

CALIDAD DE LA EDUCACIÓN Y RENDIMIENTO ACADÉMICO EN CALI 2009

MÓNICA FARLEY SÁNCHEZ MONTES

Tesis

**Juan Byron Correa Fonnegra
Asesor de Tesis**

**UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS
SANTIAGO DE CALI
2012**

CALIDAD DE LA EDUCACIÓN Y RENDIMIENTO ACADÉMICO EN CALI 2009

RESUMEN

Se realiza el análisis sobre el logro en matemática y lenguaje de los estudiantes de grado once en las pruebas Saber 11 del año 2009, mediante variables correspondientes a alumnos como situación socioeconómica y las características asociadas al plantel donde acude y la interacción entre estas. Los datos corresponden a una muestra de 16374 alumnos distribuidos en 356 colegios de jornadas completa y mañana y edades entre los 15 y 21 años. Para el análisis se utilizó modelos de regresión lineal, y modelos multinivel: nivel 1 alumno y nivel 2 colegios). Se llegó a resultados que indican diferencias significativas entre el rendimiento de los alumnos y estas se explican principalmente por factores que operan a través de diferencias entre colegios y entre alumnos.

Palabras clave: rendimiento académico, coeficiente de correlación intraclase, análisis multinivel, características socioeconómicas y del plantel educativo, Saber 11-09.

TABLA DE CONTENIDO

	Pág.
1. INTRODUCCIÓN	4
2. ESTADO DEL ARTE	5
3. MARCO TEÓRICO	8
4. OBJETIVOS	10
4.1 OBJETIVO GENERAL	10
4.2 OBJETIVOS ESPECÍFICOS	10
5. MARCO EMPÍRICO	10
5.1 DATOS	10
6. METODOLOGÍA	12
7. RESULTADOS	16
7.1 CARACTERÍSTICAS DE LOS ALUMNOS Y COLEGIOS	16
7.2 ESTIMACIÓN DEL MODELO: APORTE DE LAS VARIABLES PARA EL NIVEL ESTUDIANTE	17
7.3 ESTIMACIÓN DEL MODELO: APORTE DE LAS VARIABLES PARA EL NIVEL COLEGIO	22
7.4 ESTIMACIÓN DEL MODELO MULTINIVEL	24
8. CONCLUSIONES	31
9. REFERENCIAS BIBLIOGRÁFICAS	32
ANEXOS	

1. INTRODUCCIÓN

La educación es un motor que provee la mejora de ingresos individuales en el futuro y un factor que se encuentra directamente correlacionado con el crecimiento económico de un país. Así, la importancia de mejorar la calidad de la educación implica un importante desafío para el gobierno nacional.

De esta manera, los estudios de Lucas (1988) y Romer (1986) en los cuales la variable considerada es el capital humano, medido por los años de educación de la población, demuestran una asociación positiva entre el capital y la tasa de crecimiento del producto.

Los beneficios derivados de la educación son varios, en primer lugar la acumulación de capital humano contrarresta los efectos de los rendimientos decrecientes, permitiendo que el crecimiento persista en el largo plazo. En segundo lugar, una mejor educación de un país permite que hayan menos desigualdades y una disminución de la pobreza por cuanto la gente no queda marginada de las actividades productivas.

Dado que la calidad de la educación en Colombia se mide a través del desempeño en las pruebas saber 11, y siendo este un país de tercer mundo o en vía de desarrollo, existe una ventaja en cuanto a la inversión en capital humano, ya que gracias a la ley de rendimientos decrecientes, los beneficios obtenidos de la inversión en capital humano son mayores en los países de primer mundo.

Por estas razones es importante conocer los factores que influyen sobre los resultados obtenidos por los alumnos de Cali, tanto de la zona urbana y rural en el desempeño de las pruebas Saber 11-2009 en las áreas de matemática y lenguaje.

Las características que se evalúan en este trabajo son las socioeconómicas de alumnos y las del plantel educativo con el fin de analizar el grado de correlación que tienen y la incidencia sobre el desempeño promedio en los resultados. Para dicho motivo se emplea la técnica estadística de análisis multinivel debido a que permiten realizar estimaciones más confiables a diferencia de las técnicas tradicionales en regresiones de variables individuales dentro de un contexto de anidamiento, es decir el alumno dentro del colegio, el colegio dentro del municipio, etc. Y llegar a conclusiones más precisas de los resultados y del efecto de factores escolares y extraescolares.

En investigación del logro educativo es frecuente la disyuntiva entre el peso que tiene el efecto alumno y el efecto colegio sobre el rendimiento académico y calidad educativa. Por ello este trabajo intenta aproximarse a dilucidar un poco dicha disyuntiva en el municipio de Cali y posteriormente dar pie a otras investigaciones referentes al tema educativo y el contexto asociado a él.

2. ESTADO DEL ARTE

La calidad de la educación en Colombia es un tema ampliamente discutido no sólo por el Ministerio de educación Nacional, sino, por diversos autores interesados en abordar la educación como factor de desarrollo en una sociedad, en el sentido de que la educación contribuye a mejorar la calidad de vida, incidiendo en varios aspectos como la posibilidad de obtener mejores empleos, devengar ingresos más altos y ampliar las relaciones interpersonales.

A nivel nacional se destacan en los estudios de calidad académica, Gaviria y Barrientos (2001a), quienes estudian los principales determinantes que influyen en la calidad de la educación en Colombia, con base en las pruebas del ICFES 1999, evaluando los efectos tanto familiares y del plantel educativo sobre el rendimiento académico de los estudiantes de bachillerato para la ciudad de Bogotá. De este estudio se obtienen tres principales resultados: La educación de los padres tiene un efecto sustancial sobre el rendimiento académico, el efecto de la educación sobre los padres se transmite a sus hijos principalmente a través de la calidad de los planteles educativos, y los planteles educativos inciden de manera notable sobre el rendimiento académico. Posteriormente Gaviria (2002), indaga sobre la influencia del gasto del gobierno en educación, al evaluar la relación que tendría un aumento del gasto en educación sobre la mejora de la calidad de la educación comparando colegios públicos versus colegios privados. De este trabajo se destaca que el aporte de esta nueva variable (gasto en educación) ha estado relacionado más con la cobertura en educación en las regiones apartadas de Colombia y muy poco con la calidad. En el mismo año el trabajo de Misas (2002) debate el trabajo de Gaviria y Barrientos (2001b), cuyo trabajo se enfoca en el rendimiento académico en Bogotá. El autor plantea que en este trabajo no se hace énfasis en los procesos de educación, aprendizaje y evaluación, y argumenta que hay que hacer diferencia entre los tipos de conocimientos (discursivo y procedimental). Por tanto, el autor pone de manifiesto la importancia del tipo de conocimientos adquiridos y que influyen en los canales de transmisión de los códigos sociolingüísticos y el capital cultural heredo.

Continuando con la línea de análisis que se vincula a los determinantes involucrados en la calidad de la educación, Mina (2004) realiza su estudio a nivel municipal, y Barrientos (2008) lo lleva a cabo en el área metropolitana de Medellín. Mina (2004) tiene por objeto conocer cuáles son los factores que hacen que un municipio tenga un mejor resultado en el ICFES que otro; así mismo, trata de identificar cuáles son los más importantes para poder diseñar una adecuada política en materia de educación, observando el efecto del gasto público en la calidad de la educación. En este estudio, el autor estimó un modelo utilizando datos panel, en el cual la variable dependiente es el promedio por categoría en el examen de estado, con un panel de 897 municipios durante el período 1996-1999, obteniendo como resultados variables significativas el gasto por niño (positivo), así como la pobreza y la desigualdad de cada municipio (negativas). Las variables de entorno como la altura y la distancia del municipio con respecto a su capital ayudan a explicar los resultados en las pruebas. Por otro lado, Barrientos (2008) busca identificar cuáles son los determinantes que influyen sobre la calidad de la educación pública en Medellín (2004-2006). El autor estudia la influencia del estrato, edad del estudiante el tipo de colegio (público o privado), al igual que la planta física y observa finalmente que los atributos como la planta física del

plantel influyen en colegios privados y no en los públicos, mostrando para este último más relevancia en las aptitudes cognitivas de los estudiantes.

Por otro lado, algunos autores se han interesado por indagar el impacto que ha tenido la ejecución de la Ley 60 de 1993. Es por ello que Melo (2005) investiga el impacto de la descentralización fiscal en la provisión de la educación pública en Colombia, a través de métodos de frontera estocástica. Los resultados muestran la presencia de altos niveles de ineficiencia en costos y en producción en los diferentes departamentos del país.

De igual modo Iregui, Melo y Ramos (2006) evalúan las normas sobre descentralización educativa, y el comportamiento reciente de los indicadores de gasto público, cobertura, eficiencia, calidad. También, miden el impacto sobre el rendimiento académico de factores asociados al colegio y al entorno socioeconómico de los estudiantes. De aquí se extrae como resultados principales que las variables asociadas a la infraestructura de los colegios y al entorno socioeconómico de los estudiantes, tienen un impacto positivo y significativo sobre el logro académico.

A nivel internacional han sido diversos autores los que han abordado el tema como Hanushek (1986), el cual le da un tratamiento económico a la educación. El artículo sintetiza los resultados econométricos de 147 trabajos en los cuales se estiman los determinantes de la calidad de la educación de los planteles públicos en los Estados Unidos. Finalmente, el autor concluye que puede existir variabilidad en el sentido de que las variables pueden ser significativas en algunos de los 147 trabajos (algunos positivos y otros negativos). Hanushek (1996) evalúa las bondades de las políticas públicas que pregonan la mejora de la educación mediante una mayor asignación de dinero. Para llevar a cabo dicho objetivo, el autor realiza las estimaciones pertinentes utilizando funciones de producción en educación. Posteriormente en los trabajos de Lee y Barro (2000) Boozer y Malooney (2001) abordan en sus respectivas investigaciones la variable número de estudiantes por maestro. El primero, hace un estudio sobre la calidad de la educación, medida por los resultados en pruebas internacionales de la IEA (International Association for the Evaluation Achievement) para alumnos de secundaria en 58 países con diferente ingreso per capita, desde 1964 a 1991, demostrando que lo que permite principalmente mejorar el rendimiento académico es la reducción del tamaño de las clases y secundariamente el ingreso familiar y educación de los padres.

En el trabajo de Boozer y Malooney (2001) para la ciudad de Tennessee, emplea datos del proyecto norteamericano STAR para finales de los años noventa, experimento que evalúa los efectos del tamaño de la clase en los planteles públicos del estado. De este estudio se encuentra evidencia econométrica de los beneficios que tienen en el corto plazo la reducción del número de alumnos por docente, lo que lleva implícito que el tamaño de la clase debe irse reduciendo permanentemente.

De otro lado, Pritchett (2004) estima un índice de “carencia de educación” entendida como una “falta de preparación” para ingresar en el mercado laboral mundial competitivo del siglo XXI.

Para llevar a cabo esta tarea, el autor suma la proporción de la población de estudiantes de 15 años que obtiene una puntuación de 1 o menos en la escala PISA de lectura, y la proporción de estudiantes de 15 a 19 años que no terminó el noveno grado. Y concluye que existe una buena proporción de estudiantes que están mal capacitados para encontrar un buen empleo en el mercado laboral.

De otro lado, Altinok y Murseli (2007), analizan la relación entre la satisfacción con la educación y algunas variables sencillas a nivel individual (edad, género, zona urbana y categoría de ingresos) y a nivel de país (PIB per cápita, y cantidad y calidad de la educación). Este objetivo se lleva a cabo empleando una regresión basada en datos de la encuesta mundial de Gallup. El índice general de calidad del capital humano de Altinok y Murseli (2007) estima la calidad de la educación para los años más recientes en los que se dispone de datos. De este modo los resultados indican que los latinoamericanos tienen una percepción de sus sistemas educativos más favorable que sus homólogos en otras regiones del mundo.

Bajo la influencia de las percepciones que los países Latinoamericanos y del Caribe tienen sobre la calidad de sus sistemas educativos Hanushek y Woessman, (2007), hacen un estudio interesante con el objetivo de estimar las contribuciones de la calidad educativa al ingreso individual, a la distribución del ingreso y al crecimiento económico. Obteniendo como resultado que la calidad es un indicador más fiable de las contribuciones de la educación que los años de escolaridad al punto que la importancia de la cantidad como factor explicativo desaparece una vez que se consideran las mediciones de calidad.

De otro lado, Auguste, Echart y Franchetti (2008) en su investigación tiene como objetivo demostrar la permanente relación inversa entre calidad y cantidad de la educación para América Latina, para ello los autores comparan detenidamente varios países, principalmente en Europa Oriental, que dedican a la educación una proporción del PIB similar a la de Argentina y tienen niveles de ingreso per cápita comparables, pero obtienen mucho mejores resultados en la prueba PISA y otros programas de pruebas internacionales. Por consiguiente, los resultados a constatar es que Argentina, como otros países latinoamericanos, tiene una pirámide demográfica con una base más amplia; por lo tanto, el gasto en educación per cápita es constantemente más bajo que en los países con los que se compara, esto se traduce en que la carga para el sistema educativo sea mucho más pesada que los países con los cuales se contrasta (Hungría y Polonia)

En el trabajo de Lora (2008) analiza la relación entre la evolución de la educación y la calidad en América Latina. El autor realiza un repaso de la expansión cuantitativa de la educación en todos los países de América Latina y obtiene que las ganancias de la educación han sido menores debido a las deficiencias generales y graves de la calidad de la educación en la región, teniendo consecuencias más directas sobre la calidad de vida en la región, menor productividad, retraso en la innovación tecnológica y baja mano de obra calificada.

Para el mismo año el estudio de Flores y Herrera (2008) indagan a manera segregacional las percepciones de los padres, es decir el objetivo del estudio, es examinar de qué manera la

percepción de los padres se relaciona con los niveles de aislamiento en las comunidades chilenas. Para analizar los efectos que tiene el hecho de que los grupos aislados posiblemente no tengan buenas bases para juzgar la calidad de la educación de las escuelas. Por consiguiente, en los resultados se observaron que en las zonas no segregadas los padres de niños que asistían a escuelas con resultados promedio más altos en las pruebas de logros de aprendizaje reportaban niveles mucho más elevados de satisfacción con su escuela que en otras del mismo vecindario.

Ya para el año 2007 el estudio de Vegas y Petrow (2007) investiga la calidad de la educación en Colombia. De esta manera el informe plantea como objetivo principal identificar los factores determinantes del aprendizaje en Colombia, centrándose en la equidad, las comparaciones entre países y estudios nacionales. Aquí, se analiza el impacto de los factores relacionados con las instituciones, escuelas, estudiantes, padres y profesores en los resultados escolares de los estudiantes. De este informe se llega al siguiente resultado: en Colombia son varios los factores de los estudiantes, de las escuelas e instituciones que están asociados con el aprendizaje escolar. Y que para efectos del trabajo a desarrollar se pretende analizar cuáles son los determinantes más influyentes en la calidad de la educación y el rendimiento académico para la ciudad de Cali.

3. MARCO TEÓRICO

Los análisis de mínimos cuadrados ordinarios comúnmente utilizados suponen que cada observación tiene el mismo peso. Tal como lo han demostrado diferentes investigaciones sobre el tema, en estudios en donde las observaciones se encuentran agrupadas por niveles, como en el caso de la educación (curso y escuela, principalmente), lo que lleva a estimaciones incorrectas.

En el campo educativo, los alumnos están agrupados principalmente en salones de clases, los que a su vez forman parte de un colegio determinado. Sin embargo, cuando el grupo está definido todos sus miembros afectarán y serán afectados por el resto y tenderán a diferenciarse de otros grupos (Delprato, M., 1999). Por ejemplo, las condiciones en un colegio no oficial son diferentes a las de uno oficial de aquí que las experiencias educativas de cada alumno sean diferentes del resto. Por lo que las técnicas de análisis estadístico tradicionalmente utilizadas son de utilidad limitada e incurrir en dos tipos de errores diferenciados (Hox, J.J., 1995):

- Falacia atomista (término acuñado por Alker en 1969). Consiste en asignar el mismo peso a los alumnos, sin preocuparse por la posible variación entre los factores (socioeconómicos y escolares) del contexto escolar o del grupo.
- Falacia ecológica (término acuñado por Robinson en 1950). Se incurre en este error cuando al calcular la media de cada variable del alumno para asignársela al grupo al que pertenece, se traslada como conclusión al nivel alumno. Siendo esto sólo factible para el estudio de las relaciones de nivel colegio.

Los modelos multinivel (jerárquicos lineales) ponen fin a estos problemas trabajando con los diferentes niveles al mismo tiempo.

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 en el puntaje) y se va perfeccionando mediante la adición progresiva de variables explicativas, tanto de alumnos como de colegios. Dependiendo del efecto que la adición de las variables tenga sobre la reducción de la varianza total, éstas serán retenidas o eliminadas del modelo. Además, los niveles de agregación se representan con un submodelo, éstos expresan las relaciones que se producen entre variables dentro de un mismo nivel y especifican como las variables de un nivel influyen en otro nivel.

Otro de los motivos por los que es necesario el análisis multinivel, es porque los datos que provienen de observaciones individuales no son siempre independientes. Los alumnos de un mismo colegio tienden a parecerse entre ellos, por ejemplo: algunos colegios no oficiales atraerán principalmente a alumnos con un nivel socioeconómico elevado, mientras que otros agruparán a alumnos con un estatus socioeconómico bajo (colegios oficiales). El grado de homogeneidad de los contextos bien dado por el C.C.I (coeficiente de correlación intraclase). Las consecuencias de ignorar el C.C.I. son las siguientes (Gaviria, J.L. y Castro, M., 2005):

- La información obtenida a nivel individual no es de tanta magnitud como parece, debido a que los alumnos de los mismos colegios tienden a parecerse entre ellos. Por lo tanto, la información que proporcionan los estudiantes de una mismo colegio es menor que la que suministran los alumnos de distintos planteles educativos.
- Los errores típicos son demasiado pequeños debido a que los tests estadísticos se basan en el supuesto de independencia de las observaciones. No obstante, en esta clase de estructuras poblacionales dicho supuesto no se cumple. Como consecuencia de ello es posible confirmar la existencia de resultados significativos cuando realmente son espúreos.

En resumen, cuando las características de los alumnos y la provisión de insumos y técnicas de enseñanza son similares al 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 coeficientes estimados. Mientras que al analizar el rendimiento mediante las técnicas multinivel permiten establecer qué parte del logro es explicado por la escuela y qué parte por el estudiante.

4. OBJETIVOS

4.1 Objetivo general

Identificar los determinantes que influyen en la calidad de la educación y el rendimiento académico en Cali.

4.2 Objetivos específicos

Haciendo uso de la base de datos suministrada por el Instituto Colombiano para la Evaluación de la Educación Superior (ICFES), se plantean los siguientes objetivos específicos:

- Determinar si el grado de escolaridad de la madre influye en el desempeño académico de sus hijos.
- Detectar en qué medida las características del plantel inciden en el rendimiento académico del estudiante.
- Evidenciar si en conjunto las características familiares y el entorno educativo potencializan el rendimiento académico en Cali.

5. MARCO EMPÍRICO

5.1. Datos

Empleando los resultados del Instituto Colombiano para la Evaluación de la Educación Superior (ICFES) en el año 2009, se examinará el logro académico en la ciudad de Cali en las pruebas de matemática y lenguaje.

La Prueba Saber 11, antes conocida como examen del ICFES, es un examen de estado que evalúa a los estudiantes que están terminando su ciclo de educación media. La prueba tiene como finalidad apoyar los procesos de selección y admisión que realizan las instituciones de educación superior. Además de este propósito, la prueba busca:

- Brindar al estudiante información que contribuya a la selección de su opción profesional.
- Proporcionar información a las instituciones de educación básica y media sobre el desempeño de los estudiantes.
- Contribuir al desarrollo de estudios de tipo cultural, social y educativo.
- Servir de criterio para otorgar beneficios educativos.

El Examen de Estado se aplica en Colombia desde la década de los años 80'. Sin embargo, es sólo desde el año 2000 cuando se empieza a concebir como una evaluación centrada en las competencias de los estudiantes más que en la memorización de contenidos.

Los campos que se indagan en la Prueba Saber 11 se pueden agrupar de la siguiente manera:

- Información general del estudiante.
- Información de discapacidad del estudiante.
- Información sobre la Institución Educativa.
- Información acerca de la carrera e Institución de Educación Superior a la que desea ingresar el estudiante.
- Información de presentación del examen.
- Información familiar.
- Información socio-económica.
- Resultado de desempeño educativo.

Con respecto a los resultados de desempeño académico de las pruebas, se puede establecer una estructura similar en la organización y composición de las áreas académicas que se indagaron en las pruebas de los siguientes periodos de tiempo:

Desde 2000-1 hasta 2007-1: la estructura general constaba de núcleo común 8 áreas interdisciplinarias: lenguaje matemática, ciencias naturales (biología, física y química) y ciencias sociales (historia geografía y filosofía). Componente flexible el alumno elige una prueba de profundización (lenguaje, matemática, biología o historia) y una prueba interdisciplinar (medio ambiente, o violencia y sociedad). Idiomas, el alumno elige entre la prueba de inglés, francés o alemán.

Desde 2007-2 hasta 2009-2: la estructura general consta de núcleo común compuesto de 8 áreas interdisciplinarias: lenguaje matemática, ciencias naturales (biología, física y química) y ciencias sociales (historia geografía y filosofía). Componente flexible el alumno elige una prueba de profundización (lenguaje, matemática, biología o historia) y una prueba interdisciplinar (medio ambiente, o violencia y sociedad).

Dado que el objeto de estudio son los alumnos del municipio de Cali, el tamaño de la muestra se eligió para aquellos que asisten a planteles cuyas jornadas son completa y mañana y tienen edades entre los 15 y 21 años, debido a que tienen mayor frecuencia dentro de la población, por lo que se tiene un tamaño muestral de 16.374 alumnos para 356 colegios.

De acuerdo a la información consignada en la base de datos de las pruebas saber 11-2009 se eligieron variables pertenecientes a la información general del estudiante, institución educativa, información familiar y socioeconómica. Como lo que se busca es comparar los resultados obtenidos del área urbana respecto al área rural, las variables se cruzaron con el área, de esta manera la mencionada variable es dicotómica y se definió 1 si es zona urbana y 0 si es zona rural.

De igual modo se definió: sexo del estudiante 1 si es mujer 0 si es hombre, hogar con computador 1 si posee 0 si no posee, conexión a internet 1 si tiene 0 si no tiene. En cuanto a la información de la educación de la madre está categorizada en: 0 ninguno, 9 primaria incompleta, 10 primaria completa, 11 secundaria incompleta, 12 secundaria completa, 13 técnica o tecnológica incompleta, 14 técnica o tecnológica completa, 15 profesional incompleta, 16 profesional completa, 17 postgrado, 99 no sabe. También la ocupación del padre se encuentra categorizada en: 13 empresario, 14 pequeño empresario, 15 director o gerente general, 16 empleado de nivel directivo, 17 empleado de nivel técnico / profesional, 18 empleado auxiliar o administrativo, 19 obrero u operario 20 profesional independiente, 21 trabajador por cuenta propia, 22 hogar, 23 pensionado, 26 otra actividad u ocupación.

El estrato socioeconómico de los alumnos de Cali está entre 1 y 6. Y se agrupan como: 1 estratos 1 y 2, 3 estratos 3 y 4, 5 estratos 5 y 6. El ingreso del hogar se categoriza en 1, menos de 1 SMMLV, 2 entre 1 y menos de 2 SMMLV, 3 entre 2 y menos de 3 SMMLV, 4 entre 3 y menos de 5 SMMLV, 5 entre 5 y menos de 7 SMMLV, 6 entre 7 y menos de 10 SMMLV, 7 10 o más SMMLV. Las variables número de personas del grupo familiar e índice socioeconómico del estudiante son de carácter continuo.

Continuando con las variables que se encuentran categorizadas, la naturaleza del colegio se agrupa en oficiales y no oficiales, esta última se refiere a colegios privados y semiprivados. La categoría del colegio se describe como: 1 alto, 2 bajo, 3 inferior, 4 medio, 5 muy inferior, 6 muy superior, 7 superior. Finalmente, el índice socioeconómico del colegio se ubica en un rango de 1 a 4.

6. METODOLOGÍA

La estructura de los datos en las pruebas Saber 11-2009 es anidada, es decir, los alumnos se encuentran agrupados en aulas, que a su vez forman parte de un centro educativo determinado. Los modelos de regresión lineal no permiten tener en cuenta los efectos que pueden provocar sobre el resultado el modo en que los alumnos se asignan a los centros o a las clases dentro de los centros. Por lo que la estimación de este tipo de modelos no tendrá en cuenta la estructura jerárquica de los datos y no diferenciará entre la varianza que es debida al alumno y aquella que es debida al centro. De acuerdo a lo anterior, en el desarrollo de este trabajo se emplea los modelos multinivel o de estructura jerárquica (lineales) Murillo (2008), debido a que la regresión multinivel tiene en cuenta el carácter anidado de los datos dentro de unidades más amplias, calculando una ecuación diferente para cada nivel de agregación. De esta forma es posible diferenciar que parte del logro es explicado por el alumno y cuál es debida al centro educativo. Además es posible realizar inferencias a partir de variables que actúan en distintos niveles.

El software que permite hacer este tipo de estimaciones y que se empleó es Stata 11.2 y Mlwin 2.24.

El procedimiento inicial consiste en aplicar regresiones lineales para observar el efecto de las variables en el nivel alumno y en el nivel colegio en las pruebas de matemática y lenguaje, y posteriormente se hace la regresión del modelo multinivel con aquellas variables que representen significancia estadística en el nivel alumno y en el nivel colegio.

El modelo de regresión lineal tiene la forma:

$$Y_{ij} = \beta_{0i} + \beta X_{0i} + e_{0i}$$

Donde Y_{ij} es el puntaje del individuo i que asiste al plantel j , β_{0i} es el puntaje promedio del alumno, X_{0i} es un vector de características de los individuos que incluye, entre otras variables, el sexo, la educación de la madre la ocupación del padre, y e_{0i} es un término de error de los alumnos. De esta regresión se espera que la relación de signos entre las variables X_{0i} sea positiva respecto a Y_{ij} .

El modelo multinivel, tiene como funcionalidad ampliar el modelo de regresión lineal clásico y relacionar el modelo del primer nivel con el modelo del segundo nivel, al igual que, regresar los coeficientes del primer nivel en el segundo nivel de variables. El análisis propuesto parte de un modelo nulo o vacío al cual posteriormente se le adicionan predictores.

El modelo general queda establecido de la siguiente forma:

Para el nivel alumno (nivel 1):

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^Q \beta_{qj} X_{qij} + \varepsilon_{ij}$$

Donde:

Y_{ij} Es el rendimiento promedio de un alumno i del colegio j .

β_{0j} Es la media en rendimiento para todos los alumnos del colegio j .

β_{qj} Es el incremento en el rendimiento por cada característica de los alumnos (X) del colegio j .

X_{qij} Hace referencia a las diferentes variables relacionadas con el alumno y su entorno, donde q es una covariable del alumno i .

ε_{ij} Es el rendimiento diferencial del alumno i de la escuela j . El término aleatorio se distribuye de forma normal con media cero y varianza constante.

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

Cada coeficiente del primer nivel pasa a ser una ecuación en el segundo. Este nivel dos (colegio) quedaría formulado de la siguiente forma:

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^{Sq} \gamma_{0s} W_{sj} + \mu_{0j}$$

Donde:

γ_{00} Hace referencia al rendimiento para todas las escuelas

γ_{0s} Es el incremento sobre el rendimiento medio para cada predictor (W) del colegio j.

W_{sj} Son los predictores del nivel colegio.

μ_{0j} Es el efecto diferencial producido por la escuela j sobre aquellas que tienen sus mismas características.

$$\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{Sq} \gamma_{qs} W_{sj} + \mu_{qj}$$

Donde:

γ_{q0} Es el incremento medio de todos los colegios para cada predictor del nivel alumno.

γ_{qs} Es la aportación de cada variable de la escuela a γ_{q0} .

μ_{qj} Es el efecto aleatorio producido por el colegio j en cada uno de sus predictores.

El vector de efectos aleatorios de nivel colegio $(\mu_{1j}, \mu_{2j}, \dots, \mu_{qj})$, cuya matriz se distribuye de forma normal multivariante con media cero, varianza constante igual a $Var(\mu_{qj}) = \tau_{qq}$ para todo $q = 1, 2, \dots, n$ y para cada par de efectos aleatorios q y q' $Cov(\mu_{qj}, \mu_{q'j}) = \vartheta_{qq'}$. La matriz de varianzas – covarianzas T tiene una dimensión igual a $(Q+1) (Q+1)$.

$$\mu_{qj} \sim N(0, T)$$

$$Y_{ij} = \gamma_{00} + \sum_{s=1}^{Sq} \gamma_{0s} W_{sj} + \underbrace{\sum_{q=1}^Q \gamma_{q0} X_{qij}}_{\text{Efectos fijos}} + \underbrace{\sum_{q=1}^Q \sum_{s=1}^{Sq} \gamma_{qs} W_{sj} X_{qij} + \mu_{qj} X_{qij} + \mu_{0j} + \varepsilon_{ij}}_{\text{Efectos aleatorios}}$$

De esta forma el modelo general completamente aleatorio de dos niveles sería:

Efectos fijos incluye todos los parámetros que definen la media del rendimiento de todos los alumnos.

Efectos aleatorios muestra la estimación de la varianza en cada nivel de agregación, en este caso los alumnos (nivel 1) y los colegios (nivel 2), para cada parámetro incluido en la parte fija.

El modelo nulo para el primer nivel quedaría:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

Y_{ij} Es el rendimiento en matemática o lenguaje de un alumno i del colegio j .

β_{0j} Es el rendimiento medio en matemática o lenguaje para todos los alumnos del colegio j .

ε_{ij} Es el efecto diferencial de un alumno i del colegio j . Se distribuye de forma normal con media cero y varianza constante σ^2 .

Segundo nivel:

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

γ_{00} Es la media general para todos los colegios

μ_{0j} Es el efecto diferencial producido por la escuela j . Se distribuye de forma con media cero y varianza τ_{00}

Sustituyendo los coeficientes del segundo nivel en el primer nivel:

$$Y_{ij} = \gamma_{00} + \mu_{0j} + \varepsilon_{ij}$$

Por lo tanto la varianza del rendimiento de los estudiantes es:

$$Var(Y_{ij}) = Var(\varepsilon_{ij} + \mu_{0j}) = \sigma^2 + \tau_{00}$$

Y el coeficiente de Correlación Intraclass (C.C.I) sería:

$$\rho = \frac{\tau_{00}}{(\sigma^2 + \tau_{00})}$$

El C.C.I posibilita estimar la medida en que existe dependencia entre los individuos, es decir con este tipo de estimación se puede conocer las características que comparten los alumnos por estudiar en una mismo colegio.

Si en la estimación se tiene una correlación intraclass baja o cercana a cero significará que los alumnos dentro del mismo colegio son tan diferentes entre sí como los que pertenecen a otros colegios. Indicando que los colegios no tienen alumnos con características homogéneas y por

tanto los alumnos tienen características independientes a pesar de pertenecer a un mismo colegio, es decir existe independencia en las observaciones.

7. RESULTADOS

Después de llevar a la práctica la metodología planteada, se muestra a continuación los resultados de la estimación, y las estadísticas descriptivas de las características relevantes de los alumnos de grado once de los colegios de Cali y posteriormente los resultados de los modelos estimados.

Los resultados de las pruebas Saber 11 en las cinco regiones de Colombia (Pacífica, Caribe, Andina, Orinoquia, y Amazonia) tienen porcentajes que sobrepasan el 70% en el nivel de desempeño medio en las pruebas de matemática y lenguaje a excepción del departamento de Antioquia (Región Andina) en la prueba de matemática (52.9%). En el nivel bajo la región que presenta los porcentajes más altos es la región caribe, sin embargo el porcentaje más alto se registra en el departamento de Antioquia (38.9%). La región que mayores porcentajes tiene en el nivel de desempeño alto es la región Andina, más específicamente el departamento de Cundinamarca en un 23.5%.

En la región Pacífica, los departamentos que registran mayores porcentajes en los tres niveles es el departamento de Nariño con un 88.6% (medio), y Valle con un 18.3% (bajo) y 4,9% (alto). En la prueba de lenguaje las regiones tienen mayoritariamente desempeños medios. (Ver gráfico 1).

7.1. Características de los alumnos y colegios

Las mujeres obtienen 3.1 puntos menos en matemática que los hombres, y existe mayor variabilidad entre los puntajes de los hombres al comparar la desviación estándar para este grupo (12.0) respecto a su contrario (10.8). Mientras que para la prueba de lenguaje la diferencia es mínima (0.014) puntos y la variabilidad son casi iguales para ambos grupos 7.8 (hombres) y 7.7 (mujeres). Lo que indica que para el municipio de Cali las mujeres no tienen mayores puntajes en lenguaje que los hombres como comúnmente se piensa. (Ver tabla1).

Los alumnos Caleños reportan que el 55.5% son mujeres y el 44.5% son hombres. La zona urbana tiene el 90,6% de los estudiantes y el 9.4% están en la zona rural. Si se agrupan hombres y mujeres, se observa el porcentaje de puntajes que obtuvieron según la zona en la que viven, en la área urbana en la prueba de matemática el 88.8% de las mujeres alcanzaron un desempeño medio, los hombres un 89,0 % .En la zona rural el 88.2% de las mujeres obtuvieron un desempeño medio y los hombres el 88,9%. En la prueba de lenguaje el 98.6% de la mujeres del área urbana tiene un resultado medio similar a los hombres en 98.3%. Este resultado se repite en el área rural: 98.7% mujeres en el nivel medio y 98.0% hombres en este mismo nivel. Estos cálculos revelan que los alumnos de Cali independiente de la zona tienen resultados medios en las pruebas Saber 11-2009. (Ver tabla 2 y 3)

La información consignada en la base de datos revela que en las zonas mencionadas el nivel educativo de las madres es de secundaria completa (32.6% urbana), (25.7% rural), a nivel profesional sobresale la zona urbana en 12.7% respecto al 8% en el área rural.

Dado que hay mayor porcentaje de madres bachilleres, lo que se espera es que se relacionen con hombres de igual nivel educativo, y estos a su vez tengan empleos acordes a su formación, así el 21.9% son obreros y operarios, el 21.7% son trabajadores por cuenta propia en la urbe, en el área rural sobresalen las mismas categorías ocupacionales, obreros 22.3% y trabajador por cuenta propia el 26.2%.

En cuanto al ingreso familiar mensual devengado en estas dos áreas el 47.3% y el 51.7% reciben entre 1 y menos de 2 SMMLV respectivamente. En la zona rural el 22.5% de las familias tienen ingresos inferiores a 1SMMLV, superior al de su zona contraria (12.7%). De acuerdo al ingreso, las familias viven en estratos socioeconómicos acorde con su presupuesto, por lo que el 51.6% de la urbe viven en estratos 1 y 2, y el 40.0% viven en estratos 3 y 4. Mientras que en el área rural el 71.1% viven en estratos 1 y 2, el 23.6% viven en estratos 3 y 4.

Las familias de los estudiantes están conformadas mayoritariamente por 4 y 5 personas (32.9% y 32.5% urbana), (32.5% y 22.6% rural) respectivamente. La dotación de computador y conexión a internet de los hogares urbanos arroja que el 64.8% tiene el artículo, y el 40.5% tienen el servicio. En los hogares rurales hay una notable desventaja el 51.0% tiene computador y sólo el 28.0% tiene conexión a internet, de lo que se deduce un bajo rendimiento en las pruebas.

Conforme a las características de los planteles de la ciudad, el 62.3% de los colegios urbanos son no oficiales y el 96% son calendario B. La mayoría se ubica en categorías medias (32.8%), baja (21.0%), superior (17.2%), y sólo el 8.2% en muy superior. Este resultado es de esperarse, ya que el índice socioeconómico del colegio en su mayoría es 2 (55.5%) y 3 (29.5%). Mientras que el 60.5% de los colegios rurales son oficiales y cuenta con calendario A en un 96%. En las categorías media y alta se obtienen porcentajes muy similares (34.7% y 17.7%) respectivamente, sólo el 5% de los colegios se categorizan como muy superior. Conforme con estos resultados los índices socioeconómicos no son muy alentadores, muestran que el 34.94% tienen 1 y el 34.94% son 2. (Ver tabla 4).

7.2. Estimación del modelo: aporte de las variables para el nivel estudiante.

En los cuadros subsecuentes aparece: el valor del coeficiente, con su error típico, al igual que la varianza y desviación del logaritmo de verosimilitud (D.L.V).

Cuadro 1.

Variables	Matemática		Lenguaje	
Intercepto	46.2	(0.089)	49.5	(0.060)
Varianza	131.079	(1.448)	59.78	(0.660)
D.L.V	126303.77		113450.12	

Fuente: Cálculos de la autora

Partiendo de un modelo clásico de regresión lineal, se puede observar que el puntaje global en las pruebas de matemática y lenguaje es de 46.2 puntos y 49.5 puntos respectivamente.

Al incorporar la variable sexo del estudiante (1 si es mujer y 0 si es hombre). En el cuadro 2 se aprecia que las mujeres en matemática tienen en promedio 3.0 puntos menos que los hombres. En lenguaje el desempeño por sexo no influye demasiado, pues las mujeres sólo obtienen 0,014 puntos más que los hombres.

Cuadro 2. Aporte de la variable sexo

Variables	Matemática		Lenguaje	
Intercepto	47.8	(0.132)	49.5	(0.090)
Sexo	-3.0	(0.178)	0.014	(0.121)
Varianza	128.763	(1.423)	59.787	(0.660)
D.L.V	126011.86		113450.1	

Fuente: Cálculos de la autora

Procediendo con la adición de variables a los modelos de matemática y lenguaje, se agrega la variable educación de la madre (cuadro 3). En este modelo sólo se utilizó la educación de la madre, debido a que es ella quien influye más que el padre en los procesos cognitivos y en el desarrollo académico del alumno, motivando y alentando a que este tenga la disposición necesaria en el transcurso de su educación secundaria.

Observando el aporte de esta variable según los niveles educativos que mayor frecuencia presentan las madres en Cali. Se tiene que si pertenece al área urbana y tiene educación primaria incompleta disminuye el rendimiento promedio del alumno en matemática en -1.2 puntos. Mientras que si tiene estudios secundarios completos el rendimiento promedio en 1.9 puntos. Y aumenta en 10.0 puntos el rendimiento promedio si la madre tiene educación profesional completa respecto a las del área rural. Para el desempeño en la prueba de lenguaje se observa algo similar en los niveles educativos: disminuye en -1.1 puntos el rendimiento promedio si la madre tiene educación primaria incompleta, aumenta en 0.0773 si tiene educación secundaria completa y aumenta en 5.2 puntos si cuenta con estudios profesionales completos.

Cuadro 3. Aporte de la variable educación de la madre

Variables		Matemática		Lenguaje	
Intercepto		45.2	(0.279)	48.2	(0.194)
Sexo		-2.9	(0.167)	0.054	(0.116)
Educación de la madre	Primaria	-1.2	(0.361)	-1.1	(0.250)
	Secundaria	1.9	(0.304)	0.773	(0.211)
	Profesional	10.0	(0.359)	5.2	(0.249)
Varianza		112.740	(1.246)	54.307	(0.600)
D.L.V		123836.02		111876.12	

Fuente: Cálculos de la autora

En el cuadro 4. Se aprecia la relevancia de incluir la variable ocupación del padre, debido a que en la mayoría de los hogares son los hombres quienes aportan el ingreso en el hogar y de acuerdo a esto deciden qué proporción del mismo destinan para la educación de sus hijos. De esta manera, se tiene que los alumnos cuyos padres son pequeños empresarios incrementan en 4.5 puntos el rendimiento promedio, y si es trabajador por cuenta propia sólo aumenta en 1.0 puntos el rendimiento promedio en matemática.

Los coeficientes asociados a la prueba de lenguaje, indican que si la ocupación del padre es pequeño empresario el puntaje promedio en lenguaje aumenta en 2.8 puntos, y si es trabajador por cuenta propia el incremento es poco considerable, tan sólo de 0.5 puntos. Los resultados anteriores permiten inferir que en medida que la ocupación del padre implique una remuneración salarial mayor, el desempeño del alumno en ambas áreas del conocimiento serán también mayores respecto a los del área rural.

Cuadro 4. Aporte de la variable ocupación del padre

Variables		Matemática		Lenguaje	
Intercepto		45.2	(0.297)	47.9	(0.196)
Sexo		-2.9	(0.173)	0.097	(0.124)
Educación de la madre		0.024	(0.006)	0.013	(0.004)
Ocupación del padre	Pequeño empresario	4.5	(0.443)	2.8	(0.283)
	Independiente	1.0	(0.351)	0.5	(0.214)
Varianza		121.085	(1.338)	56.880	(0.660)
D.L.V		125005.18		102061.29	

Fuente: Cálculos de la autora

Los alumnos cuyos padres devengan menos de un salario mínimo tienen 2.2 puntos adicionales en el rendimiento promedio en matemáticas. Al incrementar el ingreso mensual del hogar el puntaje promedio aumenta en 3.2 puntos, si el hogar recibe ingresos mensuales entre 1 y 2 SMMLV y aumenta en 6.4 si el ingreso oscila entre 2 SMMLV y 3 SMMLV. (Cuadro 5).

Se puede apreciar que en la prueba de lenguaje el rendimiento promedio es directamente proporcional con el ingreso en el área urbana, así el rendimiento aumenta en 1.0 puntos si el hogar recibe mensualmente menos de 1 SMMLV, si el ingreso del hogar está entre 1 y 2 SMMLV el puntaje aumenta en 1.9 puntos, y aumenta en 3.7 puntos en relación a los alumnos del área rural si el ingreso mensual del hogar se ubica entre 2 SMMLV y 3 SMMLV.

Cuadro 5. Aporte de la variable ingreso

Variables		Matemática		Lenguaje	
Intercepto		45.2	(0.291)	48.301	(0.199)
Sexo		-2.921	(0.165)	0.093	(0.116)
Educación de la madre		0.004	(0.006)	0.002	(0.004)
Ocupación del padre		-0.151	(0.025)	-0.106	(0.018)
Ingreso del hogar	Menos de 1SMMLV	2.2	(0.649)	1.0	(0.455)
	Entre 1 y 2 SMMLV	3.2	(0.601)	1.9	(0.421)
	Entre 2 y 3 SMMLV	6.4	(0.603)	3.7	(0.423)
Varianza		110.593	(1.192)	54.364	(0.587)
D.L.V		123521.15		111893.07	

Fuente: Cálculos de la autora

El cuadro 6 indica que los alumnos del área urbana pertenecientes a estratos 1 y 2 tienen un rendimiento promedio en matemática y lenguaje menor. Obtienen -2.6 y -1.3 puntos menos, respectivamente en ambas pruebas. En la prueba de matemática los alumnos pertenecientes a estratos 3 y 4 obtuvieron 0.7 puntos más en el rendimiento promedio, sin embargo este parámetro no resulta significativo al nivel del 5% pero si al 10%, mientras que en la prueba de lenguaje resulta significativo al 5% e incrementa el rendimiento casi en un punto (0.9). Por otro lado, los alumnos pertenecientes a estratos 5 y 6 aumentaron su rendimiento promedio en matemática y lenguaje en 3.3 y 2.7 respectivamente.

Cuadro 6. Aporte de la variable estrato socioeconómico

Variables		Matemática		Lenguaje	
Intercepto		45.2	(0.283)	48.3	(0.198)
Sexo		-2.9	(0.164)	0.1	(0.114)
Educación de la madre		-0.0	(0.006)	0.0	(0.004)
Ocupación del padre		-0.1	(0.025)	-0.1	(0.018)
Ingreso del hogar		2.2	(0.082)	0.9	(0.057)
Estrato socioeconómico	1 y 2	-2.6	(0.649)	-1.3	(0.453)
	3 y 4	0.7	(0.663)	0.9	(0.465)
	5 y 6	3.3	(0.778)	2.7	(0.545)
Varianza		108.549	(1.179)	53.281	(0.589)
D.L.V		122944.84		111318.28	

Fuente: Cálculos de la autora

La variable índice socioeconómico del estudiante resultó ser positiva y significativa, lo que indica que el incremento de esta característica por alumno influye en el desempeño de matemática en 2.7 puntos más y en 1.6 puntos más en lenguaje (Cuadro 7).

Cuadro 7. Aporte de la variable índice socioeconómico del estudiante

Variables	Matemática		Lenguaje	
Intercepto	43.8	(0.261)	47.5	(0.183)
Sexo	-2.9	(0.162)	0.1	(0.114)
Educación de la madre	-0.0	(0.006)	-0.0	(0.004)
Ocupación del padre	-0.4	(0.014)	-0.2	(0.010)
Ingreso del hogar	1.3	(0.085)	0.5	(0.060)
Estrato socioeconómico	0.7	(0.092)	0.6	(0.064)
Índice socioeconómico del estudiante	2.7	(0.154)	1.6	(0.108)
Varianza	106.853	(1.182)	52.705	(0.583)
D.L.V	122589.88		111052.36	

Fuente: Cálculos de la autora

En el cuadro 8, si bien se observa que los coeficientes no llegan a la unidad, el signo negativo revela que el aumento de personas en el grupo familiar disminuye en -0.6 puntos el rendimiento promedio en matemática y en -0.3 puntos el resultado de la prueba de lenguaje.

Cuadro 8. Aporte de la variable personas del grupo familiar

Variables	Matemática		Lenguaje	
Intercepto	44.2	(0.263)	47.7	(0.185)
Sexo	-2.9	(0.162)	0.0	(0.114)
Educación de la madre	-0.0	(0.005)	-0.0	(0.004)
Ocupación del padre	-0.3	(0.016)	-0.2	(0.011)
Ingreso del hogar	1.4	(0.085)	0.6	(0.060)
Estrato socioeconómico	0.7	(0.092)	0.6	(0.064)
Índice socioeconómico del estudiante	2.8	(0.154)	1.6	(0.108)
Número de personas del grupo familiar	-0.6	(0.046)	-0.3	(0.032)
Varianza	105.872	(1.171)	52.367	(0.578)
D.L.V	122439.3		111094.3	

Fuente: Cálculos de la autora

Al incluir la variable hogar con computador en el cuadro 9, se tiene que esta variable es significativa y positiva, lo que indica que poseer computador en el hogar influye positivamente sobre el desempeño en ambas pruebas. Numéricamente el aporte de esta variable incrementa en 1.8 puntos el rendimiento promedio en la prueba de matemática y aumenta en 1.4 puntos en la prueba de lenguaje.

Cuadro 9. Aporte de la variable hogar con computador

Variables	Matemática		Lenguaje	
Intercepto	43.3	(0.277)	47.0	(0.195)
Sexo	-2.9	(0.161)	0.1	(0.113)
Educación de la madre	-0.0	(0.005)	-0.0	(0.004)
Ocupación del padre	-0.2	(0.016)	-0.1	(0.011)
Ingreso del hogar	1.3	(0.086)	0.5	(0.060)
Estrato socioeconómico	0.6	(0.092)	0.5	(0.064)
Índice socioeconómico del estudiante	2.6	(0.154)	1.5	(0.108)
Número de personas del grupo familiar	-0.5	(0.046)	-0.3	(0.032)
Hogar con computador	1.8	(0.184)	1.4	(0.129)
Varianza	105.265	(1.165)	51.992	(0.575)
D.L.V	122345.52		110829.96	

Fuente: Cálculos de la autora

En el cuadro 10, al igual que en el cuadro anterior, contar con conexión a internet aumenta el rendimiento promedio en la prueba de matemática en 1.0 puntos y en 0.7 puntos el rendimiento en lenguaje; en relación a los alumnos del área rural.

Cuadro 10. Aporte de la variable hogar con internet

Variables	Matemática		Lenguaje	
Intercepto	43.3	(0.277)	47.0	(0.195)
Sexo	-2.9	(0.161)	0.1	(0.113)
Educación de la madre	-0.0	(0.005)	-0.0	(0.004)
Ocupación del padre	-0.2	(0.017)	-0.1	(0.011)
Ingreso del hogar	1.2	(0.087)	0.4	(0.061)
Estrato socioeconómico	0.6	(0.092)	0.5	(0.065)
Índice socioeconómico del estudiante	2.5	(0.154)	1.5	(0.108)
Número de personas del grupo familiar	-0.5	(0.046)	-0.3	(0.032)
Hogar con computador	1.4	(0.205)	1.1	(0.144)
Hogar con conexión a internet	1.0	(0.218)	0.7	(0.153)
Varianza	105.130	(1.163)	51.921	(0.574)
D.L.V	122324.52		110807.85	

Fuente: Cálculos de la autora

7.3. Aporte de las variables para el nivel colegio

La adición de la variable área (cuadro 11) indica que si los alumnos asisten a colegios ubicados en el área urbana tienen un rendimiento promedio mayor en matemática a los del área rural en 2.9 putos más, de igual forma tienen resultados mayores en la prueba de lenguaje en 1.2 puntos más.

Cuadro 11. Aporte de la variable área

Variables	Matemática		Lenguaje	
Intercepto	43.5	(0.291)	48.3	(0.197)
Área	2.9	(0.305)	1.2	(0.206)
Varianza	130.380	(1.440)	59.658	(0.659)
D.L.V	126216.23		113414.88	

Fuente: Cálculos de la autora

Al incluir la naturaleza del colegio en el modelo. Los resultados muestran que los colegios no oficiales aportan 3.3 puntos más al rendimiento promedio en la prueba de matemática y 1.4 puntos más en la prueba de lenguaje respecto a los puntajes obtenidos en el área rural.

Cuadro 12. Aporte de la variable naturaleza del colegio

Variables	Matemática		Lenguaje	
Intercepto	43.5	(0.288)	48.3	(0.196)
Área	0.8	(0.325)	0.4	(0.221)
Naturaleza del colegio	3.3	(0.191)	1.4	(0.130)
Varianza	128.002	(1.414)	59.258	(0.654)
D.L.V	125914.85		113304.46	

Fuente: Cálculos de la autora

En el cuadro 13 se adiciona la variable calendario del colegio. Se tiene que los alumnos que asisten a colegios cuyo calendario académico es A aumentan al puntaje promedio en matemática

en 7.6 puntos y si es calendario B aportan 6.2 puntos. Para la prueba de lenguaje el aporte al puntaje promedio si es calendario A es de 2.7 puntos y si es calendario B es de 3.7 puntos. Estos resultados permiten inferir que si el alumno asiste a un colegio con calendario académico B obtiene mejor puntaje en la prueba de lenguaje, y si es calendario A tiene mejor desempeño en matemática.

Cuadro 13. Aporte de la variable calendario

Variables		Matemática		Lenguaje	
Intercepto		43.5	(0.287)	48.3	(0.196)
Área		2.0	(0.596)	-0.1	(0.403)
Naturaleza del colegio		-3.7	(0.193)	-1.6	(0.131)
Calendario	A	7.6	(1.587)	2.7	(1.108)
	B	6.2	(0.500)	3.7	(0.341)
Varianza		126.825	(1.401)	58.831	(0.650)
D.L.V		125763.62		113186.12	

Fuente: Cálculos de la autora

El aporte de la variable categoría del colegio a las dos pruebas, es significativo y negativo tanto en la prueba de matemática y lenguaje (cuadro 14), en la categoría baja y media se obtienen -10.2 puntos y -7.4 puntos menos en el rendimiento respectivamente. Sólo el coeficiente de los colegios con categoría muy superior es positivo y aumenta en 10.7 puntos el rendimiento en dicha prueba.

En lenguaje se presenta lo mismo que en la prueba anterior: pertenecer a la categoría bajo disminuye en -6.8 puntos y en -4.5 puntos si el colegio es categoría media, Mientras que la categoría muy superior aumenta en 4.7 puntos el rendimiento en la prueba de lenguaje.

Cuadro 14. Aporte de la variable categoría del colegio

Variables		Matemática		Lenguaje	
Intercepto		43.5	(0.254)	48.3	(0.179)
Área		-9.4	(0.984)	4.2	(0.693)
Naturaleza del colegio		-0.2	(0.191)	0.2	(0.134)
Calendario		-0.8	(0.433)	-0.1	(0.305)
Categoría del colegio	Bajo	-10.2	(0.288)	-6.8	(0.203)
	Medio	-7.4	(0.255)	-4.5	(0.180)
	Muy superior	10.7	(0.348)	4.7	(0.245)
Varianza		99.772	(1.102)	49.520	(0.547)
D.L.V		121835.19		110365.16	

Fuente: Cálculos de la autora

El aporte de la variable índice socioeconómico del colegio (cuadro 15) al puntaje promedio en matemática muestra que si el índice socioeconómico del colegio es 2 el desempeño en el puntaje de esta prueba aumenta en 2.2 puntos, si es 3 el puntaje aumenta en 6.8 puntos y si es 4 el puntaje asciende en 14.8 puntos. En la prueba de lenguaje si el índice socioeconómico del colegio es 2, el aporte en el rendimiento en la prueba de lenguaje es 2.2 puntos, si es 3 el rendimiento aumenta en 6.8 puntos y si es 4 aumenta en 14.8 el rendimiento promedio.

Cuadro 15. Aporte de la variable índice socioeconómico del colegio

Variables		Matemática		Lenguaje	
Intercepto	Área	40.3	(0.345)	46.3	(0.240)
	Naturaleza del colegio	6.0	(1.067)	2.4	(0.744)
	Calendario	0.3	(0.196)	0.7	(0.137)
	Categoría del colegio	-3.7	(0.440)	-2.2	(0.307)
		0.6	(0.047)	0.3	(0.333)
Índice socioeconómico del colegio	2	2.2	(0.348)	1.3	(0.249)
	3	6.8	(0.366)	4.6	(0.255)
	4	14.8	(0.432)	8.6	(0.301)
Varianza		107.1423	(1.184)	52.170	(0.576)
D.L.V		122904.4		111130.41	

Fuente: Cálculos de la autora

7.4. Modelo multinivel

Ahora se estima el modelo multinivel (nivel 1 alumnos y nivel 2 colegio), incorporándole variables que inciden tanto en el nivel alumno como en el nivel colegio, tales como: sexo del estudiante, educación de la madre, ingreso del hogar, estrato socioeconómico del alumno, categoría del colegio, índice socioeconómico del alumno e índice socioeconómico del colegio. Las variables mencionadas a excepción del índice socioeconómico del colegio se cruzaron con la variable área.

En el cuadro 16, aparece el coeficiente γ_{00} incluido en la parte fija del modelo y su error típico. También aparece en la tabla la varianza de cada nivel de agregación, el coeficiente de correlación intraclase (C.C.I), y la desviación del logaritmo de verosimilitud (D.L.V).

Cuadro 16. Modelo Nulo

Variables		Matemática		Lenguaje	
Efectos fijos	Media general γ_{00}	45.9	(0.377)	49.2	(0.218)
	Nivel 1 ε_{ij}	94.763	(1.058)	48.166	(0.537)
Efectos Aleatorios	Nivel 2 μ_{0i}	44.201	(3.754)	14.017	(1.251)
	C.C.I	0.318		0.225	
	D.L.V	121907.61		110684.88	

Fuente: Cálculos de la autora

La estimación de este modelo indica que existe varianza sin explicar en el rendimiento de los alumnos. El C.C.I indica que los colegios explican un 31% de la varianza del rendimiento en matemática y un 22% en lenguaje, el resto es debido a las variables del alumno y su entorno. A la vista de estos resultados el análisis jerárquico es conveniente.

El aporte de la variable sexo en el cuadro 17 es significativa y negativa, lo que indica que las mujeres del área urbana tiene un rendimiento promedio menor (-2.5 puntos) en la prueba de matemática con respecto a los hombres de la misma área. En la prueba de lenguaje, la variable sexo es significativa y positiva al nivel de significancia del 10% indicando que las mujeres del área urbana tiene un rendimiento promedio que no llega a la unidad en 0.2 puntos más que las hombres.

Cuadro 17. Aporte de la variable sexo

Variables		Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}	47.1	(0.388)	49.1	(0.226)
	Sexo	-2.5	(0.166)	0.2	(0.119)
	Nivel 1 ε_{ij}	93.427	(1.043)	48.162	(0.537)
Efectos Aleatorios	Nivel 2 μ_{0i}	44.986	(3.809)	13.984	(1.248)
	C.C.I	0.325		0.225	
	D.L.V	121685.01		110682.7	

Fuente: Cálculos de la autora

En la prueba de matemática y lenguaje se evidencia la influencia de la educación de la madre sobre el desempeño, así los hijos de madres profesionales del área urbana tuvieron 5.7 puntos y 2.4 puntos más que los hijos de las madres profesionales del área rural respectivamente. En la prueba de lenguaje el nivel educativo primaria y secundaria no resultaron significativos al 5%. (Cuadro 18).

Cuadro 18. Aporte de la variable educación de la madre

Variables			Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}		44.5	(0.438)	48.4	(0.277)
	Sexo		-2.8	(0.170)	0.1	(0.122)
	Educación de la madre	Primaria	1.1	(0.408)		
		Secundaria	2.6	(0.362)		
		profesional	5.7	(0.411)	2.4	(0.292)
Efectos Aleatorios	Nivel 1 ε_{ij}		91.823	(1.025)	47.562	(0.531)
	Nivel 2 μ_{0i}		33.048	(2.862)	9.892	(0.916)
	C.C.I		0.264		0.172	
	D.L.V		121312.54		110382.48	

Fuente: Cálculos de la autora

La variable ingreso incluida en el cuadro 19, revela que en la prueba de matemática a medida que el hogar pasa de recibir ingresos mensuales menores a 1 SMMLV a recibir entre 2 y 3 SMMLV incrementan su rendimiento promedio de 1.7 puntos a 3.8 puntos, siendo el aporte de los niveles salariales significativos y positivos. En la prueba de lenguaje solo hay un incremento positivo y significativo en 1.1 puntos si el hogar tiene ingresos entre 2 y 3 SMMLV).

Cuadro 19. Aporte de la variable ingreso

Variables			Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}		44.3	(0.445)	48.4	(0.286)
	Sexo		-2.8	(0.171)	0.1	(0.123)
	Educación de la madre		-0.0	(0.005)	-0.0	(0.004)
	Ingreso del Hogar	Menos de 1 SMMLV	1.7	(0.437)		
		Entre 1 y 2 SMMLV	2.2	(0.383)		
Entre 2 y 3 SMMLV		3.8	(0.405)	1.1	(0.287)	
Efectos Aleatorios	Nivel 1 ε_{ij}		92.447	(1.032)	47.983	(0.535)
	Nivel 2 μ_{0i}		31.792	(2.780)	9.897	(0.923)
	C.C.I		0.255		0.170	
	D.L.V		121409.7		110524.41	

Fuente: Cálculos de la autora

En el cuadro 20, la prueba de matemática arroja como resultado que los alumnos que pertenecen a estratos socioeconómicos 3 y 4 aumentan su rendimiento promedio en esta prueba en sólo 1.1 puntos, los aportes de los demás estratos no resultaron significativos en el desempeño en matemática. En cuanto a la prueba de lenguaje, pertenecer a los estratos 1 y 2 disminuyen el rendimiento promedio en -1.4 puntos. Los aportes de los demás estratos no resultaron significativos.

Cuadro 20. Aporte de la variable estrato

Variables		Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}	44.3	(0.438)	48.4	(0.277)
	Sexo	-2.8	(0.171)	0.1	(0.123)
	Educación de la madre	-0.0	(0.005)	-0.0	(0.004)
	Ingreso del Hogar	1.0	(0.086)	0.6	(0.061)
	Estrato socioeconómico			-1.4	(0.307)
Efectos Aleatorios		1.1	(0.459)		
	Nivel 1 ε_{ij}	92.350	(1.032)	47.984	(0.536)
	Nivel 2 μ_{0i}	29.450	(2.602)	8.280	(0.793)
	C.C.I	0.241		0.147	
	D.L.V	121103.99		110233.73	

Fuente: Cálculos de la autora

En el cuadro 21, se incluye la variable categoría del colegio (baja, medio, superior y muy superior). El puntaje en matemática aumenta en 4.3 puntos si el colegio del área urbana se cataloga como superior y aumenta en 11.4 puntos el rendimiento promedio si el colegio pertenece a la categoría muy superior. En la prueba de lenguaje, aumenta en 1.6 puntos si el colegio tiene categoría superior, y aumenta el rendimiento en 5 puntos si es categoría muy superior. Las categorías restantes influyen negativamente sobre el desempeño del estudiante en ambas pruebas.

Cuadro 21. Aporte de la variable categoría del colegio

Variables			Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}		44.3	(0.325)	48.4	(0.210)
	Sexo		-2.8	(0.169)	0.1	(0.120)
	Educación de la madre		-0.0	(0.005)	-0.0	(0.004)
	Ingreso del Hogar		0.8	(0.085)	0.4	(0.060)
	Estrato		0.2	(0.093)	0.3	(0.065)
	Categoría del colegio	Baja	-3.8	(0.501)	-3.8	(0.319)
		Media	-0.9	(0.466)	-1.7	(0.301)
		Superior	4.3	(0.595)	1.6	(0.383)
		Muy superior	11.4	(0.750)	5.0	(0.482)
Efectos Aleatorios	Nivel 1 ε_{ij}		93.553	(1.044)	48.277	(0.538)
	Nivel 2 μ_{0i}		4.875	(0.607)	0.982	(0.170)
	C.C.I		0.049		0.199	
	D.L.V		120862.8		109900.91	

Fuente: Cálculos de la autora

En el cuadro 22 la variable índice socioeconómico del colegio influye en el rendimiento promedio en matemática en 6.3 puntos, si el colegio tiene índice 3, y aumenta aún más si el índice del colegio es 4, aportando 11.4 puntos al rendimiento promedio en esta prueba. En la

prueba de lenguaje si bien no hay tanto aumento en los puntajes, se observa que si el índice del colegio es 3, el aporte al rendimiento es de 3.9 puntos y si el índice del colegio es 4 el rendimiento en la prueba es de 6.9 puntos.

Cuadro 22. Aporte de la variable índice socioeconómico del colegio

Variables		Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}	40.0	(0.819)	45.5	(0.497)
	Sexo	-2.9	(0.167)	-0.0	(0.120)
	Educación de la madre	-0.0	(0.005)	-0.0	(0.004)
	Ingreso del Hogar	0.8	(0.083)	0.3	(0.059)
	Estrato	0.2	(0.092)	0.2	(0.066)
	Categoría del colegio	0.3	(0.84)	0.0	(0.056)
	Índice socioeconómico del colegio	2	1.3 (0.883)	0.7 (0.536)	
		3	6.3 (0.946)	3.9 (0.576)	
Efectos aleatorios		4	11.4 (1.076)	6.9 (0.664)	
	Nivel 1 ε_{ij}	92.411	(1.033)	47.964	(0.536)
	Nivel 2 μ_{0i}	14.859	(1.440)	4.883	(0.518)
	C.C.I	0.138		0.092	
D.L.V		120825.13		110005.03	

Fuente: Cálculos de la autora

El aporte de la variable índice socioeconómico del alumno al cuadro 23 resulta ser muy despreciable en el rendimiento promedio de ambas pruebas, puesto que el coeficiente no llega a la unidad.

Una vez incluidos las variables mencionadas, se tiene que el modelo final a conseguido reducir la varianza inexplicada en ambos niveles, respecto al modelo nulo. Así, la varianza inexplicada en el nivel alumno en la prueba de matemática se redujo en un 4.9% y en un 62.6% en el nivel colegio. En lenguaje la varianza del primer nivel se redujo en 3.18% y la del segundo se redujo en 62.1%.

El modelo resulta significativo al 5% en ambas pruebas. La diferencia con el modelo nulo también resulta significativa, la resta de D.L.V con siete grados de libertad para matemática es de 1752.7 y en lenguaje es de 1380.24, ambas siguen una distribución chi cuadrado con probabilidad cero.

Cuadro 23. Aporte de la variable índice socioeconómico del alumno

Variables		Matemática		Lenguaje	
Efectos fijos	Intercepto γ_{00}	33.363	(0.777)	41.383	(0.467)
	Sexo	-2.988	(0.166)	-0.082	(0.118)
	Educación de la madre	-0.011	(0.005)	-0.008	(0.004)
	Ingreso del Hogar	0.157	(0.104)	-0.106	(0.074)
	Estrato	-0.221	(0.101)	-0.052	(0.072)
	Categoría del colegio	0.046	(0.088)	-0.143	(0.058)
	Índice socioeconómico del colegio	4.168	(0.314)	2.531	(0.193)
	Índice socioeconómico del estudiante	0.093	(0.009)	0.064	(0.006)
	Nivel 1 ε_{ij}	90.115	(1.008)	46.634	(0.521)
Efectos aleatorios	Nivel 2 μ_{0i}	16.517	(1.574)	5.307	(0.550)
	C.C.I	0.154		0.102	
	D.L.V	120154.91		109304.64	

Fuente: Cálculos de la autora

Cuadro. 24 Coeficiente de correlación intraclase

Modelo		Matemática	Lenguaje
C.C.I	Nulo	0.318	0.225
	Sexo	0.325	0.225
	Educación de la madre	0.264	0.172
	Ingreso del Hogar	0.255	0.170
	Estrato	0.241	0.147
	Categoría del colegio	0.049	0.199
	Índice socioeconómico del colegio	0.138	0.092
	Índice socioeconómico del estudiante	0.154	0.102

Fuente: Cálculos de la autora

El C.C.I calculado para el modelo nulo indica que el 31% de la variación en el rendimiento en matemática es explicada por el colegio, y en la prueba de lenguaje esta variación es explicada en un 22%. A medida que se incorporan variables del nivel alumno y colegio al modelo este coeficiente disminuye. Al incorporar el ingreso del hogar se tiene que el 25% de la variación es explicada por el colegio en matemática, y en lenguaje es explicada en un 17%. Se observa que el valor del coeficiente cae notoriamente al introducir la categoría del colegio y ahora sólo el 4.9% de la variación en matemática la explica el colegio y en lenguaje el 1.9%. Esto puede deberse a que en la mayoría de los colegios de Cali la categoría es media, lo que supone que las dotaciones escolares no son óptimas.

Alternativamente, el cuadro 24 muestra el cálculo del C.C.I, para los colegios oficiales y no oficiales en las dos pruebas, con el fin de observar que tanto en el modelo nulo como en el final, es mayor la variación explicada por el colegio en las instituciones no oficiales, y que al llegar al modelo final disminuyó el coeficiente. De esta manera en colegios privados la variación explicada por la institución es del 17% (matemática) y 10% (lenguaje), mientras que en los colegios oficiales esta variación es del 3% (matemática) y 5% (lenguaje).

Cuadro 24. C.C.I colegios oficiales y no oficiales

Modelo	Matemática		Lenguaje	
	Oficial	No oficial	Oficial	No Oficial
C.C.I				
Nulo	0.068	0.344	0.078	0.253
Índice socioeconómico del estudiante	0.033	0.170	0.052	0.101

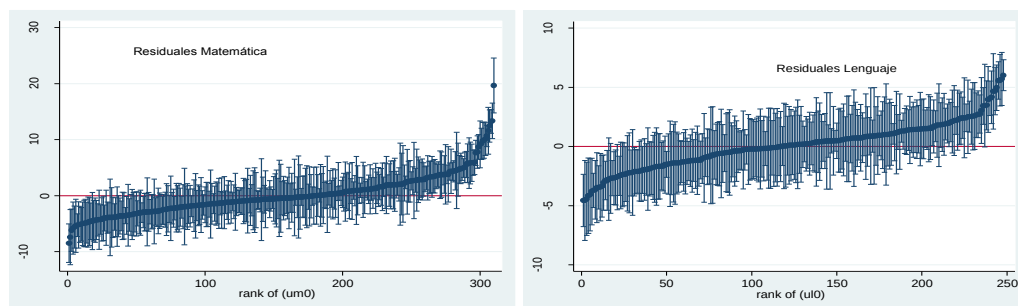
Fuente: Cálculos de la autora

Al analizar los residuales del nivel colegio para matemática (Gráfico 2) aparecen los residuos de todos los colegios, ordenados ascendentemente, con un intervalo de confianza del 99%. Aquellos cuyo intervalo de confianza se encuentra por encima del valor cero de la constante, son los que están situados por encima de su rendimiento predicho. Puede observarse que sólo el 12% aproximadamente se encuentra en esa situación, el resto de los centros educativos se encuentran situados por debajo de su rendimiento esperado

La prueba del Jarque-Bera en este nivel permite comprobar que los residuales no tienen un comportamiento normal con un valor de 7082 y un p-valor de 0.00 este valor se debe a que la asimetría dista mucho de ser cero (1.007) sugiriendo asimetría positiva y la curtosis sobrepasa el valor de tres (5.515) indicando una distribución leptocúrtica. Análogamente para la prueba de lenguaje (Gráfico 3) los residuales en este nivel no indican una distribución normal con un Jarque-Bera de 202 y un p-valor de 1.4e-44, lo que permite rechazar la hipótesis nula de normalidad.

Los valores de los residuales de los colegios se presentan gracias a los datos atípicos, es decir los alumnos de Cali presenta alta variabilidad en los resultados de las pruebas Saber 11-09, algo que es de esperarse gracias a que comúnmente los alumnos de colegios no oficiales obtienen puntajes más elevados que los alumnos que asisten a planteles oficiales. Por lo que se estimó los residuales sólo para los colegios oficiales de la ciudad, y nuevamente los residuales no presentan un comportamiento normal. Ver cuadro 25.

Gráfico 1. Residuales matemática Gráfico 2. Residuales lenguaje



Fuente: Cálculos de la autora

Cuadro 25. Residuales para colegios Oficiales y no Oficiales

Estadístico	Matemática				Lenguaje			
	Oficial y N.O	P-valor	Oficial	P-valor	Oficial y N.O	P-valor	Oficial	P-valor
Asimetría	1.007	0.00	0.171	0.00	0.169	0.00	0.156	0.00
Curtósis	5.514	0.00	2.747	0.00	3.426	0.00	3.596	0.00
Jarque Bera	7082	0.00	49.51	1.8e-11	202	1.4e-44	123.1	1.9e-27

Fuente: Cálculos de la autora

8. CONCLUSIONES

De acuerdo a las características de los alumnos de Cali, a continuación se contestan los objetivos planteados con anterioridad.

El grado de escolaridad de la madre influye notoriamente en el desempeño académico de sus hijos en las pruebas de lenguaje y matemática, si esta cuenta con estudios de educación técnico-tecnológico y superior. Sin embargo, según los resultados, tanto del área urbana y rural la mayoría de las madres caleñas apenas cuentan con el título de bachiller, por lo que priman los alumnos con desempeños medios en las pruebas.

El resultado del coeficiente de correlación intraclase del modelo nulo reveló que el 31.8% y el 22.5% de la variación en el rendimiento en matemática y lenguaje es explicado por el colegio. Al incorporarle predictores al modelo nulo se redujo el C.C.I lo que indica que ahora el 15.4% y el 10.2% del rendimiento es explicado por las características del colegio, lo que sugiere que tiene más influencia las características socioeconómicas y familiares del alumno.

El conjunto de las características del colegio y del estudiante ayudan a mejorar el desempeño del estudiante en ambas pruebas, no obstante en Cali los alumnos del área urbana y los del área rural se ubican en un desempeño medio. Aunque las condiciones de los alumnos del área rural presentan ciertas desventajas como la educación en colegios oficiales frente a los no oficiales del área urbana, el desempeño no varía mucho, gracias a que las características individuales y su background son similares.

El desempeño en las pruebas revela que los hombres obtienen mejores puntajes en matemática que las mujeres, aunque era de esperarse que las mujeres tuvieran mejor desempeño en lenguaje no fue el caso para las alumnas de Cali. Según los resultados no se presentaron mayores cambios en los puntajes.

De acuerdo a los resultados obtenidos se puede decir que el rendimiento y la calidad de la educación en Cali se sitúan en un nivel medio, muy menguado por el ingreso familiar, lo que indica que la ley de ampliación de cobertura no va estrechamente ligada con la calidad. En este caso según Rodríguez (1994) se debe hablar de calidad de la cobertura. Por otro lado, queda por fuera del alcance de este trabajo, saber cuál es el ingreso destinado a la educación en el municipio de Cali después de la mencionada ley, analizar la proporción del ingreso que las familias destinan a la educación de sus hijos, la formación de los docentes, es decir, contar con bases de datos más amplias para llegar así a conclusiones mucho más concretas.

9. REFERENCIAS BIBLIOGRÁFICAS

- ALTINOK, Nadir y HATIDJE Murseli. (2007). "International Database on Human Capital Quality". *Economics Letters*, Vol. 2, No. 96, pp., 237– 44.
- AUGUSTE, Sebastián, ECHART María y FRANCHETT Francisco i. (2008). "The Quality of Education in Argentina". *Banco Interamericano de Desarrollo*, Departamento de Investigación, Washington, DC. Documento mimeografiado.
- BARRIENTOS, Jorge (2008). "Calidad de la educación pública y logro académico en Medellín 2004-2006. Una aproximación por regresión intercuartil", *Lecturas de Economía*, No. 68, pp. 121-144. Universidad de Antioquia.
- BECKER, G. S. Y N. TOMES (1986), "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, Vol 4, No.3, pp 1-39. Reprinted as the supplement to chapter 7 in G. S. Becker, *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press, 1991.
- BECKER , MURPHY Y TAMURA (1994), "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education" ,*The National Bureau of Economic Research*, Vol ISBN: 0-226-04119-0, 3rd edition, chapter XII Human Capital, Fertility, and Economic Growth, pp. 323 – 350.
- BOOZER, M. Y MALONEY, T. (2001). "The Effects of Class Size on the Long Run Growth in Reading Abilities and Early Adult Outcomes in the Christchurch Health and Development Study". *Yale Economic Growth Center*. Discussion Paper No: 827.
- FLORES, Carolina y HERRERA María Soledad. (2008). "Understanding Quality of Life in Latin America and the Caribbean: Satisfaction, Quality of Education and Income Inequality". *Banco Interamericano de Desarrollo*. Documento de trabajo para la Red de Centros de Investigación, Washington, DC
- CUNHA, Flavio y HECKMAN James J (2007), "The Technology of Skill Formation", *IZA*, paper No. 2550 january.
- DELPARTO, M. (1999). "Determinantes del rendimiento educativo del nivel primario aplicando la técnica de análisis multinivel". Córdoba: IERA.
- GAVIRIA, A. (2002) "Los que suben y los que bajan: Educación y movilidad social en Colombia". *Fedesarrollo*, Bogotá, en coedición con Editorial Alfaomega. Capítulos 2, 3 y 4.
- GAVIRIA, Alejandro y BARRIENTOS, Jorge (2001a). "Determinantes de la Calidad de la educación en Colombia", *Archivos de Economía ISSN*, No. 159, 88p. Departamento Nacional de Planeación, Dirección de Estudios Económicos.

GAVIRIA, Alejandro y BARRIENTOS, Jorge (2001b). “Calidad de la Educación y Rendimiento Académico en Bogotá”, *Coyuntura Social*, Fedesarrollo, No. 24, pp. 111-126.

GAVIRIA, J.L. y CASTRO, M. (2005). “Modelos jerárquicos lineales.” Madrid: La Muralla.

HANUSHEK, E. (1986). “The Economics of Schooling Production and Efficiency in Public Schools” *Journal of Economic Literature*, No. 24, pp. 1141–1177.

HANUSHEK, E. (1996). “Measuring Investment in Education”, *Journal of Economic Perspectives*, Vol. 10, No. 4, pp 9 -30.

HANUSHEK, Eric A. y WOESSMANN Ludger (2007). “The Role of Education Quality in Economic Growth”. *Banco Mundial*. Policy Research Working Paper 4122. Washington, DC.

HOX, J.J. (1995). “Applied Multilevel Analysis.” Amsterdam: TT- Publikaties.

www.icfes.gov.co . (2011). “Guia de acceso a base de datos ICFES”

IREGUI ,Ana María y MELO Ligia y RAMOS, Jorge (2006). "Evaluación y análisis de eficiencia de la educación en Colombia," *Borradores de Economía*, No. 381, Banco de la Republica de Colombia.

KREMER, M. and E. MASKIN (1996), "Wage Inequality and Segregation by Skill", *NBER Working paper* No. 5718, August.

LORA, Eduardo, (2008). “Calidad de vida: más allá de los hechos”. *Inter-American Development Bank*. Vol 1, pp. 121-148.

LUCAS, R. (1988). “On the Mechanics of Development Planning” *Journal of Monetary Economics*, Vol 1, pp 22.

MARCELO, D. y ARIZA, N. 2005. “Evolución de los resultados de la educación en Colombia (1997-2003)” *Archivos de Economía*, No. 286. Departamento Nacional de Planeación.

MELO B, Ligia (2005). “Impacto de la Descentralización Fiscal Sobre la Educación Pública en Colombia”. *Borradores de Economía*. Banco de la República. <http://www.banrep.gov.co/docum/ftp/borra350.pdf>. 25/02/2011

MINA C, Alejandro (2004). “Factores Asociados al Logro Educativo a Nivel Municipal”. *Documento CEDE*, No. 15. Tesis de Magíster en Economía en la Universidad de los Andes.

MISAS, Gabriel (2002). “Calidad y Rendimiento Académico en Bogotá”. *Revista de Economía Institucional*. Vol.4, No. 7, pp. 167- 173. Universidad Externado de Colombia.

LEE, J-W. y BARRO, R. (2000). "Schooling Quality in a Cross-Section of Countries". *Economica*, No. 68, pp 465-488.

MURILLO, F, Javier (2008). "Los modelos multinivel como herramienta para investigación educativa". *Magis*, V.1.No 1, Julio, pp.45-62. Universidad Javeriana

www.rinace.net

ROMER, P. (1986). "Increasing Returns and Long-run Growth". *Journal of Political Economy*, pp 94, June.

PSACHAROPOULOS, G. y H. A. PATRINOS. (2004). "Returns to investment in education: a further update". *Education Economics*. No. 2 pp 111-134.

VEGAS, E. y J. PETROW. (2007). "Raising student achievement in Latin America: the challenge for the 21st century". *Latin American Development Forum*. Banco Mundial.Washington, D.C.

VÉLEZ, E. y G. PSACHAROPOULOS. (1992). "Schooling, ability and earnings in Colombia, 1988". *Economic Development and Cultural Change*. No. 3, pp. 629-643. (abril 1992)

VIVAS PACHECO, Harvy. (2005) "Magnitud y Distribución de los Spillovers de Capital Humano: Un modelo aplicado para las Áreas Metropolitanas de Colombia". *Tesina. Departamento de Economía Aplicada. Universidad Autónoma de Barcelona*.

ZAMBRANO, Juan Carlos (2012). "Análisis multinivel del rendimiento escolar en Matemáticas para cuarto grado de educación básica Primaria en Colombia." *Tesis. Programa de Maestría en Economía Aplicada. Universidad del Valle*.

ANEXOS

Tabla 1. Test de diferencia de medias.

Matemática						
Sexo	Obs	Media	Error estándar	Desviación estándar	Intervalo de confianza 95%	
F	9086	4.478.938	.113233	10.79342	4.456.742	4.501.135
M	7288	4.785.166	.1406087	12.00373	4.757.603	481.273
Total	16374	4.615.239	.0894752	1.144.933	4.597.701	4.632.777
Diferencia	-306.228	.1784469	-3.412.055	-2.712.504		
Lenguaje						
F	9086	4.947.462	.080265	7.6509	4.931.729	4.963.196
M	7288	4.946.032	.0917597	7.83351	4.928.045	496.402
Total	16374	4.946.826	.0604284	7.732.477	4.934.981	4.958.671
Diferencia	0.0143009	0.1215958	-.2240401	.2526419		

Fuente: Cálculos de la autora

Tabla 2. Porcentaje de hombres y mujeres

Sexo	%
F	55.5
M	44.5
Área	%
Urbano	90.6
Rural	9.4

Fuente: Cálculos de la autora

Tabla 3. Porcentaje por área y puntaje

Sexo	Área urbana puntaje en matemática %			Área rural puntaje en matemática %		
	Alto	Bajo	Medio	Alto	Bajo	Medio
Mujer	3.2	8.0	88.8	0.9	10.9	88.2
Hombre	5.3	5.7	88.9	3.3	7.8	88.9
Sexo	Área urbana puntaje en lenguaje %			Área rural puntaje en lenguaje %		
	Alto	Bajo	Medio	Alto	Bajo	Medio
Mujer	0.4	0.9	98.6	0.3	0.9	98.7
Hombre	0.6	1.1	98.3	0.4	1.5	98.0

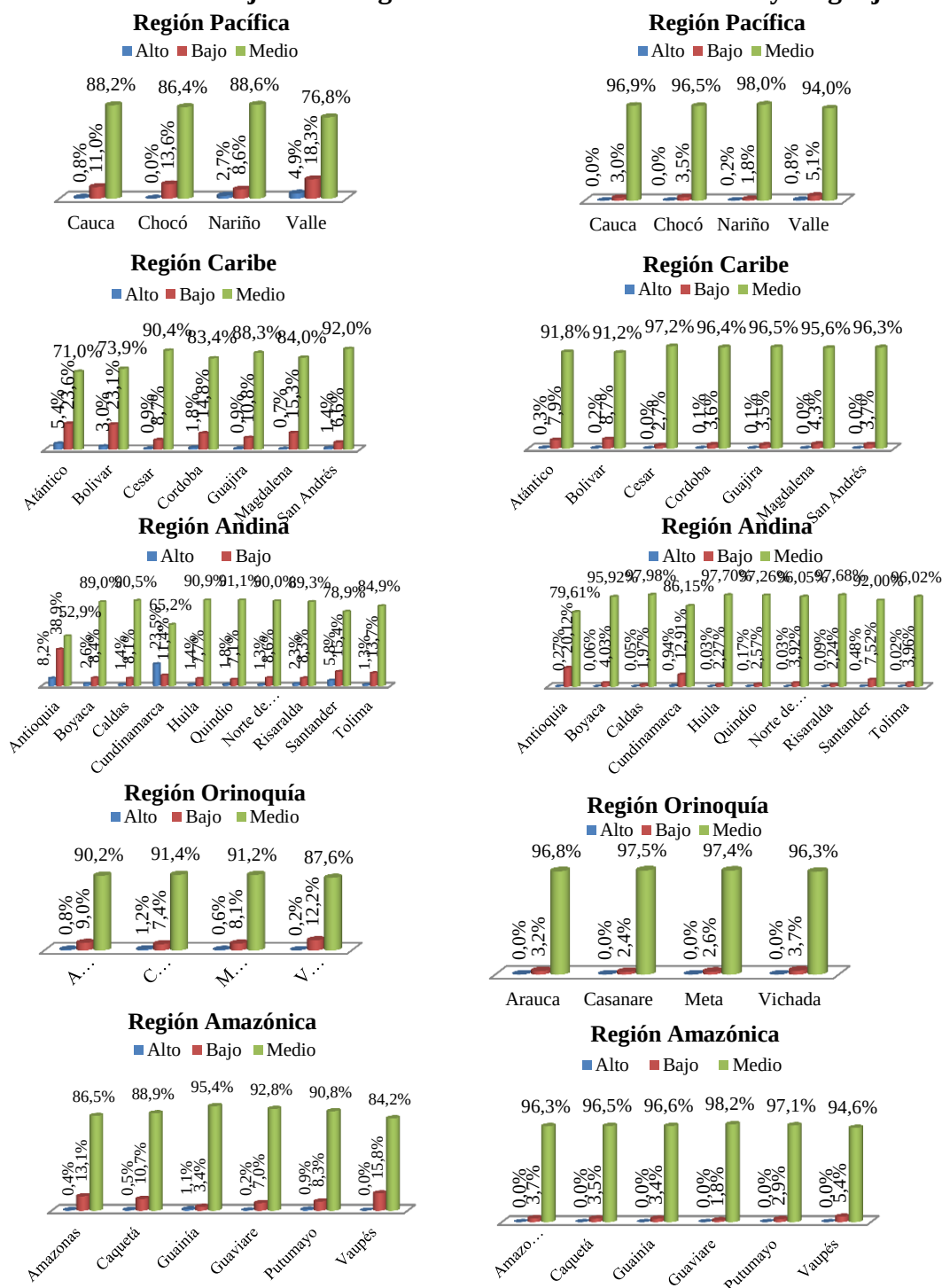
Fuente: Cálculos de la autora

Tabla 4. Descriptivas

Educación de la madre	Urbano%	Rural%	Personas del grupo Familiar	Urbano%	Rural%
			1	0.2	0.2
Ninguno	0.6	0.7	2	4.5	3.9
Primaria Incompleta	7.6	16.3	3	17.1	16.9
Primaria Completa	12.4	15.2	4	32.9	32.5
Secundaria Incompleta	17.6	20.7	5	22.6	24.9
Secundaria Completa	32.6	25.7	6	10.7	11.0
Técnica o tecnológica incompleta	1.8	2.5	7	5.2	4.6
Técnica o tecnológica Completa	6.7	4.5	8	2.9	2.38
Profesional Incompleta	2.3	2.1	9	1.4	1.4
Profesional Completa	12.7	8.0	10	1.0	0.7
Postgrado	2.9	1.3	11	0.5	0.3
No sabe	2.7	2.7	12	0.8	0.9
Ocupación del padre			¿Posee computador?		
Empresario	3.2	2.5	No	35.2	48.0
Pequeño empresario	7.4	6.6	Si	64.8	51.9
Director o gerente general	3.3	2.9	¿Posee internet?		
Empleado de nivel directivo	4.7	3.3	No	59.5	71.9
Empleado de nivel técnico / profesional	7.8	4.7	Si	40.5	28.0
Empleado auxiliar o administrativo	4.9	5.3	Naturaleza del colegio		
Obrero u operario	21.9	22.3	No Oficial	62.3	39.5
Profesional independiente	6.6	4.8	Oficial	37.7	60.5
Trabajador por cuenta propia	21.7	26.2	Calendario del colegio		
Hogar	0.8	1.2	A	0.4	
Pensionado	4.1	2.5	B	96.0	96.0
Otra actividad u ocupación	13.5	17.5	F	3.6	3.9
Ingreso familiar mensual			Categoría del colegio		
Menos de 1 SM	12.7	22.5	Alto	21.0	17.7
Entre 1 y menos de 2 SM	47.3	51.7	Bajo	17.6	30.3
Entre 2 y menos de 3 SM	18.9	13.9	Inferior	3.1	2.2
Entre 3 y menos de 5 SM	10.7	5.8	Medio	32.8	34.7
Entre 5 y menos de 7 SM	4.8	2.8	Muy Inferior	0.0	
Entre 7 y menos de 10 SM	2.4	1.8	Muy Superior	8.2	5.0
10 o más SM	3.0	1.6	Superior	17.2	10.0
Estrato socioeconómico			Índice socioeconómico del colegio		
			1	4.0	34.9
Estrato 1 & 2	51.6	71.1	2	55.5	37.9
Estrato 3 & 4	40.0	23.6	3	29.5	20.1
Estrato 5 & 6	8.4	5.3	4	11.0	7.1

Fuente: Cálculos de la autora

Gráfico 1. Puntaje en las regiones de Colombia matemática y lenguaje.



Fuente: Cálculos de la autora