

# Calidad de la educación pública y logro académico en Medellín 2004-2006. Una aproximación por regresión intercuartil

Jorge Barrientos Marín\*

**–Introducción. –I. Descripción de las cifras. –II. Análisis Estadístico. –Conclusiones. –Bibliografía.**

*Primera versión recibida en septiembre de 2007; versión final aceptada en marzo de 2008*

## Introducción

Desde la publicación de los trabajos sobre logro académico, utilizando grandes bases de datos como por ejemplo Castaño (1998), Gaviria y Barrientos (2001a, 2001b) y Restrepo y Alviár (2005) entre otros, es ya costumbre analizar la evolución del logro académico de los estudiantes, medido por el puntaje en pruebas estandarizadas como las reportadas en el examen del ICFES. En este artículo utilizamos técnicas estadísticas estándar aunque poco usadas en la práctica, de modo que la contribución de este artículo se centra en dos aspectos: primero, en el uso de información reciente de varias fuentes como el formulario de inscripción al examen ICFES (2004-2006) y la base de datos conocida como Matrícula en Línea (ML); segundo, desde el punto de vista técnico, la aplicación de regresión por nivel de rendimiento académico.

El sistema de información ML es un mecanismo de seguimiento y monitoreo de los colegios públicos de Medellín, implementado por el gobierno local a través de la Secretaría de Educación Municipal (SEM). Esta base de datos,

---

\* Jorge Barrientos Marín: investigador del Centro de Investigaciones y Consultorías y director del Departamento de Economía. Facultad de Ciencias Económicas. Universidad de Antioquia. Dirección electrónica: jbarr@economicas.udea.edu.co. Dirección postal: Ciudad Universitaria, bloque 13, A.A. 1226, Medellín, Colombia. Los errores, opiniones e interpretaciones son responsabilidad exclusiva del autor. Este artículo es derivado de la información de Matrícula en Línea, usada en la investigación Estrategia de concesión a cinco años de la prestación del servicio público educativo en plantas físicas oficiales (2002-2006).

poco utilizada, permite en gran medida mitigar la pérdida de información en el formulario de inscripción del ICFES desde 2003. Por supuesto, el sistema ML está en vía de desarrollo, sin embargo tiene potenciales usos y promete ser fuente de información más completa en futuras investigaciones en economía de la educación. Por ejemplo, Barrientos y Ríos (2007) aprovechan el sistema ML y evalúan el impacto del programa de colegios en concesión, concentrándose en el período 2004-2006. Adicionalmente, la ML contiene información de interés, por ejemplo, si un estudiante está en el régimen contributivo o recibe algún tipo de alimento en el colegio.

A partir de esta información, se construye una variable binaria para cada característica: *OUTRC* (fuera del régimen contributivo) y *ALIM* (si el estudiante recibe algún tipo de asistencia alimentaria). Calculamos el efecto de cada una de estas variables sobre el desempeño académico controlando por factores individuales y escolares. Infortunadamente, el análisis con la información sobre alimentación y cobertura en seguridad social está disponible sólo para el año 2004. Para el año 2005 y 2006 la SEM no recogió esta información.

En este trabajo pretendemos estimar el efecto de los factores individuales y escolares sobre el desempeño académico, para ello, replicamos algunas de las metodologías de uso común en investigaciones sobre el logro académico y el cálculo de la variabilidad del efecto del colegio en las pruebas ICFES (puntaje total, matemáticas y lenguaje). También es importante saber cuál tipo de estudiante está más afectado por las características individuales y los factores escolares: si aquellos con desempeño medio o aquellos con desempeño alto o bajo; el *Quantile Regression (QR)* permite responder en buena medida a esta pregunta. Como se verá, ciertos factores impactan más a estudiantes con un desempeño superior e inferior, que a aquellos con un desempeño cerca de la media de la distribución de puntajes.

De acuerdo con la revisión bibliográfica llevada a cabo, pocos trabajos usan *QR* en este contexto; Gaviria y Barrientos (2001a) hacen un primer intento para estudiantes ubicados en el 5% con mejor desempeño, y concluyen que los colegios públicos tienen un efecto adverso sobre el desempeño, especialmente apreciable sobre la cola derecha de la distribución del puntaje. Un análisis más detallado del efecto de los factores escolares sobre logro escolar, y con un espíritu más cercano al de este estudio, es hecho por Eiden y Showalter (1998) cuyos resultados sugieren que hay diferencias en el efecto de los recursos educativos y los atributos individuales sobre el logro, siendo éste diferente dependiendo de el punto de la distribución sujeto a análisis.

Otro punto que se aborda en este trabajo está relacionado con el efecto del colegio sobre el rendimiento en las pruebas ICFES; se calcula el efecto fijo colegio,

usando los respectivos puntajes como variable dependiente, en una regresión con un conjunto grande de variables binarias como regresores (una por colegio), de modo que el coeficiente de determinación de dicha regresión proporciona una idea de tal efecto. Este ejercicio se lleva a cabo cada año, y se puede observar la evolución en el tiempo del desempeño escolar (aunque independiente de un año a otro).

Claramente, se dejan de lado algunos aspectos importantes en el análisis del desempeño académico. Para nadie es un secreto que los incentivos al interior de los colegios, la promoción de los profesores y la gestión institucional en diferentes frentes (académico, administrativo, directivo y con la comunidad, etc.) juegan un papel preponderante en la calidad de educación impartida; no obstante, no hay información disponible que pueda ser cruzada con la información que existe por alumno y colegio.

Este artículo está dividido como sigue. En la sección I presentamos las estadísticas descriptivas de las variables de interés. En la sección II analizamos las estadísticas para calcular el efecto colegio y los determinantes del logro académico, tanto en la media de la distribución de la variable dependiente, como por cuantiles de rendimiento, así como el efecto de las características dentro del plantel.

## I. Descripción de las cifras

La información utilizada en este trabajo es el resultado de la concatenación de tres bases de datos diferentes: i) El sistema ML, ii) los resultados en el ICFES (2004-2006) y iii) el formulario C-100, que mide infraestructura disponible para el año 2002. Fueron eliminados los estudiantes mayores de 25 y menores de 15 años, así como aquellas observaciones que por error reportaron cero unidades en factores escolares, como metros cuadrados de aulas o unidades sanitarias, etc. Se eliminaron aquellos estudiantes en colegios del programa de Concesión, que eran 622 en 2004, 709 en 2005 y 813 en 2006. Sólo se tuvieron en cuenta los individuos matriculados en grado 11 durante el respectivo período, de modo que descartamos a quienes presentaron la prueba tiempo después de validar el bachillerato. El universo se limita a los colegios públicos del municipio de Medellín.<sup>2</sup> Teniendo en cuenta las restricciones anteriores, esto nos deja con una muestra de 7854 alumnos en 93 colegios en 2004, 7248 alumnos en 99 colegios en 2005 y 8748 alumnos en 104 colegios en 2006.

La tabla 1 presenta algunas estadísticas descriptivas de las principales variables usadas en este trabajo. Primero, presentamos las variables dependientes, a saber: el puntaje total en la prueba ICFES y el puntaje en las áreas de lenguaje

---

<sup>2</sup> La Matrícula en Línea sólo está diseñada para colegios públicos, por tanto, aunque quisiéramos incluir los colegios privados no podríamos hacerlo pues la información no está disponible.

y matemáticas; el puntaje total no incluye la prueba de inglés en ningún año. Segundo, tenemos las variables independientes que comprenden las características del alumno y las del colegio; para el año 2006 hay una media inferior en 20 puntos y 28 puntos respecto a 2005 y 2006 respectivamente, pero eso se explica (al menos en parte para el año 2006) porque en tal año el módulo de geografía e historia se juntó en uno solo (ciencias sociales).

Tabla 1. *Estadísticas Descriptivas*

Variables.	2004		2005		2006	
	Media	Desv. Est	Media	Desv. Est	Media	Desv. Est
P. Total	363,1	29,1	355,4	33,3	325,3	30,3
P. Total PI*	401,9	31,1	398,3	36,5	368,6	35,37
Lenguaje	53,1	7,53	47,5	8,1	49,4	6,3
Matemáticas	41,1	5,63	44,3	7,3	45,4	7,9
Edad	17,2	1,20	16,8	1,0	17,16	0,9
Sexo	0,39	-----	0,39	-----	0,4	-----
OUTRC	0,41	-----	N.D	N.D	N.D	N.D
ALIM	0,35	-----	N.D	N.D	N.D	N.D
Estrato1	0,084	-----	0,089	-----	0,093	-----
Estrato2	0,49	-----	0,49	-----	0,49	-----
Estrato3	0,37	-----	0,37	-----	0,36	-----
Estrato4	0,038	-----	0,032	-----	0,034	-----
Estrato5	0,005	-----	0,005	-----	0,004	-----
Estrato6	<0,0045	-----	<0,005	-----	<0,001	-----
R1<0.02	0,13	-----	0,14	-----	0,14	-----
0.02<R2<0.03	0,28	-----	0,30	-----	0,29	-----
0.03<R3<0.04	0,49	-----	0,46	-----	0,47	-----
R4>0.04	0,08	-----	0,08	-----	0,08	-----
Completa	0,004	-----	0,001	-----	0,08	-----
Mañana	0,418	-----	0,39	-----	0,43	-----
Tarde	0,57	-----	0,60	-----	0,47	-----
Planta Docent	62,0	51,7	51,6	38,4	57,6	49,7
M2aulas	1416	1257,4	1161	883,4	1332	1205
Comp www	1,1	0,37	1,1	0,42	1,1	0,42
Equip. Total	7,3	7,4	7,6	7,5	7,3	7,3
No Uni. Sanit	39,8	27,2	34,6	21,7	37,3	25,6
No Obs.	7854	7248	8748			

Fuente: ICFES, DANE y ML. Cálculos propios

N.D: Información no disponible

\* Contiene el puntaje en la prueba de Inglés

La media en lenguaje es de 53 en 2004, y alrededor de 48 en 2005 y 2006. En matemáticas la media se ubica en 41 para el año 2004 y 45 para los años 2005 y 2006. En la muestra (y para cada año) cerca del 60% son mujeres, esta cifra es consistente con los estudios de logro académico anteriores, los cuales confirman la mayor escolaridad de las mujeres desde hace ya bastante tiempo.

El promedio de edad de los bachilleres en el momento de la inscripción es de 17 años y una desviación estándar bastante pequeña. En el caso de la alimentación, el 35% de los estudiantes recibió algún tipo de alimento en el colegio, éste podía ser un vaso de leche, un desayuno o un almuerzo completo. Este programa, conocido como Almuerzo escolar está destinado para jóvenes entre 5 y 17 años de edad, residentes de Medellín y matriculados en establecimientos educativos oficiales. El alimento consiste en el suministro de un desayuno preparado (se entrega en bandeja lista para el consumo) o un almuerzo, se prevé que este alimento sea entregado durante la jornada escolar de lunes a viernes durante 185 días al año.<sup>3</sup> El cruce de la información de EPS y ALIM indica que el 40% de los estudiantes no reciben almuerzo escolar ni tienen EPS.

El 85% de los estudiantes está distribuido en el estrato 2 y 3, con un predominio del estrato 2 en cerca de 10 puntos porcentuales. Los estudiantes de estrato 1 son aproximadamente el 9% y los de estrato 4 el 3,5 %; el resto, en una proporción muy pequeña, pertenecen a estrato 5 y 6. Un poco más del 90% de los estudiantes son atendidos en colegios de jornada mañana y tarde, el resto en jornada completa y nocturna. La información de afiliación a empresas prestadoras de servicios de salud indica que el 59% reportó estar afiliado a alguna EPS.

La variable  $R_i$  ( $i=1, 2, 3, 4$ ) es una variable *dummy* que toma el valor =1 si la razón profesor alumno está en el rango indicado,  $e = 0$  en otro caso. Respecto a los insumos escolares llama la atención que, la media del número de computadores con conexión a internet es de 1,1, lo que es ciertamente una cifra muy baja. En cuanto a la planta docente, el total de equipos (*TV's, VHS's, etc*) y las unidades sanitarias, destacamos la gran variabilidad en estos factores entre colegios. En promedio, los colegios tienen una planta docente de 62 profesores, aunque una varianza relativamente grande que puede indicar déficit de profesores en algunos colegios.

La tabla 2 muestra el coeficiente de correlación para todos los años entre los distintos puntajes analizados. Es interesante notar que mientras esperaríamos

---

<sup>3</sup> Adicionalmente, para la temporada de vacaciones y fines de semana los usuarios de restaurante reciben alimentos enriquecidos con vitaminas y minerales para consumir en el hogar (normalmente leche y galleta).

altas correlaciones entre lenguaje y puntaje total, en ningún caso se supera el 0,77 y sorprende aún más que la correlación entre matemáticas y lenguaje en todos los casos está por debajo de 0,3. Este resultado se debe quizás a la naturaleza de las áreas evaluadas, es decir, mientras que la mayor parte del examen ICFES requiere de una habilidad para memorizar, otra parte de examen, como el módulo de matemática, por ejemplo, requiere habilidades y aptitudes. Por tanto, no es de extrañar la baja correlación observada entre matemáticas y los demás módulos.

Tabla 2. *Correlaciones Puntaje total, Matemáticas y Lenguaje*

Variables	Puntaje total	Matemáticas	Lenguaje
2004			
Puntaje total	1	---	---
Matemáticas	0,52	1	---
Lenguaje	0,77	0,29	1
2005			
Puntaje total	1	---	---
Matemáticas	0,48	1	---
Lenguaje	0,72	0,20	1
2006			
Puntaje total	1	---	---
Matemáticas	0,61	1	---
Lenguaje	0,64	0,28	1

Fuente: ICFES. Cálculos propios.

## II. Análisis Estadístico

### A. *Determinantes del logro académico*

Este estudio intenta hallar el efecto de los factores individuales y escolares sobre el desempeño académico. Lastimosamente, como se mencionó antes, la poca información contenida en el formulario de inscripción del ICFES es en parte remplazada por la que se halla en ML, siendo consciente de que posiblemente falta información. Para lograr el objetivo, usamos el enfoque de la función de producción. Con más exactitud se ha supuesto que el resultado académico es el producto de la interacción conjunta de factores a través de una función en el sentido matemático. Esto es, el colegio es una caja negra

donde entran los estudiantes, sus factores individuales y los insumos escolares interactúan y producen el logro académico, medido por lo general como puntajes en pruebas estandarizadas. En términos formales, este enfoque nos dice que  $Y_{ij}$  se puede relacionar con la matriz de características individuales  $X_{ij}$  y la matriz de características del plantel  $Z_j$ , a través de una función  $F(X_{ij}, Z_j)$  por lo general lineal, tal que:

$$Y_{ij} = \beta X_{ij} + \phi Z_j + \varepsilon_{ij} \quad (1)$$

Donde  $Y_{ij}$  es el puntaje del estudiante  $i$  en el colegio  $j$ ,  $X_{ij}$  es un vector que contiene atributos individuales del estudiante  $i$  (el sexo, la edad y el estrato). La matriz  $Z_j$  contiene los factores escolares como la razón profesor/alumno, el número de profesores totales (planta docente), el tipo de jornada, equipos, etc. El término de error denotado por  $\varepsilon_{ij}$  se asume con media condicionada a  $X_{ij}$  y  $Z_j$  igual a cero y varianza constante. La tabla 3 muestra los resultados de estimar el modelo (1) por mínimos cuadrados ordinarios, admitiendo la existencia de posibilidad de heteroscedasticidad a nivel de colegios. El valor del estadístico  $t$  ( $t$ -valor) es estimado usando la matriz de covarianzas de White.

Tabla 3. Características Individuales y del Plantel

Variables	2004						2005					
	Total		Lenguaje		Matemática		Lenguaje		Matemática		t-valor*	
	Coficiente	t-valor*	Coficiente	t-valor*	Coficiente	t-valor*	Coficiente	t-valor*	Coficiente	t-valor*		
Edad	-5,2	12,6	-1,1	11,0	-0,57	7,9	-1,2	12,0	-0,61	7,9		
Sexo	5,3	4,6	1,2	5,0	0,54	3,4	-0,3	0,9	0,61	3,4		
EPS	-2,1	3,0	-0,42	2,6	-0,34	3,1	-----	-----	-----	-----		
ALIM	-2,0	1,7	-0,74	2,7	-0,30	2,0	-----	-----	-----	-----		
Estrato 2	-0,69	0,5	-0,19	0,53	0,21	1,0	-0,02	0,08	0,24	1,0		
Estrato 3	3,4	2,1	0,75	1,9	0,44	1,7	1,3	3,0	0,72+	1,7		
Estrato 4	9,4	4,5	1,81	3,3	0,82	2,0	2,1	3,2	1,2	2,0		
R1<0,02	13,2	3,3	2,6	2,8	0,53	1,36	3,3	4,3	1,1	1,36		
0,02<R2<0,03	7,1	2,2	1,38	1,8	0,1	0,46	1,9	2,9	0,46	0,46		
0,03<R3<0,04	8,9	3,0	1,46	2,0	0,84	4,03	2,7	4,3	0,62	4,03		
Mañana	15,0	4,9	2,56	3,4	-0,3	1,44	9,2	15,3	3,9	1,44		
Tarde	1,6	1,16	0,26	0,95	-0,02	0,11	0,14	0,4	-0,38	0,11		
Planta Doc	0,11	2,4	0,027	2,2	0,007	1,4	0,03	2,3	0,02	1,4		
M2aulas	-0,002	2,0	-0,006	2,0	-0,0003	1,7	-0,001	2,4	-0,0006	1,7		
Comp www	2,65	1,9	0,38	1,36	0,19	0,91	0,3	1,1	0,07	0,91		
Equip. Total	0,24	2,1	0,02	0,95	0,027	2,2	0,05	2,7	0,03	2,2		
No Uni. Sanit	0,04	0,6	0,01	0,60	0,01	2,4	0,02++	1,7	0,01	2,4		
R2		0,10		0,08		0,04		0,07		0,02		

Fuente: ICFES, DANE y ML. Cálculos propios

\* Valor absoluto de la distribución t-student.



Como se ha venido observando en los variados y recientes artículos de investigación, la edad es un factor determinante en el logro y éste tiene un efecto negativo al incrementarse la edad; es decir, los estudiantes más viejos tienen más problemas para obtener buenos puntajes. Un año adicional en la edad, significa cinco puntos menos en la prueba total y un punto menos en lenguaje y matemáticas; podemos siempre argumentar, por ejemplo, que los individuos viejos tienen más dificultad de aprender o tienen altas tasas de reprobación.

A pesar de que hay más mujeres con intención de escolarizarse, su desempeño medio es inferior al de los hombres. En promedio, las mujeres obtienen aproximadamente 5 puntos menos que el hombre y en el puntaje total, medio punto menos en matemática y casi un punto menos en lenguaje. Esto se observa para ambos años; las causas sociológicas de este sesgo en contra de las mujeres, se han intentado explicar de varios modos, por ejemplo los hombres tienen una tendencia mayor a estudiar matemáticas y ciencias que las mujeres, o ellos son mejores haciendo “trampa” o incluso, se ha argumentado que las pruebas están diseñadas a la medida de los hombres.

Para el año 2004 construimos una variable binaria  $ALIM=1$  si el individuo  $i$  recibió algún tipo de alimento en el colegio  $j$  y cero en otro caso. De acuerdo con los resultados, no parece que haya un efecto positivo de los alimentos, de hecho, los resultados para estos individuos son peores, obtienen 2 puntos menos en el puntaje total y casi medio punto menos en lenguaje y matemáticas. La explicación de este resultado, aparentemente contra intuitivo, es que los jóvenes elegibles para la asistencia alimentaria llegan al sistema educativo en malas condiciones nutricionales, y quizás, de no haber recibido el alimento su desempeño promedio habría sido peor.

Dicho en otros términos, el riesgo nutricional por bajo peso es, simple y llanamente, la señal inequívoca de una corta vida llena de penurias, por lo cual el paliativo de la alimentación al interior de la escuela no es suficiente (o es quizá inútil) si se pretende un desempeño mejor que el de aquellos individuos con bajo o ningún tipo de riesgo nutricional. Es probable que individuos con muy bajo peso no sean realmente alimentados con un almuerzo o con un vaso de leche. La alimentación suministrada de este modo puede ser importante con fines de retención en el sistema educativo, pero con fines de mejoramiento del logro académico futuro, puede ser más pertinente atender la población de 0 a 5 años.

Construimos una variable binaria  $OUTRC =1$  si el individuo  $i$  está fuera del régimen contributivo. De acuerdo con el resultado reportado en la tabla 3, para 2004, no estar afiliado a EPS significó en promedio 2 puntos menos en el puntaje total y un poco menos de medio punto en lenguaje y matemáticas.

Además el 32,5% de los jóvenes con bajo riesgo nutricional (elegibles para recibir alimento) no estaban afiliados a alguna EPS y residían en los estratos más bajos, de modo que la pobreza implica bajo acceso a alimentos y una pobre atención médica, pero incluso, si tienen acceso a ésta, nada garantiza que puedan conseguir medicina o exámenes médicos. Estar fuera del régimen contributivo afecta el desempeño, pues un deficiente acceso al sistema de salud, aunado a la disminución de defensas de los jóvenes por baja nutrición, los hace más débiles, pierden la atención con facilidad y son más propensos a enfermedades.

Los resultados de la tabla 3 refrendan algo que se sospecha, el estrato de los individuos (vía ingreso familiar) es importante para garantizar un adecuado desempeño académico. Nótese que a medida que el estrato del individuo mejora, el coeficiente estimado es positivo y creciente (aunque anómalamente no es significativo para el estrato 2). El hecho de residir en determinado estrato influye en el desempeño, esto se debe a varios factores, pero sin duda el ingreso familiar permite a los padres pagar una matrícula escolar más alta, aun si hablamos de colegios públicos (muchos de ellos cobran por declaración de renta de los padres o precisamente por el estrato).<sup>4</sup> A esta conclusión también llegan Gaviria y Barrientos (2001a) pero vía educación de los padres. Este rasgo positivo del estrato se nota en los tres tipos de puntaje, el total, lenguaje y matemáticas y para ambos años.

La razón profesor-alumno tiene un efecto positivo y estadísticamente significativo, salvo en matemáticas. La razón puede ser que no importa cuántos recursos se dediquen para la enseñanza de áreas cuyo desempeño depende en gran medida de una habilidad natural. Salvo que se diseñen programas focalizados a la enseñanza de matemáticas. El puntaje total (sin la prueba de inglés) es la suma de los puntajes en lenguaje, matemáticas, las electivas y las ciencias básicas y sociales, de modo que no es sorprendente el efecto positivo al que hacemos mención. Estudiantes en la jornada de la mañana tienen un mejor desempeño que en la tarde, aunque el coeficiente de la variable *tarde* no es estadísticamente significativo. La planta docente tiene un efecto positivo aunque en matemáticas no parece ser muy importante.

Salvo el área de las aulas (medida en metros cuadrados), las demás variables de infraestructura tienen un efecto positivo. El total de equipos no es muy importante en lenguaje, pero parece tener algún efecto sobre el desempeño en matemáticas. Signos negativos de los coeficientes estimados de variables de

---

<sup>4</sup> El estrato del hogar lo define Empresas Públicas de Medellín y en la factura se especifica.

infraestructura, por lo general implican un uso poco adecuado o incluso poco eficiente de la capacidad instalada de las instituciones, así como heterogeneidades no observables entre colegios. Una de esas heterogeneidades no observables, de momento, es la gestión en todas sus dimensiones; administrativa, directiva, académica e interacción con la comunidad, etc. Estas dimensiones no se han medido y apenas se han hecho tímidos intentos por cuantificarlas, ver Barrientos *et al.* (2007).

### ***B. Regresión por cuantil de rendimiento académico***

En este punto, cabe preguntarnos si el efecto de las variables se mantiene constante cuando analizamos el desempeño por cuantil (percentil, decil, quintil etc.) de rendimiento. En otros términos, se hace la pregunta ¿para quienes o para qué tipo de estudiantes las variables independientes tienen impacto? En esta sección presentamos las estimaciones del modelo (1) para diferentes percentiles de rendimiento usando *Quantile Regression (QR)*. El modelo estadístico usado en esta sección especifica el  $\theta$ th percentil de la distribución condicionada de  $Y$  dado  $X$  y  $Z$  como una función lineal de los regresores,

$$Q_{\theta}(Y|X, Z) = \beta(\theta)X + \varphi(\theta)Z, \quad \theta \in (0,1) \quad (2)$$

La inferencia acerca de  $\beta(\theta)$  y  $\varphi(\theta)$  se realiza usando técnicas de Koenker y Bassett (1978, 1982). Las tablas 5 y 6 reportan los resultados del análisis de regresión por cuantil de rendimiento para el puntaje total en los años 2005 y 2006 y para las pruebas de lenguaje y matemáticas en 2006. Un cálculo rápido indica que existe una dispersión relativa en las variables dependientes, puntaje total, y en especial en los puntajes de lenguaje y matemáticas, los cuales indican que *QR* es una mejor estrategia empírica que simplemente mínimos cuadrados ordinarios sobre la media de la distribución. Los correspondientes coeficientes de variación de las mencionadas variables dependientes son presentados en la tabla 4.

Tabla 4. *Coefficientes de Variación Puntajes*

Variable dependiente	2004	2005	2006
Total	0,08	0,09	0,09
Lenguaje	0,14	0,17	0,12
Matemáticas	0,14	0,16	0,17

*Fuente:* cálculos propios

Tabla 5. Regresión Intercuantil  
Características Individuales y del Plantel. Puntaje Total

Variables	2005										2006									
	OLS					Quantile					OLS					Quantile				
	0,05	0,25	0,5	0,75	0,95	0,05	0,25	0,5	0,75	0,95	0,05	0,25	0,5	0,75	0,95					
Edad	-5,4*	-3,0*	-4,7*	-6,2*	-0,8*	-4,3*	-2,8*	-3,8*	-4,4*	-6,3*	-4,3*	-2,8*	-3,8*	-4,4*	-6,3*					
Sexo	2,9+	-0,7	1,2	4,6*	9,1*	7,2*	1,8+	3,4*	7,5*	13,9*	7,2*	1,8+	3,4*	7,5*	13,9*					
Estrato 2	-0,3	0,56	0,27	-0,1	5,2	0,02	-0,9	-0,73	0,76	0,01	0,02	-0,9	-0,73	0,76	0,01					
Estrato 3	5,0*	2,0	3,5*	6,1*	4,8	5,7*	1,4	2,5**	6,2*	12,6*	5,7*	1,4	2,5**	6,2*	12,6*					
Estrato 4	11,7*	10,5*	9,4*	15,0*	16,0**	11,2*	5,8+	9,5*	9,8*	18,0*	11,2*	5,8+	9,5*	9,8*	18,0*					
R1<0,02	17,9*	12,1*	12,0*	21,8*	28,2*	9,7*	10,6*	9,2*	10,8*	10,0*	9,7*	10,6*	9,2*	10,8*	10,0*					
0,02<R2<0,03	10,4*	6,13*	6,1*	14,4*	12,5**	4,8	5,1**	4,7*	6,2*	4,02	4,8	5,1**	4,7*	6,2*	4,02					
0,03<R3<0,04	13,3*	7,2*	8,1*	11,6*	20,7*	4,7	6,6*	4,2*	6,1*	3,9	4,7	6,6*	4,2*	6,1*	3,9					
Mañana	52,9*	5,5	8,7*	49,4*	55,4*	4,2**	-1,2	-1,8	-3,2**	-0,88	4,2**	-1,2	-1,8	-3,2**	-0,88					
Tarde	0,6	-1,1	0,21	0,02	1,75+	2,8	3,1**	2,9*	4,05**	7,4*	3,1**	2,9*	4,05**	5,2*	7,4*					
Planta Docente	0,2+	0,14*	0,15*	0,16*	0,17+	0,07	0,04	0,07*	0,06**	0,11**	0,07	0,04	0,07*	0,06**	0,11**					
M2aulas	-0,006*	-0,004*	-0,005*	-0,006*	-0,006**	-0,003**	-0,002**	-0,003*	-0,002*	-0,003+	-0,003**	-0,002**	-0,003*	-0,002*	-0,003+					
Comp www	2,3	3,0**	1,5	2,0+	2,9**	1,4+	0,93	1,6**	0,5	3,8**	1,4+	0,93	1,6**	0,5	3,8**					
Equip. Total	0,24**	0,10	0,24*	0,2*	0,27*	0,1	0,07	0,09**	0,15*	0,25**	0,1	0,07	0,09**	0,15*	0,25**					
No Uni. Sanit	0,09	0,08	0,08*	0,1*	0,07+	0,09+	0,09**	0,08*	0,06**	0,05	0,09+	0,09**	0,08*	0,06**	0,05					
R2	0,1	-----	-----	-----	-----	0,07	-----	-----	-----	-----	0,07	-----	-----	-----	-----					

Fuente: ICFES, DANE y ML. Cálculos propios

(\*) Significativos al 1%

(\*\*) Significativos al 5%

(+) Significativos al 10%

Tabla 6. Regresión Intercuantil Características Individuales y del Plantel 2006

Variables	Lenguaje					Matemáticas						
	OLS	Quantile				OLS	Quantile					
		0,05	0,25	0,5	0,75	0,95		0,05	0,2	0,5	0,75	0,95
Edad	-0,74*	-0,8*	-0,86*	-0,66*	-0,86*	-0,54*	-0,83*	-0,6*	-0,7*	-0,81*	-0,85*	-1,0*
Sexo	0,63*	0,06	0,65*	0,64*	0,96*	0,60**	2,8*	2,6*	2,2*	2,6*	3,4*	4,3*
Estrato 2	0,23	0,31	0,56*	0,03	0,12	0,54	-0,28	-0,6	-0,27	-0,33	0,09	-0,53
Estrato 3	1,1*	1,0**	0,56*	1,09*	0,74*	1,6*	0,8**	0,017	0,4	1,1*	1,1*	0,94
Estrato 4	1,8*	1,0	1,2*	1,7*	1,8*	1,9**	1,9*	2,7*	1,1+	2,0*	2,0*	2,2**
R1<0,02	2,1*	3,6*	1,8*	2,0*	1,68*	2,6*	0,83	0,87	1,0**	0,95+	0,12	0,23
0,02<R2<0,03	1,3**	3,1*	2,4*	1,3*	1,05*	0,6	0,27	0,01	-0,08	0,41	-0,03	0,82
0,03<R3<0,04	1,2**	2,1*	1,6*	1,2*	1,1*	0,7	0,7	0,13	0,50	0,62	0,29	1,2+
Mañana	-0,16	-0,5	1,3	-0,06	0,032	0,57	-0,46	-0,47	-0,63	-0,5	-0,15	-0,72
Tarde	0,65**	0,41	0,14*	0,5*	0,55*	0,75**	1,1*	0,6	0,67*	0,74*	1,4*	1,67*
Planta Doc	0,01+	0,02**	0,8*	0,01*	0,005+	0,03*	0,009	0,00002	0,01**	0,01**	-0,001	0,009
M2aulas	-0,003	-0,0003	0,01*	-0,0004*	0,00002	-0,0004	-0,0004	0,00002	-0,0006*	-0,0005**	-0,0001	-0,0006
Comp www	0,14	0,17	-0,0004	0,13	0,05	0,57+	0,14	0,07	0,1	-0,01	0,061	-0,40
Equip. Total	0,009	0,01	0,2**	0,01+	0,004	0,02	0,03	0,03	0,02	0,32**	0,034*	0,008
No Uni. Sanit	0,001	-0,01	0,001	0,002	0,004	-0,02	0,02**	0,01	0,02**	0,01	0,027*	0,016
R2	0,04	-----	-----	-----	-----	-----	0,04	-----	-----	-----	-----	-----

Fuente: ICFES, DANE y ML. Cálculos propios

(\*) Significativos al 1%

(\*\*) Significativos al 5%

(+) Significativos al 10%.

La variabilidad del puntaje total en el examen ICFES usando regresión por cuantil de rendimiento sólo se hizo para el año 2005 y 2006. Lo primero que se detecta es que la edad es un factor que afecta el desempeño de los alumnos a lo largo de toda la distribución de los puntajes. No importa si se mira los de mejor o peor desempeño, sin duda los alumnos de más edad siempre tienen un peor desempeño.

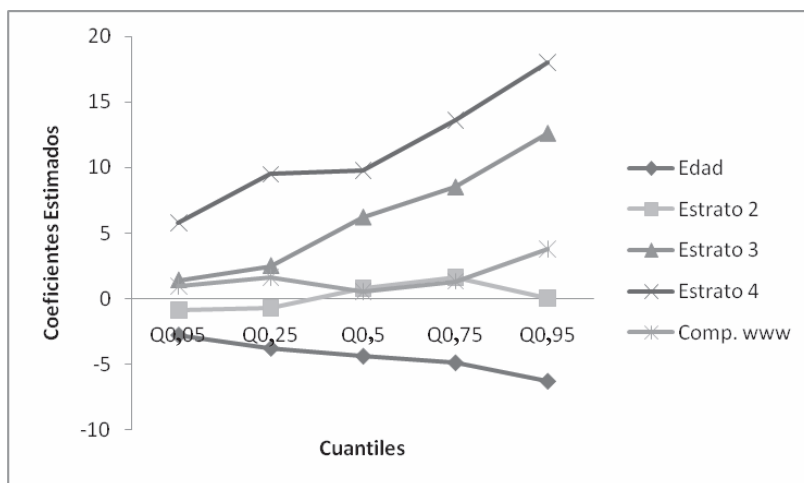
El estrato socioeconómico de los alumnos, tiene un efecto similar tanto sobre la media como sobre otros segmentos de la distribución de los puntajes; es decir, el puntaje se incrementa a través del estrato, no importa el segmento de la distribución de puntajes sujeto a análisis. Si un estudiante de desempeño bajo es puesto en estrato medio o alto, con todo lo que eso implica, tendrían gran probabilidad de mejorar su desempeño y viceversa. Al igual que el estrato, la razón profesor-alumno tiene un efecto similar sobre la media que sobre otros segmentos de la distribución de los puntajes. Quizá debamos mencionar que toda significancia estadística se pierde en el caso de matemáticas; reiterando que esta área del conocimiento parece depender en gran medida de factores externos al colegio, incluso cuando se analizan diferentes segmentos de la distribución.

Cuando miramos el efecto de los factores del colegio tenemos una pérdida de significancia estadística, salvo en la variable *área construida de las aulas* cuyo efecto es negativo (aunque la magnitud muy baja) reiterando que quizá los recursos no son adecuadamente utilizados; la planta docente, que es estadísticamente significativa y positiva en lenguaje, deja de serlo en matemáticas, otra vez, parece ser que los resultados en matemáticas parecen ser invariantes ante factores escolares.

En el caso de la planta docente tenemos un efecto positivo y creciente a través de los cuantiles, en especial hasta el tercer cuartil ( $q=0,75$ ) para el puntaje total en el año 2005. En el año 2006 el efecto es menor, incluso no muy significativo. Más interesante que observar el desempeño a través del puntaje total, es el efecto de los regresores sobre lenguaje y matemáticas para el año 2006. El efecto sobre el área de lenguaje es positivo e importante a varios niveles de significancia y en todos los cuantiles evaluados, es especialmente importante en el primer cuartil ( $q=0,25$ ). En matemáticas el efecto es menos claro, tanto en magnitud como en significancia. Las habilidades innatas de las personas son responsables del buen desempeño en matemáticas.

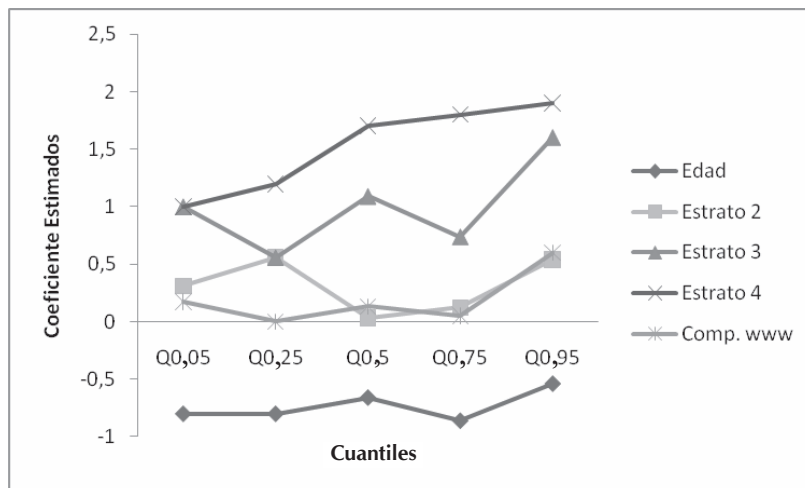
Los gráficos 1-3 muestran la distribución de los coeficientes estimados por cuantil de rendimiento en el puntaje total, lenguaje y matemáticas para el año 2006. Nótese que la estimación de los coeficientes a lo largo de la distribución de los puntajes está lejos de ser lineal y muestra un pico que indica claramente que el impacto de las variables se acentúa dependiendo del rendimiento de los alumnos. Por ejemplo, el impacto de residir en estrato 3 y 4 es mayor entre aquellos estudiantes con rendimiento superior a la media en el caso de matemáticas, y en

lenguaje el estrato 3 es sólo importante para estudiantes ubicados en el 25% con mejor desempeño. Residir en estrato 3 y 4 impacta más a aquellos estudiantes cuyo logro en el puntaje total está por encima de la mediana.



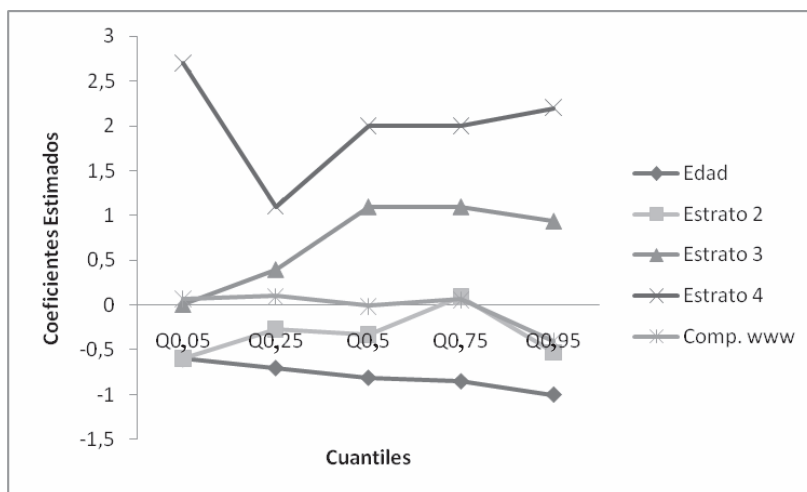
Fuente: cálculos propios

Gráfico 1. Distribución de los coeficientes estimados por nivel de rendimiento en el puntaje total, 2006



Fuente: cálculos propios

Gráfico 2. Distribución de los coeficientes estimados por nivel de rendimiento en el Lenguaje, 2006



Fuente: cálculos propios

Grafico 3. *Distribucion de los coeficientes estimados por nivel de rendimiento en el Matematicas, 2006*

### C. La importancia del colegio

Como ha sido señalado por la literatura, el logro académico depende fundamentalmente de dos grupos de factores a saber, los individuales y los escolares. Lo deseable desde el punto de vista teórico y aplicado, y como recomendación de tipo político, es determinar con la mayor precisión posible la incidencia de este grupo de factores. Por ejemplo, si el colegio (en promedio) por sí solo es “muy importante” entonces, un programa amplio en cobertura educativa bastaría para fortalecer el desempeño académico y mejorar las competencias de los alumnos. En este trabajo las características del plantel incluyen, además de la infraestructura, las características promedio de los individuos (como el estrato).

En esta sección y en la siguiente, replicamos el análisis estándar llevado a cabo por Gaviria y Barrientos (2001a, 2001b) y Restrepo y Alviar (2005). Tal análisis propone descomponer el efecto sobre el rendimiento académico. Como es bien sabido la descomposición divide la varianza del rendimiento académico en la varianza media del resultado de los individuos al interior de un mismo plantel y la varianza de los promedios de los resultados de cada colegio. Si la importancia relativa del último componente es alta, entonces mayor será el efecto del plantel sobre el desempeño y menor, claro está, la relevancia de los atributos individuales. En otras palabras, el primer componente da una medida del efecto de las características individuales.



Desde el punto de vista técnico, el efecto del colegio se calcula como el coeficiente de determinación ajustado, denotado por  $R_A$ , de una regresión cuya variable dependiente es el puntaje de un examen estandarizado y donde las variables independientes son una variable binaria que toma el valor 1 si el individuo  $i$  está en el colegio  $j$  y 0 en otro caso.

La tabla 7 muestra el  $R_A$  para el puntaje total (sin la prueba de inglés) y las áreas de matemáticas y lenguaje. La incidencia del plantel en el sector público asciende, en los tres años, al 10% aproximadamente en la prueba total, la variabilidad del rendimiento en lenguaje es de 7,5% en 2004 y 2005, pero con una caída en 2006 al 4%, a diferencia del efecto del colegio sobre el desempeño en matemáticas que presenta altibajos, es del 3% en 2004, cae al 2% y en 2006 es del 6%. Nóte que la menor incidencia del colegio en matemáticas se debe a la aptitud natural de los individuos más que a las características individuales y del plantel. Dicho en otros términos, es probable que tanto las características familiares, como el plantel, tengan poca injerencia sobre las habilidades en matemáticas, conclusión que está presente en los trabajos recientes sobre el efecto del colegio.

Tabla 7. *Incidencia del Plantel Colegios Públicos*

	Total	Lenguaje	Matemáticas
2004	0,09	0,07	0,03
2005	0,11	0,08	0,02
2006	0,09	0,04	0,05

Fuente: ICFES, DANE y ML. Cálculos propios

Para analizar si las cifras reportadas en la tabla 7 son altas o bajas, se deben comparar con los colegios no oficiales (NO) de Medellín de un lado, y de otro lado, con los colegios oficiales de Bogotá. La tabla 8 muestra la distribución de la incidencia del plantel por tipo de institución para el año 2004 en ambas ciudades. La diferencia en la incidencia del plantel en el puntaje total es de 2%, pero en el cálculo de los valores de la tabla 7 sólo se tiene en cuenta la información contenida en el formulario de inscripción y resultado del ICFES, por tal razón la diferencia en el tamaño de la muestra.

**Tabla 8. Medellín, Bogotá: incidencia del Plantel por tipo de Institución 2004**

Medellín				
	Observaciones	Total	Lenguaje	Matemáticas
Oficial	12455	0,11	0,07	0,03
No Oficial	5835	0,32	0,20	0,15
Todos	18290	0,30	0,20	0,12
Bogotá				
	Observaciones	Total	Lenguaje	Matemáticas
Oficial	32889	0,10	0,07	0,03
No Oficial	26654	0,35	0,23	0,13
Todos	59543	0,28	0,18	0,10

Fuente: ICFES. Cálculos propios

El efecto del plantel, tanto en Medellín como en Bogotá, es el mismo, con una ligera diferencia para Medellín, lo que indica que la baja influencia de los colegios públicos es generalizada y no depende en gran medida de la geografía; los estándares (bajos) de los colegios públicos en promedio son casi idénticos con independencia de la ciudad. Sin embargo, hay una diferencia de 12 y 15 puntos porcentuales respectivamente al asistir a un colegio no oficial.

La tabla 9 muestra el efecto colegio calculado con base en información del ICFES 1999, tanto para Medellín como Bogotá. No obstante, hay que hacer una anotación, a partir de 2000 el examen cambio en su fondo y en su forma; por tanto los resultados son difícilmente comparables. Sin embargo, podemos observar que el efecto del plantel es más fuerte en los colegios privados que en los públicos (independiente del tipo del examen aplicado). Esto puede explicarse por dos razones: la primera, que sea un efecto estadístico debido al nuevo examen implementado, la escala de calificación o la metodología de la prueba *per se*, y segundo, que ciertamente los atributos (individuales y socioeconómicos) estén jugando un factor aún más crucial.

**Tabla 9. Medellín, Bogotá: Incidencia del Plantel por tipo de Institución 1999**

Medellín				
	Observaciones	Total	Lenguaje	Matemáticas
Oficial	11526	0,15	0,12	0,13
No Oficial	6962	0,45	0,35	0,39
Todos	18488	0,37	0,28	0,31
Bogotá				
	Observaciones	Total	Lenguaje	Matemáticas
Oficial	25595	0,21	0,15	0,17
No Oficial	35770	0,42	0,30	0,35
Todos	61365	0,36	0,26	0,30

Fuente: ICFES. Cálculos propios

Nótese además que hay una diferencia, aunque moderada, de estos resultados y los reportados por Gaviria y Barrientos (2001a) y Restrepo y Alviar (2005), pero aún el efecto del colegio sobre el puntaje en las pruebas se mantiene en la misma dirección. Por ejemplo, cuando tenemos en cuenta todas las instituciones (oficiales y no oficiales) para Bogotá en 1999, es casi la misma incidencia que la reportada en la tabla 6 por Gaviria y Barrientos (2001a). Esta diferencia, reportada por las tablas 6 y 9, se debe a las restricciones<sup>5</sup> que se impusieron a la información para el análisis estadístico, exactamente las mismas para obtener los resultados de las tablas 7 y 8; de modo que, con las mismas variables, las mismas restricciones (salvo el tamaño de muestra) que en ambos casos se considera representativa, la única diferencia que puede ser crucial es el examen. Claramente, una conclusión definitiva sobre el efecto del cambio en el examen y posiblemente el cambio en la tendencia de la incidencia del plantel sobre el logro académico, merece un análisis más amplio y riguroso.

#### **D. Características individuales dentro del plantel**

En esta sección analizamos el efecto de las características individuales entre individuos que asisten al mismo plantel. El análisis está inspirado en Gaviria y Barrientos (2001a) y basado en el siguiente modelo empírico:

$$Y_{ij} = \beta X_{ij} + \eta_j + \varepsilon_{ij} \quad (3)$$

Con  $i=1, \dots, N$ , y  $j=1, \dots, J$ , donde  $\eta_j$  es un efecto individual no observable y  $\varepsilon_{ij}$  denota el error clásico. En este caso los  $\eta_j$  se asumen como parámetros fijos a ser estimados y se supone que  $\varepsilon_{ij}$  es independiente e idénticamente distribuido con  $E(\varepsilon | X) = 0$  y  $Var(\varepsilon | X) = \sigma_{\varepsilon}^2$  constante. Las demás variables del modelo (3) tienen la misma interpretación que en la estimación de (1). Puesto que  $\eta_j = 1$  si el individuo  $i$  asiste al plantel  $j$  y  $\eta_j = 0$  en otro caso, el modelo (3) sólo da cuenta de las diferencias en logro académico de los individuos que asisten al plantel  $j$ . Es decir, descartamos la varianza proveniente de los diferentes colegios. Expresado el modelo (3) en forma compacta nos da una idea más exacta del modelo a estimar,

$$Y = \beta X + C\eta + \varepsilon \quad (4)$$

Donde  $C$  es una matriz de *dummies* individuales de dimensión  $N \times J \times N$ , dada por  $C = I_N \otimes I$  y  $I$  es un vector de unos de dimensión  $J$  y  $\eta$  es el vector de efectos individuales a estimar.

<sup>5</sup> Individuos mayores de 15 años y menores de 27, los puntajes deben ser mayores que cero, sólo se tienen en cuenta estudiantes que estaba en la cohorte correspondiente, aquellos que eran de otras cohortes fueron descartados.

La tabla 10 muestra los resultados de la estimación del modelo (3) para los diferentes años, las diferentes áreas y el puntaje total. Las estimaciones muestran que la diferencia en edad es muy similar a las obtenidas en la estimación del modelo (1), al igual que el sexo, el sesgo positivo en desempeño hacia los hombres y los más jóvenes tiene poco o nada que ver con la diferencia entre colegios. En otras palabras, el sesgo mencionado se mantiene cuando comparamos individuos que asisten al mismo plantel y cuando comparamos individuos en diferentes planteles.

Tabla 10. *Características Individuales controlando por Colegio*

2004						
	Total		Lenguaje		Matemáticas	
	Coefficiente	t-valor*	Coefficiente	t-valor*	Coefficiente	t-valor*
Edad	-4,8	15,2	-1,0	11,8	-0,53	8,1
Sexo	8,6	11,4	1,9	9,7	1,0	6,2
EPS	-1,2	1,8	-0,24	1,4	-0,3	2,2
ALIM	-0,006	0,01	-0,14	0,5	-0,04	0,2
Estrato 2	-1,0	0,9	-0,36	1,2	0,11	0,5
Estrato 3	2,5	2,0	0,4	1,2	0,29	1,12
Estrato 4	7,5	3,7	1,18	2,2	0,57	1,5
R2	0,13		0,10		0,04	
2005						
	Total		Lenguaje		Matemáticas	
	Coefficiente	t-valor*	Coefficiente	t-valor*	Coefficiente	t-valor*
Edad	-4,6	12,3	-1,0	11,2	-0,55	6,3
Sexo	7,6	8,7	0,6	2,5	0,8	4,1
Estrato 2	-0,5	0,4	-0,03	0,1	0,1	0,3
Estrato 3	3,0	2,0	1,02	2,9	0,4	1,4
Estrato 4	7,0	2,7	1,3	2,2	1,0	1,8
R2	0,14		0,10		0,025	
2006						
	Total		Lenguaje		Matemáticas	
	Coefficiente	t-valor*	Coefficiente	t-valor*	Coefficiente	t-valor*
Edad	-3,8	12,2	-0,67	9,6	-0,7	8,9
Sexo	9,2	12,4	1,0	6,0	3,0	15,2
Estrato 2	-0,2	0,2	0,16	0,7	-0,3	1,1
Estrato 3	3,4	2,9	0,7	2,7	0,4	1,2
Estrato 4	8,8	4,2	1,3	3,0	1,4	2,7
R2	0,12		0,06		0,08	

Fuente: ICFES, DANE y ML. Cálculos propios

\* Valor absoluto de la distribución t-student

Para el año 2004 la significancia de la variable *ALIM* desaparece; esto puede interpretarse como un asunto de diferencia entre colegios. Hay colegios donde los jóvenes padecen riesgo nutricional por bajo peso y una gran mayoría de estos individuos asisten al mismo colegio (diferencias inexistentes entre individuos en el mismo colegio). Respecto a la exclusión del régimen contributivo, la evidencia sugiere que la diferencia en desempeño entre quienes están en una EPS y aquellos que no lo están, no parece que tenga nada que ver con las diferencias entre los colegios.

El estrato socioeconómico de los alumnos muestra una tendencia creciente, tanto en magnitud como en significancia, en todos los años y en cada área del conocimiento. Los estudiantes en estrato 4 tienen un mejor desempeño que los de estrato 3, y estos a su vez, un mejor desempeño que los de estrato 2. Estos resultados son muy similares cuando se estima controlando por factores escolares, lo que indica que la diferencia entre estudiantes por estrato, de nuevo, tiene poco que ver con diferencias entre colegios. La diferencia en desempeño por estrato es la misma si se comparan individuos que asisten a la misma escuela con aquellos que asisten a una diferente.

### **Conclusiones**

Varias conclusiones se pueden sacar de este análisis. Primero, que los factores escolares, en especial el estrato, los computadores y la razón profesor-alumno, tienen un efecto diferente sobre diversos segmentos de la distribución del examen ICFES de los individuos. Segundo, que el efecto del colegio parece cada vez menos fuerte, se acentúa el efecto de los atributos individuales (y probablemente de las heterogeneidades no observables). La evidencia indica que el colegio afecta más a los estudiantes en instituciones no oficiales que aquellos en instituciones públicas. Dicho de otro modo, los atributos medios del colegio (y de individuos) en los no oficiales es más relevante que en los públicos. Tercero, se observa menor incidencia del colegio en la variabilidad del rendimiento en matemáticas, que le atribuimos en buena medida a la aptitud natural de los individuos, más que a las características del plantel. Cuarto, no hay una diferencia sustancial entre ciudades al asistir a un colegio público, los individuos en colegios en Bogotá o Medellín se ven afectados en promedio del mismo modo, es decir, la baja calidad de la educación pública es general. Finalmente, respecto a la alimentación se podría pensar que ésta debería tener un efecto positivo sobre el desempeño escolar, pero la evidencia sugiere que no necesariamente es así, pues aunque los alumnos beneficiarios de la alimentación están en edad escolar, una diversa literatura sugiere que la buena alimentación es muy importante en los primeros años de la infancia (de 0 a 5 años), de modo que la desnutrición en edad temprana es determinante en cuanto a estatura y capacidades académicas en el futuro.

## Bibliografía

- ALVIAR, Mauricio y RESTREPO, Piedad P. (2006). “La Educación en Medellín: visión de futuro”, *Documento ECSIM*, Medellín.
- BARRIENTOS, Jorge y RÍOS, Paul (2007). “La Evaluación de la Gestión Privada del Servicio Público Educativo en Medellín”, *Lecturas de Economía*, No. 66, pp. 147-172.
- BARRIENTOS, Jorge; TOBÓN, David; RESTREPO, Piedad y ZAPATA, Miguel (2007). “La Evaluación del Programa de Colegios en Concesión en Medellín 2002-2006”, Grupo de Microeconomía Aplicada. Informe Final. Departamento de Economía, Universidad de Antioquia
- CASTAÑO, Elkin (1998). “El Efecto Colegio Sobre la Variabilidad del Rendimiento en Matemáticas”, *Lecturas de Economía*, No. 49, pp. 47-58.
- EIDE, Eric y SHOWALTER, Mark H. (1998). “The Effect of School Quality on Student Performance: A Quantile Regression Approach”, *Economics Letters*, Vol. 58, pp. 345-350.
- GAVIRIA, Alejandro y BARRIENTOS, Jorge (2001a). “Calidad de la Educación y Rendimiento Académico en Bogotá”, *Coyuntura Social*, Fedesarrollo, No. 24, pp. 111-126.
- GAVIRIA, Alejandro y BARRIENTOS, Jorge (2001b). “Características del Plantel y Calidad de la Educación en Bogotá”, *Coyuntura Social*, Fedesarrollo, No. 25, pp. 81-98.
- KOENKER, Roger y BASSET, Gilbert (1978). “Regression Quantiles”, *Econometrica*, Vol. 46, pp. 3-50.
- KOENKER, Robert y BASSET, Gilbert (1982). “Robust Test for Heteroscedasticity Based on Regression Quantiles”, *Econometrica*, No. 50, pp. 43-61.
- KOENKER, Robert y HALLOCK, Kevin F. (2001). “Quantile Regression”, *Journal of Economic Perspectives*, Vol. 15, No. 4, pp. 143-156.
- RESTREPO, Piedad P.; ALVIAR, Mauricio (2005). “El logro académico y el efecto colegio en las pruebas Icfes en Antioquia”, *Lecturas de Economía*, No. 60, pp. 69-95.