

Cartagena de Indias: Castillo de San Fernando

El logro académico y el efecto colegio en las pruebas Icfes en Antioquia

Piedad Patricia Restrepo
Mauricio Alviar

El logro académico y el efecto colegio en las pruebas Icfes en Antioquia

Piedad Patricia Restrepo y Mauricio Alviar

Lecturas de Economía, **60** (enero-junio 2004), pp. 67-95.

Resumen: *esta investigación muestra que, a través de un análisis del diseño institucional e implementación, el programa de bonos educativos Paces en Colombia, en el periodo 1992-1998, estuvo lejos de cumplir sus objetivos de corto y largo plazo. La no inclusión dentro del valor del bono del costo de oportunidad por los ingresos laborales perdidos, la escasa regulación y monitoreo de los colegios participantes, en especial de aquellos que surgieron a raíz del programa, son algunos de los factores que incidieron en sus resultados. Al final, se presentan algunas propuestas para mejorar el diseño de este tipo de programas que permitan alcanzar los objetivos planteados en cuanto a cobertura, calidad y equidad.*

Palabras clave: *bonos educativos, diseño institucional, incentivos, participación, focalización.*

Abstract: *this research show, through an analysis of the institutional design and the implementation of the program, that the program for educational aids Paces, in Colombia, in the 1992-1998 term, did not accomplish its short-term and long-term objectives. The fact that the opportunity cost for the lost revenues was not included in the value of the "bono" and the poor control and monitoring of the intervening schools, particularly those that were created with the program, are some of the issues that had an effect on the results. At the end, we present some proposals for a better design of this type of programs that makes possible reaching the objectives proposed as to "coveratures", quality, and equity.*

Key words: *educational aids, institutional design, "incentives", participation, focalization*

Résumé: *Cette recherche montre que, à travers une analyse de la conception institutionnelle et de la mise en oeuvre, le programme d'allocations d'éducation Paces en Colombie, dans la période allant de 1992 à 1998, il a été loin d'accomplir ses objectifs à court terme et à long terme. La non inclusion dans la valeur de l'allocation du coût d'opportunité pour les revenus de travail perdus, le faible règlement et la surveillance des écoles participantes, spécialement de ceux qui sont apparus à partir du programme, sont quelques facteurs qui ont influencé leurs résultats. Finalement, on présente quelques propositions pour améliorer la conception de ce type de programmes qui permettent d'atteindre les objectifs fixés quant à la couverture, la qualité et l'équité.*

Mots clés: *allocations d'éducation, conception institutionnelle, motivation, participation, focalisation.*

El logro académico y el efecto colegio en las pruebas Icfes en Antioquia

Piedad Patricia Restrepo y Mauricio Alviar*

-Introducción. -I. Metodología. -II. Estadísticas descriptivas de la muestra. -III. Resultados. -IV. Una primera aproximación a las diferencias subregionales. -V. Conclusiones y recomendaciones. -Anexos. -Bibliografía.

Primera versión recibida en diciembre de 2003; versión final aceptada en agosto de 2004 (eds.).

Introducción

Desde hace casi 40 años se iniciaron los estudios sobre la eficacia escolar, entendida esta como la capacidad que tienen algunos colegios de promover mejores resultados para sus estudiantes frente a otros. El estudio pionero fue el llamado informe Coleman, el cual estimó un modelo microeconómico de función de producción mediante la técnica de regresión múltiple y llegó a la conclusión de que los factores escolares no son explicativos de la variabilidad del logro de los estudiantes. Coleman encontró que un aumento del gasto en insumos, como bajar el tamaño de la clase o aumentar el salario de los profesores no era significativo para el logro de los estudiantes (Card and Krueger, 1996). A partir de allí se han llevado a cabo numerosos estudios, en especial en los países desarrollados, con resultados poco concluyentes.

* Piedad Patricia Restrepo Restrepo: Investigadora, Centro de Investigaciones Económicas –CIE–, Universidad de Antioquia. Ciudad Universitaria, Bloque 13, Apartado aéreo 1226, Medellín, Colombia. Dirección electrónica: pprestrepo@agustinianos.udea.edu.co. Mauricio Alviar Ramírez: Investigador, Centro de Investigaciones Económicas –CIE–, Universidad de Antioquia. Ciudad Universitaria, Bloque 13, Apartado aéreo 1226, Medellín, Colombia. Dirección electrónica: malviar@agustinianos.udea.edu.co.

Esta investigación fue financiada con recursos de la Universidad de Antioquia (Comité para el Desarrollo de la Investigación) y la Escuela de Gobierno y Políticas Públicas de Antioquia. Se agradece la asistencia de Ana Isabel Moreno, estudiante en formación del programa de Economía de la Universidad de Antioquia y los valiosos comentarios de Jorge Barrientos-Marín y los evaluadores anónimos del artículo. Todos los errores son responsabilidad exclusiva de los autores.

Así mismo, ha habido una evolución de los modelos, las técnicas de análisis, las variables, los factores estudiados y los instrumentos de recolección de datos. Pero la gran novedad en la investigación sobre la eficacia escolar ha sido la aparición de los modelos jerárquicos lineales o multinivel que han sido desarrollados, en gran medida, por las necesidades metodológicas de la investigación en educación (Murillo, 1999).¹

En Colombia, solo recientemente, se han llevado a cabo estudios sobre eficacia escolar. El objetivo central ha sido cuantificar el efecto del colegio sobre la variabilidad del logro académico de los estudiantes explicada por factores asociados a las características del colegio. En segunda instancia, también se ha tratado de precisar los efectos individuales de las características del estudiante, su familia y el colegio. Todos los trabajos han sido de carácter extensivo, es decir, con muestras grandes de alumnos y de centros, en su gran mayoría para Bogotá. En cuanto a la metodología, la mayoría de los trabajos ha utilizado los modelos jerárquicos lineales, a excepción del trabajo de Gaviria y Barrientos (2001) que utiliza el modelo tradicional de regresión lineal múltiple y un modelo de efectos fijos el cual absorbe un factor categórico, en este caso el colegio, para medir la correlación intraclase, y para estimar el efecto de las variables personales y familiares una vez se descuenta la varianza entre colegios.

Para el departamento de Antioquia aún no se había llevado a cabo ningún estudio de este tipo, y justamente esta investigación pretende responder los interrogantes básicos de cualquier estudio sobre calidad educativa, máxime cuando ha habido en los últimos años una preocupación creciente por el pobre desempeño de los estudiantes antioqueños en las pruebas académicas administradas por el Instituto Colombiano para el Fomento de la Educación Superior —Icfes—. Aunque el proceso educativo es complejo y aún no se tiene total claridad de su evolución en Antioquia, este estudio es un primer y valioso paso para la comprensión de los factores determinantes de la eficacia escolar en el departamento y va más allá del enfoque de función de producción planteado en el informe Coleman.²

-
- 1 Uno de los primeros trabajos aplicados que utiliza esta técnica es el de Bryk y Raudenbush (1986), en el cual estudian la efectividad relativa de las escuelas privadas y públicas, motivados principalmente por la discusión en torno a la supuesta superioridad en el logro de las primeras sobre las segundas.
 - 2 En general, el enfoque de la función de producción ha sido criticado por tener un carácter fuertemente empírico, sin apoyo en una teoría sólida. Esta es una de las razones por las cuales normalmente queda sin explicar un residuo muy amplio de las variaciones de la calidad. Adicionalmente, la mayoría de los estudios mide la calidad en un determinado momento del tiempo, a pesar de que el proceso educativo es esencialmente acumulativo y los insumos utilizados

I. Metodología

En el estudio de los determinantes de la eficacia escolar, hasta hace muy poco tiempo, la metodología utilizada era el análisis de regresión multivariado. En la literatura se reconocen dos formas de abordarlo. De un lado, la desagregación, que consiste en llevar todas las variables al nivel del individuo, y la otra opción utilizada es la agregación, que consiste en promediar todas las variables al nivel de grupo o colegio. Ninguna de estas opciones resulta óptima para hallar los parámetros de interés. En el primer caso, se viola el supuesto de independencia de las observaciones al desconocer el fenómeno de anidación del proceso educativo. Los alumnos que pertenecen al mismo grupo o al mismo colegio, comparten ciertas características que influyen en el resultado, usualmente medido a través de los puntajes en pruebas estandarizadas, como proxy de las de la eficacia escolar. Al violar el supuesto de independencia, la estimación de los errores estándar no es óptima. En el segundo caso, la agregación produce la pérdida de información valiosa.

Los modelos jerárquicos o multinivel, se han convertido en una alternativa superior en los estudios de la eficacia escolar, y en general, en fenómenos de cualquier índole en los cuales se presente anidación;³ esto es, los individuos pertenecen a grupos, y esos grupos a su vez pertenecen a unidades de mayor tamaño, y así, sucesivamente. El modelo jerárquico más complejo de dos niveles asume que:

En el primer nivel:

$$Y_{ij} = a_j^T X_{ij} + e_{ij} \quad (1)$$

Y en el segundo nivel

$$a_j = \gamma^T Z_j + \mu_j \quad (2)$$

Donde:

Y_{ij} es el puntaje del individuo i en el colegio j .

X_{ij} es una matriz de características del individuo y su entorno familiar (nivel uno).

en el pasado afectan el desempeño actual de los alumnos. Por último, es necesario recordar que el logro de los alumnos no necesariamente se mide por completo mediante los puntajes en pruebas estandarizadas. La educación debe y de hecho responde a una formación mucho más integral. Sin embargo, estas pruebas logran captar ciertas habilidades valoradas en el mercado laboral y ese es uno de los principales objetivos de la educación, preparar a las personas para ser más productivas en la sociedad mediante su vinculación al mercado laboral. Más aún, hay evidencia de que los puntajes en pruebas están altamente relacionadas al desempeño en el mercado laboral (Hanushek, 1996).

3 Las aplicaciones incluyen estudios de crecimiento, efectos organizacionales y síntesis investigativas.

α^T es el vector de coeficientes asociados a esas características que incluye la constante.

Z_j es un vector de características del colegio (nivel dos).

e_{ij} es el término de perturbación. Se asumen con media cero y varianza σ^2

γ^T es el vector de coeficientes de las variables del colegio.

μ_j es el término de error residual a nivel de colegio.

En el nivel uno se tienen las características del individuo y de su entorno familiar únicamente, pero a diferencia del modelo de regresión múltiple se asume que cada colegio tiene su propio intercepto y pendiente, por ello el subíndice j en el vector alfa. Y en el nivel dos, se reconoce que el aporte de las características del nivel uno a la variable de interés, en este caso al puntaje en las pruebas, depende de las variables del colegio.

Reemplazando la ecuación 2 en 1, para P variables de nivel uno y Q variables del nivel dos, se obtiene la ecuación de regresión completa:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} Z_{qj} X_{pij} + \mu_{pj} X_{pij} + \mu_{oj} + e_{ij} \quad (3)$$

La ecuación 3 muestra la interacción entre niveles, que expresa el efecto moderador de las variables de segundo nivel sobre la relación entre las variables X sobre Y. De otro lado, el término de error claramente es más complejo en el modelo jerárquico; μ_{pj} depende de las características de los individuos, en otras palabras no cumple el supuesto de homocedasticidad con el cual trabaja el modelo tradicional de regresión múltiple, según el cual todos los errores son independientes de todas las variables explicativas.

Las ventajas que ofrece esta metodología no solo son estadísticas sino también conceptuales frente a la tradicional estimación por mínimos cuadrados ordinarios, además de que permite explorar y responder muchas más preguntas que el análisis tradicional. Para un mayor detalle sobre los modelos tanto en la parte teórica como empírica véase Bryk y Raudenbush (1992) y Hox (1995).

II. Descripción de la muestra

La muestra utilizada corresponde a las pruebas de estado aplicadas por el Icfes (en adelante, pruebas Icfes) para el año 1999.⁴ En total, la muestra está compuesta

4 Podría llegar a pensarse que los datos están desactualizados, no obstante la calidad de los sistemas educativos cambia lentamente en el tiempo. Algunas de las razones para ello es que hay estacionalidad de la tecnología de enseñanza y los maestros cambian poco (Hanushek y Kimko, 2000).

por 26.136 estudiantes, distribuidos en 542 colegios, para un promedio de 48 estudiantes por colegio. Fueron excluidos los estudiantes mayores de 30 años, aquellos que estaban validando el bachillerato y los estudiantes provenientes de colegios que contaban con menos de 10 estudiantes en la muestra. También fueron eliminadas aquellas observaciones con valores missing para cualquier variable dentro del modelo. El mayor porcentaje de valores missing (15%) lo presentó la variable correspondiente a la profesión del padre.

La media del puntaje total fue de 252 puntos sobre 400, mientras que para las pruebas individuales de conocimiento matemático, aptitud matemática y lenguaje fueron de 51, 49 y 52, respectivamente.⁵ En cuanto a las características de los estudiantes, el 41% son hombres, la edad promedio es de 16,9 años, solo el 5,7% trabaja y el 38,5% son hermanos mayores. Dentro de las características familiares tenidas en cuenta, el número de hermanos promedio es de 2,6. La educación de padre y madre es prácticamente igual, 8 años en promedio; por su parte la educación máxima promedio para los padres alcanzó los 9,1 años. Los padres con las ocupaciones mejor remuneradas⁶ representan el 13,7% y el 35,5% de las madres trabajan.

En cuanto a los colegios incluidos en la muestra, el 76% son de carácter académico, el 18,5% de jornada completa, el 75,8% son mixtos y el 67,8% son públicos. La variable calendario no fue tenida en cuenta, puesto que el 99% de las observaciones pertenecían al calendario A (véase Tabla 1). El Valle de Aburrá representa la mayoría de la muestra con el 67%, de la cual Medellín aporta el 43%. Le siguen en su orden, las subregiones de Oriente (11%), Suroeste (6%), Urabá (4%), Norte, Nordeste y Bajo Cauca con 3% cada uno, Occidente (2%) y Magdalena Medio (1%).

III. Resultados

A. Efecto colegio o correlación intraclase

Desde el punto de vista de política pública es fundamental determinar si el colegio está aportando al rendimiento de los estudiantes, medido por los resultados en las pruebas Icfes, y de cuanto es ese aporte. Desde el informe Coleman, llevado a cabo en Estados Unidos en la década del sesenta, el cual determinó, utilizando la técnica tradicional de regresión múltiple, que el colegio no era significativo, hubo un fuerte escepticismo sobre lo que las reformas educativas a nivel escolar podrían

5 Estas pruebas individuales fueron escogidas porque son las que más se asocian con la productividad y el crecimiento de los países, además de ser las más evaluadas con fines comparativos en pruebas internacionales (Lee y Barro, 1997, pp. 8-12).

6 A ellas pertenecen los empresarios, administradores o gerentes, profesionales independientes y profesionales empleados.

Tabla 1. *Estadísticas descriptivas de la muestra*

Variable	Media	Desviación estándar	Mínimo	Máximo
Total	252,77	42,85	152	386
Aptitud matemática	51,15	9,93	22	80
Conocimiento matemático	49,38	10,05	25	80
Lenguaje	52,33	10,13	24	80
Género (hombre)	0,41			
Edad	16,9	1,8	13	30
Trabaja	0,057			
Número hermanos	2,6	1,8	0	9
Mayor hermanos	0,38			
Educación-madre	8,17	3,94	0	17
Educación-padre	8,05	4,25	0	17
Edumax	9,13	4,08	0	17
Profp.	0,137			
Madnt	0,64			
Académico	0,76			
Completa	0,18			
Mixto	0,75			
Público	0,67			

No. de observaciones: 26.136.

Fuente: Icfes. Cálculos propios.

representar.⁷ Solo casi dos décadas después con la investigación de Heyneman y Loxley (1983) se llama la atención sobre las consecuencias de generalizar resultados de países desarrollados a aquellos de menor ingreso per cápita. Ellos encuentran en estos últimos un mayor efecto de las características del colegio sobre el logro escolar, y concluyen que la principal influencia sobre el aprendizaje escolar es la calidad de los colegios y profesores a los cuales están expuestos los estudiantes.⁸

En Colombia, Castaño (1998) mide por primera vez el efecto colegio. Para ello estima el modelo jerárquico más simple —análisis de varianza de una vía con

7 Un resumen de los resultados para Estados Unidos se encuentra en Hanushek (1996).

8 Vea una mayor discusión sobre la relación entre recursos escolares y el logro escolar en Lee y Barro (1997) pp. 7-8.

efectos aleatorios— también llamado intercept only model (I.O.M), teniendo como variable dependiente el logro en matemáticas para la prueba Saber de 1993, y encuentra que este efecto es del 29%.⁹ A partir de allí, se han desarrollado otros trabajos con la misma metodología pero adicionando otras variables para llegar a modelos más completos y responder preguntas adicionales.¹⁰ Sin embargo, desde antes de la introducción de los modelos jerárquicos, la literatura había mostrado varias formas de estimar la correlación intraclase. Justamente, Gaviria y Barrientos (2001) emplean una metodología derivada del trabajo de Kremer y Maskin (1996), en la cual el Rajustado de una regresión cuya variable dependiente es el puntaje en las pruebas y cuyas variables independientes comprenden una variable ficticia por cada colegio, representa la correlación intraclase o efecto colegio. Utilizando las pruebas Icfes de 1999 para Bogotá, Gaviria y Barrientos (2001) encuentran que el efecto colegio es de 40% para la prueba total, de 34% para conocimiento matemático, 27,7% para lenguaje y de 25% para aptitud matemática. Los autores concluyen que: «Este resultado no parece depender de la metodología empleada, ni del tipo de pruebas utilizadas, ni del grado de los alumnos bajo escrutinio, lo que permite concluir que en Colombia dos tercios de la varianza del logro académico es independiente de las características del plantel».

Aplicando ambas metodologías, regresión múltiple y modelos jerárquicos, a una misma muestra se puede clarificar aun más el punto anterior. Siguiendo la metodología utilizada por Gaviria y Barrientos (2001) para la muestra de Antioquia se encuentra lo siguiente: el efecto colegio para la prueba total fue del 37%, mientras que para conocimiento matemático, lenguaje y aptitud matemática fue de 33,1%, 28,9% y 23,1%, respectivamente. Con respecto a los resultados para Bogotá obtenidos por Gaviria y Barrientos (2001) se mantiene el orden de importancia por prueba del efecto colegio. En la prueba de aptitud matemática las características del colegio tienen menos incidencia en el logro y cobran mayor importancia los factores personales o familiares.¹¹ En todas las pruebas, a excepción de lenguaje, en Antioquia es levemente menor el efecto colegio que en Bogotá. Ahora bien, al utilizar el

9 Este trabajo fue realizado en 1997 para la Misión Social del Departamento Nacional de Planeación.

10 Véase Sarmiento y González (1997), Caro (2000) y Sarmiento, Becerra y González (2000). Los resultados en la medición del efecto colegio son muy similares en cada caso.

11 Las pruebas de aptitud están estrechamente relacionadas con habilidades innatas del individuo y con habilidades promovidas en los primeros años de aprendizaje, ya sea desde el hogar o el colegio. En los últimos años de secundaria parece ser que el colegio puede aportar más para pruebas de conocimiento que de aptitud.

modelo jerárquico sin variables explicativas (I.O.M) se encuentran resultados altamente similares a los obtenidos con la primera metodología (véase Tabla 2).

Tabla 2. *Efecto Colegio, o correlación intraclase para las pruebas Icfes en Antioquia, 1999*

Pruebas	Metodología	
	Kremer y Maskin (porcentaje)	Modelo jerárquico I.O.M (porcentaje)
Total	37	36
Conocimiento matemático	33,1	32,2
Lenguaje	28,88	28,33
Aptitud matemática	23,06	22,94

Fuente: Cálculos propios.

En términos metodológicos, lo anterior evidencia que la medición del efecto colegio puede llevarse a cabo con regresión multinivel o regresión múltiple, usando variables dummy para la categoría de interés, ofreciendo resultados muy similares. Pero más importante aún, de nuevo se evidencia que las características del colegio importan en la explicación de la variabilidad del logro, con lo cual se abre un espacio para la definición de políticas educativas encaminadas a brindar las mismas oportunidades para todos los niños y jóvenes que acceden al sistema educativo.

Un estudio para Argentina (Llach, Montoya y Roldán, 1999) estableció que para la secundaria, el porcentaje explicado por las variables propias del colegio es del 38%, muy similar al encontrado para Colombia; mientras que Gaviria y Barrientos (2001) muestran un efecto colegio del 20% para Estados Unidos. Dichos resultados son acordes con el estudio de Heyneman y Loxley (1983), el cual mostraba que el efecto colegio es mucho mayor para los países de menores ingresos, en dos y hasta casi tres veces. En ellos aún existen condiciones muy diversas para los colegios, no solo en términos de insumos sino también de organización, que explican en buena medida las marcadas diferencias en el logro escolar.¹²

¹² En el caso de Estados Unidos diversas investigaciones han encontrado que los colegios católicos obtienen mejor logro promedio que los colegios públicos. Según Hanushek (1996) la razón para esta diferencia es que la provisión pública enfrenta una mínima competencia y no posee un sistema de incentivos que premie el mejor desempeño de sus miembros y permita un mejor uso de los recursos.

B. Aportes de las características personales, familiares y del colegio

La literatura sobre logro escolar reconoce cuatro grandes grupos de variables que afectan el desempeño escolar: 1) los factores personales como género, edad y habilidades innatas; 2) factores familiares como la educación de los padres, el nivel de ingreso, el número de hermanos, entre otros; 3) factores propios del colegio y de sus profesores; 4) factores propios del lugar de residencia, ya sean por país, departamento o municipio.

En este estudio se incluyen principalmente los tres primeros grupos. Aunque se pretendía incluir variables propias de los municipios, o de las subregiones en su defecto, es decir, realizar un análisis jerárquico de tres niveles, los datos disponibles no lo permitieron.¹³ No obstante, en el análisis tradicional de regresión múltiple se incluyeron variables ficticias por subregión como una forma de aproximarse a posibles diferencias subregionales en la explicación de la variabilidad del logro; o lo que es igual, para detectar la posibilidad de que existan efectos únicos y significativos asociados a las subregiones.

Todas las variables son extraídas del formulario de inscripción que completan los estudiantes para presentar las pruebas Icfes. En el primer grupo de factores se incluyeron el género, la edad, si el estudiante es el mayor de los hermanos, y si el estudiante trabaja o no. En el segundo grupo se incluyeron el número de hermanos, la educación máxima entre padre y madre, la ocupación del padre que permite inferir cuales están mejor remunerados, y si la madre es ama de casa. En el tercer grupo se incluyen la jornada (mañana, tarde, noche y completa), la modalidad (académico, técnico y normalista), el género (mixto, femenino y masculino) y la naturaleza (público y privado).

Los resultados para la regresión tradicional muestran que a excepción de si el individuo trabaja o no y la modalidad del colegio, todas las variables resultan significativas. En cuanto a la primera variable, aunque en las primeras dos estimaciones presenta un efecto negativo y significativo, en la tercera estimación pierde significancia.¹⁴ Esto implica que a los individuos que dedican algo de su

13 Más del 50% de los municipios solo contaban con un colegio dentro de la muestra, lo cual impide llevar a cabo el análisis de tres niveles, para el cual se debe garantizar un mínimo de observaciones en cada nivel que garanticen poder llevar a cabo las estimaciones con confiabilidad. En el caso de las subregiones también resulta poco adecuado, ya que ese tercer nivel solo lo conformarían nueve unidades.

14 La primera estimación incluye las variables personales, la segunda incluye además las familiares; y la tercera incluye las dos anteriores más las variables del colegio (véase Tabla 3).

tiempo a trabajar no les representa costos en términos de menor logro educativo.¹⁵ Por otro lado, si el colegio es académico o no, resulta no tener ninguna injerencia sobre los resultados de las pruebas (véase Tabla 3).

Tabla 3. *Resultados de la regresión por M.C.O para la prueba total*

VARIABLES explicativas	(1)	(2)	(3)
Género	11,66 (5,52)	10,42 (6,79)	14,08 (13,92)
Edad	-39,53 (23,23)	-31,13 (20,05)	-29,19 (21,35)
Edad ²	0,84 (20,84)	0,67 (18,37)	0,64 (19,87)
Trabaja	-6,06 (-4,39)	-2,61 (2,31)	-0,98* (0,88)
Mayor hermanos	6,46 (11,57)	1,03 (1,96)	-0,47 (2,86)
Número hermanos		-1,3 (6,78)	1,43 (2,95)
Educación máxima		2,21 (13,91)	1,28 (11,8)
Profesión padre		17,65 (9,02)	11 (7,62)
Madre no trabaja		-4,64 (8,26)	-3,94 (7,79)
Académico			1,03* (0,79)
Completa			7,73 (4,15)
Público			-16,26 (1,763)
Mixto			-20,64 (9,39)
R ²	0,13	0,24	0,32

No. de observaciones: 26.136.

Valor absoluto de los t-estadísticos en paréntesis.

* coeficiente no significativo al 10%.

Los errores estándar son los Huber-White, que corrigen por heterocedasticidad y además permiten relajar el supuesto de independencia de las observaciones.

Fuente: Cálculos propios.

15 Que en las primeras dos estimaciones resulte significativo y en la tercera no, está reflejando que su significación está recogiendo el efecto de las características del colegio que se obvian en las primeras estimaciones. Los estudiantes que trabajan tiene una mayor participación en colegios públicos, mixtos y nocturnos, en los cuales los resultados promedio son inferiores al resto de colegios.

Los hombres obtienen resultados superiores en 14 puntos, de un máximo de 400 para la prueba total, que las mujeres. Y los resultados para los estudiantes de más edad son menos favorables. Sin embargo, la relación entre esta variable y el puntaje total no es lineal, como se observa en el Tabla 3, la variable edad al cuadrado resulta significativa y positiva, esto implica que a medida que aumenta la edad del individuo se obtienen resultados menos negativos que los obtenidos si mediara una relación lineal. Ser el mayor de los hermanos resulta positivo, mas no así el número de hermanos. Lo primero reporta 1,4 puntos más que no serlo, mientras que por cada hermano adicional disminuye el logro en casi medio punto. Ser el mayor de los hermanos puede traer ventajas en cuanto los padres pueden priorizar la educación para el hijo primogénito, ya sea dedicando más tiempo y atención, o pagando un mejor colegio. Entre mayor es el número de hermanos menor es el gasto educativo per cápita por hijo, y menor también el tiempo dedicado a cada uno.

Por su parte, la educación máxima de los padres aporta por cada año adicional 1,3 puntos. Así mismo, las ocupaciones mejor remuneradas aportan 11 puntos adicionales. Mientras que los estudiantes cuyas madres no trabajan, obtienen casi cuatro puntos por debajo frente a aquellos que tienen madres con ocupación por fuera del hogar. Lo anterior podría estar mediado por el nivel educativo de las madres, esto es, a medida que las madres tienen mayor nivel educativo crece la probabilidad de ingresar al mercado laboral.¹⁶ Esto tiene dos implicaciones, de un lado, madres más educadas pueden ayudar más a sus hijos en su aprendizaje, y de otro lado, las madres que trabajan aportan ingresos al hogar que pueden destinarse a brindar a los hijos educación de mejor calidad, mediante el ingreso de sus hijos a mejores colegios. Ya que se encontró una correlación negativa aunque no muy elevada entre la educación de las madres y su ocupación (-0,33), es muy posible que el segundo mecanismo este operando más que el primero.

En cuanto a las características del colegio, la jornada completa aporta 7,7 puntos más, frente a las otras jornadas. Mientras que los estudiantes que asisten a colegios públicos y mixtos obtienen en promedio 16 y 20 puntos por debajo de los colegios privados y los femeninos y masculinos, respectivamente.

Las regresiones para las pruebas específicas en matemáticas y lenguaje muestran resultados muy similares a los de la prueba total (véase Tabla 4). En lenguaje son idénticos, esto es, las variables no significativas son si el individuo trabaja o no y la modalidad del colegio. Los hombres se siguen desempeñando mejor, y la edad, las madres amas de casa y los colegios mixtos y públicos afectan

16 Se encuentra una correlación negativa entre las variables madre no trabaja y su nivel educativo, aunque no muy grande en magnitud (-0,33).

negativamente el desempeño junto con un mayor número de hermanos. Así mismo, la educación de los padres, la ocupación del padre, ser el mayor de los hermanos y pertenecer a un colegio de jornada completa mejora el desempeño. Por su parte, para las pruebas de conocimiento matemático y aptitud matemática se mantienen estos mismos resultados a excepción de las variables número de hermanos y mayor de los hermanos, las cuales pierden importancia.

Tabla 4. *Determinantes del logro en las pruebas individuales*

Variables	Aptitud matemática	Conocimiento matemático	Lenguaje
Género	3,83 (18,62)	3,78 (14,39)	1,71 (8,15)
Edad	-6,6 (21,8)	-6,21 (19,03)	-6,44 (19,47)
Edad ²	0,14 (19,58)	0,13 (17,77)	0,14 (17,67)
Trabaja	0,39* (1,50)	-0,23* (0,91)	-0,41* (1,5)
Mayor hermanos	-0,008* (0,07)	0,13* (1,14)	0,43 (3,59)
Número hermanos	-0,03* (0,94)	0,035* (0,96)	-1,63 (4,24)
Educación máxima	0,19 (8,88)	0,23 (9,25)	0,3 (12,01)
Profesión padre	1,99 (7,01)	2,73 (7,98)	2,11 (7,13)
Madre no trabaja	-0,71 (5,61)	-0,39 (3,12)	-1,07 (8,51)
Académico	-0,18* (0,67)	0,28* (0,81)	0,03* (0,12)
Completa	1,46 (4,11)	1,64 (3,62)	1,09 (2,97)
Público	-3,04 (9,04)	-3,72 (8,8)	3,25 (8,84)
Mixto	-3,39 (8,34)	4,43 (8,36)	-4,13 (9,18)
R ²	0,22	0,26	0,26

No. de observaciones: 26.136.

Valor absoluto de los t-estadísticos en paréntesis.

*coeficiente no significativo al 10%.

Los errores estándar son los Huber-White, que corrigen por heterocedasticidad y además permiten relajar el supuesto de independencia de las observaciones.

Fuente: Cálculos propios.

Ahora bien, hasta aquí incluimos en el análisis de regresión todos los factores que la literatura educativa reconoce como determinantes del logro de los estudiantes. Se observa que sin importar cual sea la prueba los colegios públicos y mixtos están en desventaja frente a los privados, masculinos y femeninos. También se observa que la ocupación de padre y madre resultan fundamentales al igual que su educación. Es interesante indagar cómo estos factores familiares operan cuando los estudiantes pertenecen a colegios de iguales características; en otras palabras, se quiere precisar si la relevancia de estas variables se mantiene una vez descontado el efecto del colegio. Para ello se estima un modelo de efectos fijos, en el cual solo se incluyen las variables personales y familiares, y se añade una variable ficticia por cada colegio.

Los resultados para la prueba total evidencian que pierden mucho peso la educación de los padres y la profesión del padre, sus coeficientes pasan de 1,28 a 0,52 y de 11 a 2,7, respectivamente. Esto resalta aún más la importancia del plantel frente a las características socioeconómicas del estudiante, y con ello la relevancia de mejorar las condiciones de los colegios más desaventajados. Además, ganan peso el número de hermanos y el mayor de los hermanos, que pasan de -0,47 a 0,27 y de 1,43 a 2,95, respectivamente, mientras se mantienen muy estables las ventajas para los hombres y los hijos de madres que trabajan.

Lo anterior evidencia que la asistencia de hermanos al mismo colegio refuerza su logro y es aun más ventajoso para los hermanos mayores, quizás porque la concurrencia a un mismo colegio puede brindar más oportunidades de ayuda e interacción entre los hermanos que si asistieran a colegios distintos. Por su parte, el que la educación de los padres resulte menos significativa al igual que la profesión de los padres puede ser un indicio de que gran parte del aporte de estas variables reside en la alta probabilidad de que padres más educados y en ocupaciones de mas alto rango tengan el suficiente poder adquisitivo para garantizar un acceso a centros educativos de más alta calidad. En otras palabras, el efecto de esas variables se da principalmente a través del colegio cuya condición socioeconómica les permite elegir para sus hijos. En cuanto a la ocupación de la madre, su efecto se mantiene, lo cual puede estar señalando que aquellas madres que trabajan, aunque dedican menos tiempo a sus hijos, pueden aportar desde su ocupación más elementos de ayuda para el aprendizaje de sus hijos que aquellas madres que no desempeñan ninguna actividad fuera del hogar¹⁷, o, como se señaló anteriormente, el efecto

17 Es necesario hacer la salvedad que esto es válido para estudiantes que culminan su secundaria, quizás para los primeros años de aprendizaje, la presencia continua de la madre en el hogar pueda tener un efecto más benéfico.

mediador puede provenir de la mayor capacidad para pagar educación de mejor calidad a través del aumento de los ingresos del hogar obtenidos por el trabajo de la madre.

Las estimaciones para las pruebas de lenguaje y matemáticas muestran prácticamente los mismos resultados. Para conocimiento matemático las únicas diferencias apreciables se presentan en las variables número de hermanos y mayor de los hermanos que resultan significativas, aunque siguen conservando el signo positivo, mientras que para lenguaje, la variable número de hermanos, que antes tenía un efecto negativo sobre el logro, deja de ser significativa.

Ahora bien, con las consabidas desventajas de esta estimación, y los avances que proporciona la modelación jerárquica observemos que tan diferentes son los resultados y que inferencias adicionales podemos hacer.¹⁸

Para la prueba total, en lo que se refiere a las variables personales, los resultados son similares a los obtenidos por M.C.O. Los hombres siguen obteniendo puntajes más altos que las mujeres (13 puntos), la edad mantiene su efecto negativo sobre el logro (23 puntos), el mayor de los hermanos obtiene mejores resultados (2,79 puntos), y la condición laboral no implica ningún efecto sobre el logro. En cuanto a los factores de tipo familiar, el número de hermanos no resulta significativo, mientras que el resto de variables son significativas, y las madres amas de casa siguen afectando negativamente el desempeño de los estudiantes en las pruebas Icfes. En lo que se refiere a las características del colegio, el análisis multinivel tiene una gran ventaja y es que permite analizar cada una de las categorías de una variable, omitiendo una sola como punto de comparación, lo cual es ventajoso cuando la variable tiene más de dos categorías; en este caso particular, permite analizar en más detalle los efectos de cada una de las jornadas, modalidades y género (véase Tabla 5).

Permanecen las desventajas de los jóvenes que estudian en colegios públicos y mixtos. En el primer caso, los estudiantes obtienen en promedio 14 puntos menos frente a los que asisten a colegios privados; en el segundo caso, los estudiantes de

18 Se estima con máxima verosimilitud completa (Igl) y con la restringida (Rigl), produciendo resultados casi idénticos, lo cual puede ser indicio de que el tamaño muestral del nivel dos es grande. Para lograr una inferencia exacta se intentó realizar estimación bayesiana pero no fue posible, muy seguramente porque el tamaño de la información es grande y el modelo no tan sencillo. El paquete utilizado para la estimación es el MLwin, el cual no presenta los p-valores, solo los errores estándar. Siguiendo a Hox (1995) es mucho mejor presentar estos últimos ya que los p-valores, basados en errores asintóticos, son solo aproximativos. La interpretación sugerida es que son estadísticamente significativos aquellos coeficientes que exceden dos veces su error estándar (véase Tabla 5).

Tabla 5. Resultados de la regresión multinivel para la prueba total

Estimación por: Parte fija	IGLS		RIGLS	
	Coefficiente	Error estándar	Coefficiente	Error estándar
Género	13,119	0,578	13,119	0,579
Edad	-23,755	1,516	-23,691	1,524
Edad ²	0,496	0,039	0,494	0,039
Trabaja	-0,227*	0,943	-0,283*	0,944
Mayor hermanos	2,790	0,451	2,793	0,451
Número hermanos	0,148*	0,134	0,151*	0,135
Educación máxima	1,547	0,203	1,546	0,204
Profesión padre	2,197	0,716	2,190	0,716
Madre no trabaja	-3,458	0,447	-3,459	0,447
Público	-14,934	1,591	-14,921	1,601
Normalista	2,666*	3,603	-2,643*	3,614
Técnico	-4,064	0,878	-4,061	0,88
Mañana	-5,019	1,247	-5,015	1,521
Tarde	-6,894	1,476	-6,882	1,482
Noche	-12,696	2,677	-12,727	2,686
Masculino	21,281	3,499	21,308	3,523
Femenino	19,002	2,074	19,010	2,087
<i>Interacciones</i>				
Tecnico.edumaxc**	-0,557	0,160	-0,557	0,161
Mañana.edumaxc	0,513	0,179	-0,51	0,18
Tarde.edumaxc	-0,562	0,192	-0,562	0,194
Noche.edumaxc	-0,924	0,352	-0,926	0,354
Público.edumaxc	-0,559	0,175	-0,559	0,176
Noche.edadc**	1,550	0,603	-0,926	0,354
Femenino.edadc	-2,058	0,640	-2,065	0,644
<i>Parte aleatoria</i>				
Nivel uno σ^2	1028,430	9242	1028,890	9249
σ^2 Género	28783	8335	29194	8381
σ^2 Edad	1,951	0,548	2,11	0,57
σ^2 Educación Máx.	0,649	0,145	0,689	0,149
-2 loglikelihood	256820,3		256820,6	

* No significativo al 10%.

** Con fines interpretativos educación máxima y edad están centradas alrededor de sus respectivas medias muestrales.

Fuente: Icfes. Cálculos propios.

colegios femeninos y masculinos obtienen en promedio 19 y 21 puntos por encima de los que asisten a mixtos. En cuanto a la jornada se verifica también la ventaja que sostiene la completa frente al resto. La mayor desventaja la presenta la jornada nocturna con 12 puntos por debajo, seguida de la tarde con 6,8 y la mañana con 5 puntos.

Por último, esta estimación permite observar que los colegios técnicos si están en desventaja frente a los académicos, no a si los normalistas que no presentan diferencias significativas con estos últimos. Dicha diferencia no había sido establecida mediante la estimación M.C.O, la cual arrojó que la modalidad no era significativa.

Los modelos jerárquicos permiten identificar no solo si las variables de primer nivel tienen efectos fijos, esto es aquellos que no dependen de las características del colegio (en un análisis de dos niveles) sino también los efectos aleatorios, que son aquellos que varían con ciertas características de segundo nivel, y, finalmente, si presentan los dos efectos, el fijo y el aleatorio al mismo tiempo (Caro, 2000). En el caso de las variables de segundo nivel es posible capturar su aporte a la variabilidad del logro, así como el efecto de estas sobre las variables de primer nivel que presentan efectos aleatorios. Esto último se logra a través de la introducción de variables de interacción, las cuales disminuyen o acrecientan los efectos fijos del primer nivel.

En este caso, son tres las variables del primer nivel que resultaron con efectos fijos y aleatorios al mismo tiempo. El género, la edad y la educación máxima de los padres presentan efectos fijos significativos y también efectos aleatorios, que implican que dependiendo de ciertas características del colegio, los efectos de estas variables sobre el logro medio serán diferentes. Para ir más allá de un modelo de coeficientes aleatorios, y poder explicar algo de los efectos aleatorios presentes en la estimación, se incluyeron variables de interacción entre género, edad y educación máxima con cada una de las variables disponibles de segundo nivel.

En el caso del género no fue posible hallar ninguna explicación sobre sus efectos aleatorios con las variables existentes dentro del estudio. Para la edad, de ocho variables de interacción incluidas solo resultaron significativas dos.¹⁹ Los establecimientos femeninos tienen un efecto aún más negativo con respecto a la edad frente a los mixtos, así mientras para una estudiante de edad promedio (16,9 años para la muestra total) estudiar en un colegio femenino implica 19 puntos frente a un colegio mixto, por cada año adicional este efecto disminuye en 2 puntos. En los colegios

19 Las variables de interacción son: *normalista.edad*, *técnico.edad*, *mañana.edad*, *tarde.edad*, *noche.edad*, *masculino.edad*, *femenino.edad* y *público.edad*.

nocturnos, por su parte, el efecto de la edad sobre el logro se ve menguado en 1,55 puntos. Esto indica que mientras para un individuo de edad promedio (16,9 años) el efecto de estudiar en un colegio nocturno es de -12,7 puntos, por cada año adicional que tenga sobre esa edad promedio sus resultados serán menos malos en 1,55 puntos frente a los de jornada completa. En otros términos, para los estudiantes de mayor edad, estudiar en colegios nocturnos les ofrece ventajas, quizás porque estos colegios justamente enfocan sus metodologías de enseñanza a personas que por diversas circunstancias, muy vinculadas con la edad, necesitan estudiar de noche.

Para el caso de la educación máxima resultaron significativas cinco de las ocho variables de interacción incluidas. El efecto fijo de esta variable es disminuido para los colegios de jornada distinta a la completa, siendo el de mayor impacto para la noche con casi un punto por debajo, seguido de la tarde con 0,56 puntos por debajo, y la mañana con -0,51. Mientras para un individuo que estudia en un colegio de jornada completa, y cuya educación máxima de los padres es la promedio (9,13 años) el efecto de dicha variable es de 1,5 puntos adicionales, esos resultados son mucho menores para los estudiantes de las otras jornadas. Por cada año adicional de educación máxima de los padres el efecto de esta variable sobre el logro disminuye para ellos frente a los de jornada completa. Más aún, para los estudiantes cuyos padres tienen los niveles de educación más bajos, las diferencias entre estas modalidades prácticamente no existen, a excepción de la jornada nocturna que cualquiera sea el nivel de estudios de los padres sigue estando en desventaja frente a la jornada diurna y completa. Para un alumno cuyos padres tienen una educación máxima de 12 años, por ejemplo, y que estudie en colegio nocturno, se pierde el efecto del mayor nivel educativo de los padres. Lo mismo ocurre, prácticamente, con las jornadas de mañana y tarde.²⁰

20 La interpretación de los coeficientes estimados, una vez se introducen las interacciones no es igual que sin las interacciones presentes; uno de los coeficientes de las variables directas de la interacción se debe interpretar como si la otra variable involucrada fuera cero, cuando ese valor tiene significado, o como el valor promedio de la variable cuando ha sido centrada alrededor de la media muestral, como en este caso se hizo tanto para educación máxima como para edad. Así, por ejemplo, con la variable de interacción *noche*edumaxc*, el coeficiente de esta última debe interpretarse junto con el coeficiente de la interacción de *noche*edumaxc*. El primer coeficiente es de 1,55 para los estudiantes de jornada completa. Como el coeficiente de la interacción es de -0,92, ello implica que por año adicional en la educación máxima promedio, el efecto de esa variable disminuye en esos puntos para los estudiantes de la jornada nocturna. Por dos años más, que implican 11,1 años de educación máxima de los padres, el efecto negativo suma 1,8 puntos menos, con lo cual el efecto de la educación máxima se pierde (1,5 puntos es menor a 1,8 puntos). O, en otras palabras, la distancia entre el colegio nocturno frente a la jornada completa se amplía a medida que los estudiantes cuentan con padres más educados.

Los colegios públicos y técnicos también disminuyen el efecto de la educación máxima de los padres en -0,6 y -0,56 puntos por cada año adicional, respectivamente. Así que estudiantes con padres de alto nivel educativo obtienen desventajas aún mayores si asisten a colegios públicos o técnicos.

En suma, el nivel educativo de los padres es matizado por las características del colegio y su efecto es mucho menor para aquellos que estudian en colegios públicos, técnicos y en jornada diferente a la completa. Este resultado estaría indicando que mientras mayor es la educación de los padres, se podrían acortar distancias frente a efectos desfavorables de los colegios. Sin embargo, ocurre lo contrario; es decir, a medida que los padres cuentan con más educación se van ampliando las diferencias de logro educativo dependiendo de la calidad de los colegios.

Para las pruebas individuales los resultados se mantienen bastante parecidos a los de la prueba total. Todas presentan las mismas variables con efectos aleatorios: género, edad y educación máxima de los padres. En cuanto a los efectos fijos, en todas trabajar no resultó significativa, al igual que el número de hermanos, excepción hecha de conocimiento matemático. Las interacciones que resultaron significativas evidencian que la educación máxima de los padres pierde relevancia cuando se estudia en colegios públicos y técnicos, al igual que en jornadas diferentes a la completa, para cada una de las pruebas.²¹ Para aptitud matemática los resultados de las interacciones con la edad son similares a los de la prueba total. Mejoran en colegios nocturnos y disminuyen para los colegios femeninos; esto es, a medida que el estudiante tiene más edad puede obtener mejores resultados en un colegio nocturno, mientras que el resultado es más negativo si estudia en colegio femenino. Para lenguaje, resulta también significativa la interacción negativa entre mañana y edad. Por último, para conocimiento matemático los colegios femeninos y masculinos presentan resultados más negativos frente a los mixtos a medida que los estudiantes cuentan con más años mientras que los públicos aportan 0,27 puntos por cada año adicional a la edad promedio (véase Tabla 6).

21 Para conocimiento matemático, por ejemplo, resulta significativa la variable de interacción $\text{tecnico}^* \text{edumaxc}$. El coeficiente estimado es de -0,11 para esta interacción y de -0,98 para técnico. Para darle interpretación a estos resultados es necesario hacer una comparación con la modalidad académica. Así, la línea de predicción para académico es: $B_{\text{constante}} + B_{\text{edumaxc}}$ (los subíndices corresponden al lugar que ocupa la variable en la ecuación de regresión. La línea de predicción para técnico es $B_{\text{constante}} + B_{\text{edumaxc}} + B_{\text{técnico}} + B_{\text{técnico}}^* \text{edumaxc}$. y la diferencia entre esas dos líneas es el efecto de estar en un colegio técnico: $B_{\text{técnico}} + B_{\text{técnico}}^* \text{edumaxc}$. Que en este caso es de -10,99. En otras palabras, por cada año adicional en la edad promedio de educación de los padres, los estudiantes en colegios técnicos obtienen peores resultados (-0,11 puntos) frente a los colegios académicos.

Tabla 6. Resultados de la regresión multinivel para las pruebas individuales

Efectos fijos	Aptitud matemática	C. matemático	Lenguaje
<i>Parte fija</i>			
Género	3,77 (0,14)	3,32 (0,13)	1,6 (0,14)
Edad	-5,32 (0,39)	-5,41 (0,36)	0,07 (0,03)
Edad ²	0,11 (0,01)	0,11 (0,01)	0,10 (0,01)
Trabaja	0,36 (0,24)*	-0,05 (0,23)*	-0,28 (0,24)*
Mayor hermanos	0,23 (0,12)	0,42 (0,11)	0,71 (0,11)
Número hermanos	0,068 (0,035)**	0,13 (0,03)	-0,01 (0,03)*
Educación máxima	0,29 (0,05)	0,27 (0,05)	0,42 (0,05)
Profesión padre	0,44 (0,18)	0,59 (0,17)	0,57 (0,18)
Madre no trabaja	-0,63 (0,11)	-0,25 (0,109)	-0,98 (0,11)
Público	-2,77 (0,32)	-3,24 (0,38)	-3,16 (0,33)
Normalista	-0,55 (0,81)*	-1,06 (0,86)*	-0,07 (0,83)*
Técnico	-0,52 (0,21)	-0,98 (0,21)	-0,65 (0,21)
Mañana	-1,08 (0,28)	-1,2 (0,3)	-0,58 (0,29)
Tarde	-1,47 (0,32)	-1,36 (0,35)	-0,94 (0,33)
Noche	-2,21 (0,63)	-2,09 (0,65)	-1,80 (0,63)
Masculino	3,74 (0,67)	5,43 (0,88)	3,79 (0,71)
Femenino	3,05 (0,43)	3,83 (0,48)	3,96 (0,44)
<i>Interacciones</i>			
Tecnico.edumaxc	-0,13 (0,04)	-0,11 (0,04)	-0,17 (0,04)
Mañana.edumaxc	-0,12 (0,04)	-0,09 (0,04)	-0,15 (0,04)
Tarde.edumaxc	-0,13 (0,04)	-0,093 (0,05)	-0,16 (0,05)
Noche.edumaxc	-0,22 (0,09)	-0,06 (0,08)*	-0,22 (0,09)
Público.edumaxc	-0,12 (0,04)	-0,14 (0,04)	-0,11 (0,04)
Noche.edadc	0,39 (0,15)	0,22 (0,14)*	0,32 (0,15)
Femenino.edadc	-0,49 (0,16)	-0,67 (0,15)	-0,4 (0,16)
Masculino.edadc	-0,11 (0,25)*	-0,55 (0,24)	-0,06 (0,25)*
Público.edadc	0,01 (0,12)*	0,27 (0,1)	-0,08 (0,11)*
<i>Parte aleatoria</i>			
Nivel uno σ^2	69,07 (0,62)	61,58 (0,55)	67,18 (0,6)
σ^2 Género	1,09 (0,47)	1,04 (0,42)	1,05 (0,46)
σ^2 Edad	0,11 (0,03)	0,14 (0,03)	0,07 (0,03)
σ^2 Educación Máx.	0,02 (0,01)	0,03 (0,01)	0,03 (0,01)
-2 loglikelihood	185858,3	183094,9	185241,7

* No significativo al 10%

** No significativo al 5%

Se realizaron estimaciones adicionales, eliminando posibles observaciones que pudieran estar dirigiendo los resultados anteriores -outliers-, y se encontró que estos se mantienen en su mayoría. Los signos de los coeficientes y la significancia de los efectos fijos, incluyendo las interacciones para la prueba total, permanecen al igual que los efectos aleatorios para género, edad y educación máxima. Para conocimiento matemático también se mantienen todos los resultados, no así para las pruebas de lenguaje y aptitud matemática en las cuales los efectos aleatorios para edad y género pierden significancia, al igual que la interacción entre noche y edad para aptitud matemática.

IV. Una primera aproximación a las diferencias subregionales

Al analizar las estadísticas descriptivas por subregiones se evidenciaron claras diferencias entre el Valle de Aburrá y el resto de regiones del departamento.²² Como se expresó anteriormente, la estimación del modelo jerárquico de tres niveles —ya fuera el municipio o la subregión como tercer nivel—no se pudo llevar a cabo con la información disponible; sin embargo, como una aproximación se estimaron modelos jerárquicos de dos niveles para el Valle de Aburrá y para el resto de subregiones por separado.

En primer lugar, se estima el efecto colegio en los dos casos. Para el Valle de Aburrá, los resultados son levemente superiores con respecto a la muestra total, pero para el resto de subregiones la diferencia es bastante notoria (véase Tabla 7). El efecto colegio para cada una de las pruebas disminuye en casi la mitad frente al Valle de Aburrá. Lo anterior se explica porque la variabilidad entre los colegios por fuera del Valle de Aburrá es mucho menor, sus condiciones son mucho más parecidas en términos de recursos, y de composición. Mientras en el Valle de Aburrá los colegios públicos representan un 58% y los mixtos un 65%, en el resto del departamento esas participaciones son de 88% y 96% respectivamente. En síntesis, el resultado evidencia que la variabilidad de logro por fuera del Valle de Aburrá, esta mucho menos explicada por las diferencias entre colegios, los cuales son mucho más homogéneos en el resto de subregiones.

En segundo lugar, las estimaciones para determinar los efectos de las variables individuales, familiares y del colegio para la prueba total muestran que los

22 Mientras el promedio para la prueba total en el Valle de Aburrá es de 256 puntos, para el resto de subregiones es de 236 puntos. Así mismo la educación máxima promedio de los padres es de casi 10 años, para el resto está por debajo en mas de dos años; el porcentaje de padres con profesiones bien remuneradas es del 7 % en el resto de subregiones, en el Valle de Aburrá alcanza el 17%; el número de hermanos promedio es de 2,3 para el Valle mientras en el resto llega a 3,3.

Tabla 7. *Efecto Colegio o correlación intraclase para las pruebas Icfes en las subregiones. Antioquia, 1999*

Pruebas	Efecto Colegio Modelo O.I.M	
	Valle de Aburrá (porcentaje)	Resto de subregiones (porcentaje)
Total	39,9	17,9
Conocimiento matemático	35,9	17
Lenguaje	29,7	12,7
Aptitud matemática	25,7	11,9

Fuente: Cálculos propios.

resultados para el Valle de Aburrá son sumamente similares a los de la muestra total. Las únicas diferencias apreciables son que el efecto negativo de los colegios públicos aumenta en casi dos puntos, la jornada completa tiene mayores ventajas frente al resto y los colegios femeninos disminuyen su ventaja frente a los mixtos en casi tres puntos. En cuanto a las variables de interacción, para la educación máxima resultan significativas técnico, tarde y público, más no mañana y noche como en la muestra total y para la edad resultan significativas para femenino y la modalidad técnica, frente a femenino y noche de la muestra total.

Las estimaciones para el resto de subregiones revelan que las diferencias son notables. A las variables que resultan no significativas para las otras muestras se le suman algunas más. En cuanto a las variables familiares, la educación máxima y la profesión del padre no resultan significativas para explicar la variabilidad en el logro. Mientras que para las variables del colegio, la jornada de la mañana y los colegios públicos no muestran diferencias significativas con la jornada completa y los privados, respectivamente.²³ Las variables que muestran efectos aleatorios son género y la educación máxima, para la cual se encuentran efectos significativos de interacción para los colegios nocturnos y femeninos (véase Tabla 8). En general estos resultados para el resto de subregiones evidencia que el modelo pierde poder explicativo, derivado de una menor variabilidad de los factores tomados en cuenta para la estimación.

²³ Público es significativo pero a un nivel de significancia del 10%.

Tabla 8. Resultados de la regresión multinivel para la prueba total en las subregiones

Parte fija	Valle de Aburrá		Resto de subregiones	
	Coficiente	Error estándar	Coficiente	Error estándar
Género	13,26	0,75	12,85	0,89
Edad	-22,28	2,31	-28,19	1,54
Edad ²	0,46	0,06	0,62	0,04
Trabaja	-0,37*	1,4	-0,24*	1,26
Mayor hermanos	2,98	0,54	2,6	0,81
Número hermanos	0,25*	0,18	0,1*	0,2
Educación máxima	1,69	0,29	0,61**	0,38
Profesión padre	2,52	0,29	0,81	1,5
Madre no trabaja	-3,3	0,53	-3,73	0,83
Público	-16,56	2,1	-4,65**	2,6
Normalista	-7,87*	6,19	1,5*	3,88
Técnico	-3,64	1,2	-3,2	1,22
Mañana	-9,12	1,81	-2,8*	1,8
Tarde	-12,32	2,4	-3,79	1,87
Noche	-14,83	4,35	-8,2	2,67
Masculino	20,67	4	16,2	7,02
Femenino	16,41	2,4	18,04	6,35
<i>Interacciones</i>				
Tecnico.edumaxc**	-0,69	0,21	-0,43**	0,23

Continúa...

Tabla 8. *Continuación*

Parte fija	Valle de Aburrá		Resto de subregiones	
	Coficiente	Error estándar	Coficiente	Error estándar
Tarde.edu maxc	-0,73	0,31	-0,04*	0,26
Noche.edumaxc	-0,62*	0,51	-1,32	0,49
Público.edumaxc	-0,58	0,21	-0,12*	0,36
Femenino *edumaxc	0,17*	0,23	1,71	0,8
Técnico.ed.adc**	1,4	0,65		
Femenino.edadc	-2,19	0,7		
<i>Parte aleatoria</i>				
Nivel uno σ^2	1.043,18	11,35	1.001,39	15,87
σ^2 Género	23,41	9,92	36,48	14,36
σ^2 Edad	2,23	0,75		
σ^2 Educación máx.	0,54	0,17	0,57	0,22
-2 <i>loglikelihood</i>	173.593,2		83.152	

No. de observaciones de nivel uno y nivel dos para valle de Aburrá 17.649 y 307.

No. de observaciones para el resto 8.486 y 235, respectivamente.

* No significativo al 10%.

** No significativo al 5%.

Fuente: Icfes. Cálculos propios.

Nuevamente se realizan estimaciones eliminando posibles outliers. Los resultados para el Valle de Aburrá se mantienen, a excepción del efecto aleatorio del género, que está explicado prácticamente por la influencia de dos colegios. En el caso del resto de las subregiones, también pierde significancia el efecto aleatorio del género, y el coeficiente asociado a los colegios públicos resulta significativo, aunque es cuatro veces menor al asociado al Valle de Aburrá, lo que implica que aunque puede

haber diferencia en el logro medio entre los colegios públicos y privados por fuera del Valle de Aburrá, esta es claramente mucho menor que en esta subregión.

V. Conclusiones y recomendaciones

Recientes estudios en el campo de la economía de la educación han señalado la importancia notable de la calidad educativa en la explicación de la productividad y crecimiento de las economías regionales y nacionales (Hanushek y Kimko, 2000). La calidad de la educación se empieza a usar como una variable aún más relevante que las medidas usuales de capital humano —número promedio de años de educación de las personas mayores de 25 años, o la tasa de escolaridad en la educación secundaria, entre otros.

Los pobres resultados de los estudiantes antioqueños con respecto a la mayoría de departamentos del país²⁴ y las diferencias subregionales encontradas en Antioquia no permiten augurar un futuro muy prometedor en términos de igualdad de oportunidades y bienestar para los antioqueños. Aunque no se pudo establecer cuales son los mecanismos a través de los cuales opera la brecha entre logro medio en el Valle de Aburrá y el resto de subregiones, se podrían señalar los insumos escolares con que cuentan los colegios, así como el gasto educativo, la pobreza, o el nivel de actividad económica de la región. En cuanto a los primeros, es evidente que en promedio, los colegios de los municipios por fuera del Valle de Aburrá cuentan con menores recursos y ese es un factor fundamental a la hora de entender las diferencias en el logro promedio. En este sentido, será necesario un estudio posterior que corrobore esta hipótesis, y que muestre cuáles de los insumos escolares son más importantes con miras no sólo a acrecentar el gasto sino a dirigirlo mejor.

Este estudio confirma los hallazgos de los estudios que hasta el momento se han hecho sobre la incidencia de las características de los colegios en los resultados académicos de los estudiantes. Así, el efecto colegio representa casi el 40% de la variabilidad del logro, con lo cual hay un alto rango de acción para las políticas educativas que permita garantizar unas condiciones más equitativas de acceso a una educación de mayor calidad.

Al estimar por separado los modelos para el Valle de Aburrá y el resto de subregiones se encuentra que el efecto colegio es mucho menor para estas últimas,

24 Antioquia muestra resultados en las pruebas Icfes por debajo de casi todos los departamentos del país según un estudio de Alesina, Carrasquilla y Echavarría (2000) en el cual estiman un modelo de panel de datos, entre 1993 y 1998, cuya variable dependiente es un promedio para las pruebas de matemáticas y lenguaje del Icfes. Alviar y Polanía (1992) también muestran el deterioro de la educación en Antioquia durante la década del ochenta.

lo cual no implica que haya mucho menos que hacer en las subregiones en términos de reformas educativas al interior de los colegios. Por el contrario, las medias de los puntajes en las pruebas muestran una diferencia importante en los resultados, y aunque existe más homogeneidad en los colegios de fuera del Valle de Aburrá, ellos requieren de mayores esfuerzos para cortar distancias con esta subregión. En otras palabras, en el resto de subregiones el sistema educativo es más homogéneo pero menos eficiente en términos de logro promedio.

La diferencia del efecto colegio entre subregiones opera también entre países. En los de mayor desarrollo hay mayor homogeneidad entre los establecimientos educativos y eso conduce a un efecto colegio mucho menor. Pero a diferencia de estos países, en donde los resultados promedio del logro son mucho más altos que en los de bajo desarrollo, en las regiones por fuera del Valle de Aburrá la homogeneidad está acompañada de condiciones mucho más desfavorables que implican un efecto colegio mucho menor. En otras palabras, cuando la mayoría de colegios cuenta con insumos escolares abundantes y de buena calidad, esto implica con gran probabilidad un efecto colegio bajo pero acompañado de un logro promedio alto, mientras que cuando existe alta heterogeneidad entre colegios, usualmente en un contexto de bajo desarrollo, puede esperarse un efecto colegio alto junto con un logro promedio bajo, en términos relativos.

Las características de los colegios tenidas en cuenta en este estudio revelan que los colegios públicos, mixtos y técnicos muestran desventajas frente al resto de colegios. En el caso de la diferencia entre los colegios públicos y privados, algunos han argumentado que ella se debe primordialmente a la composición socioeconómica inferior de los primeros. Sin embargo, el efecto aleatorio hallado para la educación máxima de los padres —factor asociado estrechamente al nivel socioeconómico del hogar del estudiante— y la interacción negativa con la naturaleza del colegio, muestran que, al contrario de lo que podría pensarse, aquellos que cuentan con padres de alto nivel educativo y que estudian en colegios públicos, tienen una desventaja frente a los privados mayor que aquellos con padres de menor nivel educativo. Incluso las diferencias entre público y privado son prácticamente inexistentes para estos últimos.

Si la educación de los padres fuera uno de los factores más importantes para explicar esas diferencias en el logro según la naturaleza del colegio, a mayor educación cabría esperar que las diferencias se acortaran y ocurre justamente lo contrario.²⁵

25 Núñez et al (2002, p. 28) encuentran que las diferencias en logro por nivel de ingreso, una vez se controla por las características observables, son a favor de los colegios públicos para ingresos

Todo lo anterior, aunado a las interacciones negativas encontradas para los colegios nocturnos, y técnicos, corrobora la importancia del plantel en las diferencias en el logro de los estudiantes antioqueños y resalta la necesidad de reformas educativas que permitan mejorar las condiciones escolares que enfrenta una gran proporción de los jóvenes, especialmente por fuera del Valle del Aburrá.

Finalmente, en la parte metodológica, es evidente que la utilidad de los modelos jerárquicos reside principalmente en que permite responder un mayor número de preguntas y obtener mayor precisión en las estimaciones.²⁶ En este caso particular, la magnitud de los efectos fijos comparables es bastante similar a la regresión tradicional, pero permitió establecer que existen efectos aleatorios y mostrar a través de que características del colegio se producen (variables de interacción). En suma, el modelo multinivel es una versión más completa de la regresión múltiple para fenómenos en donde se presente la anidación, cuyo ejemplo más referenciado y aplicado es el educativo.

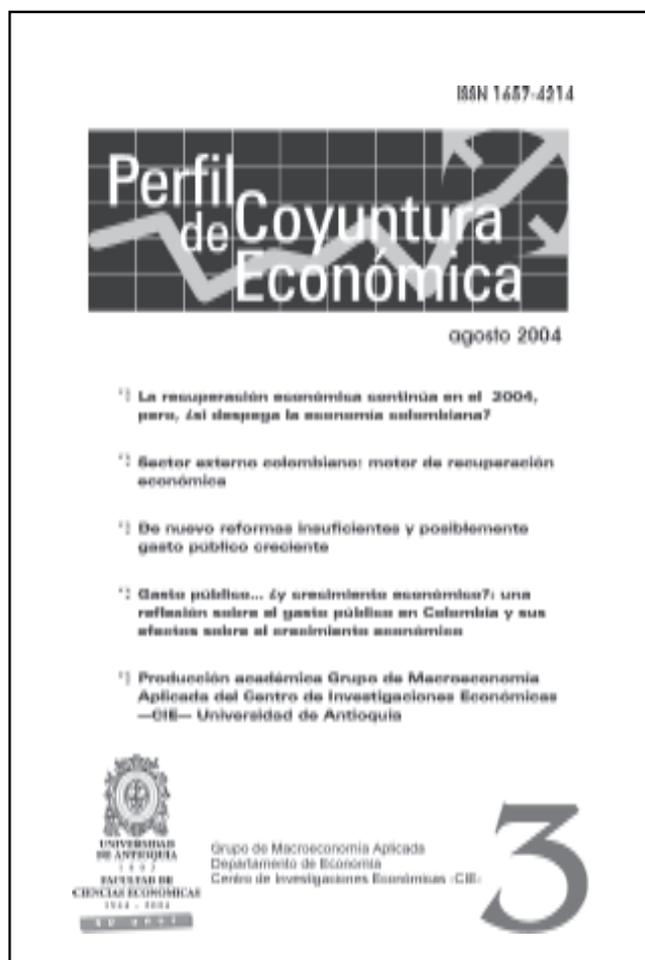
Bibliografía

- ALESINA, Alberto; CARRASQUILLA, Alberto y ECHAVARRÍA, Juan José, 2000. "Decentralization in Colombia". Working Paper No. 5.
- ALVIAR, Mauricio y POLANÍA, D., 1992. "La calidad de la educación en Colombia". Bogotá, Fedesarrollo.
- BRYSK, Anthony y RAUDENBUSH, Stephen, 1986. "A Hierarquical Model for Studying School Effects". *Sociology of Education*, Vol. 59, Issue 1, pp. 1-17.
- _____, 1992. *Hierarchical Linear Models. Applications and Data Analysis Methods*. United States, Sage Publications.
- CARD, D. y KRUEGER, A., 1996. "The Economic Return to School Quality. Assessing Educational Practices: The Contribution of Economics". Cambridge. MIT Press.
- CARO, Blanca, 2000. "Factores asociados al logro académico de los alumnos de 3º y 5º de primaria de Bogotá". *Coyuntura Social*, No. 22, mayo, pp. 65-80.
- CASTAÑO, Elkin, 1998. "El efecto colegio sobre la variabilidad del rendimiento en matemáticas". *Lecturas de Economía*, No. 49, Julio-Diciembre, 47-58
- GAVIRIA, Alejandro y BARRIENTOS, Jorge, 2001. "Determinantes de la calidad de la educación en Colombia". *Archivos de Economía*. Documento 159.
- HANUSHEK, Eric A., 1996. "Measuring Investment in Education". *Journal of Economics Perspectives* Vol. 10, No. 4.

familiares menores a un salario mínimo, pero a partir de allí las diferencias empiezan a crecer a favor de los privados.

26 Sin embargo en la medida en que el modelo tenga menos efectos aleatorios, se acerca más a un modelo de regresión tradicional.

- HANUSHEK, Eric y KIMKO, Dennis, 2000. "Schooling, Labor-Force Quality, and the Growth of Nations". *American Economic Review*. Vol. 90, No. 5.
- HEYNEMAN, S. y LOXLEY, W., 1983. "The effect of Primary-School Quality on Academic Achievement Across twenty-nine High and Low-income Countries". *American Journal of Sociology*, No. 6, Vol. 88.
- Hox J., J., 1995. *Applied Multilevel Analysis*, Amsterdam, TT-Publikaties.
- KREMER, M. y MASKIN, E., 1996. "Wage Inequality and Segregation by Skill". *NBER Working Paper*, No. 5718.
- LLACH, Juan José; MONTOYA, Silvia y ROLDÁN, Flavio, 1999. *Educación para todos*. Buenos Aires, IERAL.
- LEE, Jong-Wha y BARRO, Robert, 1997. "Schooling Quality in a Cross Section of Countries". *NBER Working Paper*, No. 6198.
- MURILLO T., F. Javier. "Los modelos jerárquicos lineales aplicados a la investigación sobre eficacia escolar". *Revista de Investigación Educativa*. Vol. 17, No. 2, p. 453-460.
- NÚÑEZ, Jairo; STEINER, Roberto; CADENA, Ximena y PARDO, Renata, 2002. " Cuáles colegios ofrecen mejor educación en Colombia?". Bogotá, Documento CEDE 2002-3.
- RASBASH, *et al.*, 2000. "A user's Guide to MLwin". Centre for Multilevel modelling. Institute of Education . University of London.
- SARMIENTO, Alfredo y GONZÁLEZ, Jorge Iván, 1997. "La calidad de la educación y el logro de los planteles educativos". *Planeación y Desarrollo*. Vol. 28, No. 1 p. 25-57.
- SARMIENTO, Alfredo; BECERRA, L. y GONZÁLEZ, Jorge Iván, 2000. "La incidencia del plantel en el logro educativo del alumno y su relación con el nivel socioeconómico". *Coyuntura Social*, No. 22, mayo, p. 53-64.



Distribución nacional e internacional

Ecoe Ediciones. Calle 32 No. 17-22, tels.: 288 2556 y 288 5651, Fax: 283 1673, Santafé de Bogotá, Colombia. Dirección electrónica: ecoe@col1.telecom.com.co.
Editorial Universidad de Antioquia, tels.: 210 5010 - 210 5056, Medellín, Colombia.
Dirección Electrónica: mercadeo@editorialudea.com.

Precio (ejemplar)

\$10.000.00, sujeto a posteriores reajustes.
Valor de suscripción (dos números al año) \$18.000.00

Correspondencia, canje y suscripciones

Perfil de Coyuntura Económica, Facultad de Ciencias Económicas, Universidad de Antioquia.
Apartado aéreo 1226, teléfonos: 210 5842 - 210 5844. Telefax: (574) 210 5843.
Medellín, Colombia, Suramérica.