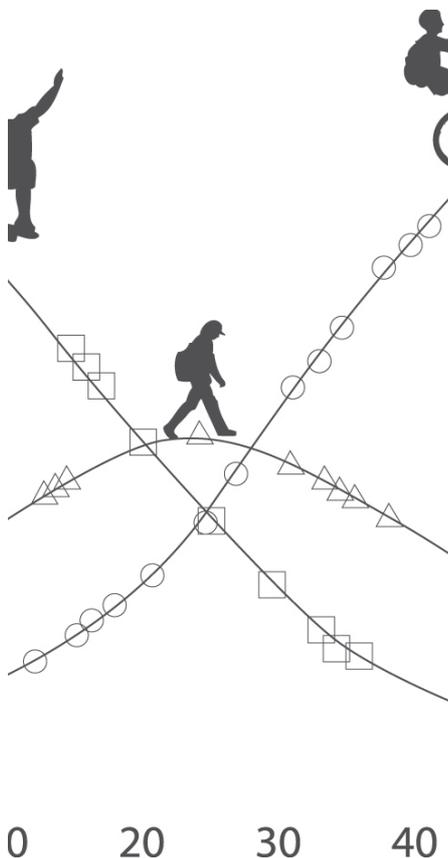


Determinantes del rendimiento académico en Colombia.

Pruebas ICFES - Saber 11°, 2009*



Sandra Milena Chica Gómez, Ing.

Ingeniera Administradora, Universidad Nacional.
Estudiante de Maestría en Economía, Universidad EAFIT.
Coordinadora de Proyecto, Secretaría de Educación
de la Alcaldía de Medellín.
smchica@gmail.com

Diana Maritza Galvis Gutiérrez, Ec.

Economista, Universidad de Antioquia.
Estudiante de Maestría en Economía, Universidad EAFIT.
Asistente de Investigación, Universidad EAFIT.
dmgalvis@gmail.com

Andrés Ramírez Hassan, Msc.

Economista, Máster en Economía y
Doctor(c) Ciencias-Estadística, Universidad Nacional.
Máster en Finanzas, Universidad EAFIT.
Docente e Investigador, Universidad EAFIT.
aramir21@eafit.edu.co

* Los autores agradecen las recomendaciones que hicieron los dos evaluadores anónimos del artículo. Sin duda, contribuyeron enormemente a mejorar la calidad del mismo.

Recepción: 03 de febrero de 2010 | Aceptación: 31 de octubre de 2010

Resumen

El trabajo que se expone está basado en los resultados que obtuvieron los estudiantes en las áreas de matemáticas y lenguaje de las pruebas ICFES Saber 11°, del segundo semestre de 2009. Mediante tales datos se pretende identificar los determinantes del rendimiento académico en Colombia. Con este fin, se utiliza el modelo *Logit* Ordenado Generalizado. Los resultados que se analizan enseñan la relevancia que tienen las variables socioeconómicas en el desempeño en ambos campos del saber. En particular, las variables *nivel de ingreso* y *nivel de escolaridad* de los progenitores. Tales variables presentan un impacto positivo y significativo en el resultado de las pruebas. Igualmente se encuentra una señal reveladora en la jornada académica: los estudiantes de jornada completa obtienen puntajes más altos en comparación con los alumnos que asisten a otras jornadas. Este resultado debe tomarse en consideración a la hora de establecer políticas de cobertura fundamentadas en la partición de la jornada y no en la creación de nuevas instituciones.

Academic performance determinants in Colombia: the ICFES tests

Abstract

This work uses the results from the mathematics and language areas achieved by students in the 2009 second semester ICFES Saber 11° test to identify the academic performance determinants in Colombia. The Generalized Ordered Logit model is used. The results show the relevance of socio-economical variables have in performance on both areas. In particular income and scholarly levels of parents have a significant positive impact in the test results. Daily school time had a significant impact on the results: students studying full-time have higher scores. This results should be considered in the establishment of new education coverage policies based on reducing the time of students in school rather than creating new institutions.

Palabras clave

Rendimiento académico
Calidad en la educación
Factores socioeconómicos
Logit Ordenado Generalizado
clasificación JEL: C35, I20, I21

Key words

Academic performance
education quality
Socio-economic factors
Generalized Ordered Logit

Introducción

Los desafíos actuales son cada vez mayores: la globalización ha llevado a borrar aquellos límites que separaban las culturas, el mundo se viene unificando y las distancias se han acortado cada vez más. Frente a estos grandes retos, la educación constituye un instrumento indispensable para que la humanidad pueda progresar hacia los ideales de paz, libertad y justicia social (Delors, 1996, 7). Con base en esta premisa, durante los últimos años se han realizado esfuerzos cada vez mayores por aumentar la cobertura educativa en Colombia; más aun, se han enfocado acciones para lograr una mayor pertinencia de la educación que imparten las diferentes instituciones.

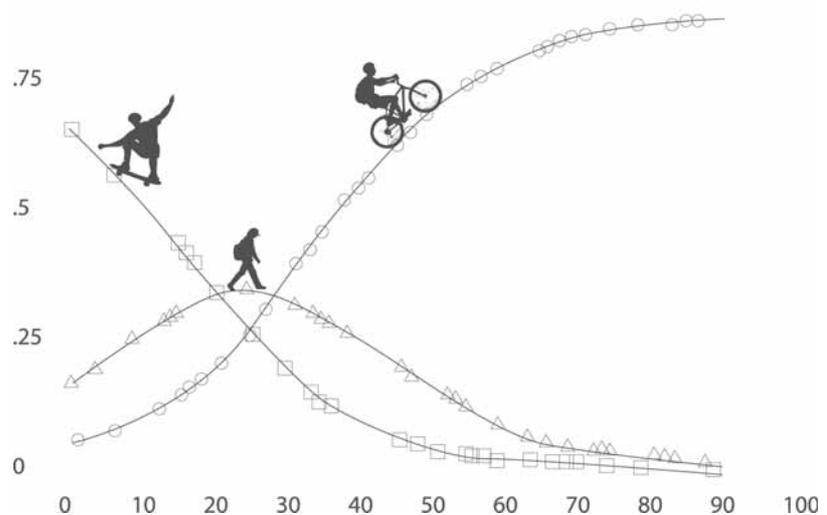
El marco actual pondera enormemente el rol del capital humano en la sociedad. Por eso es vital determinar cuáles son los factores más importantes en el desempeño académico de los jóvenes. En tal sentido, el presente trabajo tiene como principal objetivo identificar los determinantes del rendimiento académico en Colombia, mediante el análisis de los resultados de las áreas de lenguaje y matemáticas en las pruebas Saber 11° del segundo semestre de 2009, áreas que son consideradas como fundamentales en el aprendizaje.

La preocupación por conocer los determinantes del rendimiento académico de los estudiantes cobró relevancia tras la publicación del Informe Coleman (1966). En este se concluye que el rendimiento escolar en los Estados Unidos estaba influenciado en gran medida por las características socio-económicas de los estudiantes y que las variables asociadas a la institución educativa tenían poco o ningún efecto sobre las diferencias en el desempeño escolar. Los anteriores resultados suscitaron gran controversia, ya

que muchos estudiosos del tema no entendían cómo las variables asociadas al plantel educativo no tenían influencia en el rendimiento académico, lo cual motivó innumerables estudios que buscaron corroborar o refutar la hipótesis planteada por dicho informe.

De otra parte, en Colombia se han realizado varios estudios similares con el fin de determinar los factores que influyen en el rendimiento académico de los estudiantes. Se destaca el realizado por Gaviria y Barrientos (2001a y b). Allí, los autores analizan los resultados de las pruebas de Estado. Ellos encontraron que las características asociadas al plantel educativo inciden de manera significativa en el rendimiento y lo hacen en mayor medida que las variables socioeconómicas. Sin embargo, no desconocen que el nivel de educación de los padres juega un papel fundamental en el desempeño. Adicionalmente, evidencian que existe una brecha pronunciada entre los resultados para instituciones oficiales y privadas: estas últimas alcanzan mayores logros en las pruebas. Todo lo anterior deja en entredicho la aplicación de la hipótesis de Coleman (1966) para Colombia.

La relevancia del presente artículo radica en la identificación de los principales factores que determinan el rendimiento académico de los estudiantes en el país. Se constituye, entonces, en un importante instrumento de análisis para orientar



políticas educativas y sociales que busquen el mejoramiento de la calidad educativa. Asimismo, el conocer los resultados para las áreas de lenguaje y matemáticas permite explorar las fortalezas y las falencias de los estudiantes, lo cual puede servir de guía para que las instituciones educativas enfoquen sus planes de trabajo.

Los datos que se utilizaron para el análisis provienen de la información obtenida del formulario de registro de cada estudiante, el archivo de resultados y una base de datos institucional de las pruebas ICFES Saber 11° para el segundo semestre del 2009. El formulario de registro brinda variables de contexto que permiten caracterizar al estudiante en un entorno socioeconómico, el institucional proporciona información referente al plantel educativo y los resultados permiten conocer el desempeño obtenido por cada alumno en las áreas de interés. Para el estudio se consideró como variable dependiente el resultado de los jóvenes en las áreas de matemáticas y lenguaje. Los datos se agruparon en los niveles establecidos por el ICFES: bajo, medio y alto. En ese sentido, la metodología escogida para el estudio está basada en diferentes modelos donde la variable dependiente es discreta y jerarquizada.

El trabajo se expone en cinco secciones adicionales a esta. En la primera se realiza una revisión de la literatura existente en cuanto a los determinantes del rendimiento académico, en la segunda se describe la metodología utilizada para el estudio, en la tercera se presentan los datos y una breve descripción de algunas variables objeto de estudio, en la cuarta se exponen los resultados de la metodología econométrica implementada y, por último, se presentan las principales conclusiones y recomendaciones.

1. Marco Teórico

Diferentes trabajos han tratado de explicar los determinantes del rendimiento académico de acuerdo con los resultados obtenidos por los estudiantes en su proceso de aprendizaje. El

primer documento encontrado que da forma a este tipo de estudios es el Informe Coleman (1966) para Estados Unidos. En este se examina el efecto de los insumos escolares¹ sobre las diferencias en el desempeño escolar. Los resultados que publica este estudio causaron controversia en el mundo académico, pues se encontró que tal efecto era prácticamente nulo y destacó que las diferencias en los resultados escolares se debían principalmente al origen social del estudiante.

Años más tarde Jencks, et al. (1972) realizó un trabajo similar con el cual se reafirmaron los resultados expuestos por Coleman (1966); su principal conclusión afirma que variables como los recursos financieros de la escuela, sus políticas y las características de los maestros influyen de manera secundaria en la explicación del rendimiento escolar. En un caso más específico, Alexander y Simmons (1975) examinaron los determinantes del desempeño académico en países del Tercer Mundo y dedujeron que las características asociadas al entorno socioeconómico de los estudiantes eran las más relevantes, confirmando la hipótesis planteada por Coleman en 1966 y Jencks, et al. en 1972.

En la década siguiente, Delfino (1989) realizó un estudio para Argentina y se basó en las pruebas de lenguaje, matemáticas y ciencias sociales para estudiantes de grado séptimo. Allí encontró que los alumnos de las escuelas privadas y urbanas obtenían mejores calificaciones, que el nivel económico de la zona influía positivamente y que el tamaño de la escuela, contrario al tamaño de la clase, tiene un efecto positivo sobre el rendimiento.

De manera adicional, se encontró una recopilación de investigaciones sobre el tema, realizada por Wolff, Schiefelbein y Valenzuela (1993), en la cual se argumenta como conclusión principal que los insumos educativos sí ejercen una influencia en las habilidades cognitivas de los alumnos,

¹ Se refiere a variables como gasto por alumno, docentes por alumno, calidad del profesorado, disponibilidad de libros y otros recursos didácticos.

independientemente de las características del medio familiar y económico. Esta posición va en sentido contrario a los estudios ya mencionados.

Más tarde, en un trabajo realizado por Espínola y Martínez (1996) para la CEPAL se presentan las diferentes variables que influyen en el desempeño de la enseñanza en primaria. El estudio se llevó a cabo con base en un cuestionario que tenía como propósito encontrar cuál de las siguientes variables tenía importancia en el logro académico: características institucionales y organizacionales de la escuela, familia, sexo, características personales y prácticas del profesor para instruir. Los investigadores encontraron que los factores *familia y prácticas educativas del profesor* eran claramente las variables más importantes en términos de su efecto sobre el logro educativo de los alumnos de educación básica. Este resultado resalta la importancia de tener presentes las variables individuales tanto como las institucionales.

Mella y Ortiz (1999), por su parte, realizaron otro estudio para comprobar los determinantes del rendimiento académico en el sistema escolar chileno. En este trabajo utilizaron un modelo de ecuaciones estructurales y tuvieron en cuenta variables como el resultado promedio en matemática y castellano, expectativas de la madre del nivel educativo que alcanzará su hijo, escolaridad de la madre, escolaridad del jefe del hogar e ingreso mensual familiar. Los autores encontraron que la variable *ingresos familiares y nivel educativo del jefe de hogar* tienen poca significancia en el rendimiento, mientras que el nivel de escolaridad de la madre y sus expectativas tienen un efecto positivo en el desempeño académico.

De otro lado, en una investigación más reciente, Toutkoushian y Curtis (2005) desarrollaron su labor para los colegios públicos en New Hampshire (Estados Unidos) con el fin de demostrar cómo el desarrollo socioeconómico de la zona puede explicar los resultados académicos de los estudiantes; para ello recurrieron a modelos multivariados. Los resultados que obtuvieron muestran que los factores

socioeconómicos como la alimentación gratuita, la escolaridad de los padres y la tasa de empleo del distrito explican en gran medida el rendimiento académico y la propensión de los alumnos a seguir con estudios universitarios.

En cuanto a la literatura colombiana al respecto, se hallaron trabajos que han generado diversas conclusiones. Gaviria y Barrientos han sido dos de los autores que más han dedicado sus investigaciones a estudiar los determinantes del rendimiento académico en el país. Ellos (Gaviria y Barrientos, 2001a) analizaron determinantes del rendimiento académico de los estudiantes en la ciudad de Bogotá y plantean tres conclusiones principales: primero, la educación de los padres tiene un efecto sustancial sobre el rendimiento académico; segundo, el efecto de la educación de los padres se transmite principalmente a través de la calidad de los planteles educativos, y, tercero, estos inciden de manera notable sobre el rendimiento en esta ciudad.

En el mismo año, Gaviria y Barrientos (2001b) presentaron un trabajo más amplio donde se examinan los determinantes de la calidad de la educación secundaria para el caso colombiano con base en las pruebas ICFES 1999. Su análisis se centró en tres aspectos: efecto de las características familiares sobre el rendimiento académico, efecto de las características del plantel sobre la calidad y efecto del gasto público sobre la calidad relativa de los planteles públicos respecto a los privados. Los resultados principales del estudio corroboran algunas evidencias empíricas encontradas en otros países, tales como la influencia positiva del nivel educativo de los padres en el rendimiento y el mayor rendimiento académico derivado de los planteles privados después de controlar por variables de tipo socioeconómico. Además, al evaluar las características de los planteles, mostraron que la educación de los docentes, el número de profesores por alumno y la infraestructura de la institución tienen un efecto positivo sobre los resultados en las pruebas y que el pertenecer a jornada única (completa) representa también un aumento en el rendimiento.

Barrientos (2008) también estudió estos determinantes para el caso de la ciudad de Medellín a partir de las pruebas ICFES para el periodo 2004 y 2006. Por medio de una regresión por aproximación intercuartil, el autor descubrió que, para el caso de esta ciudad, el efecto del colegio parece cada vez ser menos fuerte que variables individuales. Además, encontró evidencia para afirmar que las variables relacionadas con el colegio afectan más a los estudiantes en instituciones privadas que en las públicas. Por otra parte, en su documento compara también sus resultados con los correspondientes a Bogotá y concluyen que son similares, es decir, la baja calidad de la educación pública es general.

Otro estudio en tal sentido es el elaborado por Correa (2004). Este autor estima un modelo multinivel para la ciudad de Cali —planteles y estudiantes— con el fin de incorporar características individuales y las asociadas a la institución. Tal modelo permite tener en cuenta la relación existente entre el alumno y el plantel donde estudia. El autor concluye que las variables asociadas al plantel son muy significativas a la hora de determinar el rendimiento en la ciudad de Cali.

2. Metodología

La metodología econométrica que se usa en este artículo está fundamentada, en primera instancia, en los modelos ordenados, de acuerdo con lo planteado por Cameron y Trivedi (2005). Se inicia con los modelos *Logit* Ordenados, los cuales parten de la siguiente configuración:

$$y = \begin{cases} 1 & \text{si } x\beta + \varepsilon \leq \mu_1 \\ 2 & \text{si } \mu_1 < x\beta + \varepsilon \leq \mu_2 \\ & \vdots \\ J & \text{si } \mu_{J-1} < x\beta + \varepsilon \end{cases}$$

Donde y es la variable observada que denota, para el caso que se analiza, el nivel de calificación obtenido por el individuo; por J se designa el número de niveles de calificación estipulados por

el ICFES (bajo, medio y alto); μ_j son umbrales a ser estimados; x es un vector de características de los individuos y atributos de los planteles educativos; los β son parámetros a estimar, y ε son perturbaciones estocásticas. En particular, se asume la siguiente especificación que corresponde a un modelo *Logit* Ordenado:

$$(2) \quad \Pr[y_i = \text{bajo}] = \frac{e^{\mu_1 - x_i\beta}}{1 + e^{\mu_1 - x_i\beta}}$$

$$(3) \quad \Pr[y_i = \text{medio}] = \frac{e^{\mu_2 - x_i\beta}}{1 + e^{\mu_2 - x_i\beta}} - \frac{e^{\mu_1 - x_i\beta}}{1 + e^{\mu_1 - x_i\beta}}$$

$$(4) \quad \Pr[y_i = \text{alto}] = 1 - \frac{e^{\mu_2 - x_i\beta}}{1 + e^{\mu_2 - x_i\beta}}$$

Lo anterior implica el supuesto de riesgos proporcionales o regresiones paralelas, es decir, se impone a priori la restricción que los parámetros poblacionales son constantes entre los diferentes umbrales. Esta hipótesis puede ser corroborada a partir de la prueba de Brant (1990) o utilizando una prueba de razón de verosimilitud.

Una posible alternativa al modelo *Logit* Ordenado es la estimación de un *Continuation Ratio Model*². En particular, este procedimiento es bastante interesante en el presente contexto porque el objeto de análisis en este modelo es la probabilidad de transición hacia un estado superior dado que se está en un estado igual o superior al evidenciado por el agente, es decir, se analiza $\Pr[(y_i > j | y_i \geq j)]$. Entre tanto, en el modelo *Logit* Ordenado, el análisis se concentra en $\Pr[y_i \leq j]$. Específicamente, para la aplicación expuesta en este artículo, los parámetros estimados a partir de un *Continuation Ratio Model* pueden ser utilizados para cuantificar los principales factores que determinan el avance

2 Los autores agradecen enormemente la observación elaborada por uno de los pares evaluadores con respecto al supuesto de regresiones paralelas y el mecanismo de estimación a través de un *Continuation Ratio Model*.

en términos de logros educativos de los estudiantes medidos a través de una mayor calificación.

Dicha cuantificación también se puede obtener utilizando los parámetros de un modelo *Logit* Ordenado, aunque se deben realizar algunas manipulaciones algebraicas adicionales fundamentadas en el teorema de Bayes (Xing, 2010). No obstante, la metodología basada en el *Continuation Ratio Model* también engendra el supuesto de riesgos proporcionales. Es así como, si no se cumple este supuesto, bien sea bajo un modelo *Logit* Ordenado o un *Continuation Ratio Model*, una

alternativa plausible es estimar un modelo *Logit* Ordenado Generalizado (Fu, 1998). Este no impone la restricción a priori. El modelo es equivalente a una serie de regresiones logísticas binarias donde las categorías de la variable dependiente están combinadas según el orden intrínseco que esta tiene. Dado $J=3$, la categoría uno se compara con una que reúne al resto de categorías $j=2,3$; cuando $j=2$, la comparación es entre las dos primeras categorías con la número 3.

Para este trabajo, el modelo se especifica como muestra la siguiente ecuación:

$$(5) \quad x^j \beta = \beta_{0,j} + \beta_{1,j} \text{sexo} + \beta_{2,j} \text{edad} + \sum_{i=0}^5 \beta_{3+i,j} \text{estrato} + \beta_{9,j} \text{trabaja} + \\ \beta_{10,j} \text{área vive} + \sum_{l=0}^8 \beta_{11+l,j} \text{nivel padre} + \sum_{h=0}^8 \beta_{20+h,j} \text{nivel madre} + \sum_{k=0}^3 \beta_{29+k,j} \text{nivel Sisben} + \\ \beta_{33,j} \text{hacinamiento} + \beta_{34,j} \text{tiene computador} + \beta_{35,j} \text{internet} + \beta_{36,j} \text{reproductor DVD} + \\ \beta_{37,j} \text{servicio TV} + \sum_{n=0}^5 \beta_{38+n,j} \text{ingresos mensuales} + \sum_{s=0}^3 \beta_{44+s,j} \text{jornada} + \\ \sum_{p=0}^2 \beta_{48+p,j} \text{caracter} + \sum_{q=0}^1 \beta_{51+q,j} \text{población}$$

3. Datos

Los datos que sirven como fundamento para este trabajo corresponden a la información que aparece en el sitio web del Instituto Colombiano para la Evaluación de la Educación Superior (ICFES). Allí se encuentran las bases de datos correspondientes a los resultados de las pruebas de Estado ICFES Saber desde el año 2000. Para el estudio se tiene en cuenta la prueba de Estado Saber 11° aplicada en el segundo semestre de 2009, pues esta es la base más reciente. Con tal información se construyó una base de datos general que cruza la de tres archivos: el registro de inscripción de estudiantes, el registro de las instituciones y el informe de resultados. La variable de interés para el presente estudio es el resultado de la evaluación en las principales áreas del conocimiento: matemáticas y lenguaje. Los logros se agruparon en las categorías que propone el ICFES (2009): bajo (0-30 puntos), medio (30,01 a 70) y alto (70,01 a 100).

La base de datos contiene 519.686 registros en total; sin embargo, solo se tuvieron en cuenta aquellos que presentaron información completa para todas las variables de estudio, es decir, 427.858. Cada registro contiene la información de un estudiante que presentó la prueba de Estado Saber ICFES 2009/2. Los datos se pueden agrupar en tres categorías: personales, información académica (variables propias de la institución educativa) e información familiar o del hogar (variables de contexto). Las variables que se usaron en la estimación se exponen en la tabla 1; en la última columna se señala, para el caso de los factores, aquel que se escogió como referencia.

Tabla 1: Diccionario de variables

Variable		Descripción de la variable	Variable omitida
NM		Nivel matemáticas	
NL		Nivel lenguaje	
Sexo	0	Femenino	
	1	Masculino	
Edad		Años	
Estrato	1	Estrato 1	
	2	Estrato 2	
	3	Estrato 3	
	4	Estrato 4	
	5	Estrato 5	
	6	Estrato 6	X
	8	Vive en una zona rural donde no hay estratificación socioeconómica	
Trabaja	0	No	
	1	Si	
Área donde vive	0	Área Rural	
	1	Cabecera Municipal	
Nivel padre/Nivel madre	0	Ninguno	
	1	Primaria Incompleta	
	2	Primaria Completa	
	3	Secundaria (bachillerato) Incompleta	
	4	Secundaria (bachillerato) Completa	
	5	Educación técnica o tecnológica incompleta	
	6	Educación técnica o tecnológica Completa	
	7	Educación Profesional Incompleta	
	8	Educación Profesional Completa	
	9	Postgrado	X
Nivel Sisben	1	Nivel 1	
	2	Nivel 2	
	3	Nivel 3	
	4	Está clasificada en otro nivel	
	5	No está clasificada por el Sisben	X
Hacinamiento	0	Igual o inferior a 3 personas por dormitorio	
	1	Mayor a 3 personas por dormitorio.	
Tiene computador	0	No	
	1	Si	
Internet	0	No	
	1	Si	
Reproductor DVD	0	No	
	1	Si	

Continúa en la siguiente página...

...continuación de la Tabla 1

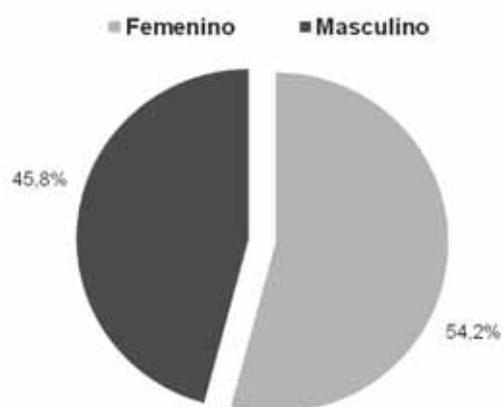
Variable		Descripción de la variable	Variable omitida
Servicio TV	0	No	
	1	Si	
Ingresos mensuales	1	Menos de 1 SM	
	2	Entre 1 y menos de 2 SM	
	3	Entre 2 y menos de 3 SM	
	4	Entre 3 y menos de 5 SM	
	5	Entre 5 y menos de 7 SM	
	6	Entre 7 y menos de 10 SM	
	7	10 o más SM	X
Jornada	1	Completa u Ordinaria	
	2	Mañana	X
	3	Noche	
	4	Sabatina-Dominical	
	5	Tarde	
Carácter	1	Académico	
	2	Técnico	
	3	Académico y Técnico	
	4	Normalista	X
Género de la población de la institución*	0	Femenino	
	1	Masculino	
	2	Mixto	X

*Esta variable será utilizada como proxy de tipo de institución.

Fuente: elaboración propia

A continuación se muestra la estadística descriptiva de las variables presentes en el análisis. En este sentido, en el gráfico 1 se puede observar que el 54,2% de los estudiantes son mujeres y en el gráfico 2, que en los estratos 1, 2 y 3 se encuentra concentrada gran parte de la muestra. Asimismo, en la tabla 2 se puede apreciar la situación laboral del estudiante que presenta la prueba. Se halló que solo el 11,3% pertenece al mercado laboral. En el gráfico 3 se expone que el 79,1% pertenece al área urbana.

Gráfico 1. Género

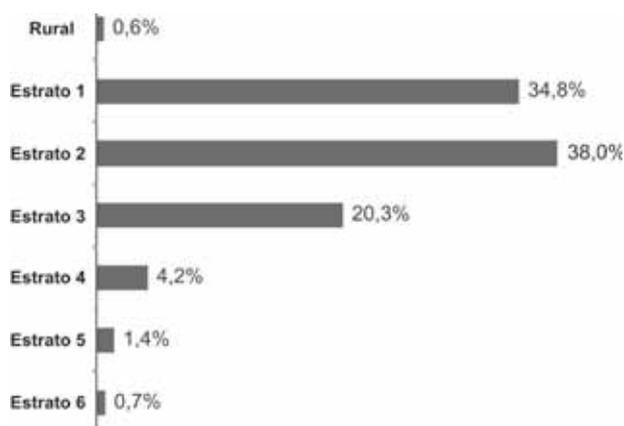


Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

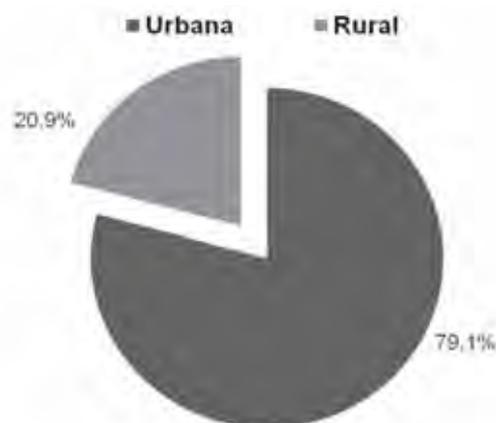
Tabla 2. Situación laboral

Trabaja			
	Frecuencia	Porcentaje	Acumulado
No	379.336	88,7	88,7
Si	48.522	11,3	100

Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

Gráfico 2: Estrato socioeconómico

Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

Gráfico 3: Área donde reside

Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

Una variable que toma principal relevancia, según los estudios revisados en la literatura, es el nivel educativo de los progenitores. Para el caso de la muestra en estudio, se puede observar en la tabla 3 que cerca del 80% de los padres poseen un nivel educativo igual o inferior a la secundaria completa y que solo alrededor del 10% de ellos tienen estudios superiores (profesional completa y postgrado).

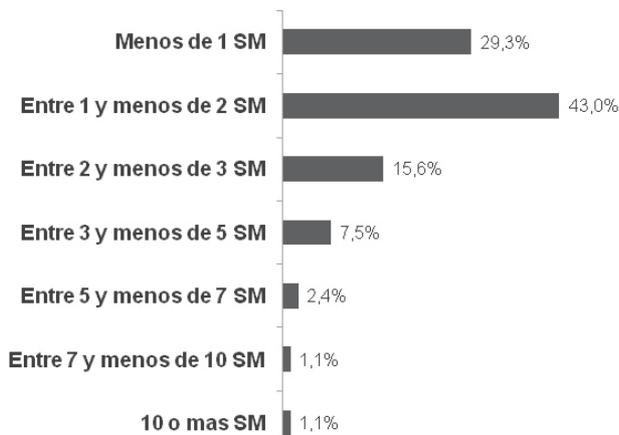
Tabla 3. Nivel de educación de los padres

Nivel de educación	Padre		Madre	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Ninguno	18.197	4,3	10.602	2,48
Primaria Incompleta	89.304	20,9	77.165	18,04
Primaria Completa	86.187	20,1	84.825	19,83
Secundaria Incompleta	67.162	15,7	78.023	18,24
Secundaria Completa	95.443	22,3	103.241	24,13
Técnico/Tecnólogo Incompleto	4.875	1,1	6.429	1,5
Técnico/Tecnólogo Completo	16.229	3,8	19.375	4,53
Profesional Incompleta	6.687	1,6	7.206	1,68
Profesional Completo	36.572	8,6	33.995	7,95
Posgrado	7.202	1,7	6.997	1,64
Total	427.858	100	427.858	100

Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009

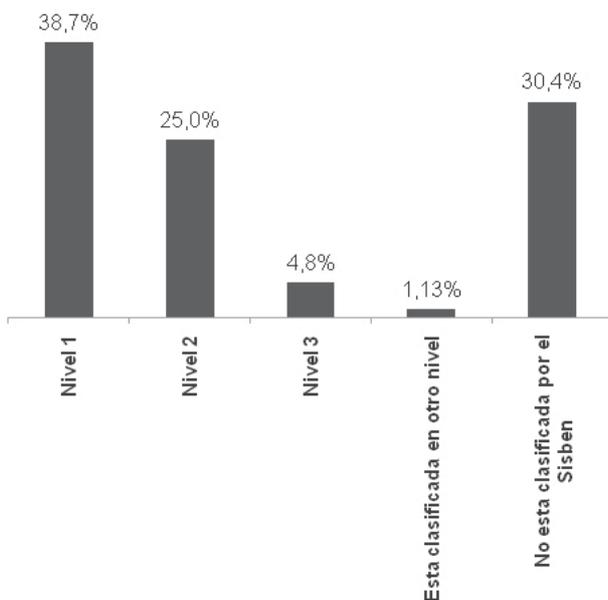
La base de datos explorada también brinda información sobre los ingresos mensuales de los hogares y si estos hacen parte del Sisben. Por tanto, corroborando las variables expuestas anteriormente, el gráfico 4 muestra que el 72,3% de los estudiantes tiene un ingreso familiar inferior a dos salarios mínimos; en el gráfico 5 se observa que el 63,7% se encuentra en los niveles 1 y 2 del Sisben.

Gráfico 4: Ingresos mensuales del hogar



Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

Gráfico 5. Nivel del Sisben



Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

Como control, se incluyeron variables sobre la posesión de computador y servicio de internet en tanto facilidades que tiene el estudiante actual para reforzar su aprendizaje. Tales variables se exponen en la tabla 4, donde el 42,3% de los alumnos posee al menos un computador y el 27,8% tiene acceso a Internet.

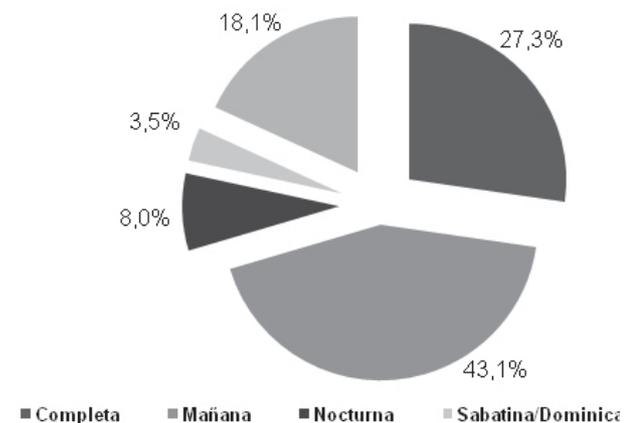
Tabla 4. Servicios

Servicios	Si (%)	No (%)
Computador	42,3	57,74
Internet	27,8	72,2

Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

De igual manera, se tuvieron en cuenta algunas variables características de la institución educativa: la *jornada* del plantel y el tipo de *población* que atiende. La primera se exhibe en el gráfico 6, donde el 43,1% de los estudiantes pertenecen a la jornada de la mañana y el 18,1% a la tarde. Ahora bien, en cuanto a la población que atiende el colegio, el gráfico 7 revela que más del 90% de las instituciones educativas colombianas son mixtas. En este trabajo no se cuenta con una variable que caracterice el tipo de institución como oficial o privada, ya que el ICFES solo estandariza para su registro dos tipologías, oficial y no oficial; esta última incluye colegios privados de licitación y contratación, por lo cual no se puede hacer un análisis directo entre las instituciones oficiales propiamente y las privadas. Por lo anterior, se recurrió a la variable *población* que atiende el colegio, ya que en Colombia las instituciones privadas pueden reservarse el derecho de concentrar el servicio exclusivamente en hombres o mujeres.

Gráfico 6. Jornada institución educativa



Fuente: Pruebas ICFES Saber 11°, segundo semestre de 2009.

Tabla 5. Resultados

Nivel	Matemáticas		Lenguaje	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Bajo	21.026	4,9	10.469	2,45
Medio	401.457	93,8	417.272	97,53
Alto	5.375	1,3	117	0,02
Total	427.858	100	427.858	100

Fuente: Pruebas ICFES Saber 11º, segundo semestre de 2009.

Finalmente, se exponen los datos correspondientes a los niveles asignados a cada área evaluada: lenguaje y matemáticas. En efecto, en la tabla 5 se puede apreciar que en matemáticas hay una mayor población que se encuentra en el nivel bajo y en nivel alto con respecto a lenguaje, concluyendo que en matemáticas hay una mayor dispersión de los datos. Por lo demás, se resalta que en lenguaje la cantidad de estudiantes pertenecientes al nivel alto es despreciable, debido al número de datos que se utilizaron en este estudio.

4. Resultados

En primera instancia se procedió a trabajar los modelos *Logit Ordenado* y *Continuation Ratio*, tanto para los resultados obtenidos en matemáticas como en lenguaje. Luego se realizaron la prueba de razón de verosimilitud y la prueba de Brant; en ambas se encontró evidencia estadística en contra de la hipótesis de riesgos proporcionales.

A continuación se procede a enseñar los resultados obtenidos a partir del Modelo *Logit Generalizado*³.

4.1 Matemáticas

Para analizar el modelo *Logit Ordenado Generalizado* que se expone en la tabla 6 se debe tener en cuenta que en la segunda columna se compara

el nivel bajo en la calificación con respecto a los niveles medio y alto; por su parte, en la tercera columna se comparan los niveles bajo y medio con el nivel alto en la calificación. Además, las últimas tres columnas enseñan los efectos marginales asociados a cada nivel.

La interpretación de las dos primeras columnas se realiza de la siguiente manera: los coeficientes positivos significan que, ante valores más altos de la variable explicativa, es mayor la probabilidad de ubicarse en una categoría superior a la actual; por tanto, los coeficientes negativos muestran que, ante aumentos de la variable explicativa, la probabilidad de estar en la categoría actual o en una más baja es mayor.

En particular, se observa que los hombres tienen mayor probabilidad de obtener calificaciones altas en las pruebas de matemáticas. Por otra parte, más edad implica una mayor probabilidad de ubicarse en el nivel bajo. Al respecto de la categorización socioeconómica, se observa que en general los bachilleres de estratos uno a cinco y los ubicados en el área rural, evidencian, en términos estadísticos, una mayor probabilidad de ubicarse en el rango medio de la evaluación comparándolos con los logros educativos de los estudiantes del estrato seis, los cuales presentan mayor probabilidad de ubicarse en el nivel alto.

De otra parte, se observa que los bachilleres que trabajan y que están ubicados por fuera de la cabecera municipal tienen una mayor probabilidad de encontrarse en el rango inferior de la prueba,

³ Los resultados de estos ejercicios no se publican en el artículo por limitaciones de espacio, pero están disponibles para los lectores interesados mediante comunicación con los autores o la revista.

comparados con los alumnos que no presentan estas características. Tal vez uno de los factores más relevantes en el desempeño académico de los bachilleres es la educación de los progenitores. Como se puede evidenciar en la tabla 6, los estudiantes que son hijos de personas con un bajo nivel de formación académica evidencian, sistemáticamente, mayor probabilidad de ubicarse en el nivel bajo de la prueba. Específicamente, la evidencia estadística indica que los hijos de personas con posgrado evidencian una mayor probabilidad de ubicarse en el nivel alto de la prueba.

Este fenómeno también se demuestra con la variable *ingreso*, puesto que para los estudiantes que provienen de familias con bajas entradas se observa una mayor probabilidad de encontrarse en el nivel bajo de la prueba. Al respecto de los hogares que pertenecen a la encuesta Sisben, se halló que estos presentan mayor probabilidad de ubicarse en el nivel bajo de la prueba cuando son comparados con bachilleres provenientes de

familias no encuestadas. De acuerdo con otras variables, los bachilleres que poseen computador tienen mayor probabilidad de obtener una calificación alta, aunque la tenencia de internet no resultó estadísticamente significativa.

Cuando se analizan variables propias de la institución se observa que los bachilleres que toman clases en horas de la tarde, noche y fines de semana evidencian una mayor probabilidad de ubicarse en los niveles bajos de la prueba, mientras que los estudiantes de jornada completa exhiben mayor probabilidad de tener un nivel alto de la calificación. Esto es un resultado fundamental, puesto que se está indicando que aumentar el mecanismo de cobertura a partir de la división de la jornada académica implica presumiblemente una reducción en los logros académicos de los estudiantes. Igualmente, las instituciones cuya población estudiantil es solo masculina o femenina evidencian una mayor probabilidad de obtener resultados superiores comparados con las restantes instituciones.

Tabla 6. Estimación modelo *Logit* Ordenado Generalizado

Variable dependiente: Nivel matemáticas					
Variables	Modelo <i>Logit</i> Ordenado Generalizado Bajo	Modelo <i>Logit</i> Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Sexo	0,205***	1,066***	-0,00832***	0,00462***	0,00370***
	(0,0148)	(0,0359)	(0,000599)	(0,000622)	(0,000174)
Edad	-0,0219***	-0,287***	0,000896***	7,84e-06	-0,000904***
	(0,00150)	(0,0164)	(6,17e-05)	(7,79e-05)	(4,80e-05)
Estrato_1	0,242	-1,006***	-0,00964*	0,0125**	-0,00288***
	(0,148)	(0,114)	(0,00573)	(0,00573)	(0,000314)
Estrato_2	0,401***	-0,508***	-0,0158***	0,0173***	-0,00152***
	(0,147)	(0,0901)	(0,00560)	(0,00560)	(0,000264)
Estrato_3	0,428***	-0,312***	-0,0157***	0,0166***	-0,000899***
	(0,147)	(0,0831)	(0,00479)	(0,00479)	(0,000220)
Estrato_4	0,504***	-0,0448	-0,0168***	0,0169***	-0,000139
	(0,153)	(0,0812)	(0,00408)	(0,00408)	(0,000246)
Estrato_5	0,647***	-0,0327	-0,0201***	0,0202***	-0,000102
	(0,182)	(0,0886)	(0,00414)	(0,00414)	(0,000271)

Continúa en la siguiente página...

...continuación de la Tabla 6

Variable dependiente: Nivel matemáticas					
Variables	Modelo Logit Ordenado Generalizado Bajo	Modelo Logit Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Estrato_8	-0,102	-1,036**	0,00439	-0,00234	-0,00205***
	(0,164)	(0,461)	(0,00735)	(0,00736)	(0,000530)
Trabaja	-0,123***	-0,198***	0,00526***	-0,00468***	-0,000581***
	(0,0212)	(0,0743)	(0,000949)	(0,000968)	(0,000202)
Área vive	0,107***	0,481***	-0,00450***	0,00316***	0,00134***
	(0,0176)	(0,0572)	(0,000762)	(0,000775)	(0,000145)
Nivel padre 0	-0,629***	-1,137***	0,0338***	-0,0315***	-0,00225***
	(0,131)	(0,182)	(0,00895)	(0,00895)	(0,000223)
Nivel padre 1	-0,658***	-0,806***	0,0325***	-0,0304***	-0,00207***
	(0,128)	(0,0910)	(0,00751)	(0,00751)	(0,000205)
Nivel padre 2	-0,663***	-0,818***	0,0329***	-0,0309***	-0,00208***
	(0,127)	(0,0835)	(0,00759)	(0,00759)	(0,000188)
Nivel padre 3	-0,639***	-0,705***	0,0323***	-0,0305***	-0,00178***
	(0,127)	(0,0746)	(0,00783)	(0,00783)	(0,000166)
Nivel padre 4	-0,635***	-0,733***	0,0309***	-0,0290***	-0,00193***
	(0,127)	(0,0625)	(0,00726)	(0,00726)	(0,000156)
Nivel padre 5	-0,440***	-0,446***	0,0221**	-0,0209**	-0,00114***
	(0,150)	(0,119)	(0,00906)	(0,00906)	(0,000248)
Nivel padre 6	-0,308**	-0,440***	0,0144**	-0,0133*	-0,00114***
	(0,136)	(0,0724)	(0,00718)	(0,00718)	(0,000159)
Nivel padre 7	-0,454***	-0,263***	0,0228**	-0,0221**	-0,000732***
	(0,148)	(0,0821)	(0,00896)	(0,00895)	(0,000203)
Nivel padre 8	-0,503***	-0,229***	0,0251***	-0,0244***	-0,000658***
	(0,128)	(0,0531)	(0,00765)	(0,00765)	(0,000141)
Nivel madre 0	-0,715***	-1,020***	0,0403***	-0,0382***	-0,00207***
	(0,123)	(0,273)	(0,00919)	(0,00920)	(0,000337)
Nivel madre 1	-0,609***	-0,891***	0,0301***	-0,0279***	-0,00219***
	(0,118)	(0,105)	(0,00694)	(0,00694)	(0,000212)
Nivel madre 2	-0,609***	-0,668***	0,0298***	-0,0281***	-0,00176***
	(0,117)	(0,0877)	(0,00679)	(0,00679)	(0,000203)
Nivel madre 3	-0,595***	-0,724***	0,0293***	-0,0274***	-0,00186***
	(0,117)	(0,0766)	(0,00683)	(0,00683)	(0,000174)
Nivel madre 4	-0,496***	-0,544***	0,0230***	-0,0215***	-0,00151***
	(0,116)	(0,0624)	(0,00610)	(0,00610)	(0,000164)
Nivel madre 5	-0,509***	-0,236**	0,0263***	-0,0256***	-0,000667**
	(0,133)	(0,108)	(0,00844)	(0,00844)	(0,000272)
Nivel madre 6	-0,281**	-0,305***	0,0130**	-0,0121*	-0,000842***
	(0,124)	(0,0693)	(0,00639)	(0,00639)	(0,000170)

Continúa en la siguiente página...

...continuación de la Tabla 6

Variable dependiente: Nivel matemáticas					
Variables	Modelo Logit Ordenado Generalizado Bajo	Modelo Logit Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Nivel madre 7	-0,230*	-0,190**	0,0104	-0,00988	-0,000547**
	(0,139)	(0,0813)	(0,00696)	(0,00696)	(0,000215)
Nivel madre 8	-0,293**	-0,200***	0,0135**	-0,0129**	-0,000581***
	(0,120)	(0,0556)	(0,00613)	(0,00613)	(0,000150)
Nivel Sisben 1	-0,0916***	-0,353***	0,00379***	-0,00270**	-0,00108***
	(0,0269)	(0,0723)	(0,00112)	(0,00114)	(0,000218)
Nivel Sisben 2	-0,0606**	-0,384***	0,00252**	-0,00141	-0,00111***
	(0,0252)	(0,0526)	(0,00106)	(0,00107)	(0,000144)
Nivel Sisben 3	-0,0803**	-0,307***	0,00340*	-0,00256	-0,000847***
	(0,0404)	(0,0669)	(0,00177)	(0,00177)	(0,000164)
Nivel Sisben 4	-0,178**	-0,265**	0,00791**	-0,00717**	-0,000737**
	(0,0724)	(0,120)	(0,00348)	(0,00348)	(0,000294)
Hacinamiento	-0,0930***	-0,459***	0,00395***	-0,00275**	-0,00120***
	(0,0239)	(0,112)	(0,00105)	(0,00108)	(0,000244)
Tiene computador	0,201***	0,450***	-0,00811***	0,00663***	0,00149***
	(0,0231)	(0,0527)	(0,000919)	(0,000936)	(0,000189)
Reproductor DVD	-0,00181	-0,261***	7,40e-05	0,000779	-0,000853***
	(0,0157)	(0,0382)	(0,000644)	(0,000656)	(0,000133)
Servicio TV	0,0109	-0,182***	-0,000445	0,00103	-0,000582***
	(0,0162)	(0,0396)	(0,000664)	(0,000675)	(0,000130)
Internet	0,0190	-0,0140	-0,000777	0,000821	-4,41e-05
	(0,0273)	(0,0434)	(0,00111)	(0,00112)	(0,000136)
Ingresos mensuales 1	-0,418***	-1,241***	0,0186***	-0,0153**	-0,00326***
	(0,149)	(0,0965)	(0,00719)	(0,00719)	(0,000246)
Ingresos mensuales 2	-0,398***	-0,977***	0,0168***	-0,0138**	-0,00299***
	(0,148)	(0,0741)	(0,00648)	(0,00648)	(0,000254)
Ingresos mensuales 3	-0,353**	-0,803***	0,0163**	-0,0143*	-0,00197***
	(0,148)	(0,0691)	(0,00762)	(0,00762)	(0,000150)
Ingresos mensuales 4	-0,216	-0,559***	0,00964	-0,00823	-0,00141***
	(0,149)	(0,0645)	(0,00724)	(0,00724)	(0,000138)
Ingresos mensuales 5	0,0446	-0,356***	-0,00179	0,00275	-0,000954***
	(0,163)	(0,0676)	(0,00644)	(0,00643)	(0,000156)
Ingresos mensuales 6	0,270	-0,0419	-0,00983	0,00996	-0,000130
	(0,200)	(0,0714)	(0,00643)	(0,00642)	(0,000216)
Jornada completa	0,0303	0,269***	-0,00123	0,000332	0,000903***
	(0,0189)	(0,0307)	(0,000767)	(0,000774)	(0,000114)
Jornada nocturna	-0,562***	-1,917***	0,0288***	-0,0256***	-0,00313***
	(0,0246)	(0,232)	(0,00154)	(0,00155)	(0,000174)

Continúa en la siguiente página...

...continuación de la Tabla 6

Variable dependiente: Nivel matemáticas					
Variables	Modelo Logit Ordenado Generalizado Bajo	Modelo Logit Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Jornada sabatina/ dominical	-0,514***	-1,636***	0,0264***	-0,0237***	-0,00269***
	(0,0327)	(0,356)	(0,00206)	(0,00207)	(0,000258)
Jornada tarde	-0,0328	-0,372***	0,00136	-0,000310	-0,00105***
	(0,0210)	(0,0526)	(0,000876)	(0,000885)	(0,000136)
Carácter académico	-0,119**	0,327***	0,00482**	-0,00582***	0,000995***
	(0,0531)	(0,105)	(0,00212)	(0,00214)	(0,000311)
Carácter técnico	-0,107*	0,302***	0,00453*	-0,00559**	0,00106**
	(0,0549)	(0,111)	(0,00240)	(0,00243)	(0,000429)
Carácter académico y técnico	-0,121**	0,195*	0,00516**	-0,00582**	0,000655*
	(0,0549)	(0,111)	(0,00242)	(0,00245)	(0,000396)
Población femenina	0,411***	0,446***	-0,0143***	0,0126***	0,00173***
	(0,0478)	(0,0520)	(0,00140)	(0,00141)	(0,000246)
Población masculina	0,355***	0,407***	-0,0125***	0,0109***	0,00157***
	(0,100)	(0,0521)	(0,00299)	(0,00299)	(0,000252)
Constante	4,598***	1,730***			
	(0,222)	(0,323)			
Observaciones	427.858	427.858	427.858	427.858	427.858

Errores estándar en paréntesis *** p<0,01, ** p<0,05, * p<0,1

Fuente: elaboración propia.

El análisis anterior permite conocer la tendencia de cómo las variables explicativas afectan el desempeño en el área de las matemáticas; sin embargo, es importante cuantificar estos efectos y conocer cuáles son los más representativos. Esto se realiza a través de la lectura de los efectos marginales expuestos en las últimas tres columnas de la tabla 6. Se debe tener en cuenta que la interpretación de las diferentes columnas no es la misma para sus indicadores. Así, en la primera columna se muestran los efectos de las variables explicativas sobre la probabilidad de permanecer en el nivel bajo, por lo que signos positivos reflejan un efecto de permanencia y signos negativos uno de

cambio hacia las demás. En las columnas restantes, el signo negativo representa la probabilidad de pasar a otras categorías más bajas.

Como ya se mencionó, una de las variables que más afecta el resultado es el nivel de educación de los padres; por ejemplo, un estudiante con un padre que no tiene ningún nivel de educación tiene una probabilidad de 3,38% mayor de permanecer en el nivel bajo con respecto al que tiene padre con posgrado. Se observa además que la magnitud de la probabilidad de permanecer en el nivel bajo disminuye a medida que aumenta la escolaridad del padre. Ocurre lo mismo para los niveles

medio y alto, donde el signo negativo expresa la probabilidad de estar en un nivel inferior (la magnitud con signo negativo disminuye cuando el padre es más educado). Este comportamiento es similar para el nivel de escolaridad de la madre. Al mismo tiempo, si se comparan las magnitudes de las probabilidades asociadas al nivel de escolaridad del padre versus el de la madre, se obtiene que la educación del padre tiene un mayor impacto en los resultados de la prueba, a excepción de cuando los padres no tienen ningún nivel de escolaridad.

Otra variable que presenta un impacto representativo es el nivel de ingreso mensual. A manera de ilustración, se aprecia que los estudiantes en cuyo hogar se percibe un ingreso mensual inferior a un salario mínimo tiene una probabilidad mayor (de 1,86%) de permanecer en un nivel bajo con respecto a los estudiantes en donde sus hogares cuentan con más de 10 salarios mínimos. En la cuarta columna se señala que la probabilidad que un estudiante con un nivel de ingresos más bajo pase del nivel medio al bajo es mayor en 1,53%, si se compara con un bachiller perteneciente al nivel de ingresos más alto, mientras la probabilidad asociada de pasar del nivel alto a un nivel inferior es tan solo de 0,33% más, lo cual puede explicarse por la poca cantidad de estudiantes que aparecen en nivel alto y pertenecen al grupo de ingresos 1.

La variable *jornada* también presenta resultados interesantes para las categorías nocturna y sabatina-dominical. La probabilidad de permanencia en el nivel bajo de un alumno que estudia en jornada nocturna es de 2,88% mayor que la de un estudiante que asiste a la jornada de la mañana;

la de un bachiller perteneciente a jornada sabatina y dominical es de 2,6% más con respecto a este mismo estudiante. En otras palabras una persona de cualquiera de estas dos jornadas tiende a tener desempeños inferiores en la prueba comparados con los estudiantes de la jornada de la mañana.

Asimismo, la variable relacionada con la *población* que asiste a la institución educativa evidencia que el hecho de estudiar en un establecimiento con población exclusivamente femenina o masculina disminuye la probabilidad de pertenecer a un nivel bajo (1,43% y 1,25%, respectivamente), con respecto a la población mixta.

4.2 Lenguaje

En la tabla 7 se exponen los resultados de la estimación a través de la metodología *Logit* Ordenado Generalizado. A nivel general, la gran diferencia entre las áreas de lenguaje y matemáticas es que no se identificó una característica que lleve a los estudiantes a obtener un nivel de puntaje alto, lo cual se determina al observar que los efectos marginales asociados al nivel bajo y al nivel medio son iguales en magnitud pero con signo contrario, mientras que los efectos marginales del nivel alto no resultaron estadísticamente significativos. Esto se explica por el bajo porcentaje de estudiantes que se encuentra en dicho rango (0,02%, según la tabla 5). Al respecto de las variables explicativas, las mujeres obtienen mejores resultados en lenguaje, mientras que los hombres obtienen mejores resultados en matemáticas. Las demás variables se comportan de manera similar que para el área de matemáticas.

Tabla 7. Estimación modelo *Logit* Ordenado Generalizado

Variable dependiente: Nivel lenguaje					
Variables	Modelo <i>Logit</i> Ordenado Generalizado Bajo	Modelo <i>Logit</i> Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Sexo	-0,192***	-0,355*	0,00373***	-0,00373***	-6,76e-08
	(0,0204)	(0,213)	(0,000400)	(0,000400)	(1,17e-05)
Edad	-0,0221***	-0,257**	0,000425***	-0,000425***	-4,96e-08
	(0,00188)	(0,112)	(3,63e-05)	(3,70e-05)	(8,60e-06)
Estrato_1	0,456**	-2,394***	-0,00830**	0,00830**	-4,16e-07
	(0,191)	(0,663)	(0,00333)	(0,00333)	(7,20e-05)
Estrato_2	0,657***	-1,881***	-0,0119***	0,0119***	-3,34e-07
	(0,190)	(0,449)	(0,00329)	(0,00329)	(5,78e-05)
Estrato_3	0,682***	-2,046***	-0,0109***	0,0109***	-2,51e-07
	(0,190)	(0,384)	(0,00254)	(0,00254)	(4,35e-05)
Estrato_4	0,619***	-1,084***	-0,00917***	0,00917***	-1,33e-07
	(0,198)	(0,329)	(0,00221)	(0,00221)	(2,31e-05)
Estrato_5	0,640***	-0,973***	-0,00925***	0,00925***	-1,21e-07
	(0,235)	(0,377)	(0,00246)	(0,00246)	(2,10e-05)
Estrato_8	0,0922	-16,29	-0,00170	0,00170	-2,11e-07
	(0,214)	(4,281)	(0,00376)	(0,00376)	(3,63e-05)
Trabaja	-0,302***	-14,57	0,00652***	-0,00652***	-1,01e-06
	(0,0270)	(731,2)	(0,000652)	(0,000663)	(0,000153)
Área vive	0,116***	1,000*	-0,00230***	0,00230***	1,51e-07
	(0,0245)	(0,516)	(0,000502)	(0,000502)	(2,61e-05)
Nivel padre 0	-0,950***	-14,85	0,0285***	-0,0285***	-3,63e-07
	(0,213)	(1,425)	(0,00930)	(0,00930)	(5,88e-05)
Nivel padre 1	-0,967***	-0,805	0,0251***	-0,0251***	-1,26e-07
	(0,209)	(0,693)	(0,00719)	(0,00719)	(2,19e-05)
Nivel padre 2	-0,994***	0,101	0,0263***	-0,0263***	2,00e-08
	(0,209)	(0,469)	(0,00739)	(0,00739)	(3,47e-06)
Nivel padre 3	-0,932***	-0,929*	0,0250***	-0,0250***	-1,35e-07
	(0,209)	(0,527)	(0,00760)	(0,00760)	(2,34e-05)
Nivel padre 4	-0,982***	-0,802**	0,0253***	-0,0253***	-1,27e-07
	(0,208)	(0,384)	(0,00705)	(0,00705)	(2,20e-05)
Nivel padre 5	-0,679***	-0,286	0,0182**	-0,0182**	-4,81e-08
	(0,240)	(0,770)	(0,00863)	(0,00863)	(8,33e-06)
Nivel padre 6	-0,674***	-0,185	0,0177**	-0,0177**	-3,27e-08
	(0,219)	(0,418)	(0,00762)	(0,00762)	(5,67e-06)
Nivel padre 7	-0,756***	-0,712	0,0210**	-0,0210**	-9,93e-08
	(0,236)	(0,564)	(0,00904)	(0,00904)	(1,72e-05)
Nivel padre 8	-0,728***	-0,538*	0,0190***	-0,0190***	-8,40e-08
	(0,210)	(0,289)	(0,00722)	(0,00722)	(1,46e-05)

Continúa en la siguiente página...

...continuación de la Tabla 7

Variable dependiente: Nivel lenguaje					
Variables	Modelo Logit Ordenado Generalizado Bajo	Modelo Logit Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Nivel madre 0	-0,793***	-15,12	0,0223***	-0,0223***	-2,80e-07
	(0,191)	(1,603)	(0,00750)	(0,00750)	(4,73e-05)
Nivel madre 1	-0,738***	-3,152***	0,0181***	-0,0181***	-3,26e-07
	(0,185)	(1,099)	(0,00570)	(0,00570)	(5,65e-05)
Nivel madre 2	-0,775***	-3,816***	0,0190***	-0,0190***	-4,02e-07
	(0,184)	(1,079)	(0,00570)	(0,00570)	(6,96e-05)
Nivel madre 3	-0,700***	-1,834***	0,0169***	-0,0169***	-2,26e-07
	(0,184)	(0,513)	(0,00553)	(0,00553)	(3,92e-05)
Nivel madre 4	-0,650***	-1,367***	0,0149***	-0,0149***	-2,00e-07
	(0,183)	(0,370)	(0,00499)	(0,00499)	(3,46e-05)
Nivel madre 5	-0,758***	-1,111	0,0211***	-0,0211***	-1,32e-07
	(0,202)	(0,768)	(0,00778)	(0,00778)	(2,28e-05)
Nivel madre 6	-0,338*	-1,214***	0,00757	-0,00757	-1,43e-07
	(0,194)	(0,444)	(0,00501)	(0,00501)	(2,48e-05)
Nivel madre 7	-0,313	-0,347	0,00697	-0,00697	-5,69e-08
	(0,216)	(0,433)	(0,00554)	(0,00554)	(9,85e-06)
Nivel madre 8	-0,440**	-0,608**	0,0102**	-0,0102**	-9,22e-08
	(0,188)	(0,281)	(0,00515)	(0,00515)	(1,60e-05)
Nivel Sisben 1	-0,150***	0,241	0,00292***	-0,00292***	4,78e-08
	(0,0381)	(0,496)	(0,000756)	(0,000756)	(8,28e-06)
Nivel Sisben 2	-0,130***	-0,0121	0,00259***	-0,00259***	-2,32e-09
	(0,0360)	(0,373)	(0,000737)	(0,000737)	(4,09e-07)
Nivel Sisben 3	-0,120**	-0,510	0,00243**	-0,00243**	-7,89e-08
	(0,0575)	(0,603)	(0,00122)	(0,00122)	(1,37e-05)
Nivel Sisben 4	-0,106	-15,22	0,00215	-0,00215	-2,28e-07
	(0,107)	(1,441)	(0,00227)	(0,00227)	(3,93e-05)
Hacina- miento	-0,00768	-14,98	0,000148	-0,000148	-6,21e-07
	(0,0344)	(1,158)	(0,000664)	(0,000671)	(9,18e-05)
Tiene computador	0,269***	0,798**	-0,00508***	0,00508***	1,69e-07
	(0,0332)	(0,366)	(0,000613)	(0,000613)	(2,93e-05)
Reproductor DVD	0,000163	-0,168	-3,13e-06	3,16e-06	-3,31e-08
	(0,0220)	(0,272)	(0,000422)	(0,000422)	(5,73e-06)
Servicio TV	-0,00902	-0,320	0,000173	-0,000173	-6,35e-08
	(0,0226)	(0,273)	(0,000434)	(0,000434)	(1,10e-05)
Internet	-0,0148	-0,729**	0,000285	-0,000284	-1,22e-07
	(0,0392)	(0,286)	(0,000759)	(0,000759)	(2,11e-05)
Ingresos mensuales 1	0,0296	-1,256*	-0,000566	0,000566	-2,01e-07
	(0,187)	(0,702)	(0,00355)	(0,00355)	(3,48e-05)

Continúa en la siguiente página...

...continuación de la Tabla 7

Variable dependiente: Nivel lenguaje					
Variables	Modelo Logit Ordenado Generalizado Bajo	Modelo Logit Ordenado Generalizado Medio	Efectos marginales nivel bajo	Efectos marginales nivel medio	Efectos marginales nivel alto
Ingresos mensuales 2	-0,00201	-0,795*	3,87e-05	-3,86e-05	-1,49e-07
	(0,186)	(0,477)	(0,00357)	(0,00357)	(2,58e-05)
Ingresos mensuales 3	-0,00203	-0,358	3,90e-05	-3,90e-05	-6,14e-08
	(0,186)	(0,418)	(0,00358)	(0,00358)	(1,06e-05)
Ingresos mensuales 4	0,0758	-0,502	-0,00141	0,00141	-7,90e-08
	(0,187)	(0,388)	(0,00339)	(0,00339)	(1,37e-05)
Ingresos mensuales 5	0,356*	0,547	-0,00584**	0,00584**	1,39e-07
	(0,211)	(0,334)	(0,00292)	(0,00292)	(2,40e-05)
Ingresos mensuales 6	0,522**	-0,342	-0,00792**	0,00792**	-5,61e-08
	(0,266)	(0,438)	(0,00312)	(0,00312)	(9,72e-06)
Jornada completa	0,00883	0,287	-0,000169	0,000169	5,93e-08
	(0,0277)	(0,203)	(0,000530)	(0,000530)	(1,03e-05)
Jornada nocturna	-0,846***	-14,46	0,0234***	-0,0234***	-5,94e-07
	(0,0319)	(998,4)	(0,00121)	(0,00121)	(9,20e-05)
Jornada sabatina/ dominical	-0,827***	-14,11	0,0235***	-0,0235***	-3,12e-07
	(0,0410)	(1,564)	(0,00164)	(0,00164)	(5,14e-05)
Jornada tarde	-0,0196	-0,329	0,000378	-0,000378	-5,73e-08
	(0,0309)	(0,391)	(0,000601)	(0,000601)	(9,92e-06)
Carácter académico	-0,354***	1,338	0,00654***	-0,00654***	2,34e-07
	(0,0885)	(1,012)	(0,00158)	(0,00158)	(4,06e-05)
Carácter técnico	-0,271***	0,927	0,00568***	-0,00569***	2,51e-07
	(0,0909)	(1,057)	(0,00208)	(0,00208)	(4,36e-05)
Carácter académico y técnico	-0,361***	1,064	0,00780***	-0,00780***	3,03e-07
	(0,0907)	(1,044)	(0,00219)	(0,00219)	(5,25e-05)
Población femenina	0,712***	-0,256	-0,0103***	0,0103***	-4,41e-08
	(0,0881)	(0,278)	(0,000913)	(0,000913)	(7,64e-06)
Población masculina	0,278**	0,273	-0,00470**	0,00470**	6,03e-08
	(0,128)	(0,394)	(0,00190)	(0,00190)	(1,04e-05)
Constante	5,689***	-1,377			
	(0,319)	(2,299)			
Observaciones	427.858	427.858	427.858	427.858	427.858

Errores estándar en paréntesis *** p<0,01, ** p<0,05, * p<0,1

Fuente: elaboración propia.

Al respecto de los efectos marginales, se destaca que al igual que en el área de matemáticas, la variable asociada al nivel de escolaridad de los padres tiene un efecto significativo, tanto estadísticamente como

en magnitud, sobre el resultado en el área. Por ejemplo, un estudiante con un padre que no tiene ningún nivel de educación tiene una probabilidad mayor (2,85%) de permanecer en el nivel bajo con

respecto al que tiene padre con posgrado. En este caso, también ocurre que la magnitud del impacto es superior en el nivel del padre que en el de la madre; el impacto para esta última variable es de 2,23% para el caso mencionado.

La variable *jornada* también presenta efectos con magnitudes significativas en cuanto al área de lenguaje. Cuando la jornada es nocturna o sabatina y dominical hay una probabilidad mayor (2,34% y 2,35%, respectivamente) de permanecer en el nivel bajo en comparación con la jornada de la mañana. Para la variable *población* atendida por la institución, se obtienen resultados similares a los correspondientes al área de matemáticas. El estudiar en una institución femenina aumenta la probabilidad en un 1,03% de pasar del nivel bajo al nivel medio con respecto a una institución con población mixta, mientras que al analizar el caso para la institución masculina se obtiene una probabilidad asociada de 0,47% más de subir de nivel.

4.3 Análisis de escenarios

Como se puede observar en los resultados anteriores, la evidencia estadística indica que para el área de matemáticas hay mayor número de variables significativas, por lo cual, en esta sección, se analizarán diferentes escenarios para este campo del conocimiento, con base en los resultados suministrados por el modelo *Logit* Ordenado Generalizado. El interés se centra en conocer qué variables objeto de política disminuyen la probabilidad de que un estudiante con diferentes características se ubique en un nivel bajo.

Primer escenario. Aquí se analiza cómo afecta su permanencia en el nivel bajo el tener o no tener computador un individuo con condiciones socioeconómicas desfavorables. *Individuo particular:* hombre, 17 años, estrato 1, no trabaja, vive en cabecera, nivel de educación del padre es secundaria, nivel de educación de la madre es primaria, pertenece a nivel Sisben 1, no está en condición de hacinamiento, **computador (0/1)**,

nivel de ingreso 2, jornada tarde y carácter de la institución académica.

Tabla 8. Computador

Características	Probabilidad asociada al nivel bajo
No tiene computador	8,02%
Tiene computador	7,21%

Fuente: elaboración propia.

Como se aprecia en la tabla 8, para un individuo con las características descritas, el tener un computador afecta positivamente el resultado en el área de matemáticas. El no tener computador tiene asociada una probabilidad de permanencia en el nivel bajo de 8,02%, mientras que si posee el instrumento, la probabilidad disminuye a 7,21%.

Segundo escenario. En este se examina el impacto que tiene el nivel de escolaridad de los padres, entre primaria y secundaria, sobre la permanencia en el nivel bajo. *Individuo particular:* hombre, 17 años, estrato 1, no trabaja, vive en cabecera, *el nivel de escolaridad de ambos (primaria/secundaria)*, nivel Sisben 1, no está en condiciones de hacinamiento, sí tiene computador, nivel de ingresos 1, jornada de la tarde y carácter de la institución académica.

Tabla 9: Nivel de los padres

Características	Probabilidad asociada al nivel bajo
Los dos padres tienen primaria	7,59%
Los dos padres tienen secundaria	6,71%

Fuente: elaboración propia.

En la tabla 9 se expone que el nivel de escolaridad de los padres disminuye la probabilidad de que el puntaje obtenido en la prueba para matemáticas esté en nivel bajo, para un individuo con las características descritas. Así, la probabilidad de permanencia en el nivel bajo para un estudiante donde ambos padres tiene nivel de educación

primaria es de 7,59%, mientras que si ambos padres tienen nivel de escolaridad secundaria, dicha probabilidad se reduce a 6,71%.

Tercer escenario. Para este caso, se examina la probabilidad de permanencia en el nivel bajo de un individuo con características socioeconómicas favorables, teniendo en cuenta si estudia en jornada de la tarde o jornada completa. *Individuo particular:* hombre, 17 años, estrato 5, no trabaja, vive en cabecera, nivel profesional para ambos padres, no tiene Sisben, no presenta condiciones de hacinamiento, tiene computador, nivel de ingreso 6, jornada escolar (tarde/ completa).

De la tabla 10 se puede concluir que un individuo con las características descritas tiene menor probabilidad de permanecer en un nivel bajo si estudia en jornada completa (2,92%), mientras que para un estudiante que asiste a la jornada de la tarde la probabilidad aumenta (3,18%). Este resultado tiene serias implicaciones sobre una política de aumento en la cobertura fundamentada en el acortamiento de la jornada académica, puesto que tal política podría implicar una reducción en los resultados que obtienen los bachilleres en las pruebas.

Tabla 10. Jornada

Características	Probabilidad asociada al nivel bajo
Jornada tarde	3,18%
Jornada completa	2,92%

Fuente: elaboración propia.

Conclusiones y recomendaciones

En los últimos años las políticas educativas en Colombia han centrado su atención en mejorar la calidad de la educación. Para ello se han adelantado diversos programas basados en mejoras institucionales: aumento de gasto en educación, calidad del profesorado y mejor infraestructura, entre otras. No obstante, muchos estudios, tanto a nivel nacional como internacional, han puesto en evidencia que las variables socioeconómicas juegan un papel primordial en la calidad de la educación; medida obtenida a través del desempeño en pruebas estándar.

Para el presente trabajo se utilizó la metodología *Logit* Ordenado Generalizado con el fin de encontrar los principales factores que determinan el rendimiento académico de los bachilleres en el país: individuales, socioeconómicos e institucionales. Con este objetivo se analizaron los resultados de las áreas de lenguaje y matemáticas en las pruebas Saber 11° del segundo semestre de 2009. Dicha metodología se utiliza para hacer estimaciones cuando la variable dependiente es discreta, jerarquizada y se incumple el supuesto de riesgos proporcionales. Para este caso, la variable dependiente se categorizó en los tres niveles propuestos por el ICFES: bajo, medio y alto.

En general, tener condiciones económicas favorables afecta en gran medida la probabilidad de no ubicarse en el nivel bajo para ambas áreas. Lo anterior se puede explicar debido a que un entorno socioeconómico favorable le permite al estudiante contar con las condiciones esenciales para dedicarse sin inconvenientes a sus estudios. Un individuo con nivel socioeconómico alto puede acceder

a mejor alimentación, transporte, infraestructura y herramientas informáticas, entre otros aspectos. En tal sentido, se destaca lo encontrado para la tenencia de computador, puesto que la presencia de esta herramienta en el hogar implica una mayor probabilidad de obtener buenos resultados en la prueba.

Una variable que se resalta en el análisis es el nivel de escolaridad de los progenitores, ya que su significancia estadística y su magnitud son relevantes. En efecto, a medida que el nivel de escolaridad de los progenitores aumenta, la probabilidad de estar en un nivel bajo disminuye. Este resultado refleja que padres con mayor nivel educativo pueden apoyar favorablemente el proceso de enseñanza-aprendizaje de sus hijos de manera directa.

En cuanto a las variables propias del estudiante, los resultados arrojaron evidencia a favor de un mejor desempeño en el área de matemáticas para los hombres, mientras que las mujeres obtienen mejores resultados en el área de lenguaje. En cuanto a la variable *edad*, en la medida que esta aumenta hay mayor probabilidad de permanecer en un nivel bajo en ambas áreas.

Con respecto a las variables institucionales y en referencia a las áreas de matemáticas y lenguaje se evidenció que los estudiantes de jornada completa tienen mejor desempeño en la prueba que aquellos que asisten a las demás jornadas. En otras palabras, los individuos con mayor probabilidad de ubicarse en el nivel bajo son aquellos que estudian en la jornada nocturna o en la sabatina y dominical. Esto, probablemente, está asociado con las características de este segmento de la población estudiantil: individuos que normalmente no disponen de tiempo para realizar labores académicas por fuera de las aulas de clase.

Tales resultados son bastante relevantes desde el punto de vista de las implicaciones de política económica, puesto que se debe pensar seriamente en una estrategia de aumento de la cobertura educativa sin que ello implique un detrimento de los resultados estudiantiles en las pruebas de Estado. Con relación a la variable correspondiente a la población que atiende el colegio, se encuentra que el hecho de estudiar en un establecimiento con población exclusivamente femenina o masculina disminuye la probabilidad de pertenecer a un nivel bajo.

En el trabajo se plantearon diferentes escenarios y de su examen se pueden extraer algunas recomendaciones en materia de política económica. Las pruebas mejorarían, sin duda, a partir de una política encaminada a dotar con un computador a los estudiantes con condiciones económicas desfavorables. Del escenario planteado para la variable *jornada*, se puede inferir la conveniencia de implementar políticas con las cuales se garantice que los estudiantes permanezcan una mayor parte de su tiempo en actividades de tipo escolar.

Por último, del escenario analizado para el nivel de educación de los progenitores se puede concluir que uno de los objetivos para aumentar la calidad de la educación debe contemplar la ejecución de programas dirigidos a que la población en general tenga mayores niveles de escolaridad. Es decir, se debe tratar de romper el círculo vicioso en el sentido que padres menos educados implican hijos menos educados, quienes, a su vez, serán los padres del mañana.

Los resultados generales motivan a la reflexión sobre varios aspectos. A pesar de los efectos significativos que las variables institucionales presentan sobre el desempeño de los estudiantes en las pruebas, las variables de tipo socioeconómico tienen un mayor efecto sobre las competencias, en particular el nivel de ingresos y de escolaridad de los progenitores. Por tanto, es importante que los hacedores de política económica centren su atención en subsanar las falencias asociadas al nivel socioeconómico de los estudiantes, en realizar programas que garanticen una buena alimentación, salud, facilidad de transporte, entre otros factores; estas son variables determinantes para que los estudiantes dispongan de condiciones básicas, esenciales, para llevar a cabalidad su formación académica y puedan obtener mejores resultados que reflejen un aumento de la calidad educativa.

Bibliografía

Alexander, L. y Simmons, J. (1975). "The determinants of school achievement in developing countries: the educational production function", *Cuadernos de Trabajo*, 201. Washington: Banco Mundial, pp. 1-73.

Barrientos, J. (2008). "Calidad de la educación pública y logro académico en Medellín 2004-2006. Una aproximación por regresión intercuartil", *Lecturas de Economía*, 68. Universidad de Antioquia, pp. 121-144.

Brant, R. (1990). "Assessing proportionality in the proportional odds model for ordinal logistic regression", *Biometrics*, 46. University of Calgary, Canada, pp. 1171-1178.

Cameron, C. y Trivedi, P. (2005). *Microeconomics, methods and applications*. New York: University of Cambridge.

Coleman, J. (1966). Informe Coleman: *Equality of educational opportunity*. Washington: US Government Printing Office.

Correa, J. (2004). "Determinantes del rendimiento educativo de los estudiantes de secundaria en Cali: un análisis multinivel", *Revista Sociedad y Economía*, 6. Cali, pp. 81-105.

Delfino, J. (1989). "Los determinantes del aprendizaje", *Ensayos en Economía de la Educación*. Buenos Aires: Serie discusión, 4, Universidad Nacional de Córdoba, pp. 287-316.

Delors, J. (1996). "La educación encierra un tesoro". Informe a la Unesco. Comisión Internacional sobre la Educación para el siglo XXI, [HTTP://WWW.UNESCO.ORG/EDUCATION/PDF/DELORS _ S.PDF](http://www.unesco.org/education/pdf/DELORS_S.PDF). (1 de febrero de 2011).

Espínola, V. y Martínez, R. (1996). "Importancia relativa de algunos factores del rendimiento educativo", En COHEN, Ernesto. *Educación, eficiencia y equidad*. Santiago de Chile : Ediciones SUR, 1997; 1ª edición, pp. 85-91, <http://www.sitiosur.cl/r.php?id=711>. (5 de febrero de 2011)

Fu, V. (1998). "Estimating Generalized Ordered Logit Models". En Stata Technical Bulletin 44: Reprints, vol 8, pp. 160-164. College Station, TX: Stata Press.

Gaviria, A. y Barrientos, J. (2001a). "Calidad de la educación y rendimiento académico en Bogotá", *Coyuntura Social*, 24. Fedesarrollo, pp. 111-126.

Gaviria, A. y Barrientos, J. (2001b). "Determinantes de la calidad de la educación en Colombia", *Archivos de Economía*, 159. Departamento Nacional de Planeación, Bogotá, DC, pp. 1-88.

Instituto Colombiano para la Evaluación de la Educación (ICFES). (2010). "Bases de datos pruebas Saber", <ftp://ftp.icfes.gov.co/>. (1 de febrero de 2011)

Instituto Colombiano para la Evaluación de la Educación (ICFES). (2009). "Guía de orientación Examen de Estado 2009", <http://www.icfes.gov.co/>. (1 de febrero de 2011)

Jencks, C. et al. (1972). "Inequality: a reassessment of the effect of family and schooling in America", *New York: Basic Books*, New York, pp. 400.

Mella, O. y Ortiz, I. (1999). "Rendimiento escolar. Influencias diferenciales de factores externos", *Revista latinoamericana de estudios educativos*, 1(XXIX). Centro de estudios educativos, Distrito Federal, México, pp. 69-92.

Tobón, D.; Posada, H. y Ríos, P. (2008). "Determinants of the performance of the Schools in Medellin in the High-School Graduation-Year Test (ICFES)", *Cuadernos de Administración*, 38(22). Pontificia Universidad Javeriana de Colombia, pp. 311-333.

Toutkoushian, R. y Curtis, T. (2005). "Effects of socioeconomic factors on public high school outcomes and rankings", *The Journal of Educational Research*, 5(98). Indiana University and Carleton College, pp. 259-271, <http://www.jstor.org/stable/27548088> (8 de febrero de 2011)

Wolff, L.; Schiefelbein, E. y Valenzuela, J. (1993). "Mejoramiento de la calidad de la educación primaria en América Latina y el Caribe: hacia el siglo XXI", *Banco Mundial, América Latina y el Caribe, Programa de Estudios Regionales*, 28. Banco Mundial, pp. 1- 153.

Xing, L. (2010). "Ordinal regression analysis. Fitting the Continuation Ratio Model to educational data using STATA". NERA Conference Proceedings 2010. Eastern Connecticut State University, Paper 35, pp. 1-25, http://digitalcommons.uconn.edu/nera_2010/35/ (9 de febrero de 2011).