) se hallan hoy puestos en una encrucijada: no pueden ni retornar a un anterior estado de equilibrio ni alcanzar uno nuevo. “Como actores principales dentro de esos sistemas, las universidades públicas y privadas han ingresado a una época de turbulencias para la cual no se prevé término. La actual encrucijada tiene su origen en un simple hecho: las demandas ejercidas sobre las universidades superan su capacidad de respuesta”³

¹ Véase Brunner, 1990

² Como se ha visto con los nuevos procesos de acreditación y homogenización de los sistemas educativos a nivel global

³ Clark, B. R. (1998) *Creating Entrepreneurial Universities. Organizational Pathways of transformation*; IAU Press, Pergamon, Great Britain.

Gráfico 1: Numero de estudiantes inscritos, admitidos y matriculados en el programa de Economía, universidad Eafit



Fuente: Eafit, 2007

Tal como lo muestra el gráfico 1 no se puede ver una tendencia clara sobre el ingreso de estudiantes a la carrera de economía en la universidad y en general el crecimiento poblacional a través del tiempo no es constante. Basados en las matriculas de cada semestre en el periodo de 2001-1 2006-1 Se busca identificar la demanda de los estudiantes por el programa de economía dentro de la Universidad Eafit teniendo en cuenta el costo de la misma y analizar a qué se deben las posibles variaciones de ésta.

El trabajo se desarrollará en varias partes, la primera será un breve marco teórico enfocado a la teoría de la demanda convencional y los determinantes de la misma, seguido a esto tendremos algunas evidencias o hallazgos en la educación superior para el caso de Colombia, luego se desarrollará un análisis al modelo econométrico incluyendo el análisis de la elasticidad precio en la demanda por cupos en Economía y por último las pruebas para el modelo econométrico. Para

realizar este análisis se utilizaron metodologías econométricas de mínimos cuadrados, apoyados por el software econométrico Eviews 5.1.

En última instancia se presentarán dos modelos de respuesta discreta utilizados como modelos de elección con respecto a la carrera; la poca evidencia que se ha encontrado con respecto a estos modelos limita el análisis y aplicación de este modelo al caso de los estudiantes de Economía de la universidad EAFIT, sin embargo se plantea el modelo y se presentan las posibles variables que explicarían el comportamiento de la demanda hacia el programa de Economía. Para finalizar se muestran los principales resultados y conclusiones acerca de las aplicaciones utilizando herramientas econométricas y las propuestas que se deben tener en cuenta para posteriores investigaciones.

II. MARCO TEORICO.

1. El consumidor y sus preferencias

La teoría del consumidor asume que el individuo elige su canasta de consumo de acuerdo a las alternativas de bienes y servicios disponibles, de manera que la satisfacción derivada de su consumo sea la máxima posible, sujeta a la restricción presupuestaria del individuo. Lo anterior implica que el consumidor está consciente de las alternativas que enfrenta y es capaz de evaluarlas, es decir, se comporta racionalmente. La información relacionada con la satisfacción del consumidor respecto a todos los bienes demandados, está contenida en la función de utilidad, la cual se puede representar como:

$$U = f(q_1, \dots, q_n) \quad (1)$$

Donde el consumidor decide comprar una combinación de bienes (q_n) que le otorgue el mayor nivel de satisfacción y, por lo tanto, maximice su utilidad (U). A partir de la función de utilidad se pueden obtener las funciones de demanda no compensada (o demanda Marshalliana) y compensada (o demanda Hicksiana), donde la primera incluye tanto el efecto precio como el efecto ingreso, mientras que la segunda sólo incluye el efecto precio (Varian, 1994). La demanda de un bien depende de diversas variables que pueden representarse por la expresión (2).

$$Q^d = Q^d (P, I, P_s, P_c, G) \quad (2)$$

Donde la variable dependiente Q^d representa la demanda del bien, y las variables explicativas son el precio (P), el ingreso (I), el precio del bien sustituto (P_s), el precio del bien complementario (P_c), y los gustos y preferencias (G).

A partir de la expresión (2), pueden obtenerse las elasticidades de demanda como:

$$\eta_{Q^d, VE_i} = \frac{\partial Q^d}{\partial VE_i} \frac{VE_i}{Q^d} \quad (3)$$

Donde η_{Q^d, VE_i} mide el grado de sensibilidad de la variable dependiente (Q^d) frente a un cambio porcentual en alguna de las variables explicativas (VE_i). La demanda puede clasificarse como inelástica al precio cuando $-1 < \eta_{Q^d, p}$, de elasticidad unitaria cuando $\eta_{Q^d, p} = -1$, y elástica cuando $\eta_{Q^d, p} < -1$. Por otra parte, el bien Q puede considerarse como bien inferior cuando $\eta_{Q^d, i} < 0$, bien normal cuando $0 < \eta_{Q^d, i} < 1$, y bien superior cuando $\eta_{Q^d, i} > 1$ (Henderson y Quandt, 1980)

2. Antecedentes: evidencia empírica

Los estudios empíricos hechos en España sobre demanda de Educación Superior se han basado, tradicionalmente, en dos metodologías distintas: la estructural y la analítica. La primera centra su atención en los flujos de estudiantes a través del sistema educativo con independencia de las causas que generan dichos flujos. Mientras que, por su parte, los métodos analíticos explican la demanda mediante la valoración de los factores que afectan a dicho sistema (Mora, 1989). La demanda por programas en la educación superior en Colombia puede realizarse desde varios puntos de vista ya sea desde la perspectiva de las preferencias individuales o desde la perspectiva de la oferta de plazas en el sistema educativo ya sea en niveles superiores de estudio o de niveles de pregrado en primeros semestres. En nuestro caso se puede considerar el punto de vista de la matrícula de alumnos en primer semestre por primera vez (los primíparos) o desde la estructura de graduados en las diversas áreas o programas. En este trabajo se tomará la Matrícula para primer semestre en primer curso teniendo en cuenta la estructura de la demanda de cupos para educación superior en el país, desde la perspectiva de la matrícula en las instituciones oficiales.

Con datos oficiales del ICFES para el año 2000, indica que las carreras tecnológicas son el área a la cual más estudiantes se matriculan en primer semestre y por primera vez, representando el 26% de la matrícula de los primíparos⁴; la segunda área de demanda son: ingeniería y arquitectura, que representan el 22.5% de esta matrícula. El tercer lugar de demanda es para las ciencias sociales (economía, administración, derecho, contaduría) con el 17.6%. Le siguen las ciencias de la educación con una demanda del 14.9% y las ciencias

⁴ Datos tomados de la página del ICFES, 2007

de la salud con el 10%. Adicionalmente se observa que las carreras técnicas profesionales tienen muy baja demanda en términos de matrícula, pues tan solo representa el 1.3% del total.

Graduados del sector oficial en el país Desde la perspectiva de los graduados del sector oficial en el país, el área de las tecnologías es la que mayor representación tiene, al representar el 30% del total. El segundo lugar, para graduados, lo ocupa el área de ciencias de la educación (25.8%), el tercer lugar es para las ciencias sociales (17.4%), el cuarto lugar es para las técnicas profesionales (10.9%) y en quinto lugar de la estructura para los graduados de ingeniería y arquitectura con el 7%. Matrícula para primer semestre en primer curso – sector no oficial en el país La estructura que presenta la matrícula para primíparos del sector privado en el país, por áreas de conocimiento es diferente. Las ciencias sociales tienen el primer lugar en la demanda con relación a la oferta; equivale al 32.6% de la matrícula. El segundo lugar es para el área de las ingenierías y arquitecturas (24.7%). El tercer lugar es para el área de la salud (13%). Las tecnologías tienen un peso del 9.7% y las técnicas profesionales un 9.7%.

Graduados del sector no oficial en el país Desde la perspectiva de los graduados del sector no oficial en el país, existe coincidencia entre la estructura de la matrícula y de graduados. Las ciencias sociales tienen el primer lugar (40.8%). El segundo lugar es para ingenierías, arquitecturas y afines (aunque como el sector oficial, se gradúan menos de los que se matriculan, la diferencia es bastante menor, de 25% de matrícula, se gradúa el 17%) (16.8%). El tercer lugar es para las tecnologías (12.9%). Ciencias de la salud (12.2%), el cuarto. Las ciencias de la educación representan el 8.2% y las técnicas profesionales el 7.3%. A diferencia del comportamiento en el sector oficial, la comparación entre matriculados y graduados en el área de ingeniería establece una diferencia menor, aunque también hay menos graduados que matriculados. En las demás áreas, el

comportamiento es bastante homogéneo, es decir, una proporción similar a la que ingresa al programa se gradúa.

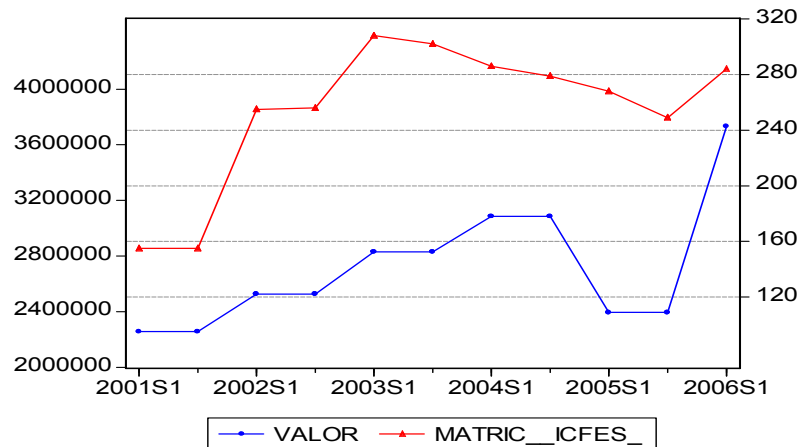
III. MODELO ECONOMETRICO

1. Comportamiento de las variables

El desarrollo de un modelo de demanda puede realizarse como mínimo con dos variables, el precio y las unidades económicas demandadas; además de estas dos variables, en el estudio de demanda puede incluirse variables como ingreso, precios de sustitutos o inclusive variables dicotómicas que diferencien productos (Variant, 1994).En el caso del estudio presentado en este informe se tomará solo el precio o costo del semestre y la cantidad de estudiantes matriculados para el semestre correspondiente; luego se incluirá el PIB per cápita⁵ como indicador del ingreso durante el periodo y así poder medir la elasticidad ingreso dentro del modelo. Como en toda función de demanda, es de esperarse un comportamiento inverso entre las dos variables, es decir que la demanda caiga mientras más aumente el precio (ley de demanda) y una relación positiva con respecto al ingreso ya que las personas prefieren mas el consumo del bien, en este caso mas matriculas siempre y cuando aumenten sus ingresos. En el grafico 2 se observa la trayectoria de las dos series durante el periodo 2001 a 2006. Es de anotar que el valor del primer semestre es el mismo que para el segundo semestre del mismo año ya que el costo de la matricula no varia durante todo el año.

⁵ El PIB per cápita es tomado de la base del Banco de la Republica y se encuentra en millones de pesos a precios constantes de 1994, además se aclara que el valor o precio de la matricula esta también a precios constantes, un buen ejercicio mas adelante sería deflactar las variables

Gráfico 2: Número de estudiantes matriculados y valor del semestre para el periodo 2001-2006



Fuente: ICFES- EAFIT,2007

En el eje derecho del grafico se tiene el numero de estudiantes matriculados para cada semestre, mientras que el eje izquierdo se encuentra la escala en pesos del valor de la matricula a precios constantes, este valor fue tomado de la base de datos de la universidad EAFIT, mientras que el numero de matriculados se extrajo de la base de datos del ICFES.

En el caso de los modelos de demanda, la revisión de la literatura indica que en la mayoría de las estimaciones se utiliza el método de MCO, aunque existen otras técnicas alternativas de estimación, tales como el método de máxima verosimilitud (MV) o método de grandes muestras (Gujarati, 2000). Las variables que normalmente explican la demanda son el precio del bien, el ingreso, el precio de los sustitutos, la tendencia y una variable dicótoma de cambio estructural. En las estimaciones de demanda internacional, los estudios como los realizados por MCFADDEN (1974) incluyen como variable explicativa el tipo de cambio. La

mayoría de los modelos incorporan las variables explicativas con periodos de rezago.

El modelo a trabajar en este estudio esta representado como:

$$\text{LnMatr}_t = \alpha + \beta \text{LnCosto}_t + \mu_t \quad (4)$$

En donde LN es el logaritmo neperiano o natural, la variable Matr es el numero de estudiantes matriculados cada semestre y Costo es el valor de la matricula. Dado este modelo, la elasticidad precio demanda corresponde al parámetro β y se puede observar que es una elasticidad constante para todo el periodo. Los resultados presentados a continuación se encuentran en el anexo 1:

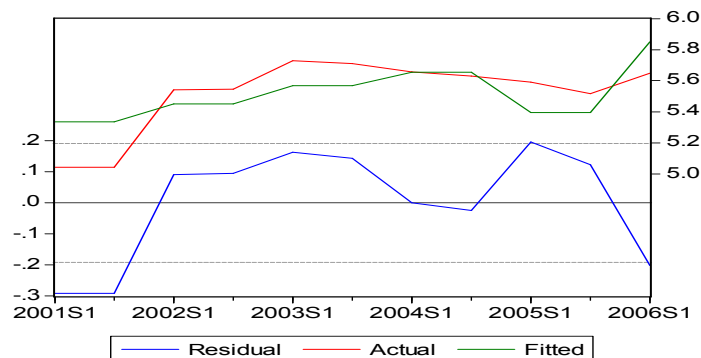
$$\begin{aligned} \text{Ln } \hat{\text{Matr}}_t &= -9.65 + 1.02 \text{LnCosto}_t \\ &\quad (-1.69) \quad (2.66) \end{aligned} \quad (5)$$

$$R^2 = 44.12$$

En el modelo se observa un nivel de ajuste del 44.12 % lo que indica que el costo de la matricula explica la demanda por el programa de economía en un 44%, sin embargo faltaría un 56% que estaría siendo explicado por otras variables como por ejemplo preferencias de los individuos, efectos de las otras universidades en el mismo programa académico entre otras. Obsérvese que la elasticidad precio esta dada por el segundo coeficiente y es de alrededor de 1.02, por lo tanto estamos hablando de una elasticidad unitaria, entre paréntesis se encuentran los estadísticos t de cada parámetro estimado, la interpretación de la constante no tiene un significado económico coherente y que no se puede hablar de demandas negativas independientemente del precio de las matriculas.

Al hallarse un 44% de ajuste, se observa en el grafico 3 unas brechas muy grandes entre los valores ajustados y el verdadero comportamiento de las matriculas, estos resultados pueden arrojar unos residuos que no cumplan a cabalidad con los supuestos clásicos. De otro lado la elasticidad precio es de 1.02, casi una elasticidad unitaria, además se observa que este parámetro es positivo indicando aumentos de la demanda frente a aumentos del precio esto contradice la racionalidad económica, aunque puede encontrar explicacion en que para los demandates estas variables (ingreso y precio) no son tan determinantes a la hora de su eleccion universitaria y consideran otras variables mas importantes que estas las cuales trataremos de revelar posteriormente

Gráfico 3: Gráfica del modelo logarítmico ajustado de demanda y su trayectoria durante el periodo 2001- 2006



El cumplimiento de los supuestos de homoscedasticidad y de normalidad de los residuos, garantizan un modelo estable en el tiempo y permiten en u momento dado ampliar el estudio tomando mas años o incluir mas variables. A continuación incluiremos la variable ingreso para mejorar el ajuste y poder observar el comportamiento de la elasticidad ingreso dentro de la función de demanda.

$$LnMatr_t = \alpha + \beta_1 LnCosto_t + \beta_2 LnPIB_t + \mu_t \quad (6)$$

Para este modelo, el segundo coeficiente β es el estimador de la elasticidad ingreso y es de esperarse un valor positivo por la teoría de demanda en donde es racional pensar que si aumenta el poder adquisitivo, la demanda aumenta. El modelo ajustado se encuentra en el anexo 2 y los resultados son:

$$\begin{aligned} \ln \hat{M}atr_t = & -9.65 + 1.02 \ln Costo_t + 0.43 \ln PIB_t \\ & (-0.67) \quad (1.89) \quad (0.24) \end{aligned} \tag{7}$$

$$R^2 = 44.5$$

En este nuevo modelo el coeficiente de determinación no mejoró mucho y adicional a esto se observa que el ingreso es una variable poco explicativa por lo tanto la inclusión del PIB no es tan importante dentro del modelo. Obsérvese que la elasticidad precio sigue oscilando alrededor de uno lo que indica nuevamente que frente a un 1 % en el aumento del precio de la matrícula, la demanda aumenta en un 1 %, es decir que hay una relación casi a uno a uno entre la demanda y el precio de la matrícula

2. Pruebas sobre el modelo logarítmico

Al nuevo modelo se le desarrollará una batería de pruebas teniendo en cuenta las tres variables. Para probar la validez de los supuestos del modelo lineal se realizaron pruebas estadísticas de autocorrelación, heterocedasticidad, multicolinealidad y normalidad de los errores. Según Judge *et al.* (1988), para poder hacer inferencias sobre los verdaderos estimadores debe cumplirse que las variables no consideradas en la regresión estén distribuidas aleatoriamente; la suma de estas variables (error de un modelo de regresión) debe tener distribución normal, con media cero, covarianza cero y varianza constante. Greene (2000), sin embargo, sostiene que aunque la normalidad no es necesaria para obtener la

mayoría de los resultados que se utilizan en el análisis de la regresión múltiple, ésta permite obtener algunos resultados estadísticos exactos. En esta investigación se realizaron pruebas de normalidad para inferencias y conclusiones.

Tabla 1: Autocorrelación de los residuales del modelo logarítmico

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.230963	Probability	0.356489
Obs*R-squared	3.200357	Probability	0.201861

A pesar de que el estadístico Durbin Watson se encuentre muy cercano a cero (anexo 2) la prueba general de autocorrelación que se ve en la tabla 1 muestra que la hipótesis nula de incorrelación no es rechazada, por lo tanto se puede afirmar que los residuos no están autocorrelacionados. Hay que recordar que la escasez de datos le quitan potencia a este tipo de pruebas además no aseguran un resultado fiable (Enders, 1995).

Para comprobar el hecho de que no hay correlación de los residuos, se puede observar en la tabla 2 que el correlograma de estos residuales del nuevo modelo no tienen correlaciones fuertes y además posee al parecer un comportamiento ruido blanco⁶

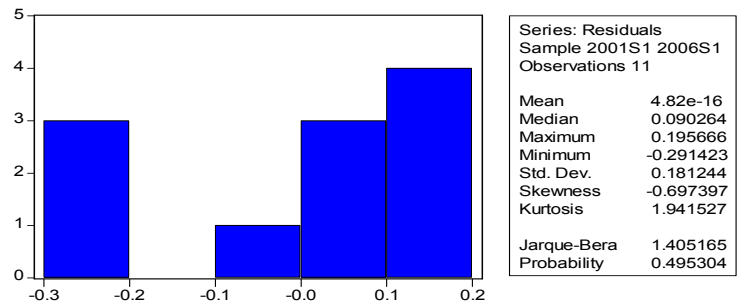
⁶ Gujarati, 2004. este término es utilizado para denominar las perturbaciones que no están correlacionadas y que se distribuyen normales de media cero y varianza constante.

Tabla 2: correlación de los residuos del modelo logarítmico

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			0.377	0.377	2.0281	0.154
2			-0.082	-0.261	2.1344	0.344
3			-0.207	-0.088	2.9003	0.407
4			-0.288	-0.226	4.5951	0.331
5			-0.136	0.025	5.0374	0.411
6			0.000	-0.056	5.0374	0.539
7			0.000	-0.079	5.0374	0.655
8			0.000	-0.051	5.0374	0.754
9			0.000	-0.028	5.0374	0.831

La ultima prueba que se realizará es para comprobar que definitivamente los residuos del modelo se distribuyen normales. Para ello se puede observar que en el grafico 4 aparece el estadístico Jarque Bera que identifica si una serie sigue la distribución normal o no. En síntesis se encuentra un estadístico Jarque Bera de 1.4, considerablemente pequeño por lo tanto no se va a rechazar la normalidad de los residuos.

Gráfico 4. Histograma y prueba de normalidad de los residuos



La realización de las anteriores pruebas garantizan que el modelo sea aplicable bajo otros criterios y otros periodos de tiempo, sin embargo hay que recordar que solo 6 años no es suficiente para este tipo de modelos logarítmicos, aunque el modelo sea aplicable o no, se debe tener un muestreo mas grande que mejore la potencia de las pruebas sobre el mismo.

IV. DEMANDA DE EDUCACION SUPERIOR SEGÚN LOS CONDICIONAMIENTOS DE LOS ESTUDIANTES EN LA ELECCION DE LA CARRERA DE ECONOMIA

Un segundo modelo que se puede ofrecer para identificar la demanda de los estudiantes por la carrera de economía, se basa en los criterios de elección por estudiar o no en este programa, este modelo resumido en un modelo probabilístico debe hacerse con una cohorte de nuevos estudiantes en un año determinado, de allí se deben realizar bajo ciertos criterios o variables el análisis del porcentaje de alumnos que ingresan a la carrera como primera opción, teniendo en cuenta las limitaciones que tienen los individuos en su elección según los orígenes sociales, escolares y condiciones objetivas.

Los condicionamientos individuales corresponden mas específicamente: los orígenes escolares que son la trayectoria académica que puede ser definida por la institución inmediatamente anterior a la que pertenecía el estudiante; los orígenes sociales que pueden estar representados por su lugar de nacimiento o por su nivel socioeconómico; por ultimo, están las condiciones objetivas que puede estar representado por una variable ficticia que identifique si el individuo se encuentra trabajando, está en búsqueda del mismo o simplemente no trabaja.

Para realizar este modelo de corte transversal se necesitan otras variables como: edad del individuo al momento de presentar el examen, el género, nivel académico del padre y de la madre. De igual forma en el origen académico es necesario incluir la calificación promedio durante la secundaria lo cual representa el desempeño del estudiante durante su bachillerato.

Para este tipo de modelos hay que tener en cuenta que hay varios factores que condicionan la elección de la carrera por parte de los estudiantes; hay que recordar que este es un análisis microeconómico donde la elección de los agentes

suele ser muchas veces subjetiva. Uno de estos factores es que los estudiantes ven la educación como una inversión donde el coste se verá compensado por la obtención de un título y una inserción profesional; pero también existe otro factor contextual en la elección y es la oferta de plazas lo cual condiciona la racionalidad de los estudiantes en la elección de su carrera.

En síntesis al usar un «modelo de respuesta cualitativa», podemos predecir la probabilidad de que una respuesta sea elegida. Si un graduado de secundaria debe optar entre estudios universitarios de mayor o menor duración, podemos predecir la probabilidad de que elija cada una de las alternativas. La probabilidad de elegir cada una de las opciones está relacionada con las características personales de cada individuo, como la habilidad escolar en la etapa preuniversitaria, los ingresos familiares, el nivel educativo de los padres, pero también depende de las características de las alternativas, como el atractivo de los estudios, los derechos de matrícula, entre otros.

1. El «modelo logit binomial» aplicado a las elecciones educativas⁷

Si asumimos que un estudiante elige la alternativa más atractiva de entre dos opciones, la elección que nosotros observamos revela las preferencias del estudiante. Si dicho estudiante (que denotamos con i), elige una carrera universitaria de ciclo largo, esto implica que: $U_{i1} > U_{i0}$, donde U_{i1} y U_{i0} son las utilidades que i asocia con una carrera universitaria de ciclo largo y ciclo corto, respectivamente. La utilidad U_j que le reporta al individuo i la alternativa j ($j = 1$: estudiar una carrera universitaria de ciclo largo; $j = 0$: estudiar una carrera universitaria de ciclo corto), se descompone aditivamente en un componente

⁷ Salas Velasco (2000)

sistemático que depende de un vector de atributos X (habilidad escolar, background social, etc.), y otro aleatorio ε_{ij} :

$$U_{ij} = \bar{U}_{ij} + \varepsilon_{ij} \quad (8)$$

Sin embargo, la utilidad U_{ij} no es observable. Lo que observamos es la decisión Y_i , que vale 1 si el individuo i elige una carrera de ciclo largo, y vale 0 si elige una carrera de ciclo corto. Si un individuo racional elige la alternativa que le reporta mayor utilidad, tendríamos que:

$$\begin{aligned} \text{probabilidad } [Y_i = 1] &= \text{probabilidad } [U_{i1} > U_{i0}] \\ \text{probabilidad } [Y_i = 0] &= \text{probabilidad } [U_{i0} > U_{i1}] \end{aligned}$$

McFadden (1974) demuestra que, en este caso, la probabilidad de que el estudiante i elija la alternativa 1 es:

$$\text{prob}[Y_i = 1] = \frac{e^{X_i' \beta}}{1 + e^{X_i' \beta}} \quad (9)$$

En la ecuación (9) se observa la forma reducida para el «modelo logit binomial» (o «modelo logístico»), donde el vector fila X_i' de variables explicativas para el individuo i -ésimo contiene las variables independientes o explicativas consideradas en el apartado precedente (incluyendo una constante), y donde se supone que los ε 's no observados siguen una distribución de probabilidad logística⁸

En síntesis este modelo es presentado como una posibilidad de medir el comportamiento o probabilidad de que un estudiante en un año determinado elija

⁸ El «modelo logístico» puede ser visto como un caso especial de un modelo general de maximización de utilidad Salas cobos (2006).

estudiar Economía en vez de otra carrera; este manejo de datos de corte transversal se debe elaborar con variables explicativas que dictaminen o caractericen al estudiante: nivel económico, procedencia, nivel académico en cuanto al promedio en la secundaria.

2. Un «modelo logit multinomial» de elección de titulación universitaria

El estudiante que termina sus estudios de secundaria y decide acudir a la universidad debe elegir, en realidad, una de entre más de dos opciones alternativas. El «modelo logit binomial» descrito anteriormente, en el que se simplificaba la elección limitándola a optar entre estudiar o no economía, puede generalizarse incrementando el número de alternativas –es decir, incluyendo como opciones las diferentes titulaciones universitarias. Las alternativas posibles (Medicina, Enfermería, Derecho, otras Económicas, etc.) no están ordenadas. El modelo de elección múltiple entre alternativas no ordenadas puede venir generado también por modelos de utilidad aleatoria. Cuando el i -ésimo individuo ha de elegir entre J posibilidades (titulaciones), la utilidad de escoger la j -ésima opción es (Greene, 2000):

$$U_{ij} = \beta' z_{ij} + \varepsilon_{ij} \quad (10)$$

Si el individuo escoge la opción j -ésima es porque, de entre las J utilidades diferentes, la máxima es U_{ij} . El modelo estadístico se construye, por tanto, teniendo en cuenta que escoger la alternativa j -ésima equivale a decir que (Greene, 2000):

$$\text{prob}(U_{ij} > U_{ik}) \text{ para cualquier otro } k = j \quad (11)$$

Si definimos Y_i como una variable aleatoria que indica la alternativa escogida, y si las variables explicativas contienen sólo información sobre características específicas de los individuos (x_i), entonces la probabilidad de elegir la opción j -ésima viene dada por (McFadden, 1974):

$$prob [Y = j] = \frac{e^{\beta_j' x_i}}{1 + \sum_{k=1}^j e^{\beta_k' x_i}} \quad (12)$$

Este es el denominado «modelo LOGIT multinomial». Para desarrollar el modelo expresado en la ecuación (12) se debe contar con un buen conjunto de variables explicativas X_i que permitan explicar el grado de decisión de los estudiantes, es decir que sí determinen de verdad la probabilidad de que la elección si sea por la carrera que se busca: Economía; es un desafío para el investigador que en el modelo ajustado las variables sí sean significativamente importantes y que los coeficientes estimados posean el signo esperado.

En conclusión este segundo modelo determina más a fondo la preferencia de los estudiantes por las diversas carreras en general, es así como por ejemplo en la universidad EAFIT se puede ajustar un modelo LOGIT con el objeto de identificar las preferencias en la demanda de carreras universitarias en los estudiantes que ingresan a la misma universidad; para realizar un tipo de estudio con modelos de variable dependiente discreta (PROBIT o LOGIT).

2.1 Objetivo del modelo logístico

Ajustar un modelo en donde las variables explicativas sean las características de los estudiantes permite identificar cuál de ellas es la que más explica o influye en la elección del estudiante por la carrera, es decir, define si el nivel educativo del padre o la madre son más importante en la elección del estudiante o si el género

influye de manera significativa en esta decisión, aunque es de esperarse que el ingreso de cada individuo sea la variable que caracteriza más la decisión del estudiante ya que estamos hablando de que la educación superior en la universidad EAFIT hace parte de un bien totalmente excluyente en cuanto al costo que las mismas carreras conllevan.

2.2 Variables propuestas bajo la metodología Logit

A continuación se presenta un modelo con variable dependiente discreta (ver Anexo 5) que parte de los anteriores modelos probabilísticos y que poco han sido utilizados para este tipo de investigaciones⁹. Partiendo de una población de aspirantes nuevos que son admitidos a la carrera de los cuales algunos son matriculados y otros no (variable discreta dicotómica que indica matriculado=1, no matriculado=0) se proponen las siguientes variables que caracterizan a cada uno de los aspirantes a estudiar economía:

- Nivel educativo del padre: es una variable cualitativa que debe determinar exactamente hasta que grado, ya sea básico o estudio superior, realizó el padre.
- Nivel educativo de la madre: esta variable también debe caracterizar el grado de estudio exacto de la madre.
- Tipo de colegio de procedencia del aspirante: corresponde a una variable dicotómica en la cual se aclara si el estudiante proviene de un colegio público o privado.

⁹ Este tipo de modelos de elección discreta fue utilizado por el docente e investigador Andres Ramírez Hassan, (2007, EAFIT) quien aplicó por medio d una encuesta realizada en el 2007 un modelo para la demanda por cupos universitarios en Colombia.

- Rango de ingresos: es una variable determinada puede ser en tres o cuatro categorías de rangos que ayude a diferenciar los ingresos de los aspirantes: bajos, bajos medios, medios y altos¹⁰
- Genero: variable dicotómica de dos valores: 0 si es mujer y 1 si es hombre
- Estrato socioeconómico: es una variable cualitativa de un rango de 1 hasta 6.
- Preferencias: el aspirante debe dictaminar por qué desea estudiar Economía, ya sea por posibilidades laborales, influencia familiar, costo de la matricula, prestigio de la universidad, opciones de financiación, infraestructura física, ubicación geográfica, prestigio de la carrera u otras opciones¹¹; para tomar esta variable se debe enunciar una serie de razones por las cuales el estudiante decide ingresar a determinado programa, solo puede responder la más asertiva.

Obsérvese que las anteriores 7 variables que se incluirán como variables explicativas del modelo reúnen las características antes vistas de: orígenes escolares y orígenes sociales, además se selecciona una variable más subjetiva que busca identificar las preferencias de los aspirantes con lo que se logra el objetivo de identificar esos factores que están condicionando la demanda.

En síntesis, para alcanzar el objetivo de este estudio bastaría con contar con estas variables para la implementación del modelo, sin embargo el avance más significativo de este informe es la propuesta de estas nuevas variables en la determinación de la demanda y poder identificar cuál de ellas genera efectos mayores en el componente `probabilístico, es decir en la elección de estudiar economía.

¹⁰ Estos rangos deben ser dictaminados dependiendo del año en que se tomen y de acuerdo al comportamiento del salario mínimo,

¹¹ Esta variable permite identificar los gustos o intereses del estudiante por la carrera y ayuda a determinar la demanda de la misma.

V. ANALISIS Y CONCLUSIONES

- Dentro del análisis de la función de demanda es necesario incluir más variables que ayuden a explicar el comportamiento de la misma, sin embargo el estudio debe contar con un buen número de datos para realizar un buen ajuste. Se deben tener en cuenta variables como sustitutos a las matriculas de Eafit sin perder de vista que la universidad que se tome como punto de referencia debe ser homogénea, es decir en cuestión de infraestructura y de criterios académicos deben seguir un mismo patrón, empezando por su carácter de universidad privada.
- En el modelo lineal que se formuló, se observa que el termino de error deja mucho en el modelo que se explica por otros factores exógenos, dentro de estos se podrían explicar factores como reformas al pensum, procesos de acreditación durante los seis años de estudio, lo que hace variar mucho las preferencias de los agentes; de otro lado hay que tener en cuenta que el modelo solo abarca una parte microeconómica del sistema de educación

superior, por lo tanto, en estudios futuros se deberá tener en cuenta comportamiento de otros programas educativos relevantes o complementarios al programa de Economía y que ayuden a identificar mejor la elección de los agentes de matricularse en Este programa, es decir que se deben compartir otras variables complementarias como el valor de matriculas sustitutas o de otros sustitutos cercanos que el individuo pueda tomar a cambio del programa de economía.

- En Economía, como en otras ciencias sociales, se han propuesto varias teorías para explicar la demanda de educación, y todas ellas tienen en cuenta los costes, la renta disponible, las expectativas de ingresos futuros y el desempleo. En la práctica, las perspectivas de empleo determinan, probablemente, la demanda de educación. La notable influencia del capital humano adquirido en el sistema educativo podría ser un argumento para explicar la fuerte expansión de la demanda de estudios superiores que ha experimentado otros países.
- El *background* social se puede considerar como factor determinante de las elecciones educativas. En sus teorías, el nivel educativo y el nivel ocupacional de los padres afectan a las opciones educativas de los hijos. Esta conclusión es algo que es de esperarse en el modelo LOGIT propuesto en la última parte del informe; Parece evidente que el nivel educativo y ocupacional de los padres, ya que los valores que se transmiten a los hijos y la existencia de un mayor o menor ambiente cultural pueden ejercer una importante influencia sobre la demanda de Educación Superior (Mora, 1990). El nivel de educación de los hijos de padres con un *status* socioeconómico más alto es mayor.

- En síntesis, las posibles causas de la demanda o no en el programa de economía no esta justificada solo por el precio de la matricula, ya se observó que se pueden tener otras variables relevantes para el estudio que sirvan como herramientas para explicar el comportamiento de la demanda por el programa de economía, como lo muestra el modelo sencillo de demanda explicado en el trabajo
- En un modelo micro-económico de demanda educativa, debemos contemplar la influencia que los gastos de la escolarización universitaria ejercen sobre la cantidad de Educación Superior que deciden acumular los individuos o sus familias como por ejemplo gastos escolares que prevean las familias, además es importante en trabajos futuros incluir también la variable genero que determina de un modo o de otro la demanda por este programa educativo.
- Con respecto al modelo probabilístico, queda abierta la discusión para trabajos futuros, con base en este modelo y con la información necesaria, es posible dictaminar la probabilidad de que un estudiante tome la decisión de seguir la carrera de economía: Un método directo que permite aproximarse al proceso de toma de decisiones educativas en el ámbito universitario es preguntar directamente a los estudiantes acerca de los motivos que les empujaron o animaron a seguir «educándose» una vez concluida la enseñanza secundaria.
- Se propone el modelo LOGIT para desarrollos futuros en estudios de demanda en donde se cuente con la información suficiente y adicionalmente se sugiere la posibilidad de incluir las variables como una información adicional que se solicite a los aspirantes nuevos en el momento de diligenciar un formulario de ingreso al campus universitario.

- Obviamente, modelos como los planteados en este artículo pueden presentar algunas limitaciones -debidas, por supuesto, a la carencia de información. En situaciones prácticas, las imperfecciones del mercado, la incertidumbre de obtener ingresos futuros y la habilidad escolar influyen también en las elecciones educativas. Las imperfecciones del mercado pueden impedir o dificultar que algunas personas "cosechen" los beneficios que se derivan de la educación. Por otra parte, la elección de carrera que efectúan los individuos al terminar sus estudios medios también está determinada por las anticipaciones que hacen sobre sus oportunidades de éxito académico. Elegir una carrera de ciclo largo supone invertir más años en educación y, para los individuos con peor trayectoria académica en los niveles preuniversitarios, asumir un mayor riesgo de fracaso escolar¹²

VI. BIBLIOGRAFIA

- BECKER, W. E. (1990): «The Demand for Higher Education», en HOENACK, S. A.; COLLINS, E. L. (eds.): The Economics of American Universities. Albany, State University of New York Press, pp. 155-188.
- BISHOP, J. (1977): «The Effect of Public Policies on the Demand for Higher Education», en Journal of Human Resources, 12, pp. 285-307.
- BRUNNER, J.J. (1990) Educación Superior en América Latina: Cambios y Desafíos; Fondo de Cultura Económica, Santiago de Chile, 1990.

¹² Este limitante dentro del análisis es citado por Salas como conclusión del estudio realizado por Latiesa (1989) para la Universidad Autónoma de Madrid en donde toma en cuenta que la edad determina el riesgo del estudiante de valorar más o menos la carrera que vaya a estudiar.

- CLARK, B.R. (1998) Creating Entrepreneurial Universities. Organizational Pathways of Transformation; IAU Press, Pergamon, Great Britan.
- CRAMER, J. S. (1991): The Logit Model. An Introduction for Economists. Londres, Edward Arnold.
- ENDERS, Walter, Applied econometric time series, 1st edition, Iowa State University, 1995
- GREENE, William H. Econometrics Análisis. Four Edition. Prentice Hall. 2000.
- GUJARATI D. Econometría. McGraw - Hill. Cuarta Edición. 2004.
- HENDERSON, J.M., and R.E. Quandt. 1980. The theory of consumer behavior. p. 28, 39-40, 60. *In* J.M. Henderson and R.E. Quandt (eds.) Microeconomic theory: a mathematical approach. 3rd ed. McGraw-Hill, New York, USA.
- JIMÉNEZ, J. D.; SALAS VELASCO, M. (2000): «Modeling Educational Choices. A Binomial Logit Model Applied to the Demand for Higher Education», en Higher Education, 40, pp. 293-311.
- JUDGE, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T. C. Lee. Introduction to the Theory and Practice of Econometrics. John Willey and Sons. Second Edition. 1988.
- KODDE, D. A. (1986): «Uncertainty and the Demand for Education», en Review of Economics and Statistics, 68, pp. 862-872.

- LATIESA, M. (1989): «Demanda de Educación Superior: evaluaciones y condicionamientos de los estudiantes en la elección de carrera», en REIS, 46, pp. 101-139.
- MADDALA. G.S. Introducción a la econometría. Prentice Hall. Segunda edición. 1996.
- MCFADDEN, D. L. (1974): «Conditional Logit Analysis of Qualitative Choice Behavior», en ZAREMBKA, P. (ed.): Frontiers in Econometrics. New York, Academic Press, pp. 105-142.
- MORA, J. G. (1989): «La demanda de Educación Superior: una revisión de estudios empíricos », en *Revista de Educación*, 288, pp. 351-375.
- _____ (1990): *La demanda de Educación Superior*. Madrid, Consejo de Universidades.
- NOVALES. A. Econometría. McGraw - Hill. Segunda Edición. 1993.
- PENA J. Bernardo T., J. Estabillo, M. Galindo, M. Leceta, M. Zamora. Cien ejercicios de econometría. Pirámide. 1999.
- RAMÍREZ Hassan, Andrés. Modelos de elección discreta: una aplicación a la demanda por cupos universitarios en Colombia. Dirección de Investigación y Docencia EAFIT, 2007
- SALAS Velasco, Manuel y Martin-Cobos Puebla, Manuel. (2006). “La demanda de educación superior: un análisis microeconómico con datos de corte transversal”. En: *Revista de Educación*, 339 (2006), pp. 637-660.
- STEWART M. , Wallis K. Introducción a la Econometría. Alianza Editorial. 1984.

- WOOLDRIDGE, J.M. Introducción a la Econometría. Un enfoque moderno. Thomson. Primera edición. 2001.
- VARIAN, Hal R, Microeconomía intermedia, 3ra edición, Barcelona 1994
- <http://icfes.gov.co>.(2007)

VII. ANEXOS

ANEXO 1 Modelo logarítmico con una sola variable

Dependent Variable: LOG(MATRIC_ICFES_)				
Method: Least Squares				
Date: 12/08/07 Time: 10:00				
Sample: 2001S1 2006S1				
Included observations: 11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9.658700	5.691401	-1.697069	0.1239
LOG(VALOR)	1.024910	0.384422	2.666107	0.0258
R-squared	0.441276	Mean dependent var	5.514403	
Adjusted R-squared	0.379195	S.D. dependent var	0.242474	
S.E. of regression	0.191048	Akaike info criterion	-0.309620	
Sum squared resid	0.328494	Schwarz criterion	-0.237275	
Log likelihood	3.702910	F-statistic	7.108124	
Durbin-Watson stat	1.006304	Prob(F-statistic)	0.025786	

ANEXO 2 Modelo logarítmico con dos variables regresoras

Dependent Variable: LOG(MATRIC_ICFES_)
--

Method: Least Squares				
Date: 12/13/07 Time: 06:11				
Sample: 2001S1 2006S1				
Included observations: 11				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-14.82902	22.07164	-0.671858	0.5206
LOG(VAOR)	0.952581	0.503276	1.892761	0.0950
LOG(PIB)	0.432874	1.777974	0.243465	0.8138
R-squared	0.445385	Mean dependent var		5.514403
Adjusted R-squared	0.306731	S.D. dependent var		0.242474
S.E. of regression	0.201890	Akaike info criterion		-0.135184
Sum squared resid	0.326077	Schwarz criterion		-0.026667
Log likelihood	3.743512	F-statistic		3.212211
Durbin-Watson stat	0.977441	Prob(F-statistic)		0.094616

ANEXO 3 Datos tomados en el modelo logístico

año	matriculas (icfes)	Valor matriculas	PIB percapita
2001-1	155	2256081	1751959
2001-2	155	2256081	1751959
2002-1	255	2526811	1754734
2002-2	256	2526811	1754734
2003-1	308	2830028	1791792
2003-2	302	2830028	1791792
2004-1	286	3084730	1848250
2004-2	279	3084730	1848250
2005-1	268	2393203	1905507
2005-2	249	2393203	1905507
2006-1	284	3732524	2002910

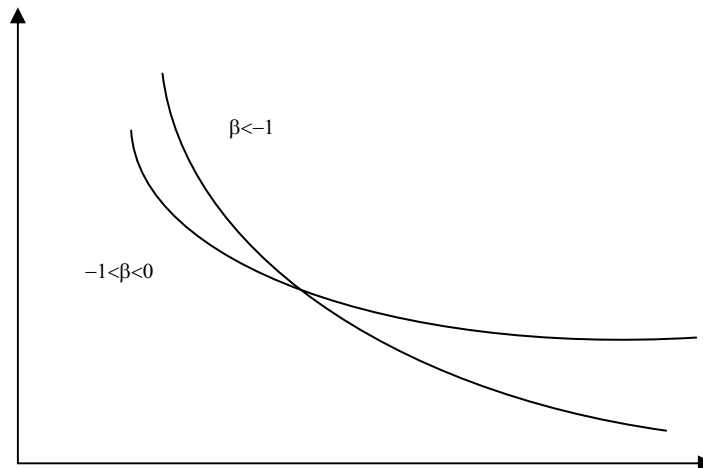
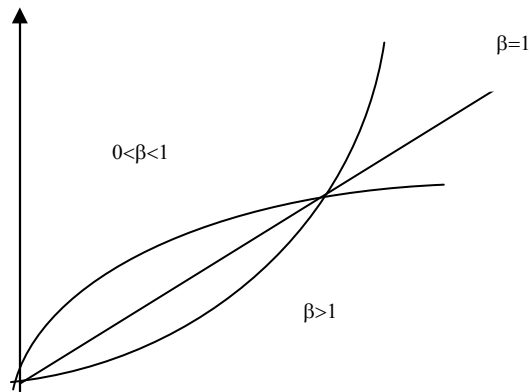
Fuente: DANE - ICFES

ANEXO 4 Transformación de variables y modelos logarítmicos

Transformación log-log

Una forma funcional que se puede presentar en un modelo bivalente es la siguiente:

$$Y_t = A X_t^\beta \quad X_t > 0$$



A la expresión anterior se le puede realizar la siguiente transformación:

$$\ln(Y_t) = \alpha + \beta \ln(X_t) \quad \text{donde } \alpha = \ln(A)$$

La cual convierte el modelo en lineal y por lo tanto se le puede aplicar la metodología de mínimos cuadrados ordinarios para estimar los parámetros α y β .

De manera general, dada una función $Y=f(X)$, la elasticidad de Y con respecto a X está dada por:

$$\eta = \frac{X_t}{Y_t} \frac{dY_t}{dX_t} : \text{cambio del } \eta\% \text{ en } Y \text{ ante un cambio del } 1\% \text{ en } X$$

Para el modelo log-log se tiene que

$$\beta = \frac{X_t}{Y_t} \frac{dY_t}{dX_t}$$

Por lo tanto se habla de un modelo de elasticidad constante.

ANEXO 5 Modelos de probabilidad: Modelo de probabilidad logístico (logit)

En este anexo se desarrolla un modelo estadístico para el cual las metodologías OLS y 2SLS no son apropiadas. Además este modelo no es lineal en los parámetros y por lo tanto las estimaciones provienen de un proceso algorítmico. Consideremos el siguiente caso:

A una variable aleatoria W que tiene por función de densidad

$$f(w) = \frac{e^{-w}}{(1+e^{-w})^2} \quad w \in \mathbb{R}$$

Se le denomina variable aleatoria logística. Esta variable aleatoria tiene por media cero, la varianza está dada por $\pi^2/3$ y además es simétrica con respecto a su media. La gráfica de f es muy similar a la de una "t" con siete grados de libertad y por lo tanto también a la de una normal estándar (tiene colas más pesadas que la normal). La función de distribución está dada por

$$F(a) = \Lambda(a) = \frac{e^a}{1+e^a}$$

El modelo de probabilidad logístico (logit) está dado por

$$P(y_i = 1) = F(\mathbf{X}_i\boldsymbol{\beta}) = \Lambda(\mathbf{X}_i\boldsymbol{\beta}) = \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})}$$

La estimación de este modelo es similar al Probit. En (1) simplemente se reemplaza Φ por Λ y σ por 1

Observaciones

1. El contraste global de regresión para estos modelos, tiene por estadístico de prueba a:

$$2[l(\hat{\alpha}, \hat{\boldsymbol{\beta}}^*) - l(\hat{\alpha}, \mathbf{0})] \stackrel{a}{\sim} \chi_{k-1}^2$$

Donde α y $\boldsymbol{\beta}^*$ son los parámetros asociados al intercepto y a las regresoras, respectivamente. $l(\hat{\alpha}, \hat{\boldsymbol{\beta}}^*)$, $l(\hat{\alpha}, \mathbf{0})$ son los valores de los logaritmos de la función de verosimilitud para el modelo completo estimado y para el modelo bajo la restricción ($\boldsymbol{\beta}^* = \mathbf{0}$), $k-1$ es el número de regresoras en el modelo. Las pruebas marginales para los parámetros están basadas en el estadístico denominado "t", pero se desarrollan con base en la normal (0, 1)

2. En los modelos Probit y Logit se tiene que

$$\frac{\partial P(Y_i = 1)}{\partial (X_{ji})} = \frac{\partial E(y_i)}{\partial (X_{ji})} = \beta_j f(\mathbf{X}_i\boldsymbol{\beta})$$

Donde $f(\mathbf{X}_i\boldsymbol{\beta})$ es la función de densidad de la normal o de la logística. Lo anterior nos permite decir que β_j no tiene la misma interpretación que en el modelo LP, pero el signo de β_j sí indica la dirección de la relación entre X_j y $E(y_i)$. Si $\beta_j > 0$, un aumento de una unidad en X_j , incrementa la probabilidad de que $Y_i = 1$. La magnitud del cambio está dada por $\beta_j f(\mathbf{X}_i\boldsymbol{\beta})$. Es bueno

observar que esta interpretación no es válida para el caso en el cual la regresora es discreta.

3. Con base en la función de verosimilitud puede definirse un estadístico que tiene cierta semejanza con el coeficiente de determinación usual. En este caso se habla del R^2 de McFadden, el cual se define

$$R_{MC}^2 = 1 - l(\alpha, \hat{\beta}_*) / l(\alpha, \mathbf{0})$$

Si la capacidad explicativa del modelo es muy buena se debe tener que $l(\alpha, \hat{\beta}_*)$ debe ser muy superior a $l(\alpha, \mathbf{0})$ y por lo tanto este coeficiente podría estar cerca de uno.

4. No es fácil decidir, entre los modelos logit y probit, cual escoger. Los estadísticos R_{MC}^2 , el logaritmo de la verosimilitud (valores máximos), AIC, el criterio de Schwars (valores más pequeños) pueden ayudar en esta decisión.
5. Es posible que se presenten problemas de heteroscedasticidad en este tipo de modelos. Como es de esperar este es un problema relativamente complejo, sin embargo Davidson y MacKinnon en 1993 desarrollan un test para algunas situaciones particulares de heteroscedasticidad. En el texto de Greene, se muestra como desarrollar esta prueba.