



SEDE ACADÉMICA DE ECUADOR

**FACULTAD LATINOAMERICANA DE CIENCIAS  
SOCIALES**

**SEDE ECUADOR**

**PROGRAMA DE POLÍTICAS SOCIALES Y GESTIÓN**

**MAESTRIA EN POLÍTICAS SOCIALES**

**EVALUACIÓN DE IMPACTO PROGRAMA DE  
ALIMENTACIÓN ESCOLAR**

**Lenin Cadena Minotta**

**Director: Dr. Juan Ponce J.**

**QUITO, MARZO 2004**

# INDICE

## INTRODUCCIÓN

### CAPITULO 1

#### INTRODUCCIÓN A LA EVALUACIÓN DE IMPACTO

- 1.1) Introducción
- 1.2) La estimación de impactos: conceptos y formalización inicial
- 1.3) Diseños de evaluación
- 1.4) Diseños experimentales de evaluación
  - 1.4.1) Formalización de un diseño experimental de evaluación
- 1.5) Diseños no experimentales de evaluación
- 1.6) Técnica de Variables instrumentales
  - 1.6.1) Metodología de evaluación
  - 1.6.2) Supuestos en la utilización de IV
  - 1.6.3) Problemas econométricos
- 1.7) Emparejamiento por puntos de propensión
  - 1.7.1) Características del emparejamiento
  - 1.7.2) Metodología de evaluación
    - 1.7.2.1) Vecino más cercano o emparejamiento uno a uno
    - 1.7.2.2) Grupo de 5 vecinos más cercanos
    - 1.7.2.3) Emparejamiento Kernell
  - 1.7.3) Supuestos en la utilización del emparejamiento
  - 1.7.4) Problemas econométricos
- 1.8) Conclusiones

## **CAPITULO 2**

### **DESCRIPCIÓN DEL PROGRAMA DE ALIMENTACIÓN ESCOLAR PAE**

- 2.1) Introducción
- 2.2) Objetivos del programa
- 2.3) Antecedentes
- 2.4) Características de la ración diaria
- 2.5) Estrategia operativa
- 2.6) Costos administrativos del programa
- 2.7) Costo unitario
- 2.8) Focalización del programa
- 2.9) Análisis de la eficiencia de Focalización
- 2.10) Impacto Distributivo
- 2.11) Conclusiones

## **CAPITULO 3**

### **ANÁLISIS DE DATOS Y VARIABLES**

- 3.1) Fuente de datos
- 3.2) Revisión de la muestra
- 3.3) Calculo de variables
- 3.4) Conclusiones

## **CAPÍTULO 4**

### **VARIABLE INSTRUMENTAL**

- 4.1) Introducción
- 4.2) Modelo Teórico de asistencia escolar
- 4.3) Problemas econométricos
  - 4.3.1) Creación de variable instrumental

- 4.3.2) Pruebas de especificación del modelo
- 4.4) Especificación econométrica del modelo de asistencia escolar
  - 4.4.1) Revisión de variables
- 4.5) Resultados econométricos
  - 4.5.1) Resultados de la evaluación del programa
  - 4.5.2) Resultados del modelo de asistencia
- 4.6) Conclusiones

## **CAPÍTULO 5**

### **EMPAREJAMIENTO POR PUNTAJES DE PROPENSIÓN**

- 5.1) Introducción
- 5.2) Especificación econométrica modelo de participación
- 5.3) Resultados del modelo de de participación
- 5.4) Creación de un soporte común
- 5.5) Resultados de la evaluación del impacto
  - 5.5.1) Emparejamiento uno a uno.
  - 5.5.2) Grupo de vecinos más cercanos
  - 5.5.3) Emparejamiento Kernell
- 5.6) Prueba de no observables
- 5.7) Conclusiones

## **CAPITULO 6**

### **CONCLUSIONES Y RECOMENDACIONES**

- 6.1) Conclusiones Generales

## CAPITULO 3

### ANÁLISIS DE DATOS Y VARIABLES

#### 3.1) Fuente de datos

Esta investigación utiliza información de la Encuesta de Condiciones de Vida de 1999 (Living Standards Measurements Survey LSMS) de donde se obtuvo la muestra de estudiantes fiscales de nivel primario con las características socioeconómicas de sus hogares. La segunda fuente es la Base para investigadores del Sistema Integrado de Indicadores Sociales del Ecuador (SIISE 1999), que facilitó la información sobre oferta educativa y criterios de focalización del programa. Adicionalmente, del Censo del magisterio fiscal (2000) se incluyen la categoría económica docente y número de profesores contratados por la comunidad, calculados como promedios cantonales.

La ECV99 es una encuesta con una muestra de 5081 hogares, con cobertura urbana y rural sobre dos regiones del Ecuador continental (Costa y Sierra)<sup>22</sup>. La encuesta está dividida en diversos módulos<sup>23</sup>, que registran en detalle las características de los encuestados. A nivel de individuos la muestra alcanza 25980 casos. En el módulo de educación existe información sobre niños en edad escolar que asisten a escuelas fiscales, y una pregunta para identificar a los beneficiarios del Desayuno Escolar, que sirve de base para esta evaluación.

---

<sup>22</sup> La ECV 1999 no tiene muestra de la Región Amazónica.

<sup>23</sup> Tales como salud, educación, condiciones de empleo, entre otras.

### 3.2) Revisión de la muestra

Para definir la muestra de evaluación se siguieron los siguientes pasos:

- a) Cálculo de una variable dummy que identifique a los beneficiarios del Desayuno escolar.
- b) Cálculo de una variable que identifica a los niños matriculados en primaria fiscal.
- c) Cálculo de una variable que indica la frecuencia con que se recibe el beneficio (todos los días, 1 a 4 días a la semana, rara vez).

Para asegurar que el universo este constituido por niños que recibían el Desayuno escolar todos los días, en edad y año escolar adecuado, se eliminaron los casos cuya edad no se ajustaba al rango adecuado para educación primaria (entre 6 y 12 años); y se eliminaron los casos que no recibían el beneficio todos los días (el 3% del total de la muestra).

Finalmente, con la combinación de estas variables obtuvimos nuestro universo de evaluación. A continuación una revisión de la muestra.

#### Cuadro No.7

##### Composición de la muestra de beneficiarios y no beneficiarios

<i>Condición</i>	<i>niños en primaria fiscal</i>	<i>Porcentaje</i>
no beneficiarios	1791	64.1
beneficiarios	1003	35.9
Total	2794	100

El universo que servirá para esta evaluación esta conformado por 2794 niños en edad escolar que asisten a escuelas fiscales. La muestra de

beneficiarios del programa alcanza 1003 individuos. La muestra de los no beneficiarios esta formada por los restantes 1791 individuos.

Adicionalmente, la muestra sirve para evaluar algunos de los criterios de focalización del programa (Capítulo 2).

### Cuadro No.8

Composición de la muestra según Area, Región Natural y Condición de pobreza

Area	no beneficiario	beneficiario	Total	Porcentaje
Rural	874	797	1671	59.81
Urbano	917	206	1123	40.19
Total	1791	1003	2794	100.00
<b>Región Natural</b>				
Costa	896	351	1247	44.63
Sierra	895	652	1547	55.37
Total	1791	1003	2794	100.00
<b>Pobreza</b>				
No pobre	659	197	856	30.66
Pobre	1132	804	1936	69.34
Total	1791	1003	2794	100.00

Contamos con un importante tamaño de muestra para el área rural (59.81% de los casos), y de población en condiciones de pobreza (el 69.34%). También se cuenta con tamaños significativos por región. Los siguientes cuadros identifican la edad y el grado en que se matricularon los grupos considerados.

### Cuadro No.9

Composición de la muestra según grado en que se matriculó en primaria fiscal

Grado en primaria fiscal	no beneficiario	beneficiario	Total	Porcentaje
1	403	226	629	22.51
2	310	218	528	18.90
3	315	178	493	17.64
4	296	160	456	16.32
5	248	123	371	13.28
6	219	98	317	11.35
Total	1791	1003	2794	100.00

## Cuadro No.10

### Composición de la muestra según edad

Edad	no beneficiario	beneficiario	Total	Porcentaje
6	236	104	340	12.17
7	270	161	431	15.43
8	290	165	455	16.28
9	255	169	424	15.18
10	286	152	438	15.68
11	264	148	412	14.75
12	190	104	294	10.52
Total	1791	1003	2794	100.00

Acorde con el cuadro 9, la recepción del beneficio se reduce a medida que aumentan los grados de primaria. La muestra se reduce a la mitad entre el primer y el último grado (de 22% a 11%). En cambio, por edad no hay mayor variación, concentrándose los datos entre los 7 y 11 años (cuadro No.10).

### 3.3) Calculo de variables

Para evaluar el impacto del programa utilizamos información relevante relacionada con la asistencia escolar<sup>24</sup>. Las variables calculadas en base a la ECV99 son: beneficiario del programa, grado, matriculado en primaria fiscal, área, región, consumo del hogar, instrucción del jefe del hogar, instrucción de la madre, número de dependientes en el hogar menores de 10 años, y número de dependientes de entre 11 y 15 años.

Las variables obtenidas del Censo del Magisterio Fiscal son: Categoría económica docente promedio en el cantón, y Número de profesores contratados por la comunidad. De la Base para Investigadores del SIIE 1999 utilizamos el promedio de alumnos por profesor, el promedio de

<sup>24</sup> Los modelos teóricos fueron definidos para cada tipo de evaluación. Revisar Capítulos 4 y 5.



alumnos por aula, ambas variables a nivel cantonal, y el Índice de Vulnerabilidad Social (IVS).

A continuación un cuadro con las variables creadas incluyendo la fuente, el método de construcción y una breve descripción.

**Cuadro No.11**

Nombre	Nombre en el modelo	tipo	Fuente	Método de construcción	Descripción
edad	edad	continua	ECV99	Consta en la encuesta pd03	edad del niño
grado	grado	continua	ECV99	Creada a partir de pe32 y pe33	grado en que se matriculó
Matriculado en primaria fiscal	primfiscal	continua	ECV99	Combinación de preguntas pe12 y pe15	valores entre 1 y 6
Beneficiario del programa	d_pae	dummy	ECV99	Creada a partir de pe13	1=beneficiario 0=no beneficiario
Asistencia escolar	asiste_1	continua	ECV99	Calculada a partir de pe20 y pe22	Se consideran únicamente los casos entre 0 y 20 días al mes
Area de residencia	areal	dummy	ECV99	Creada a partir de area (consta en la encuesta)	1=área urbana 0=área rural
región	region	dummy	ECV99	Creada a partir de región (consta en la encuesta)	1=Costa 0=Sierra
consumo del hogar	lncns_hog	continua	ECV99	Construida a partir de adjcap (consumo per cápita)	Transformada en su logaritmo natural para normalizarla
instrucción del jefe de hogar	ins_jefe1	continua	ECV99	Calculada usando pd04, pe32 y pe33	años de estudio aprobados por el jefe del hogar
-instrucción de la madre	ins_mam	continua	ECV99	Calculada usando pd04, pe32 y pe33	años de estudio aprobados por la madre
Número de dependientes en el hogar de entre 0-10 años	menos10	continua	ECV99	Calculada usando pd04 y edad	dependientes en el hogar de entre 0-10 años
Número de dependientes en el hogar de entre 11-15 años	entre10_15	continua	ECV99	Calculada usando pd04 y edad	dependientes en el hogar de entre 0-10 años
Costo directo de educación	lnced_hog	continua	ECV99	Construida con variables pe24, pe25, pe26, pe27, pe28, pe29 y pe30.	Transformada en su logaritmo natural para normalizarla
Trabajo infantil	d_trabaja	dummy	ECV99	Construida con variables pa20 y d_pae	Identifica niños que reciben el beneficio y trabajan
Tiempo en minutos	ttmin	continua	ECV99	Construida a partir de pe18b	Tiempo en minutos trasladado desde la casa a la escuela
Traslado a pie	camina	dummy	ECV99	Construida a partir de pe19	=1 camina hasta la escuela =0 en animales o bicicleta

Nombre	Nombre en el modelo	tipo	Fuente	Método de construcción	Descripción
Categoría económica docente promedio en el cantón	cateco_p	continua	Censo del Magisterio	Construida con Censo del Magisterio y llevada a ECV99	Utilizada como una aprox. a experiencia docente
Número de profesores contratados por la comunidad (a nivel cantonal)	nprof_com	continua	Censo del Magisterio	Construida con Censo del Magisterio y llevada a ECV99	
Promedio de alumnos por profesor a nivel cantonal	alupro98	continua	SIISE 1999	Procesada por SIISE 1999 y llevada a ECV99	
Promedio de alumnos por aula a nivel cantonal	aluau98	continua	SIISE 1999	Procesada por SIISE 1999 y llevada a ECV99	
Índice de Vulnerabilidad Social	ivs	continua	SIISE 1999	Procesada por SIISE 1999 y llevada a ECV99	

En el siguiente cuadro presentamos estadísticos descriptivos de las variables utilizadas en los modelos, diferenciando entre grupo de tratamiento y grupo de control.

## Cuadro No.12

### Estadísticos descriptivos de variables utilizadas en los modelos

Variable	Grupo de no beneficiarios		Grupo de beneficiarios	
	Media	Desv. Stand.	Media	Desv. Stand.
asistencia escolar	17.48353	5.145224	16.72782	5.846736
grado	3.18593	1.691101	3.02991	1.627822
area	0.5120045	0.4999955	0.2053838	0.4041834
region	0.4997208	0.5001396	0.6500499	0.4771919
ln(costo de educación)	10.77646	0.9714139	10.61808	0.806539
ln(consumo del hogar)	13.8197	0.6002609	13.57457	0.6023498
menores de 10 años	2.061418	1.346057	2.338983	1.429915
entre 10 y 15 años	1.311558	1.07229	1.438684	1.073626
instrucción del jefe del hogar	6.336125	4.111005	4.616152	3.210701
instrucción de la madre	5.997057	4.087744	4.148577	3.278888
el niño trabaja	0.1094361	0.3122727	0.1635095	0.3700144
camina hasta la escuela	0.7788945	0.4151073	0.9121756	0.2831806
tiempo en minutos distancia a la escuela	14.18035	10.67872	13.20858	10.66422
categoría económica promedio en el cantón	8.340281	0.6189745	8.127936	0.7199888
número de profesores contratados	71.74707	83.5308	37.27717	59.87703
porcentaje de alumnos por profesor	16.75249	8.660102	16.97683	9.919703
porcentaje de alumnos por aula	24.25794	7.876566	23.67416	7.789875
índice de vulnerabilidad social	34.64975	19.35103	45.04985	22.39597

En general, el grupo de beneficiarios presenta inferiores condiciones al grupo de control. Revisando el cuadro 12, la asistencia escolar promedio del grupo de beneficiarios es inferior en 0.68 días al grupo de control. El trabajo infantil tiene mayor incidencia en el grupo de tratamiento. Incluso la educación del jefe de hogar y de la madre es inferior entre los beneficiarios. El costo directo de educación y el consumo del hogar entre ambos grupos no presentan mayores diferencias. En el grupo de variables de oferta educativa, la categoría económica de los profesores no presenta mayores diferencias entre grupos, y el promedio de profesores contratados por las comunidades en los beneficiarios es casi la mitad de los contratados por el grupo de control.

### **3.4) Conclusiones**

1. En resumen nuestra muestra tiene tamaños adecuados que permiten realizar una evaluación del impacto de recibir Desayuno escolar en asistencia, controlando por varios aspectos, entre los que se cuentan: características del niño, del hogar, de la oferta educativa, y el costo de educación; y también diferenciar impactos por diversas categorías: área y región. Hasta aquí una aproximación a las variables nos permiten encontrar las primeras diferencias entre beneficiarios y quienes no reciben el beneficio que buscaremos comprobar en los capítulos posteriores.

## CAPÍTULO 4

### VARIABLE INSTRUMENTAL

#### 4.1) Introducción

En este capítulo presentamos una evaluación del impacto del Programa en asistencia escolar utilizando la técnica de variables instrumentales. El problema principal para la estimación de este impacto es la presencia de regresor endógeno, esto es, la dummy de participación en el programa que debería captar su efecto, no puede ser incluida directamente en el modelo porque no es un factor exógeno dado que la decisión de participación en el programa no fue aleatoria. Para la corrección de este problema utilizamos un instrumento correlacionado con la participación en el programa, pero no con sus resultados.

Inicialmente presentamos un modelo teórico de asistencia escolar del que derivamos nuestras especificaciones econométricas. Esta investigación presta mucha atención a demostrar la validez del instrumento utilizado para la corrección del regresor endógeno, y de su forma de cálculo. Incluimos una revisión de los problemas econométricos y pruebas para las especificaciones utilizadas. Finalmente presentamos los principales resultados y conclusiones<sup>25</sup>.

#### 4. 2) Modelo Teórico de asistencia escolar

Para nuestra estructura analítica utilizamos un modelo desde la perspectiva del capital humano, que establece que la decisión de escolaridad

---

<sup>25</sup> En esta investigación presentamos únicamente los resultados econométricos más importantes. Un reporte adicional están a disposición con solicitud al autor.

depende de su costo individual y familiar privado y los beneficios esperados.

La consideración de base es que las variaciones en asistencia escolar entre hogares sugiere que los padres tienen una valoración diferenciada de los beneficios de la educación. Esta valoración es hecha sobre la base de un balance entre los costos y los beneficios esperados de enviar a sus hijos a la escuela.

Los beneficios de tener hijos mejor educados son a la vez materiales (apoyo económico en la vejez) y de satisfacción porque alcancen un mejor nivel de vida. A su vez, los costos de asistencia escolar se diferencian en costos directos (dinero para alimentación y transporte; materiales; entre otros), y costos indirectos valorados como la contribución económica que haría el niño en caso de trabajar. Este costo podría ser mayor, por ejemplo, en casos de crisis económicas generalizadas o en época de cosecha en las zonas agrícolas. Finalmente la decisión de enviar a sus hijos a la escuela en un día en particular es una consecuencia de que en esta comparación los beneficios sean mayores a los costos.

Para formalizar estas consideraciones presentamos una adaptación a nuestros requerimientos de Bedi y Marshall (2000). Consideremos que el año escolar tiene  $n$  días; definiendo  $i$  como uno de esos días. Asumimos que cada hogar tiene una función de utilidad definida por los beneficios asociados con la asistencia escolar un día  $i$ , simbolizado por  $h_i$ , y el consumo del hogar en un día  $i$ , representado por  $c_i$ .

El hecho de que el niño asista a la escuela representa costos para el hogar que reducen la disponibilidad de consumo. Acorde con esto, la utilidad

para el hogar en un día  $i$  condicionada a la asistencia escolar (definida con el subíndice 1), esta dada por:

$$U_{i1} = U(b_i, c_{i1})$$

Como hemos señalado, el ingreso del hogar estaría restringido:

$$y_i = c_{i1} + p_i$$

donde  $y_i$  es el ingreso del hogar  $i$ , y  $p_i$  son los costos totales de la asistencia escolar. A su vez, la utilidad para el hogar en un día  $i$  condicionada a la decisión de no enviar a sus hijos a la escuela (definida con el subíndice 0), esta dada por:

$$U_{i0} = U(c_{i0})$$

En este caso, el presupuesto del hogar destinaría todo su ingreso al consumo diario  $y_i = c_{i0}$ .

Utilizando las utilidades asociadas a cada opción ( $U_{i1}, U_{i0}$ ), los hogares decidirán por la opción que les permita maximizar su utilidad:

$$U_i^* = \max(U_{i1}, U_{i0})$$

donde  $U_i^*$  es la máxima utilidad. Si definimos la asistencia con una variable dummy, donde  $a_i = 1$  si un niño asiste a la escuela y viceversa, la decisión de asistencia escolar  $a_i = 1$  es posible para el caso en que  $U_{i1} > U_{i0}$ .

Para nuestra investigación asumimos que la recepción del beneficio reduce los costos directos de asistencia escolar ( $p_i$ ), ampliando la disponibilidad de ingreso en el hogar:

$$y_{i*} = c_{i1} + (p_i - d_i)$$

donde  $d_i$  es el valor monetario del desayuno escolar. Si  $y_{i*}$  es el ingreso de los hogares de los niños que reciben el desayuno, y  $y_i$  el caso contrario, tenemos que  $y_{i*} > y_i$ . Este resultado mejora el consumo de los hogares cuyos hijos reciben el beneficio dado que  $c_{i1} = y_{i*} - p_i$ , mejorando la utilidad derivada de la decisión de enviar a sus hijos a la escuela. Por lo tanto, la recepción del beneficio es un aliciente para la asistencia escolar en un día  $i$ .

La presentación de esta estructura teórica de análisis<sup>26</sup>, tuvo por objetivo explicar los factores que inciden en la decisión de los hogares de enviar a sus hijos a la escuela, asumiendo que el efecto de recibir el desayuno escolar reduce el costo directo de asistencia diaria y aumenta la probabilidad de que los hogares decidan enviar a sus hijos a la escuela.

Aunque para esta investigación es central una justificación teórica de efectos esperados en un programa transferencias en especie, su objetivo no es el cálculo de la probabilidad de asistencia, sino estimar el efecto directo en días de asistidos a la escuela de recibir el programa utilizando especificaciones de tipo lineal bajo Mínimos Cuadrados Ordinarios, con la técnica de variable instrumental.

---

<sup>26</sup> Para ejemplos de su utilización ver el clásico trabajo de Gertler y Glewwe (1989); o para una aplicación en el Ecuador revisar Younger, Villafuerte y Jara (1997).

### **4.3) Problemas econométricos**

El principal problema para la estimación del impacto del programa es la creación de un instrumento válido que permita corregir la endogeneidad presente en la variable de participación en el programa. Adicionalmente, es necesario comprobar una adecuada especificación para el modelo de asistencia escolar utilizado.

#### **4.3.1) Creación de variable instrumental**

Siguiendo las recomendaciones ya presentadas en el Capítulo 1, instrumentos adecuados para corregir el sesgo en la variable del programa son los criterios de focalización utilizados para la asignación del beneficio. Dadas las limitaciones de información, los criterios que pudieron ser calculados son: Índice de vulnerabilidad social (*ivs*); población bajo la línea de pobreza (*pobrez99*); porcentaje de población indígena en el cantón (*ptje\_ind*); y número de planteles unidocentes en el cantón (*red\_plan*).

Acorde con la ecuación de corrección del sesgo, el procedimiento necesario es regresar la variable del programa contra los criterios de focalización. Los valores así calculados son el instrumento a utilizar. En la literatura relacionada es frecuente la recomendación de utilizar un modelo tipo probit binomial. Sobre esta recomendación Angrist y Krueger (2001) critican la utilización de un modelo tipo probit para corregir el sesgo en regresor endógeno dado que podría tener resultados contrarios a lo esperado, en caso de que el modelo esté mal especificado o tenga una forma funcional incorrecta; por lo que recomiendan su comparación con estimaciones de tipo lineal MCO. Los resultados obtenidos nos permitieron considerar como mejor opción un modelo probabilística y como único instrumento el *ivs\_cant*. Los valores calculados (predichos) con este modelo



son almacenados en la variable citada, creando de esta forma la variable de participación corregida por sesgo ( $dpae\_prob$ )<sup>27</sup>

Corno indicamos en el Capítulo 1, Angrist y Krueger (2001) señalan que el sesgo que se busca corregir se podría elevar proporcionalmente al número de regresores en exceso que se utilicen para corregir la variable endógena, por lo que es necesario ser cuidadoso con la decisión de escoger el o los mejores instrumentos.

Para asegurar la adecuada elección del instrumento utilizamos una prueba propuesta por Baker (2000) utilizando únicamente  $ivs\_cant$ . Esta prueba comprueba mide la calidad de instrumento introduciendo los residuales del modelo de participación (b) en el modelo de asistencia escolar (a); en caso de que resulten significativos la estimación de impacto no podría realizarse dado que la correlación entre residuales sería positiva, invalidando uno de los supuestos básicos de un buen instrumento.

Los resultados al incluir los residuales del modelo de participación en el modelo de asistencia escolar si permitieron realizar la estimación y no resultaron estadísticamente significativos, por lo que se ratifica la validez de utilizar el Índice de Vulnerabilidad Social como único instrumento.

#### **4.3.2) Pruebas de especificación del modelo de asistencia escolar**

Antes de presentar las diferentes modelos para captar el impacto del programa, es necesario verificar si los patrones de asistencia escolar son similares a nivel nacional, o se ven afectados por la ubicación geográfica del hogar, en cuyo caso es necesario presentar las especificaciones diferenciando por área de residencia y región. Para comprobar

---

<sup>27</sup>

aplicaremos una extensión de la Prueba de significancia global de una regresión múltiple basada en el Análisis de Varianza <sup>28</sup>. Este test sigue una distribución F:

$$F = \frac{\text{Suma explicada de los cuadrados} / (k-1)}{\text{Suma de residuales al cuadrado} / (n-k)} = \sim F_{(k-1) : (n-k)}$$

Dado el modelo de regresión con k variables es el modelo A:

Regresión A: Nivel nacional (restringido)

$$\text{asistencia} = \beta_1 + \beta_2 x_2 + \dots + \beta_{16} x_{16} + \varepsilon$$

Regresión B: Área urbana (no restringido)

$$\text{asistencia} = \alpha_1 + \alpha_2 x_2 + \dots + \alpha_{16} x_{16} + \varepsilon$$

Regresión C: Área rural (no restringido)

$$\text{asistencia} = \lambda_1 + \lambda_2 x_2 + \dots + \lambda_{16} x_{16} + \varepsilon$$

Siendo  $H_0$ : los coeficientes de pendientes de las regresiones A B y C son simultáneamente iguales (por lo que sólo sería necesario un modelo a nivel nacional); y la hipótesis alternativa  $H_1$ : no todos los coeficientes de pendiente son simultáneamente iguales. La regla de decisión del test es: si  $F_{\text{obtenido}} > F_{\text{tabla}}$  entonces rechazamos  $H_0$ . Los valores  $F_{\text{obtenido}}$  (2.65) y  $F_{\text{de la tabla}}$  (1.57) al 95% de confiabilidad nos permiten rechazar la hipótesis nula  $H_0$ , por lo que es necesario utilizar un modelo por cada área.

También realizamos el mismo cálculo sobre las regiones del país, Costa y Sierra. En base a sus resultados es necesario diferenciar los impactos por

<sup>28</sup> Prueba F. Econometría Básica. Damodar Gujarati. Tercera Edición 1997. Capítulo 8.

región<sup>29</sup>. En base a los resultados de la prueba F presentamos cuatro modelos diferentes: Área urbana, Área rural, Región Sierra y Región Costa.

Finalmente utilizamos una prueba de correcta especificación del modelo. Este test desarrollado por Ramsey es conocido como Prueba del error de especificación en regresión<sup>30</sup>, esta basado en la Prueba F ya presentada, pero utilizando los  $R^2$  para decidir sobre  $H_0$ : el modelo esta mal especificado. Utilizando igual regla de decisión que en la prueba F, los valores obtenidos fueron: F calculado (11.24) y el F de la tabla de distribución (5.63) al 95% de confiabilidad. Por lo tanto  $H_0$  fue rechazada. El modelo utilizado si está bien especificado<sup>31</sup>.

#### 4.4) Especificación econométrica del modelo de asistencia escolar

Para definir de manera analítica el modelo de impacto del programa sobre la asistencia escolar, planteamos la siguiente relación teórica:

$$\begin{aligned} \text{Asistencia} = & \sum \alpha_i (\text{Variables del alumno}) + \sum \alpha_j (\text{Variables del hogar}) + \\ & \sum \alpha_k (\text{Variables de oferta}) + \sum \alpha_l (\text{Costo de la educación}) \end{aligned}$$

Donde la asistencia escolar es una función lineal de las características del niño, del hogar, de la oferta educativa y del costo de la educación. El modelo econométrico utilizado con Mínimos Cuadrados Ordinarios es el siguiente:

$$\begin{aligned} \text{asiste\_l} = & \alpha_1 \text{dpae} + \alpha_2 \text{grado} + \alpha_3 \text{trabaja} + \alpha_3 \text{camina} + \alpha_4 \text{ttmin} + \alpha_5 \text{areal} + \alpha_6 \text{dregion} \\ & + \alpha_7 \text{lnced\_hog} + \alpha_8 \text{lncons\_hog} + \alpha_9 \text{menos10} + \alpha_{10} \text{entre10\_15} + \alpha_{11} \text{ins\_jefl} \\ & + \alpha_{12} \text{ins\_mam} + \alpha_{13} \text{cateco\_p} + \alpha_{14} \text{nprof\_com} + \alpha_{15} \text{alupro98} + \alpha_{16} \text{alual98} + \mu_i \end{aligned}$$

<sup>29</sup> F obtenido (4.94) y F tabla (1.51). Anexo 3.

<sup>30</sup> Gujarati. Capítulo 13.

<sup>31</sup> Resultados en Anexo 3.

Donde consideramos válidos los supuestos de la regresión normal clásica sobre el comportamiento de los errores  $\mu_i$ , es decir, los errores siguen una distribución normal, con media 0 y varianza  $\sigma^2$ , simplificando:

$$\mu_i \sim N(0, \sigma^2)$$

#### **4.4.1) Revisión de variables**

##### ***Variable dependiente***

La asistencia escolar es una variable continua que asume valores entre 0 y 20, que es el rango de días posibles de asistencia en el último mes.

##### ***Variables explicativas***

###### **Características del niño**

El grupo de variables del alumno incluye una dummy que indica si recibe o no el programa; forma de traslado a la escuela (en la que captamos el efecto de trasladarse caminando); y la distancia en minutos hasta el establecimiento.

###### **Características del Hogar**

Las variables del hogar incluyen área de residencia; región; consumo, que es utilizado como su logaritmo natural para normalizar su comportamiento; dos variables de dependientes del hogar diferenciándolos por su edad (menores de 10 años y entre 10 y 15 años); la escolaridad del jefe del hogar; la escolaridad de la madre.

###### **Características de la oferta educativa**

Todas estas variables fueron calculadas como promedios a nivel cantonal. Utilizamos categoría económica docente; el número de profesores contratados por la comunidad (como una aproximación de la influencia

de los padres en el manejo de la escuela); la relación alumnos por profesor; y la relación alumno-aula.

**Cuadro No.13**

**VARIABLES UTILIZADAS EN EL MODELO DE ASISTENCIA ESCOLAR**

Nombre	Nombre en el modelo	tipo	Fuente	Método de construcción	Descripción
Participación sin corregir	d_pae	dummy	ECV99	Pregunta pe13	=1 recibe el programa =0 no
Participación corregida	dpae_prob	continua	Creada	Valores predichos de modelo de participación	
edad	edad	continua	ECV99	Consta en la encuesta pd03	edad del niño
grado	grado	continua	ECV99	Preguntas pe32 y pe33	grado en que se matriculo
Matriculado en primaria fiscal	primfiscal	continua	ECV99	Preguntas pe12 y pe15	valores entre 1 y 6
Beneficiario del programa	d_pae	dummy	ECV99	Creada a partir de pe13	1=beneficiario 0=no beneficiario
Asistencia escolar	asiste_1	continua	ECV99	Preguntas pe20 y pe22.	Casos entre 0 y 20 días al mes
Area de residencia	area1	dummy	ECV99	Creada a partir de area	1=área urbana 0=área rural
region	region	dummy	ECV99	Creada a partir de región	1=Costa 0=Sierra
consumo del hogar	lncons_hog	continua	ECV99	A partir de adjcap (consumo per cápita)	logaritmo del consumo para normalizar
instrucción del jefe de hogar	ins_jefe1	continua	ECV99	Preguntas pd04, pe32 y pe33	años de estudio aprobados por el jefe del hogar
instrucción de la madre	ins_mam	continua	ECV99	Calculada usando pd04, pe32 y pe33	años de estudio aprobados por la madre
Número de dependientes de entre 0-10 años	menos10	continua	ECV99	Calculada usando pd04 y edad	dependientes en el hogar de entre 0-10 años
Número de dependientes de entre 11-15 años	entre10_15	continua	ECV99	Calculada usando pd04 y edad	dependientes en el hogar de entre 0-10 años
Costo directo de educación	lnced_hog	continua	ECV99	Construida con variables desde pe24 hasta pe30	logaritmo del costo para normalizar
Trabajo infantil	d_trabaja	dummy	ECV99	Construida con variables pa01 y pa18	=1 el niño trabaja =0 no trabaja
Tiempo en minutos	ttmin	continua	ECV99	Construida a partir de pe18b	Tiempo en minutos traslado a la escuela
Traslado a pie	camina	dummy	ECV99	Construida a partir de pe19	=1 camina =0 otra opción
Categoría económica docente	cateco_p	continua	Censo del Magisterio	Censo del Magisterio	Promedio cantonal
Número de profesores contratados por comunidad	nprof_com	continua	Censo del Magisterio	Censo del Magisterio	Promedio cantonal
Promedio de alumnos por profesor	alupro98	continua	SIISE 1999	SIISE 1999	Promedio cantonal
Promedio de alumnos por aula	aluaul98	continua	SIISE 1999	SIISE 1999	Promedio cantonal

## **4.5) Resultados econométricos**

En el cuadro 14 presentamos los resultados de las estimaciones de los cuatro modelos considerados. Su discusión se presenta en dos partes. Primero el efecto promedio del programa, y después los resultados de los determinantes de la asistencia a la escuela por grupos de variables.

### **4.5.1) Resultados de la evaluación del programa**

La estimación del impacto es captada por la variable de participación en el programa (*dpae\_prob*). Su estimación presenta un parámetro negativo pero no significativamente diferente de cero para tres de los modelos considerados. Únicamente en el caso de la región costa el programa presenta un impacto estadísticamente significativo pero negativo (-8.26 días promedio). Dada la inestabilidad del coeficiente y de su nivel de significancia, podemos sugerir que para 1999 el PAE no presenta impacto alguno en asistencia escolar.

### **4.5.2) Resultados del modelo de asistencia**

#### ***Variables del niño***

La variable de forma de traslado al establecimiento presenta un efecto negativo promedio de -1.24 días en la región sierra, pero ningún efecto en los otros modelos. La distancia al establecimiento medida en tiempo de traslado en minutos, no presenta efecto en los días asistidos en ningún modelo. En cambio, la asistencia escolar presenta una relación positiva y significativa con el grado en que esta inscrito el niño (contribución de 0.24 región sierra, 0.09 región costa, y 0.19 en área urbana, en días promedio). Dado que este estudio no analiza los factores de matriculación escolar, este hallazgo positivo sobre grado y asistencia escolar podría indicar que, entre la población que efectivamente asiste a la escuela, la ampliación de los contenidos relacionada con grados superiores, motiva la asistencia

### ***Variables del hogar***

En el modelo del área urbana el coeficiente de región (=1 región Sierra) presenta un promedio de -3.75 días de asistencia para niños. Este resultado no es aislado. Como hemos señalado, la forma de traslado al establecimiento (caminando) para esta región también es negativa y significativa; por lo que asumimos que estas variables en conjunto explican que en esta región existiría un importante efecto de factores climáticos para la decisión de asistir al establecimiento.

El consumo del hogar presenta un comportamiento inesperado. Su incidencia es negativa y significativa para la región sierra y el área urbana, y no significativa para los otros modelos. En la región Sierra su efecto es una reducción de -1.25 días en promedio, y en el modelo para el área urbana es de -0.76 días. Otro resultado inesperado es el nulo efecto de la escolaridad de la madre y el jefe del hogar. Estudios sobre la demanda de la educación primaria para el Ecuador (Younger, Villafuerte y Jara 1997; Salazar (2002)) si encuentran significativos estos factores. Pero esta investigación no analiza probabilidades de matrícula, y sus resultados indican que mayores niveles de consumo y de escolaridad del jefe del hogar y de la madre, no contribuyen a explicar un día de asistencia adicional entre población que esté efectivamente matriculada en la primaria fiscal.

### ***Variables de oferta educativa***

Las relaciones alumno-profesor y alumno-aula presentaron valores significativos antes de introducir los demás grupos de variables, pero en los cuatro modelos presentados no tienen efecto alguno. La variable más relevante de este grupo es la categoría económica promedio de los profesores, que ha sido incluida como una aproximación de la

acumulación de experiencia docente. Su efecto resultó negativo y significativo para la región sierra; un resultados inicialmente no esperado que nos permite sugerir la existencia del agotamiento docente, dado que, generalmente durante sus años de trabajo, su dedicación ha estado centrada en una sola área temática. Lamentablemente no contamos con la información necesaria para sostener este criterio, pero es un primer indicio que requiere ser analizado posteriormente. El efecto que presentó para la región Sierra fue de -1.44 días promedio. Para la región costa (-0.65 días promedio (al 99% de confiabilidad); y para el área rural de -1.02 días en promedio. Para el área urbana su resultado también fue negativo pero no estadísticamente válido.

### ***Costo de la educación***

El costo directo y la variable utilizada como aproximación al costo indirecto presentan importantes resultados en los modelos por área de residencia, no así por región. Acorde con la literatura relacionada, el costo de oportunidad disminuye la asistencia escolar en las áreas agrícolas o rurales. El modelo del área rural confirma esta consideración. En promedio, si un niño del área rural trabaja, asiste 0.76 días menos. Para el área urbana esta variable no tiene efecto significativo. El resultado negativo que se esperaba para el costo directo de asistir a la escuela no se confirmó en el área urbana (efecto positivo de 0.7 días promedio), dado que la investigación se concentró en los últimos 20 días de asistencia en una muestra tomada a mediados del año escolar, y no al inicio donde los costos directos son más altos (matrícula, uniformes, entre otros).



## Cuadro No.14

## Resultados del impacto en asistencia escolar del PAE 1999

	Región Sierra	Región Costa	Área Urbana	Área Rural
<b>Variable dependiente</b>				
asistencia escolar último mes	Coef.	Coef.	Coef.	Coef.
constante	44.33911 (9.120699) ***	25.27229 (3.373829) ***	29.77405 (8.017088) ***	37.42544 (7.073833)
<b>Variables del niño</b>				
dpa <sub>a</sub> _prob (participación corregida)	-8.565739 (7.794462)	-8.263142 (3.411938) **	-1.280715 (6.128787)	-6.283999 (8.286224)
gracia	0.2482465 (0.107481) **	0.0922239 (0.046597) **	0.1901615 (0.074374) **	0.1264024 (0.089146)
camina hasta la escuela	-1.243782 (0.6696618) *	0.4510283 (0.574630)	-0.4095111 (0.530231)	-0.4245869 (0.795326)
tiempo en minutos	0.0148515 (0.0233517)	0.000485 (0.019329)	0.0279516 (0.022146)	-0.0016843 (0.021735)
<b>Variables del hogar</b>				
area de residencia	0.4424545 (1.38092)	0.1740182 (0.266860)		
región			-3.759562 (1.019338) ***	-2.47812 (1.919122)
consumo del hogar	-1.254392 (0.5033522) **	0.0360588 (0.179545)	-0.7626409 (0.370325) **	-0.4910494 (0.444737)
dependientes menores de 10 años	-0.1550743 (0.2304956)	-0.0902039 (0.124270)	-0.1650139 (0.120399)	-0.0720738 (0.186542)
dependientes entre 10 y 15 años	-0.1413798 (0.298343)	0.0143556 (0.090109)	-0.2617702 (0.216009)	0.1144228 (0.209897)
escolaridad jefe de hogar	-0.0219227 (0.0958864)	0.0433433 (0.029072)	0.0085414 (0.050995)	0.0157936 (0.084019)
escolaridad de la madre	0.0332408 (0.0883097)	-0.0092948 (0.040129)	0.0011539 (0.055408)	0.0518751 (0.073720)
<b>Variables de oferta educativa</b>				
categoría económica docente	-1.446223 (0.6328527) **	-0.6574716 (0.247594) ***	-0.918018 (0.633716)	-1.028803 (0.536689)
número de profesores de la comunidad	-0.0111321 (0.0099968)	-0.0029891 (0.001715) *	-0.0015006 (0.004917)	-0.0022971 (0.013873)
relación alumno/profesor	-0.1622514 (0.1335081)	-0.0058316 (0.015408)	0.0639641 (0.063487)	-0.0369667 (0.031800)
relación alumno/aula	0.1164592 (0.1073857)	0.0325163 (0.016941)	-0.0115405 (0.048741)	0.0606342 (0.068923)
<b>Variables de costo de educación</b>				
el niño trabaja (costo indirecto)	-0.8741458 (0.5853139)	-0.2373393 (0.279782)	0.2713818 (0.558109)	-0.762115 (0.428051)
costo de la educación	0.4351231 (0.3824756)	-0.0195373 (0.125250)	0.7024399 (0.238634) ***	-0.2367684 (0.289701)
R2	0.0432	0.0325	0.1625	0.1186
Observaciones	1594	1397	1195	1796

Errores estándar corregidos por diseño muestral en paréntesis

\*significativo al 1%, \*\* al 5%. \*\*\* al 10%

#### 4.6) Conclusiones

1. Esta evaluación basada en la técnica de variable instrumental ha tratado de corregir el principal problema econométrico conocido en la literatura de evaluación, esto es, el sesgo de selección, e incluye pruebas de validez del instrumento utilizado. Sus resultados permiten sugerir que el impacto del Programa de Alimentación Escolar en asistencia escolar es estadísticamente nulo.
2. La corrección del sesgo de selección por características observables fue posible utilizando el criterio de focalización Índice de vulnerabilidad social. El modelo de mejor ajuste para la corrección de la variable de participación fue escogido en una comparación entre especificaciones MCO y Probit. Las pruebas realizadas al instrumento demostraron que no presentaba correlación con la asistencia escolar ni con sus residuales validando su utilización. Una contribución de este estudio es la comprobación empírica de que para minimizar este sesgo es necesario que el número de regresores (instrumentos) sea igual al número de variables endógenas que se pretende corregir (Angrist y Krueger 2001).
3. El modelo de asistencia escolar utilizado consideró cinco grandes grupos de factores explicativos, estos son: características del niño; características del hogar; variables de la oferta educativa; y el costo de la educación, (que utilizó los costos directos y una dummy de trabajo infantil como aproximación al costo de oportunidad). Se utilizó la prueba Reset para confirmar su adecuada especificación. Como resultado de la adaptación de una prueba de especificación de modelos (Prueba F), la ecuación de asistencia escolar se presentó en

cuatro modelos diferentes: región costa, región sierra, área rural y área urbana.

4. Del modelo de asistencia escolar los factores de mayor incidencia positiva son: el grado en que está matriculado el niño, y el costo de la educación en el área urbana. Entre los factores de incidencia negativa tenemos: la forma de traslado a la escuela (en el modelo de la región sierra); la antigüedad de los profesores (en los modelos por región y el área rural); y el trabajo infantil (en el modelo de área rural).