

## Determinantes del rendimiento educativo de los estudiantes de secundaria en Cali: un análisis multinivel

John Jairo Correa<sup>1</sup>

### Resumen

Se analiza el impacto que sobre el logro educativo de los estudiantes de Cali de grado once tienen, de un lado, variables correspondientes a los alumnos, y al plantel, y a la interacción entre ellas, del otro. Los datos corresponden a 16,998 alumnos de 427 colegios. Para el análisis se utilizaron modelos lineales de niveles múltiples con dos niveles (alumno y colegio). Se encontró que existen diferencias significativas entre el rendimiento de los alumnos y estas se explican principalmente por factores que operan a través de diferencias entre escuelas y entre alumnos.

### Abstract

This paper analyses the effect of several variables on the school performance of students of Cali in the eleventh grade (the higher school level). The variables analyzed are the following: personal characteristics of the students, school variables, and their interaction. The data were collected from 427 schools of Cali for a sample of 16,998 students. Multilevel lineal modeling with two levels was used for the analysis. It was found that there are significant differences between the school performance of the students, and these are explained mainly by factors that operate through differences between schools and students

**Palabras Claves:** Logro escolar, Educación secundaria, Análisis multinivel, Contexto socioeconómico, Antecedentes familiares, Segmentación educativa, ICFES, Cali.

---

<sup>1</sup> Egresado del programa de Economía de la Universidad del Valle [[j2correa@hotmail.com](mailto:j2correa@hotmail.com)]. Este artículo es un resumen del trabajo de grado realizado junto con Germán Darío Martínez. Se agradece al profesor Boris Salazar, Jefe de Departamento de Economía y director del trabajo, sus ideas y valiosos comentarios. Igualmente, a Luis Piñeros y Libardo Barrera su apoyo técnico. Sin embargo, quien firma el artículo es el único responsable de lo expuesto en él.

Después de varias décadas dedicadas al problema de la cobertura, las autoridades educativas colombianas han comenzado a mostrar, en los últimos años, un mayor interés por los problemas de la calidad y equidad de la educación<sup>2</sup>, convirtiéndolas en objetivos de la política y en instrumentos estratégicos de su desarrollo. Reconocido el vínculo entre el mejoramiento de la calidad y el conocimiento de las circunstancias en que se desarrolla el aprendizaje, se hace indispensable investigar el por qué de las variaciones en el desempeño de los estudiantes de secundaria.

Según el Ministerio de Educación Nacional, en Colombia se presenta una grave crisis en la educación básica secundaria. Los estudios de esta entidad concluyen que una cuarta parte de los estudiantes de los colegios públicos no logran superar el nivel mínimo de las pruebas en lenguaje y menos del 20% alcanzan el nivel óptimo en séptimo y noveno grado. A nivel agregado se identificó que el 81% de los Departamentos tienen bajo desempeño en el examen del ICFES y sólo algunos colegios privados alcanzan niveles superiores (Ministerio de Educación Nacional, 2001). Los resultados muestran la existencia de disparidades sociales en los logros educacionales, lo que alimenta la polémica sobre calidad y equidad.

Desde esta perspectiva, ¿cuál es el objetivo fundamental del sistema educativo? El reto es lograr, al mismo tiempo, calidad y equidad en el servicio educativo. Una educación es de calidad si desarrolla en el individuo la capacidad de responder de manera efectiva a las situaciones que lo rodean, así como a ambientes culturales y tecnológicos en continua transformación. En este contexto, el desempeño representa una forma de aproximación al ideal de calidad en la educación. Es decir, la calidad puede ser medida como el rendimiento observado de los estudiantes. La equidad puede entenderse como la compensación de cada uno de los colegios con respecto al nivel socio-económico de los estudiantes (Bryk y Raudenbush, 1992; Brandsman y Knuver, 1989). Ahora bien, la responsabilidad de toda sociedad justa es asegurar que sus miembros con niveles similares de talento y capacidad, e independientemente de cualquier criterio de discriminación, tengan las mismas posibilidades de acceso al sistema educativo (Rawls, 1979). El Estado debe garantizar a todos los individuos el acceso a una educación de calidad.

Una forma de abordar el análisis de la distribución equitativa de la calidad educativa es asociar las características familiares, personales y del colegio al que asiste el alumno, por un lado, y su rendimiento escolar, por el otro. Esto implica, al mismo tiempo, aproximarse a la determinación de la importancia relativa de los factores escolares y extra-escolares del aprendizaje. Si la distribución de los

---

<sup>2</sup> El acceso a una educación secundaria de calidad puede ser un buen mecanismo de predicción de las posibilidades futuras de las personas (Bossier, Knight, y Sabot, 1985; Bishop, 1989, 1992). La calidad de la educación secundaria afecta las probabilidades de ingreso a la universidad; la formación académica incide también en el desempeño de las personas en el mercado laboral a través de su efecto sobre algunas habilidades básicas: la expresión oral y escrita y la capacidad de formular y resolver problemas complejos (Lockheed, 1991: 5). En síntesis, la calidad de la educación secundaria es un factor de suma importancia en el logro socio-económico de las personas y un factor determinante del acceso a oportunidades y a la movilidad social.

aprendizajes no se ajusta estrictamente a la de los recursos socio-económicos y familiares de los alumnos, ¿es posible que las características del colegio no tengan efectos importantes sobre la distribución de los resultados o, por el contrario, éste cumple un papel efectivamente redistributivo?

De ello se deriva el siguiente interrogante: ¿quiénes pueden recibir una educación de calidad? Si sólo los estudiantes que tienen una mejor situación socio-económica pueden acceder a una educación secundaria de buena calidad, tendremos que aceptar el vínculo entre la calidad de la educación y el nivel socio-económico y las implicaciones que tiene ésta conexión sobre la distribución equitativa de oportunidades y la movilidad social en la región. Es decir, si se reconoce la existencia de ventajas y de oportunidades (ciertas cualidades personales adquiridas anteriormente o fuera de la escuela) que afectan al aprendizaje escolar, se violaría el principio de la igualdad de posibilidades. En términos normativos, la distribución de las probabilidades en los resultados educativos deben ser independientes del nivel socio-económico, de forma que sólo la capacidad, el talento, la aspiración y el esfuerzo deberían generar diferencias en el logro de los estudiantes.

Por otro lado, si son los planteles educativos los que garantizan a los estudiantes recibir una educación de calidad (Card y Krueger, 1996a, 1996b; Altonji y Dunn, 1996), aquellos planteles con mejores equipos, calidad en los maestros y directivos e infraestructura –es decir, los colegios mejor dotados– ofrecerían una mejor educación, lo que llevaría a tener que considerar políticas que propendan por la igualación en la calidad educativa.

El objetivo de este trabajo es conocer el peso relativo que tienen los factores escolares e individuales en la determinación del rendimiento medido por la prueba del ICFES. Es decir, se trata de estimar la variación del rendimiento entre colegios y estudiantes de educación media en Cali. De esta forma, se espera determinar qué proporción de las desigualdades en el rendimiento es explicada principalmente por factores que operan a través de diferencias entre escuelas y entre alumnos. Si, por ejemplo, se establece que las características familiares son las más importantes, ello implicará que todo esfuerzo orientado a igualar las oportunidades tendrá que ir más allá de igualar el acceso a cierta calidad de educación. Si se establece lo contrario, el acceso deberá convertirse en el principal objetivo de cualquier esfuerzo encaminado a igualar las oportunidades.

Ahora bien, en Colombia son escasos los estudios sobre los determinantes del desempeño escolar que utilicen técnicas de análisis estadístico distinto a la de Mínimos Cuadrados Ordinarios. Pese a que el Ministerio de Educación ha realizado estimaciones sobre la base de los resultados de las pruebas del ICFES, en ellas no se han tenido en cuenta los avances metodológicos experimentados a lo largo de esta década: el perfeccionamiento del *software* especializado permite una re-estimación de aquellos resultados. Es por eso que aquí se propone una técnica diferente: el Análisis Multinivel, utilizando el Software MLwiN (Rasbash, *et al.*, 2000). Para nuestro análisis se utilizan modelos lineales de niveles múltiples con

dos niveles (alumno y escuela): la variación del rendimiento de cada estudiante en torno al rendimiento promedio de su colegio y la variación del rendimiento promedio de cada colegio en torno al rendimiento promedio de todos los colegios de su misma ciudad.

Frente a la técnica de Mínimos Cuadrados Ordinarios, el análisis multinivel ofrece varias ventajas. El principal punto es que en estas poblaciones las observaciones individuales no son, en general, del todo independientes. En efecto, los alumnos de un mismo colegio tienden a parecerse entre sí debido a un proceso de selección (homogenización de los estudiantes por clases sociales) y una historia común que los alumnos comparten por el hecho de concurrir al mismo colegio. De esta forma la correlación promedio (conocida como correlación intra-clase) entre las variables de los alumnos del mismo colegio será mayor que la correlación de las mismas variables medidas entre los alumnos de colegios distintos. Las pruebas estadísticas descansan sobre el supuesto de la independencia de las observaciones, y como este supuesto es violado en esta clase de estructuras poblacionales, los errores estándar estimados de las pruebas estadísticas tradicionales serán bastante reducidos, y esto conducirá a que la mayoría de los resultados sean significativamente espurios (Hox, 2002).

Este texto se divide en tres secciones. En la primera se describen los estudios sobre el rendimiento educativo de los estudiantes, subrayando sus hipótesis y conclusiones. En la segunda parte se estiman una serie de modelos de niveles múltiples, incluyendo las características relacionadas con el estudiante y con los colegios, captando el efecto de cada variable sobre el rendimiento educativo de forma individual, y actuando en su conjunto. En la última sección se encuentran las conclusiones.

## Revisión bibliográfica

En las últimas cuatro décadas de investigación sobre la educación, se ha descubierto una asociación entre las características del alumno y su logro escolar. La investigación de Coleman *et al.* (1966), una referencia importante al respecto, trataba de encontrar el efecto específico que los insumos escolares tenían sobre la distribución del aprendizaje, previo control de las características socio-económicas de la familia del alumno. La investigación concluyó que estos insumos tenían, aparentemente, poco efecto sobre las diferencias en el desempeño escolar; es decir, que las diferencias obedecían, casi exclusivamente, a la distinción en el origen social del alumno, mientras que las variables escolares, al menos las consideradas en el estudio, producían un efecto prácticamente nulo.

Existen numerosas críticas acerca de la metodología y la técnica usada en este estudio, tales como la simplicidad del modelo insumo-producto empleado; la ausencia de mediciones sobre los aspectos del proceso de enseñanza-aprendizaje, denominado “caja negra”; la forma de llevar a cabo el análisis de regresión (*step by step*) ya

que, existiendo colinealidad entre las variables de la escuela y las familiares, cuando éstas entran primero, no dejan varianza para explicar a los factores propiamente escolares (Bowles y Levin, 1968); el carácter aditivo del modelo de insumo-producto y la exclusión de cualquier probable interacción entre ellos. Finalmente, el estudio tiene una limitación metodológica importante dada la forma en que se midieron los insumos escolares (por ejemplo, no era posible relacionar a cada grupo de estudiantes con su respectivo maestro).

Sin embargo, este estudio alentó a otros investigadores, quienes intentaron replicar o contradecir sus hallazgos empleando, para ello, otras metodologías y técnicas de medición y análisis. Poco después, Jencks (1972) presentó un trabajo que reafirmaba los hallazgos de Coleman *et al.*: lo más importante en la explicación del rendimiento escolar serían las características del estudiante.

A fines de los años ochenta, investigaciones que usaron la técnica del meta-análisis se propusieron probar la validez de algunos modelos de determinación del rendimiento (Walberg, 1986; Fraser *et al.*, 1987). Resumidamente, concluyeron que: *a*) las relaciones entre rendimiento y ciertos factores son bastante persistentes; *b*) no hay un factor que haga desaparecer totalmente los efectos de los otros; y, por tanto, *c*) el rendimiento es explicable a través de ciertos factores que interactúan entre sí y participan con una pequeña pero importante parte en la comprensión de la variancia total (Fraser, 1989).

Alexander y Simmons (1975) llegaron a las mismas conclusiones que Coleman y Jencks. Sin embargo, basados en la revisión y crítica de los estudios realizados con anterioridad y en sus propias investigaciones, Heyneman (1976: 200-211; 1980: 403-406) y sus colaboradores (Heyneman y Loxley, 1983: 13-21) concluyeron que los errores de medición, los procedimientos metodológicamente cuestionables y la incorrecta especificación de los modelos de rendimiento utilizados explican la ambigüedad en los resultados obtenidos por las investigaciones precedentes. La hipótesis básica de Heyneman era que las sociedades menos desarrolladas son más homogéneas que las industrializadas y, por tanto, que los antecedentes sociales de los alumnos ejercen menos influencia sobre el rendimiento.

Edmonds (1979), por su parte, se muestra interesado no sólo en describir procesos y asociaciones entre el logro y la calidad de las escuelas, sino en su creación en zonas urbanas pobres; eso hace que este tipo de estudios tuviera un sesgo hacia la equidad. Otra revisión contemporánea la constituyen los trabajos de Schiefelbein y Simmons (1980), en los cuales informaban que, de acuerdo con la mayoría de las investigaciones consideradas, en los países del Tercer Mundo la clase social ayuda significativamente a predecir el logro del alumno y que las características familiares son más relevantes cuando se trata del aprendizaje en “lenguaje” comparado con “ciencias”. En esta misma perspectiva, Lockheed, Fuller y Nyirondo (1989: 239-256) advierten que la verificación de un supuesto predominio de los factores escolares en el Tercer Mundo se debe, exclusivamente, a problemas conceptuales y metodológicos, tales como incluir sólo el aspecto material en la

medición de clase social, dejando por fuera aspectos culturales, o el hecho de usar mediciones que sólo son apropiadas o confiables en el primer mundo (educación, ocupación). Si se tienen en cuenta estos aspectos, el efecto del origen social sobre el rendimiento es mayor.

A la hora de explicar el desempeño de los estudiantes, Scheerens (1992) tiene en cuenta tanto las variables de insumos físicos como de procesos, al tiempo que incorpora técnicas estadísticas avanzadas. La autora hace un análisis del estado de la investigación sobre el tema y de la posibilidad de su unión con estudios de mejoramiento escolar en los países en desarrollo. El análisis de ocho estudios latinoamericanos sobre factores de eficacia escolar en primaria (reseñados por Vélez, Schielfelbein y Valenzuela, 1993), estimando funciones de producción educacionales mediante el empleo de técnicas de análisis tradicionales (mínimos cuadrados ordinarios), muestra que el porcentaje del logro explicado por la escuela oscila entre el 49% y el 60%.

Fuller y Clarke (1994) resumen los resultados de las investigaciones realizadas durante la primera mitad de la década de los ochenta con relación a las asociaciones estadísticamente significativas entre las variables de insumos y las variables de procesos; ellos encuentran un mayor número de estudios sobre educación primaria, y recomiendan la incorporación de aspectos relacionados con los maestros.

Haveman y Wolfe (1995), por su parte, examinan el rendimiento y los logros de los hijos en su vida estudiantil, incluyendo desde muestras de niños en escuelas primarias hasta estudiantes de la educación superior. Ellos concluyen que se encuentran numerosos resultados, aunque no siempre son consistentes. En la misma línea, Henevel y Craig (1996) presentan un listado de los factores que tienen mayor incidencia en el rendimiento académico con base en estudios realizados en el África Sub-sahariana y advierten que tales modelos pretenden ser generales, en el sentido de que podrían aplicarse tanto a los países desarrollados como a los subdesarrollados.

Hanushek (1995) muestra la escasa consistencia de los estudios sobre logro educativo con relación a ciertos insumos físicos de la escuela: en la relación alumno-docente, encuentra que en el 72% de los casos su efecto sobre el desempeño de los estudiantes ha sido nulo, en el 28% restante la mitad arroja efectos positivos, mientras que en la otra mitad los efectos son negativos, lo cual hace difícil entender su verdadero efecto. Scheerens (1997) incluye un factor adicional: la infraestructura física de las escuelas. La incorporación de los estudios de eficacia escolar realizados en los países en desarrollo lleva al mismo autor, en 1999, con base en los meta-análisis realizados por Hanushek (1995) y Fuller y Clark (1994) a las siguientes conclusiones: por un lado, que persiste la concentración de la atención de los analistas en aspectos relacionados con la provisión de insumos físicos, pero muy poco con factores organizacionales; por otro, que es escasa la utilización de técnicas de análisis multinivel como herramienta analítica.

En Chile, Mizala y Romaguera (1998) encuentran que los establecimientos privados obtienen 19 puntos más en la prueba SIMCE (prueba estandarizada de

rendimiento) que los públicos. Reynolds *et al.* (2000) afirma que no se incluyen las medidas adecuadas de las variables de la escuela y su proceso: esta exclusión contribuye a subestimar la magnitud de los efectos de la escuela sobre el rendimiento educativo. Willms y Somers (2001) introducen el análisis de los efectos asociados con las variables de la escuela, una vez controlado el efecto de las variables del contexto familiar de los estudiantes, y concluyen que las mejores escuelas son aquellas que poseen altos niveles de recursos escolares.

### **Antecedentes en Colombia**

En 1975 se creó el Programa Nacional de Mejoramiento Cualitativo de la Educación y, ya en los años ochenta, el Instituto SER realizó los primeros intentos de evaluación del rendimiento educativo, estableciendo las diferencias entre la escuela nueva y la escuela rural tradicional. En 1991 se inició el Sistema Nacional de Evaluación de la Calidad de la Educación (SABER) con la aplicación de pruebas de logro en las áreas de matemáticas, ciencias y lenguaje para los grados 3°, 5°, 7° y 9°. Estas encuestas permitieron recolectar información sobre factores presumiblemente asociados con el desempeño en las pruebas, procesar la información y elaborar análisis y, por consiguiente, aclarar algunas situaciones del sistema educativo nacional.

Entre 1994 y 1995, Colombia participó en el tercer estudio comparativo sobre matemáticas y ciencias (TIMSS), concluyéndose que: 1) en Colombia existe gran diferencia entre el currículo propuesto, el currículo desarrollado en el aula y el logrado por los alumnos; 2) nuestro currículo estaba a la par con los currículos de los países desarrollados; 3) el rendimiento de los estudiantes colombianos estaba muy por debajo del promedio internacional; y, 4) a nivel nacional, el puntaje en ciencias era mayor al alcanzado en matemáticas y los estudiantes con mayor rendimiento estaban localizados en la región nororiental y en el sector privado.

La literatura colombiana sobre la calidad educativa es relativamente reciente. Aldana *et al.* (1992) determinaron y analizaron a nivel estadístico los niveles de calidad alcanzados por los estudiantes de 3° y 5° de primaria, para las áreas de lenguaje y matemáticas, y encontraron que existen diferencias en el rendimiento a nivel territorial y cultural. La Misión Social (DNP) (1992) analizó el problema desde la perspectiva de los docentes, en lo que se refiere a su cantidad, a algunas de sus características, a su motivación salarial y a su situación laboral; se encontró que la forma como los Municipios y Departamentos han cubierto las necesidades de la población en materia educativa van en detrimento de la calidad de la educación y de las remuneraciones de los docentes.

El Ministerio de Educación Nacional y SABER (1995) evaluaron el efecto de las características de los alumnos, del hogar, del aprendizaje, de los planteles y de los docentes sobre el rendimiento académico de los estudiantes de 3° y 5°. Sarmiento y Caro (1997), a partir de datos de primaria y secundaria entre 1973 y 1994, realizaron

un análisis estadístico sobre los cambios en cobertura, escolaridad, eficiencia y equidad del sistema educativo, encontrando mejoras en eficiencia y eficacia, pero muy lentas, ineficientes e inequitativas. La Misión Social (DNP) (1997) con base en las encuestas y los resultados de las pruebas SABER de 1993 para el calendario A (3° y 5° de primaria) analizaron el impacto de los establecimientos escolares y concluyeron que los factores asociados al plantel educativo pueden afectar en un 29% el logro escolar (para hallazgos similares sobre el rendimiento en matemáticas de alumnos de 3° de primaria, cfr. Castaño, 1998; para estudiantes de 11°, cfr. Moreno y Piñeros, 1998). A partir de las mismas pruebas, Sarmiento, *et al.* (2000) presentaron un informe sobre la incidencia del plantel en el logro educativo del alumno y su relación con el nivel socio-económico: se encontró que el logro promedio de los planteles públicos es superior al de los privados, puesto que, al filtrar el nivel socio-económico, los primeros tienen un porcentaje mayor de estudiantes pobres que los segundos.

Caro (2000) estudió cuáles eran los factores escolares y extra-escolares que afectaban el desempeño académico de los estudiantes de 3° y 5° elemental de Bogotá D.C. Observó que el sector público tiene mayor homogeneidad, menor soporte administrativo y menor autonomía de los rectores sobre sus establecimientos educativos que su contraparte, el sector privado. Casas *et al.* (2002) investigaron cómo el cambio en la contextualización del examen del ICFES, a partir del año 2000, afectó al papel de la escuela en la predicción del logro educativo. Basándose en las pruebas ICFES de 1997, 1999 y 2000, encontraron que la importancia bruta de la escuela, que entre 1997 y 1999 se situaba en niveles similares (entre el 27% y el 37%), se reduce considerablemente con el nuevo examen, ubicándose entre el 10% y el 27% para el 2000. Además, la importancia neta del colegio (que oscilaba entre el 13% y el 20% en 1999) pasó a estar entre el 6% y el 11% en el 2000.

El DNP (2002) resalta la eficiencia, la calidad y la equidad como los objetivos de mayor importancia para el lineamiento de la educación en Colombia. Nuñez *et al.* (2002) determinaron, mediante simulación, las diferencias en las pruebas del ICFES entre los alumnos de colegios públicos y privados. Después de controlar las características de los individuos, del hogar, de la infraestructura del colegio y del nivel educativo de los docentes, tanto a nivel nacional como a nivel urbano y rural, encontraron una diferencia significativa a favor de los colegios privados. Gaviria y Barrientos (2001) encontraron que el rendimiento es mejor en los colegios privados y que, en las instituciones públicas, los aumentos en la educación de los docentes y las mejoras de la infraestructura física no están asociados con mayores niveles de rendimiento de los estudiantes.

## El modelo multinivel

En nuestro sistema educativo, los estudiantes se encuentran agrupados en salones de clases, y estos, a su vez, forman parte de un colegio. Este agrupamiento hace

que los alumnos de un curso tengan características o experiencias educativas diferentes de los estudiantes pertenecientes a otros cursos; las diferencias pueden ser aún mayores entre los colegios. Por ejemplo, las condiciones en un colegio privado son diferentes a las de uno oficial. Este agrupamiento de las variables hace que las técnicas de análisis estadístico usadas tradicionalmente sean de utilidad limitada: los análisis de mínimos cuadrados ordinarios (OLS) suponen que cada observación tiene el mismo peso.

Para medir el logro académico de los estudiantes se deben tener en cuenta los factores de la familia (socioeconómicos) y los del colegio (jornada, naturaleza, etc.), buscando medir la proporción de la variación en el resultado debida al alumno y la proporción debida al colegio. Para ello, los modelos convencionales de regresión lineal no parecen ser los más apropiados. Han sido consideradas varias opciones con el objetivo de que, utilizando las técnicas de mínimos cuadrados ordinarios, los estimados resultantes sean lo menos sesgados posible.

En primer lugar, a cada estudiante se le pueden asignar las condiciones del colegio. Acá se estaría incurriendo en un problema de dependencia debido a que todos los estudiantes de un mismo colegio compartirían la información asociada a éste, con lo cual se estaría violando el supuesto de independencia entre las observaciones, subestimándose el error estándar y omitiendo la naturaleza jerárquica del fenómeno.

Si, por otro lado, se resumen las características de los estudiantes a nivel de cada colegio, se estaría omitiendo el análisis dentro del plantel, perdiendo con ello gran parte de la explicación de la variabilidad en el rendimiento y obteniendo frecuentemente mayores asociaciones entre las variables debido a la reducción de la varianza, resultante de la agregación. Así, no sería posible allegar resultados individuales a partir de resultados obtenidos de los agregados.

Una tercera solución es realizar un análisis por separado de cada colegio, con lo cual se suprimiría el análisis de la variabilidad intergrupala (entre planteles), descontando el análisis de las causalidades más importantes debido al desconocimiento del efecto del agrupamiento y subestimando el efecto de los factores familiares asociados al estudiante.

En resumen, cuando las características de los alumnos y la provisión de insumos y técnicas de enseñanza son similares en el interior de los colegios, pero diferentes entre éstos, la utilización de las técnicas tradicionales de estimación se traduce en distorsiones del término de error, del error estándar y, por consiguiente, de los niveles de significancia de los estimados. Sin embargo, aplicando las técnicas multinivel se puede establecer qué parte del logro es explicado por la escuela y qué parte por el estudiante.

La estimación de un modelo jerárquico exige un análisis sistemático partiendo del más simple de los modelos (el Modelo Vacío, que no incluye predictores y que solamente estima la media global del logro) y se va perfeccionando mediante la adición progresiva de variables explicativas, tanto de alumnos como de planteles.

Dependiendo del efecto que la adición de las variables tenga sobre la reducción de la varianza total, éstas serán retenidas o eliminadas del modelo. A continuación se presentan estos pasos sistemáticos.

### **Técnica de análisis**

El *análisis por niveles múltiples* (o *modelos jerárquicos lineales*) se aplica a datos con una “estructura de agrupamiento jerárquico” (Aitkin y Longford, 1986; Bryk y Raudenbush, 1992; Goldstein, 1987), como es el caso de los datos del sector educativo (los alumnos forman parte de un aula, que a su vez es parte de un colegio, situada en un municipio, dentro de un departamento, etc.). Esta técnica correlacional es apropiada para analizar variaciones en el rendimiento de los estudiantes que son miembros de un colegio, permitiendo la descomposición de una variable (*rendimiento*) en sus componentes dentro del grupo (intra-escuela) y entre grupo (inter-escuela) y el análisis de la asociación entre variables en esos diferentes niveles de agregación (características del estudiante, el aula o el colegio). Para la estimación se utilizó el método de Mínimos Cuadrados Generalizados Iterativos (IGLS); básicamente es un refinamiento secuencial del procedimiento basado en Mínimos Cuadrados Generalizados (GLS). El método es aplicable a todos los modelos multinivel normales. Este procedimiento produce estimadores sesgados de los parámetros aleatorios debido a que no tiene en cuenta la varianza muestral de la parte fija del modelo.

### **Método de estimación de un modelo multinivel de dos niveles<sup>3</sup>**

El modelo está compuesto por dos sub-modelos, uno en el nivel 1 y otro en el nivel 2. Como el problema de investigación es sobre datos de estudiantes anidados dentro de escuelas, el modelo del nivel 1 representa las relaciones de las variables a nivel del estudiante y el modelo del nivel 2 captura la influencia de los factores escolares. Formalmente, hay  $i = 1, \dots, n_j$  unidades del nivel 1 (estudiantes) dentro  $j = 1, \dots, J$  unidades del nivel 2 (escuelas).

#### *Modelo del nivel 1:*

En este modelo se representa la variable dependiente para el caso (alumno)  $i$  dentro de la unidad  $j$  como:

$$\begin{aligned}
 Y_{ij} &= \beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + \dots + \beta_{Qj} X_{Qij} + e_{ij} \\
 &= \beta_{0j} + \sum_{q=1}^Q \beta_{qj} X_{qij} + e_{ij}
 \end{aligned}$$

<sup>3</sup> Basado en Bryk, Raudenbush y Congdon (1996).

donde:

$\beta_{qj}$  ( $q = 0, 1, \dots, Q$ ) son los coeficientes del nivel 1;

$X_{qij}$  es el predictor  $q$  del nivel 1 para el caso  $i$  en la unidad  $j$ ;

$e_{ij}$  es el efecto aleatorio del nivel 1;

$\sigma^2$  es la varianza de  $e_{ij}$ , la varianza del nivel 1.

El término aleatorio se distribuye en forma normal:

$$r_{ij} \sim N(0, \sigma^2)$$

*Modelo del nivel 2:*

Cada uno de los coeficientes definidos en el modelo del nivel 1, se convierten en variables dependientes del modelo del nivel 2,

$$\begin{aligned} \beta_{qj} &= \alpha_{q0} + \alpha_{q1}W_{1j} + \alpha_{q2}W_{2j} + \dots + \alpha_{qS_q}W_{S_qj} + u_{qj} \\ &= \alpha_{q0} + \sum_{s=1}^{S_q} \alpha_{qs}W_{sj} + u_{qj} \end{aligned}$$

donde:

$\alpha_{qs}$  ( $q = 0, 1, \dots, S_q$ ) son los coeficientes del nivel 2;

$W_{sj}$  es el predictor del nivel 2; y

$u_{qj}$  es el efecto aleatorio del nivel 2.

Se supone que, para cada unidad  $j$  del nivel 2, el vector  $u_{qj}$  se distribuye como normal multivariante, y cada elemento de  $u_{qj}$  tiene una media cero y una varianza:

$$Var(u_{qj}) = \tau_{qq}$$

para cada par de efectos aleatorios  $q$  y  $q'$  se tiene

$$cov(u_{qj}, u_{q'j}) = \tau_{qq'}$$

Los componentes de la varianza y covarianza se agrupan en una matriz de dispersión,  $T$ , cuya dimensión es  $(Q+1) \times (Q+1)$ .

Los coeficientes del nivel 1 se pueden modelar en el nivel 2 de tres formas distintas:

1. Coeficientes del nivel  $i$  fijo,

$$\beta_{qj} = \alpha_{q0}$$

2. Coeficiente del nivel 1 con variación no aleatoria en las unidades del nivel 2,

$$\beta_{qj} = \alpha_{q0} + \sum_{s=1}^S \alpha_{qs} W_{sj}$$

3. Coeficiente del nivel 1 con variación aleatoria en las unidades del nivel 2,

$$\beta_{qj} = \alpha_{q0} + u_{qj}$$

O con variables del nivel 2,

$$\beta_{qj} = \alpha_{q0} + \sum_{s=1}^S \alpha_{qs} W_{sj} + u_{qj}$$

La dimensión de T depende del número de coeficientes del nivel 1 especificados como aleatorios.

### Modelo vacío

La estimación del modelo vacío, o de efectos fijos, constituye el punto de partida en todo análisis de tipo multinivel. Los únicos parámetros que son estimados son las constantes o interceptos de los dos niveles y su utilidad consiste en calcular el logro promedio de los alumnos y de las escuelas. El modelo está dado por las ecuaciones siguientes:

**Tabla 1. Características de los modelos vacíos**

Definición	Supuestos	Varianza
$Y_{ij} = \alpha_{0j} + e_{ij}$	$U \sim N(0, \sigma_u^2)$	$\text{var}(Y_{ij}) = \sigma_u^2 + \sigma_e^2$
$Y_{ij} = \alpha_{0j} + U_{ij}$		
$Y_{ij} = \alpha_{0j} + U_{0j} + e_{ij}$	$e \sim N(0, \sigma_e^2)$	$CCI = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$

El modelo vacío no contiene ningún tipo de regresor o variable explicativa ni del estudiante ni de los colegios, por consiguiente, el coeficiente  $\beta_{0j}$  equivale al intercepto o media, en nuestro caso, del desempeño de los estudiantes, la cual se aproxima a la media aritmética de la misma. Los términos de error de la escuela ( $e$ ) y del estudiante ( $U$ ) siguen una distribución normal con media cero y varianzas iguales  $\sigma_e^2$  y  $\sigma_u^2$  respectivamente. La varianza total es la suma de  $U$  y  $e$ , y la importancia del colegio corresponde a la proporción de la varianza total atribuible a ese nivel, aunque sin ningún control respecto a las variables de contexto.

En este modelo, el coeficiente de correlación representa el tamaño relativo a la varianza entre colegios. El coeficiente varía de 0 a 1: cuando este presenta un valor de cero, se dirá que los alumnos están homogéneamente distribuidos entre escuelas y que el desempeño del alumno es independiente del colegio que frecuenta y que toda la varianza en el logro escolar será debida a la variabilidad entre alumnos; la responsabilidad del colegio en el resultado de la prueba ICFES será nula. Por

otra parte, cuando el coeficiente de correlación intra-escuela toma el valor de uno, toda la variabilidad en el desempeño de los alumnos se debe exclusivamente a la diferencia entre las escuelas; en este caso las características individuales del alumno no afecta en nada su logro académico.

### ***Población estudiada***

Las variables que hacen referencia a las características individuales, familiares y del plantel fueron extraídas del formulario de inscripción del ICFES aplicado en marzo del año 2001. Para efectos del estudio se realizó la estandarización de algunas variables del individuo. El archivo se conforma con 16998 estudiantes de 427 colegios de Cali.

### ***Definición de variables***

La variable dependiente, o criterio, es el puntaje obtenido por el alumno en la prueba del ICFES, previamente estandarizado<sup>4</sup>. Las variables independientes son las características del alumno y del colegio, y se pueden organizar así:

- a) Variables individuales del alumno que se refieren al capital económico familiar, capital cultural familiar, además de otras variables que lo caracterizan.
  - Sexo: variable *dummy* con código cero para mujeres y uno para hombres, incluida porque la mayoría de los estudios concluyen que los hombres alcanzan mejores puntajes.
  - Z\_años: variable estandarizada que indica la edad de los estudiantes.
  - Z\_mayor\_n: variable estandarizada que corresponde al mayor nivel educativo de los padres; se toma el mayor nivel de educación de uno de los padres.
  - Hermano mayor (herm\_may): indica la posición del individuo en la familia; toma el valor de 1 si es el hermano mayor y el valor de 0 en caso contrario.
  - Z\_nuh: variable estandarizada que refleja el número de hermanos; observamos que por cada hermano adicional el puntaje disminuye en promedio 0.4 puntos.
  - Trabaja (trb): variable *dummy*; 1 si trabaja y 0 si no lo hace.
  - Z\_ingf: variable estandarizada que indica el total del ingreso de la familia del estudiante.
- b) Variables que corresponden al nivel del colegio; todas son categóricas:
  - a. Naturaleza (nat): representa si el colegio es oficial o no público
  - b. Tipo (tip): según el tipo de bachillerato, con tres categorías: académico, técnico, y académico-técnico.

---

<sup>4</sup> Se trabaja con valores estandarizados con el objeto de tornar directamente visibles (en proporciones) los cambios que se producen en la varianza no explicada del rendimiento y el peso relativo de cada una de las variables en la explicación de la variación del rendimiento.

- c. Valor de la pensión (vrp): con 7 categorías de acuerdo al valor en pesos de la pensión de los estudiantes.
- d. Jornada (jor): incluye tres categorías: mañana, tarde y noche.

## Análisis de los resultados

### Modelo 1. Variación intra-escuela e inter-escuela

Se estima la variación alrededor de la media global del rendimiento (parte fija) y, simultáneamente, los correspondientes a cada nivel de agregación (parte aleatoria). Este modelo constituye el punto de partida del análisis porque define una línea de base que sirve como patrón de referencia comparativo para calcular el rendimiento promedio y las dos varianzas (inter-escuela, inter-alumno), se regresan los rendimientos escolares de cada alumno ( $Y_{ij}$ ) sobre una constante (constante) que asume un valor de uno para todos los individuos. Formalmente la operación se expresa de la siguiente manera:

$$z\_ptotal_{individua,codcolegio} \sim N(XB, \Omega)$$

$$z\_ptotal_{individua,codcolegio} = \beta_{0individua,codcoleg} \text{constante}$$

$$\beta_{0individua,codcoleg} = -0,115(0,031) + u_{0codcoleg} + e_{0individua,codcoleg}$$

$$[u_{0codcoleg}] \sim N(0, \Omega_u) : \Omega_u = [0,358(0,028)]$$

$$[e_{0individua,codcoleg}] \sim N(0, \Omega_e) : \Omega_e = [0,642(0,007)]$$

$$-2 * \log\text{likelihood(IGLS)} = 41842,870 \text{ (16998 of 16998 cases in use)}$$

Dónde  $z\_ptotal_{ij}$  es el puntaje del ICFES (rendimiento estandarizado) del alumno  $i$  que asiste al colegio  $j$ ; (Constante) Es una Constante igual a uno (1);  $\hat{\alpha}_{0ij}$  es el rendimiento promedio estimado (parte fija);  $e_{0ij}$  es el “residuo” a nivel alumno (nivel 1), para el  $i$ -ésimo alumno en la  $j$ -ésima escuela;  $u_{0j}$  es el residuo a nivel de escuela (nivel 2), para la  $j$ -ésima escuela. Las últimas dos variables ( $e_{0ij}, u_{0j}$ ) son cantidades aleatorias con una media igual a cero, no correlacionadas y con una distribución normal (parte aleatoria). Por lo tanto, podemos estimar sus varianzas, indicadas como  $\sigma_e^2$ ,  $\sigma_u^2$  respectivamente, quedando así especificados los porcentajes de la variación total del rendimiento que se debe a las diferencias entre las escuelas y los alumnos.

**Tabla 2. Distribución proporcional de la variación del rendimiento según nivel de agregación**

Efecto Fijo	Estimación	Error Estándar
$\beta_{0ij}$ = Intercepto	-0.115	0.031
<b>Efecto Aleatorio</b>		
Varianza de Nivel 2		
$\sigma_u^2 = \text{var}(U_{0j})$	0.358	0.028
Varianza de Nivel 1		
$\sigma_e^2 = \text{var}(e_{ij})$	0.642	0.007
-2*loglikelihood	41842.870	

De los resultados del cuadro se infiere que:

- a. La varianza inter-escuela de los puntajes promedios de las escuelas es significativa. Alrededor del 36% de las variaciones del rendimiento se debe a diferencias entre los establecimientos educativos. Es decir, existen diferencias en la calidad educativa de las instituciones: los colegios se diferencian entre si respecto del nivel de aprendizaje que alcanzan sus alumnos. En otras palabras, más de un tercio de las desigualdades en los niveles de rendimiento en el ICFES son atribuibles a la existencia de una fuerte diferencia institucional en la educación media; esto, por el peso relativo que tienen los factores de los colegios en la explicación de la variación del rendimiento.
- b. La variación entre-alumno (o intra-escuela) constituye más de la mitad del total del logro educativo. La explicación de las desigualdades entre los alumnos en los colegios debe buscarse ahora en sus características individuales más que en factores vinculados con el agrupamiento del sistema educativo. Esto es así porque la variación en el rendimiento educativo es explicado en un 64% por las variables asociadas al individuo, lo que implica que las diferencias en las características del estudiante (habilidades, capacidad de aprender, el numero de hermanos, el sexo, entre otras) son determinantes en los resultados finales de la prueba, aunque no son capaces de explicar en su totalidad la variación en el logro educativo.

En resumen, pertenecer a un determinado colegio implica una mayor (o menor) probabilidad de acceder a un nivel de logro educativo más alto (o más bajo) que el esperado en otros colegios. Sin embargo, se observó que, de los dos niveles de agrupamiento (colegio, individuo) el mayor peso en la variación del logro se sitúa en el individuo. Aparentemente, las características del individuo y de sus familias son mucho más relevantes que las diferencias entre planteles cuando se trata de explicar las desigualdades en el rendimiento de los estudiantes. Lo ideal sería que

las características del plantel tuvieran un impacto pequeño o nulo en el logro educativo, pues esto nos diría que se ha logrado una buena homogeneidad tanto en infraestructura como en calidad en los diferentes planteles que imparten educación media secundaria. Pero, como se dijo anteriormente, las características del individuo no son definitorias, por lo que es necesario observar el impacto de las variables del plantel en el logro educativo.

El coeficiente de correlación intra-escuela es igual a:

$$\rho = \frac{\sigma_u^2}{\sigma_e^2 + \sigma_u^2} = \frac{0.358}{0.358 + 0.642} = 0.36$$

Este coeficiente, que equivale a la proporción de la varianza total explicada por este nivel, por su parte, se puede interpretar como un indicador de la diferencia que representa para un estudiante asistir a un determinado colegio. Debe tenerse en cuenta que se parte para la estimación de la varianza total o bruta, la cual incluye las diferencias de contexto y otras variables de entrada de los estudiantes (nivel socio-económico, valor pensión, entre otras). Para los países de la OCDE (Organización para la Cooperación y Desarrollo Económico) la varianza entre escuelas o Coeficiente de Correlación Intra-Escuela no representa más del 10% o el 15% de la varianza total del rendimiento de los estudiantes. En los países en vías de desarrollo, esas diferencias son mayores, haciendo que el CCI se sitúe entre el 30% y el 40%.

Los resultados obtenidos en el modelo 1, son interesantes pues se puede constatar que el 36% de la varianza total del rendimiento se debe a diferencias entre las escuelas, o sea, que la variabilidad en el desempeño de los alumnos está claramente influenciada por las diferencias entre colegios: la responsabilidad del colegio en el resultado de la prueba del ICFES es bastante significativa. Los factores escolares, en un sentido amplio, tiene capital importancia en la explicación de la variación de los resultados en la prueba ICFES para los estudiantes de Cali en el año 2001.

Para efectos de nuestro trabajo se deben conservar los valores de la prueba de máxima verosimilitud (41842.870) y la variación intra-escuelas (0.36 ó 36%) ya que son valores de referencia para evaluar la fuerza de la relación de cada una de las variables estudiadas con el rendimiento.

## **Modelo 2. Efectos sobre el rendimiento por separado de las variables asociadas al plantel**

Para determinar si las variables asociadas al plantel inciden significativamente en el logro escolar se evalúa el grado de asociación de cada medición individual con el rendimiento. Es decir, se trata de saber cuáles de las variables individuales utilizadas explican una proporción significativa de la variación del rendimiento escolar.

Para ello, se debe incluir individualmente los indicadores en el modelo vacío y analizar las estimaciones que resultan para cada uno de ellos.

Con excepción de las variables naturaleza y valor pensión 3, todos los indicadores tienen eficacia predictiva propia y determinable sobre el logro educativo (ver la Tabla 3); es por eso que todas estas variables deben de ser tenidas en cuenta si se quiere mejorar la estimación de la relación entre rendimiento y factores asociados al plantel. Cada una de las variables por separado logra tener una pequeña relación en el logro educativo del alumno; es decir, el porcentaje de la variación del plantel no sufre grandes cambios a medida que se adiciona una variable.

El conjunto de variables del valor de la pensión han conseguido que la variación inter-escuela no explicada descienda del 35.8% al 25.9% en el modelo vacío; es decir, estas variables definen alrededor del 30% de las diferencias de rendimiento promedio existente entre escuelas. De forma diferente, la variación inter-alumno no presenta cambios significativos cuando se adicionan las variables asociadas al plantel.

Ninguna de las variables del plantel actuando por separado tienen el peso suficiente para disminuir de manera significativa la varianza intra-escuela (Coeficiente de Correlación Intra-escuela CCI); es decir, el porcentaje de la variación del plantel no sufre grandes cambios a medida que se adiciona una variable. La mayor influencia la ejerce la variable jornada que explica el 5% de la varianza entre las escuelas que no es explicada.

**Tabla 3. Parámetros estimados para las diferentes variables asociadas al plantel**

Variables y Niveles	Modelo 0 estimación	e.s	Modelo 1 estimación	e.s	Modelo 2 estimación	e.s	Modelo 3 estimación	e.s	Modelo 4 estimación	e.s
<b>Naturaleza</b>			0,004 *	0,04						
<b>Tipo</b>										
Técnico					-0,388 **	0,068				
Académico y técnico					0,084 **	0,076				
<b>Jornada</b>										
Mañana							0,106 **	0,074		
Tarde							-0,072 **	0,096		
Noche							-0,42 **	0,091		
<b>Valor Pensión</b>										
Pensión 2									-0,05 **	0,028
Pensión 3									0,034 **	0,037
Pensión 4									0,148 **	0,047
Pensión 5									0,23 **	0,049
Pensión 6									0,34 **	0,057
Pensión 7									0,324 **	0,07
<b>Niveles</b>										
Colegio	0,358	0,028	0,357	0,027	0,328	0,025	0,299	0,023	0,259	0,021
Alumno	0,642	0,007	0,642	0,007	0,642	0,642	0,642	0,007	0,645	0,007
CCI	0,36		0,36		0,34		0,31		0,33	
<b>Test de Verosimilitud</b>	41842,87		41842,28		41811,25		41775,16		41797,57	

\*\* p £ 0.05

### **Modelo 3. Efectos sobre el rendimiento en conjunto de las variables asociadas al plantel**

Al determinar si las variables asociadas al plantel en su conjunto tienen alguna incidencia en el logro escolar, se evalúa el grado de asociación de las variables en su conjunto con el rendimiento; por tal motivo, se incluyen todos los indicadores en el modelo vacío.

El efecto conjunto de las variables del plantel consiguen descender la varianza intra-escuela no explicada (coeficiente de correlación intra-escuela) a 23%; es decir, las variables asociadas al plantel explican al menos el 13% de las diferencias en el logro educativo promedio entre las escuelas. Las variables del plantel actuando en su conjunto disminuyen la varianza inter-escuela del 35.8% al 19.3%. En definitiva, el comportamiento de los datos analizados nos deja observar claramente que existe cierta relación entre el logro educativo y las anteriores variables asociadas al plantel, y que en conjunto logran explicar en un alto porcentaje la varianza no explicada en el logro entre las escuelas.

### **Modelo 4. Efectos sobre el rendimiento de las variables asociadas al individuo**

En este modelo se explica la proporción de la varianza total del logro educativo a partir de cada una de las variables asociadas al individuo y posteriormente se evalúa el peso de los factores asociados al individuo en su conjunto.

Al igual que las variables asociadas al plantel, las asociadas al individuo tienen poca relación con el logro académico cuando se estiman por separado, teniendo un escaso impacto sobre la varianza entre escuelas no explicada. El mayor nivel educativo de los padres y el ingreso familiar son las variables que ejercen mayor influencia sobre dicha varianza (ver la Tabla 4).

Puede observarse que las estimaciones de todas las mediciones son significativas cuando se las considera individualmente. La variable mayor nivel educativo, ha conseguido que la variación inter-escuela no explicada descienda del 35.8% al 27.6% en el modelo vacío; es decir, esta variable define casi el 30% de las diferencias de rendimiento promedio existente entre las escuelas. La variación inter-alumno disminuye cuando se introduce la variable ingreso familiar, al pasar del 64.2% al 62.2%.

Ninguna de las variables del individuo actuando por separado tienen peso suficiente para disminuir de manera significativa la varianza intra-escuela (Coeficiente de Correlación Intra-escuela CCI); es decir, el porcentaje de la variación del plantel no sufre grandes cambios a medida que se adiciona una variable. La mayor influencia la ejerce la variable mayor nivel educativo que explica el 6% de la varianza no explicada entre escuelas.

Al igual que en los modelos anteriores, la variación entre alumnos (o intra-escuela) constituye más de la mitad del total del logro educativo. La explicación de

las desigualdades debe buscarse en variables que expliquen de una manera más precisa las características individuales.

**Tabla 4 Parámetros estimados para las diferentes variables asociadas al individuo**

Variables y niveles	Modelo 0 estimación	e.s	Modelo 1 estimación	e.s	Modelo 2 estimación	e.s	Modelo 3 estimación	e.s	Modelo 4 estimación	e.s	Modelo 5 estimación	e.s	Modelo 6 estimación	e.s
Años			-0,102 **	0,008										
Hno Mayor					0,08 **	0,013								
Ing. Familiar							0,091 **	0,009						
Nivel Educ.									0,031 **	0,002				
Núm. Hnos.											-0,035 **	0,006		
Sexo													0,202 **	0,015
<b>Niveles</b>														
Colegio	0,358	0,028	0,331	0,026	0,354	0,027	0,295	0,019	0,276	0,022	0,35	0,027	0,361	0,028
Alumno	0,642	0,007	0,636	0,007	0,64	0,007	0,622	0,007	0,635	0,007	0,641	0,007	0,634	0,007
CCI	0,36		0,34		0,36		0,31		0,3		0,35		0,36	
<b>Test de Verosimilitud</b>	41842,87		41667,61		41802,67		41743,01		41556,25		41814,3		41651,53	

\*\* p £ 0.05

### Modelo 5. Modelo total individuo

En este modelo se explica el efecto de las variables relacionadas con el individuo en su conjunto. El modelo es el siguiente:

$$z\_ptotal_{ij} \sim N(XB, \Omega)$$

$$z\_ptotal_{ij} = \beta_{ij} constante + 0,198(0,014)sexo_{ij} + 0,104(0,008)zmayor\_n_{ij}$$

$$+ -0,03(0,007)z\_muj_{ij} + 0,062(0,009)z\_ing_{ij} + -0,090(0,008)z\_años_{ij} + 0,049(0,014)herm\_may_{ij}$$

$$\beta_{0ij} = -0,199(0,027) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = [0,234(0,019)]$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = [0,622(0,007)]$$

$$-2*\loglikelihood(IGLS)= 41174,850 (16998 of 16998 cases in use)$$

En los resultados obtenidos se puede constatar que el efecto conjunto de las variables relacionadas con el individuo tienen un gran impacto sobre la varianza entre las escuelas no explicada. El coeficiente de correlación intra-escuela es de 27%, logrando de esta manera explicar el 9% de la varianza no explicada en el logro educativo entre las escuelas. Esto, comparado con las variables introducidas en forma individual, consigue una mejor explicación del rendimiento de los alumnos en los diferentes planteles.

El efecto inter-escuela disminuye significativamente cuando se incorporan todas las variables del individuo en su conjunto, pasando de un 35.8% al 23.4%. Por consiguiente, estas variables explican alrededor del 40% de las diferencias de rendimiento promedio existente entre las escuelas

### Modelo 6. Efecto sobre el logro educativo de las variables asociadas al plantel y al individuo en su conjunto

Aquí se pretende evaluar el grado de asociación de cada variable con el rendimiento, considerándolas simultáneamente: los dos grupos de variables (características del individuo y del plantel) son analizados conjuntamente. El resultado del modelo es el siguiente:

$$z\_ptotal_{ij} \sim N(XB, \Omega)$$

$$z\_ptotal_{ij} = \beta_{0ij}constante + 0.198(0.014)sexo_{ij} + 0.104(0.008)zmayor\_m_{ij} + -0.03(0.007)z\_muh_{ij} + 0.062(0.009)z\_ing_{ij} + -0.090(0.008)z\_años_{ij} + 0.049(0.014)herm\_may_{ij}$$

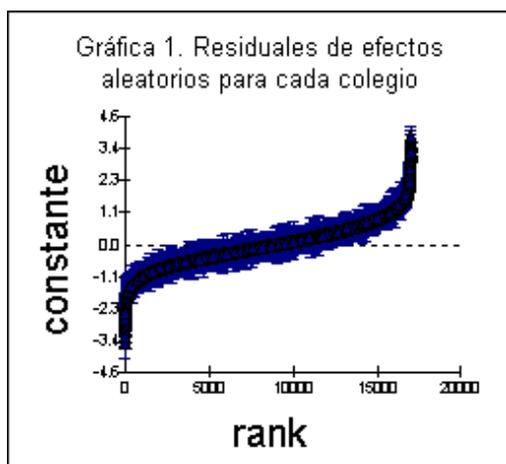
$$\beta_{0ij} = -0.199(0.027) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \\ e_{0ij} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.234(0.019) \\ 0.622(0.007) \end{bmatrix}$$

$$-2*\text{loglikelihood(IGLS)} = 41039,050 \text{ (16998 of 16998 cases in use)}$$

Se puede observar que el efecto conjunto de las variables plantel y del individuo consiguen disminuir la varianza entre las escuelas no explicada al 19%; es decir, logran explicar en su conjunto el 17% de la varianza. Lo que nos lleva a pensar en la importancia que tienen estas variables sobre el logro educativo del estudiante.

El efecto inter-escuela disminuye significativamente cuando se incorporan todas las variables del individuo en su conjunto, pasando de 35.8% a 14.7%. Estas variables explican más del 50% de las diferencias de rendimiento promedio existente entre las escuelas.



Finalmente, la Gráfica 1 muestra los residuos para cada uno de los establecimientos (oficiales y privados) contenidos en la población y sus respectivos intervalos de confianza. Observando estos intervalos se puede concluir que un grupo de establecimientos se sitúan por debajo de la línea de corte (0,0), otros se encuentran sobre dicha línea y otros se encuentran por encima de la misma. Si se tiene en cuenta que los residuos representan la separación entre la media poblacional estimada por los parámetros, podríamos afirmar que un número considerable de colegios se encuentran por debajo de la media, a un nivel de significancia estadística del 5%.

## **Conclusiones**

El análisis de los datos confirma que, en Cali, el efecto de las variables asociadas al plantel sobre el rendimiento educativo es significativo, lo que implica que las políticas educativas tienen que estar encaminadas a mejorar la calidad de los planteles. La asistencia de un alumno a un colegio cualquiera del sistema implica una determinada probabilidad de acceder a una mayor (o menor) calidad en la educación, en comparación con el que obtendría en otra escuela, y con ello, a un más alto (o bajo) rendimiento. Por consiguiente, existen colegios que ofrecen una mayor probabilidad de alcanzar mayor aprendizaje. Pero además, se estableció otra característica de la dinámica de la escuela: la fuerza de la relación entre el logro escolar y las características individuales y familiares varía positiva y significativamente entre los establecimientos; es decir, existen diferencias notables entre los colegios con respecto a su capacidad de compensar desigualdades sociales.

El acceso a la educación superior depende de las condiciones del logro educativo de los niveles anteriores, particularmente de los resultados del ICFES y de los factores que lo condicionan. En este sentido, así se busque universalizar las posibilidades de acceso a la educación superior, si no se logra incidir sobre los factores que determinan los resultados del examen de estado, se continuará con la tendencia de la sobre-selección de estudiantes con mayores capitales culturales y económicos, y por tanto, no se lograrán mayores impactos en términos de equidad en dicho nivel.

Como en algunos otros trabajos (Ridell, 1993; 1997), los resultados obtenidos en el modelo vacío indican que el 36% de la varianza total del rendimiento se debe a diferencias entre las escuelas. La capacidad explicativa por separado de las variables de esta varianza es baja, y características como el valor de la pensión, el sexo o la naturaleza del plantel no ejercen una gran influencia, a diferencia del modelo de regresión lineal, debido a que en el modelo multinivel se tiene en cuenta la estructura jerárquica en que se articulan los datos. Pero, a medida que se incrementa el número de variables en el modelo, estas explican de una mejor manera la varianza no explicada del logro educativo de los estudiantes, permitiendo disminuirla en un 17%. Es decir, en conjunto lograr explicar el 17% de la variación

no explicada del logro educativo entre los estudiantes.

Si se desea tener una apreciación más precisa del efecto que tienen las variables del plantel y del individuo sobre el rendimiento por separado, es necesario incorporar una nueva serie de indicadores complementarios que se acerquen de manera más precisa a la descripción de las características del plantel y de los individuos y que sean capaces de aclarar en buena medida la varianza no explicada en el rendimiento de los alumnos.

La obvia relevancia de los temas tratados aquí contrasta llamativamente con la escasez de trabajos publicados sobre los determinantes de la calidad de la educación secundaria en Colombia aplicando la técnica multinivel. Es por eso que este estudio puede verse como una primera aproximación al problema de la cuantificación del impacto de distintos determinantes socioeconómicos y del plantel sobre el logro educativo. Difícilmente un único estudio puede dar cuenta de todos los aspectos relevantes para abordar un tema tan importante y urgente. Se espera que las ideas, metodologías y resultados de este trabajo motiven investigaciones futuras que ahonden en la problemática abordada. A tal efecto, existen varias líneas posibles de investigación futura:

1. Este estudio utiliza como punto de partida la educación secundaria, que es donde probablemente se den algunos de los aspectos más ricos de la interrelación entre características familiares y escolaridad. Una obvia dirección de investigación consiste en extender los resultados de este estudio a otros niveles, como la educación universitaria.

2. Varios de los resultados de este tipo de análisis chocan con la limitación natural de la calidad y cantidad de información disponible. Otra interesante línea de investigación consiste en enriquecer el análisis a través del estudio de otras fuentes de información sobre colegios o sobre hogares; la información de la Encuesta Nacional de Hogares (como la proporcionada por los formularios c-100 y c-600 del DANE), si bien no permite explorar la dimensión temporal del problema (existen una o dos ondas de cada una), permitiría incorporar información más detallada acerca de las características del colegio y los hogares.

3. Con todo, las encuestas mencionadas no están diseñadas específicamente para estudiar la problemática educativa. De modo que una importante línea de investigación consiste en desarrollar encuestas diseñadas con este propósito específico.

4. Algunos resultados del estudio llaman la atención sobre fenómenos de carácter marcadamente interdisciplinario. Por ejemplo, es importante entender por qué la educación de la madre es más importante que la del padre, o cómo el tamaño de las familias incide negativamente en la escolaridad. El aporte de otros enfoques sociales complementarios puede contribuir notoriamente al entendimiento de estos fenómenos.

## Bibliografía

- AITKIN, M. y LONGFORD, N. (1986) "Statistical modelling issues in school effectiveness", en *Journal of the Royal Statistical Society*, 149 (A), pp. 1-42.
- ALDANA, E.; RODRIGUEZ, P. y HEDERICH, C. (1992) "El logro educativo en matemáticas y lenguaje de los alumnos de tercer grado", en *Coyuntura Social* n° 7, Bogotá, Fedesarrollo, pp. 67-83.
- ALEXANDER, L. y SIMMONS, J. (1975) "The determinants of school achievement in developing countries: the educational production function", en *Cuadernos de trabajo* n° 201, Washington, Banco Mundial.
- ALTONJI, J. y DUNN, T. (1996) "Using siblings to estimate the effects of school quality on wages", en *Review Of Economics And Statistics* n° 78, pp. 665-671.
- BISHOP, J. (1989) "Is the test score decline responsible for the productivity grow decline?", en *American Economic Review*, vol. 75, n° 3, pp. 1016-30.
- BISHOP, J. (1992) "The impact of academic competencies on wages, unemployment, and job performance", en *Carnegie-Rochester Conference Series on Public Policy* n° 37, USA, pp. 127-194.
- BOSSIERE, M.; KNIGHT, JB. y SABOT, RH. (1985) "Earnings, Schooling, Ability, and Cognitive Skills", en *American Economic Review*, American Economic Association, vol. 75, n° 5, 1016-1030.
- BOWLES, S. y LEVIN, H. (1968) "The determinants of scholastic achievement. An appraisal of some recent evidence", en *Journal of Human Resources*, vol. 3, n° 1, pp. 3-24.
- BRANDSMAN, H.P. y KNUVER, J.W.M. (1989) "Effects of school and classroom characteristics on Pupil Progress in language and arithmetic's", en *International Journal of Educational Research*, n° 13, pp. 777-788.
- BRYK, A.S. y RAUDENBUSH, S.W. (1992) *Hierarchical Linear Models: applications and data analysis methods*, Newbury Park, California, Sage.
- BRYK, A.S.; RAUDENBUSH, S.W. y CONGDON, R. (1996) *HLM: Hierarchical Linear and Nonlinear Modeling with the HLM/2L and HLM/3L*, Chicago, Scientific Software International, Inc.
- CARD, D. y KRUEGER, A. (1996a) "Labor market effects of school quality: theory and evidence", *NBER working paper*, n° 5450, National Bureau of Economic Research, Inc. Michigan, pp. 1-69.
- CARD, D. y KRUEGER, A. (1996b) "School resources and students outcomes: an overview of the literature and new evidence from North and South California", en *Journal of Economic Perspectives*, n° 10 (Fall), pp. 31-50.
- CARO, B. (2000) "Factores asociados al logro académico de los alumnos de 3° y 9° de primaria en Bogotá", en *Coyuntura Social*, n° 22, Bogotá, Fedesarrollo, pp. 1-16.
- CASAS, A.; GAMBOA, L. Y PIÑEROS, L. (2002) "El Efecto Escuela en Colombia, 1999-2000", *Borradores de Investigación*, Universidad del Rosario, n° 27, Bogotá, pp. 3-37.
- CASTAÑO, E., (1998) "El Efecto Colegio sobre la variabilidad del rendimiento en matemáticas", en *Lecturas de Economía*, n° 49, Medellín, Universidad de Antioquia, pp. 49-57.
- COLEMAN, J.; CAMPBELL, E.; HOBSON, C.; Mc PARTLAND, I.; MODD, A.; WENFELD, F. y YORK, R. (1966) *Equality of educational opportunity*, Washington: U.S. Department of Health/ Education and Welfare/Office of Education/Government Printing Office.
- DEPARTAMENTO NACIONAL DE PLANEACIÓN (2002) *Bases para un Plan de Desarrollo 2002-2006. Hacia un Estado Comunitario*, Bogotá.
- EDMONDS, R. (1979) "Effective schools for the urban poor", en *Educational Leadership*, vol. 37, n° 10, pp. 15-24.
- FRASER, B. (1989) "Research syntheses on school and instructional effectiveness", en *International Journal of Educational Research*, vol. 13, n° 7, pp. 707-719.

- FRASER, B.; WALBERG, H.; WELCH, W. y HATTIE, J. (1987) "Syntheses of educational productivity research", en *International Journal of Educational Research*, vol. 11, n° 2, pp. 147-252.
- FULLER, B. y CLARKE, P. (1994) "Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules and pedagogy", en *Review and Educational research* n° 64, pp. 119-157.
- GAVIRIA, A. (2001) "Movilidad social en Colombia: un análisis comparativo", en *Ensayos Sobre Colombia y América Latina*, Madrid, BBVA, pp. 273-291.
- GAVIRIA, A. y Barrientos, J. (2001) "Características del Plantel y Calidad de la Educación en Bogotá", en *Coyuntura Social*, n° 25, Bogotá, Fedesarrollo, pp. 81-98.
- GOLDSTEIN, H. (1987) *Multilevel Models in Educational and Social Research*. Londres, Griffin.
- HANUSHEK, E. (1995) "Interpreting recent research and schooling in the developing countries", en *The World Bank Research Observer* n° 10, pp. 227-246.
- HAVEMAN, R y WOLFE, B. (1995) "The effect of the primary school avail on academia achievement across twenty-Nine high and low-income countries", en *American Journal of Sociology*.
- HENEVEL, W. y CRAIG, H. (1996) "Schools count: World Bank Project design and the quality of Primary Education in Sub-saharian Africa", en *World Bank Technical Paper*, n° 303, Washington D.C., The World Bank.
- HEYNEMAN, S. (1976). "Influences on academic achievement: A comparison of results from Uganda and more industrialized societies", en *Sociology of Education*, 49. vol. 11, no. 2, pp. 245-259.
- HEYNEMAN, S. (1980) "Difference between developed and developing countries: Comments on Simmons and Alexander's determinants of school achievement", en *Economic Development and Cultural Change*, vol. 28, núm. 2, pp. 403-406
- HEYNEMAN, S. y LOXLEY, W. (1983) "The effect of primary school quality on academic achievement across 29 high and low-income countries", en *American Journal of Sociology*, vol. 88, n° 6, pp. 1162- 1194.
- HOX, J. (2002) *Multilevel analysis techniques and applications*, Utrecht University, The Netherlands.
- JENKS, C. S. (1972) *Inequality: a reassessment of the effect of family and schooling in America*, New York, Basic Books.
- LOCKHEED, M.E. (1991) *Improving primary education in developing countries*, Oxford, Oxford University Press for the World Bank.
- LOCKHEED, M.E.; FULLER, B. y NYIRONDO, R. (1989) "Family effects on students achievement in Thailand and Malawi", en *Sociology of Education*, vol. 62, pp. 239-256.
- MINISTERIO DE EDUCACIÓN NACIONAL – SABER (1995), *La calidad de la educación: un asunto de todos*. Serie Documentos Especiales.
- MINISTERIO DE EDUCACION NACIONAL. COLOMBIA. *Estadísticas de Educación*. 2001.
- MISIÓN SOCIAL (1992); "El Logro Educativo en Matemáticas y Lenguaje de los Alumnos de Tercer Grado", Departamento Nacional de Planeación, Núm. 7, Pág. 67-83.
- MISIÓN SOCIAL (1997) "La Calidad de la educación y el logro de los planteles educativos", en *Planeación y Desarrollo* vol. 28, n° 1, Bogotá, Departamento Nacional de Planeación, pp. 25-62.
- MIZALA, A. y ROMAGUERA, P. (1998) "Desempeño escolar y elección de colegios: la experiencia chilena", *Documentos de Trabajo* 36, Centro de Economía Aplicada, Universidad de Chile, pp. 1-43.
- MORENO, H.; PIÑEROS, L. (1998) *Factores asociados al logro en matemáticas en el examen del ICFES: una comparación entre colegios oficiales y no oficiales*, Bogotá, Ministerio de Educación Nacional, policopiado.
- NUÑEZ, J.; STEINER, R.; CADENA, X. y PARDO, R. (2002) *¿Cuáles Colegios Ofrecen Mejor Educación en Colombia?* Documento de trabajo CEDE, n° 3, Bogotá, Universidad de los Andes, pp. 1-56.
- RASBASH, J.; BROWNE, W.; GOLDSTEIN, H.; YANG, M.; DRAPER, D.; LANGFORD, I. y TOBY, L. (2000) *A user's guide to MLwiN*. Londres, Center for Multilevel Modelling, Institute of Education, University of London

- RAWLS, J. (1979) *Teoría de la Justicia*, Madrid, Fondo de Cultura Económica.
- REYNOLDS, D.; TEEDLIE, C.; CREEMERS, B.; SCHEERENS, J.; TOWNSEND, T. (2000) "An introduction to school effectiveness research", en *The International Handbook of School Effectiveness Research*, New York, Falmer Press.
- RIDELL, A. (1993) "The evidence on public/private educational trade-offs in developing countries", en *International Journal of Educational Development*, n° 13, pp. 373-386.
- RIDELL, A. (1997) "Assessing designs school effectiveness research and school improvement in developing countries", en *Comparative Education Review*, vol. 41, n° 2, pp. 178-204.
- SARMIENTO, A, Y CARO, B. (1997) "El avance de la educación en Colombia: lento, insuficiente e inequitativo", en *Planeación y Desarrollo*, vol. 28, n° 1, Bogotá, Departamento Nacional de Planeación, pp. 11-24.
- SARMIENTO, *Becerra, L.*, (2000) "La Incidencia del Plantel en el logro educativo del alumno y su relación con el nivel socioeconómico", en *Coyuntura Social*, n° 22, Bogotá, Fedesarrollo, pp. 53-62.
- SCHEERENS, J. (1992) *Effective schooling: theory and practice*, London, Casell.
- SCHEERENS, J. (1997) *The foundation of educational effectiveness*, Oxford, Pergamon.
- SCHEERENS, J. (1999) "School Effectiveness in Developed and Developing Countries; A Review of the Research Evidence". Washington DC: The World Bank, pp. 1-48
- SCHIEFELBEIN, E. y SIMMONS, J. (1980) "Los determinantes del rendimiento escolar. Reseña de la investigación para los países en desarrollo", en *Educación Hoy*, n° 60, pp. 5-12.
- VÉLEZ, E., SCHIEFELBEIN, E. Y VALENZUELA, J. (1993) "Factors Affecting Achievement in Primary Education: A Review of Literature for Latin America", Banco Mundial.
- WALBERG, H.J. (1986) "Synthesis of research on teaching", en M.C. Wittrock (ed.) *Handbook of research on teaching*, Washington, American Educational Research Association, ed. M. C. Wittrock. New York: Macmillan, pp. 214-229.
- WILLMS, D. y SOMERS, M. (2001) "Family, classroom and school effects on children`s educational outcomes in Latin American School Effectiveness and School Improvement", vol. 12, n°4, pp. 409-445.