FORMATO DE INFORMES TÉCNICOS FINALES
DE PROYECTOS DE INVESTIGACIÓN

Programa Nacional de Estudios Científicos en Educación
COLCIENCIAS

Título: "*La incidencia del entorno familiar y la calidad de los entornos escolares en los logros educativos de los estudiantes: un estudio aplicado con las pruebas SABER y TIMSS*"	
Código: 1106-518-2818	**Contrato No.** 254-2011
Entidad ejecutora: Universidad del Valle – Facultad de Ciencias Sociales y Económicas	**Grupo de investigación:** GERA – Grupo de Economía Regional y Ambiental
Director del proyecto: Harvy Vivas Pacheco	**Fecha de inicio:** 3 de mayo de 2011

Tabla de contenido

	Pág.
1. Sinopsis divulgativa	3
2. Sinopsis técnica (en español e inglés)	4
3. Resumen técnico	6
3.1 Importancia y componentes del estudio	6
3.2 El problema de investigación y su relevancia	7
3.3 Dimensión metodológica del estudio	8
3.4 Resultados concomitantes con los objetivos del estudio	12
3.5 Anexos – Resumen técnico	16
4. Cuadro de resultados de generación de conocimiento	29
4.1 Cuadro de otros resultados de generación de conocimiento no comprometidos contractualmente	32
5. Cuadro de otros resultados	33
5.1 Cuadro de otros resultados no comprometidos contractualmente	36
6. Descripción del impacto actual o potencial de los resultados	41
Anexos	43
Anexo A Artículo "Potencial de logro educativo, entorno socioeconómico y familiar: una aplicación empírica con variables latentes"	
Anexo B Artículo "Logros académicos, experiencia y prácticas pedagógicas de los docentes en educación para la ciudadanía"	
Anexo C Artículo "Colombia en las pruebas de Educación Cívica de la IEA"	
Anexo D Constancia de formación de recurso humano	

Anexo E Constancia de capacitación (entrenamiento)
Anexo F Constancia de curso organizado
Anexo G Participación en eventos científicos
Anexo G1 Resumen de ponencia “Formación docente y logros académicos en educación cívica y formación para la ciudadanía”
Anexo G2 Resumen de ponencia “Colombia en las pruebas de Educación Cívica de la IEA”
Anexo H Constancia de formación de recurso humano
Anexo I Publicación divulgativa “Gran acuerdo nacional por una educación de calidad”
Anexo J Participación en eventos científicos
Anexo J1 Resumen de ponencia “Influencia del clima democrático en el salón de clase sobre los resultados en conocimiento cívico en América Latina: un estudio con las pruebas ICCS (2009) de la IEA”
Anexo J2 Programación conferencista invitado al Congreso de Educación
Anexo J3 Certificación participación en espacio académico local
Anexo J4 Programación conferencista invitado evento nacional
Anexo K Certificado capacitación ICFES

1. Sinopsis divulgativa

A pesar de los enormes esfuerzos del país en las últimas dos décadas para aumentar los niveles de capital humano y ampliar la cobertura educativa, aún persisten diferencias importantes en la provisión de servicios educativos de alta calidad. Por tal razón, bajo la hipótesis de que la apropiación efectiva de los recursos públicos está condicionada por la calidad de los entornos escolares y familiares, particularmente el nivel educativo de los padres, se hizo necesario auscultar en mayor detalle cuáles son los factores específicos a estas dos dimensiones que afectan los niveles de proficiencia obtenidos en las pruebas nacionales e internacionales de competencias en matemática, lecto-escritura y formación para la ciudadanía.

Por tales razones, se estudiaron los factores asociados al rendimiento académico de los estudiantes colombianos, particularmente los que corresponden a la incidencia del entorno familiar y a la calidad de los entornos en los que los estudiantes realizan sus actividades escolares.

La investigación se ubica en el campo de la Economía de la Educación y sus aplicaciones empíricas se realizaron con la información de los microdatos proporcionada por el ICFES en las pruebas SABER, además de los resultados disponibles en la plataforma estadística de la IEA (*International Association for the Evaluation of Educational Achievement*), entre las que se destacan los resultados de TIMSS (matemática), PIRLS (lecto-escritura) y los microdatos de la pruebas de competencias en educación cívica (ICCS) liberados recientemente.

Entre los hallazgos más relevantes cabe destacar que se corroboró la hipótesis de que los factores de los entornos familiares y escolares inciden de manera importante en las diferencias de proficiencia. Para el caso de las pruebas en matemática, competencia lectora y educación cívica, a partir de las pruebas TIMSS, PIRLS, ICCS y SABER, se identificaron diferencias importantes en los puntajes obtenidos por género, localización de la escuela (urbano/ rural), naturaleza jurídica del establecimiento educativo (oficial/privada).

Así mismo, en el caso particular de las competencias lectoras, se corroboró que los desarrollos o dificultades observados en la proficiencia lingüística y en las competencias en lectura en los primeros años guardan estrecha relación con los resultados obtenidos en las pruebas. De acuerdo con los resultados de las contrastaciones con las pruebas Saber 11 y PIRLS, se aportó evidencia sobre la enorme segregación escolar que aún persiste en nuestro país, lo cual se pudo apreciar en la variabilidad explicada de las diferencias en lenguaje, la cual se atribuye fundamentalmente a la diferencia entre colegios. Estos resultados, al controlar por indicadores de ingreso monetario medio de los colegios aportaron evidencia adicional sobre la presencia de un patrón socioeconómico de segregación escolar.

2. Sinopsis técnica (en español e inglés)

La pregunta que inspiró la investigación se centró en cómo inciden el entorno familiar y la calidad de los entornos escolares en el rendimiento académico de los estudiantes colombianos.

Mediante la explotación de los microdatos de las pruebas de competencias nacionales e internacionales, se identificó la manera como las características de la familia y el ambiente escolar actúan como predictores de los logros de los escolares.

A partir de los microdatos de las pruebas nacionales e internacionales la investigación permitió estudiar las implicaciones de los entornos de aprendizaje sobre la calidad educativa, los logros y las competencias académicas de los estudiantes. Para dilucidar los factores de mayor relevancia en el rendimiento, el estudio se centró en los *scores* o los resultados de proficiencia y la interacción con los factores asociados al *background* familiar, al entorno escolar y, en algunos casos a las prácticas pedagógicas de los docentes.

La aproximación metodológica se hizo mediante la utilización de técnicas estadísticas multinivel en la línea de los trabajos de Goldstein (1999), Hox (2002) y Raudenbush (1992).

Después de los diversos ejercicios se pudo corroborar que efectivamente los factores del entorno escolar juegan un papel preponderante a la hora de explicar la variabilidad de los resultados en las pruebas de competencias. Adicionalmente, después de introducir otros factores asociados a las prácticas pedagógicas, el estudio encontró que éstas también juegan como otro factor importante en la dispersión de los resultados entre colegios, países e individuos. Para el caso de las pruebas en matemática, competencia lectora y educación para la ciudadanía, a partir de las pruebas TIMSS, PIRLS, ICCS y SABER, se identificaron diferencias importantes en los puntajes obtenidos por género, localización de la escuela (urbano/ rural), naturaleza jurídica del establecimiento educativo (oficial/privada).

En el caso particular de las competencias lectoras se corroboró que los desarrollos o dificultades observados en la proficiencia lingüística y en las competencias en lectura en los primeros años guardan estrecha relación con los resultados obtenidos en las pruebas. De acuerdo con los resultados de un modelo incondicional de dos niveles con las pruebas Saber 11 la correlación intraclase llega a 27.3% que indica la variabilidad explicada de las diferencias en lenguaje por la diferencia entre colegios y que en nuestro caso se puede interpretar como un índice de segregación educativa. Al introducir un modelo incondicional al primer nivel y condicional al segundo nivel con un indicador socioeconómico de colegio proveniente de las pruebas Saber 11, el índice de correlación intra-clase cae a un valor de 11.58%, aportando evidencia fuerte de la presencia de *sorting* socioeconómico en los resultados de proficiencia.

Technical synopsis

This study researched the influence of the familiar background and the quality of the school environment in the academic achievement of the Colombian students.

Research used disaggregated dataset of the national and international test of competences and investigated the characteristics of the family and the school environment to predict the achievements of the pupils.

We focused on the scores and we study the interaction with the familiar background, the school environment and, in some cases with the pedagogic practices of the teachers.

The techniques that we use were the statistical multilevel models suggested in the works of Goldstein (1999), Hox (2002) and Raudenbush (1992).

The results that we obtained demonstrated that effectively the school background plays a preponderant role to explain the variability of the results in the achievements of the pupils. The pedagogic practices have also influence of important way in the explanation of the differences between schools, countries and individuals. The tests of mathematics, reading and writing, education for citizenship, (TIMSS, PIRLS, ICCS and SABER 11) allowed identifying important differences in the scores obtained for the students according to sex, the school location, juridical organization of the educational establishment.

In the particular case of the reading tests and writing, the study confirmed that the difficulties observed in the tender age of childhood has close relationship with the results obtained in the achievements of the students. According to the results of the unconditional model of two-levels with the SABER 11, the intra-class correlation was 27.3 %. This result suggests the big variability in language explained by the difference between schools. These coefficients suggest the huge educational segregation of our country. The unconditional model to the first level and conditional to the second level with a socioeconomic indicator by school in SABER 11 show a intra-class correlation of 11.58%. This value show strong evidence of the presence of socioeconomic sorting in the tests.

3. Resumen técnico

A continuación se presenta la síntesis técnica del estudio. La primera parte destaca la importancia de la investigación, la segunda, el problema analizado y su relevancia, la tercera resume la dimensión metodológica del estudio y finalmente se presentan los principales resultados y hallazgos concomitantes con los objetivos del estudio.

3.1. Importancia y componentes del estudio

El estudio investigó los factores asociados al rendimiento académico de los estudiantes colombianos, particularmente los que corresponden a la incidencia del entorno familiar y la calidad de los entornos escolares.

La hipótesis fundamental que inspiró el estudio apuntaba a que la apropiación efectiva de los recursos públicos orientados hacia la financiación de los servicios educativos está condicionada por la calidad de los entornos escolares y por el *background* familiar. Para contrastar la hipótesis el estudio intentó identificar, estimar y analizar en detalle a partir de las diferentes pruebas nacionales e internacionales de competencias, los factores específicos a estas dos dimensiones que limitan o potencian los niveles de proficiencia obtenidos por los estudiantes colombianos.

Tal y como se destaca en diversos estudios, a pesar de los enormes esfuerzos del país en las últimas dos décadas para aumentar los niveles de capital humano y ampliar la cobertura educativa, aún persisten diferencias importantes en la provisión de servicios educativos de alta calidad. Por tal razón, bajo la hipótesis de que la apropiación efectiva de los recursos públicos tiene como condicionante a la calidad de los entornos escolares y el *background* familiar, se hizo necesario auscultar en mayor detalle cuáles son los factores específicos a estas dos dimensiones que afectan los niveles de proficiencia obtenidos en las pruebas nacionales e internacionales.

Como se anotó en un informe de avance presentado en el año 2011, la investigación se ubicó en el campo de la Economía de la Educación y sus aplicaciones empíricas se realizaron con la información de los microdatos proporcionada por el ICFES en las pruebas SABER, además de los resultados disponibles en la plataforma estadística de la IEA (International Association for the Evaluation of Educational Achievement), entre las que se destacan los resultados de TIMSS, PIRLS, y los microdatos de la pruebas de competencias en educación cívica (ICCS) liberados recientemente que permiten contrastar y controlar algunos hallazgos de las demás pruebas.

En la línea del proyecto inicialmente formulado, el marco teórico se apoyó en los postulados de interacción social en las instituciones educativas y en la necesidad de afrontar los problemas metodológicos de identificación (*the reflection problem*) de los efectos endógenos, contextuales y correlacionados a los que se enfrentan las investigaciones sobre logros educativos. El estudio aquí propuesto introdujo algunas

variantes a la metodología de estimación de estos efectos que corrientemente aparecen entremezclados a la hora de evaluar los logros agregados de los estudiantes.

3.2. El problema de investigación y su relevancia

La pregunta que inspiró la investigación fue la siguiente:

¿Cómo inciden el entorno familiar y la calidad de los entornos escolares en el rendimiento académico de los estudiantes colombianos?

Los objetivos que orientaron el trabajo fueron:

OG: Dilucidar los factores familiares y del entorno escolar que inciden en mayor proporción en la calidad educativa, medida a través de los indicadores de proficiencia de las pruebas de competencias nacionales e internacionales.

OE1: Identificar a través de métodos de análisis multinivel los factores de los entornos familiares y escolares que inciden en mayor proporción en las diferencias de proficiencia.

OE2: Estimar los impactos diferenciales que sobre los resultados de las pruebas SABER tienen por regiones los factores asociados a los entornos familiares y escolares.

OE3: Comparar los resultados obtenidos en las exploraciones multinivel para Colombia con una muestra de países de América Latina incluidos en los estudios de competencias.

OE4: Simular, a partir de las estimaciones econométricas con modelos lineales jerárquicos, el impacto que podrían tener algunas políticas orientadas al mejoramiento de los entornos escolares en los que se forman los estudiantes.

Mediante la explotación de los microdatos de las pruebas de competencias nacionales e internacionales, se identificó la manera como las características de la familia y el ambiente escolar operan como predictores de los logros de los escolares y como condicionantes de la efectividad de los recursos públicos comprometidos en la financiación de los servicios educativos.

A partir de la información de los microdatos es posible aproximarse al estudio de las implicaciones de los entornos de aprendizaje sobre la calidad educativa, los logros y las competencias académicas de los estudiantes, tal y como lo sugieren algunas investigaciones internacionales y nacionales. Para dilucidar los factores de mayor relevancia en el rendimiento, el estudio se centró en los *scores* o los resultados de proficiencia y la interacción con los factores asociados al *background* familiar, al entorno escolar y, con base en algunas de las pruebas, las prácticas pedagógicas de los docentes, a través de la utilización de técnicas estadísticas multinivel, que tal y como lo señalaran Goldstein (1999), Hox (2002) y Raudenbush (1992), son las apropiadas para este tipo de análisis.

Después de los diversos ejercicios y una vez alcanzados los principales objetivos del estudio trazados inicialmente se pudo corroborar que efectivamente los factores del entorno escolar juegan un papel preponderante a la hora de explicar la variabilidad de los resultados en las pruebas de competencias. Adicionalmente, después de introducir otros factores asociados a las prácticas pedagógicas, el estudio encontró que éstas también

juegan como otro factor importante en la dispersión de los resultados entre colegios, países e individuos.

El conocimiento profundo y detallado de los factores de mayor incidencia en el rendimiento de los educandos aporta nuevos elementos para la asignación eficiente de los recursos destinados a la financiación de los servicios educativos en el orden centralizado y descentralizado del país.

Finalmente, es importante comentar que la propuesta se articuló a la línea de investigación Economía y Sociología de la Educación del Grupo de Economía Regional y Urbana –GERA- **COL0004049,** grupo de investigación de la Facultad de Ciencias Sociales y Económicas de Univalle.

3.3 Dimensión metodológica del estudio

De acuerdo con la disponibilidad de los microdatos de las pruebas SABER y las bases de datos internacionales de la IEA (PIRLS, TIMMS, ICCS), el diseño metodológico de los diferentes componentes presentó una matriz disciplinaria regular que le dio unidad conceptual y técnica a los diversos productos y que se formaliza a través de algunos modelos analíticos que pudieron contrastarse mediante técnicas econométricas. La matriz disciplinaria de toda la investigación se fundamentó en los postulados de interacción social sugeridos por Manski (1993, 1995) que se resumen a continuación:

Sea R un escalar que denota el resultado en alguna prueba y que condensa la proficiencia del individuo i.

El vector G resume el conjunto de características asociadas al grupo de referencia de los individuos, por ejemplo grupos étnicos o clases.

F indica el conjunto de variables relacionadas con los atributos socioeconómicos y corresponde al vector exógeno de atributos.

U la habilidad no observable directamente.

El modelo analítico y sus implicaciones metodológicas

El modelo lineal se puede escribir de la siguiente manera[1]:

$$R = \alpha + \beta E(R \mid G) + E(F \mid G)'\gamma + F'\lambda + U$$
$$E(U \mid G, F) = G'\delta$$

Donde E, operador de esperanza matemática, denota el valor esperado.

La regresión de R sobre (G, F) genera la ecuación fundamental de interacción social:

[1] De acuerdo con la especificación de Vivas, H., 2008, capítulo 3.

$$E(R|G,F) = \alpha + \beta E(R|G) + E(F|G)'\gamma + F'\lambda + G'\delta$$

Esta expresión corresponde al valor de proficiencia esperado, dados los vectores grupales y familiares.

(R,G,F) se observan directamente pero las realizaciones de la variable U no.

(α,β,γ,λ,δ) es el vector de parámetros.

Así, la especificación intenta capturar los efectos endógenos (β), contextuales (γ), los efectos correlacionados (δ) y los directos que genera el background familiar (λ).

Obsérvese que si δ es significativo los efectos correlacionados importan y en consecuencia los individuos que conforman el grupo tienden a comportarse o exhibir resultados similares en el atributo no observado U. Si por el contrario, δ=0, los resultados individuales en logros son independientes de G. Si δ es muy grande o tiende a infinito, todos los miembros del grupo tienden a tener idéntica distribución del atributo no observado U.

Adicionalmente es preciso anotar que si E(R|G), E(F|G), F y G fuesen tratamientos exógenos (Manski, 1995, p. 130) entonces el vector de parámetros (β,γ,λ,δ) sencillamente capturaría efectos puros de tratamiento; pero, tanto E(R|G) como E(F|G) dependen de G funcionalmente y por tal razón aparece un problema de identificación en el sentido que le confiere Manski.

Al aplicar esperanza matemática respecto a F a la ecuación de interacción social anterior se obtiene que:

$$E(R|G) = \alpha + \beta E(R|G) + E(F|G)'\gamma + E(F|G)'\lambda + G'\delta$$

Y, en consecuencia,

$$E(R|G) = \frac{[\alpha + E(F|G)'(\gamma + \lambda) + G'\delta]}{(1-\beta)},$$

Esto último, si y solo si β es diferente de la unidad.

Esta ecuación reducida es la solución única de equilibrio social que indica la dependencia funcional que tiene R de G, de tal modo que E(R|G) es perfectamente colineal con [1, E(F|G), G].

Esto último implica que los parámetros no quedan identificados y así entonces los efectos endógenos puros no pueden distinguirse de los contextuales ni de los correlacionados. Al sustituir la solución anterior en la ecuación de interacción social, la forma reducida de esta última puede escribirse de la siguiente manera:

$$E(R|G,F)=\frac{\alpha}{(1-\beta)}+E(F|G)'[\frac{\gamma+\lambda\beta}{(1-\beta)}]+G'\frac{\delta}{(1-\beta)}+F'\lambda$$

Aquí los parámetros quedan identificados cuando no hay colinealidad perfecta. La única dificultad es que no se puede distinguir entre efectos endógenos puros y contextuales, pero la virtud de la ecuación radica en que permite capturar la incidencia de los factores sociales como conjunto (Manski, 1995, p. 132).

Obsérvese que cuando $[\frac{\gamma}{(1-\beta)}+\frac{\lambda\beta}{(1-\beta)}]$ es significativo una implicación es que γ y β también deberían ser diferentes de cero.

Ahora bien, si se supone en la ecuación reducida que δ=γ=0 (efectos contextuales y correlacionados nulos), el modelo que se deriva es endógeno puro:

$$E(R|G,F)=\frac{\alpha}{(1-\beta)}+E(F|G)'\frac{\lambda\beta}{(1-\beta)}+F'\lambda$$

Bajo λ diferente de cero (que captura el efecto directo del background familiar), entonces el modelo queda identificado.

Al identificar el valor del parámetro compuesto $\frac{\lambda\beta}{(1-\beta)}$, entonces se pueden obtener β y α.

Bajo tal supuesto tendríamos una aproximación indirecta a la magnitud de los efectos endógenos (β) y directa a los que genera el *background* familiar (λ).

Una especificación compacta y adecuada para la realización de los ejercicios econométricos que se presentarán más adelante expresa la forma reducida como sigue:

$$R_{ics}=\frac{\alpha}{(1-\beta)}+\lambda F_{ics}+\phi Sch_{cs}+\frac{\gamma+\lambda\beta}{(1-\beta)}\bar{F}_{(-i)cs}+\mu_{cs}+\varepsilon_{ics}$$

En donde R_{ics} corresponde a los puntajes, valores plausibles o *scores* de las pruebas y que tal como lo anotamos con anterioridad intentan capturar la proficiencia del estudiante i que pertenece a la clase c de la escuela s. F_{ics}: contiene el *background* familiar y las características individuales. $\bar{F}_{(-i)cs}$: representa el *background* medio de los compañeros de clase (excluyendo al individuo i) o algún indicador que defina el perfil de los pares escolares. Sch_{cs}: corresponde al conjunto de atributos asociados a la escuela s (entorno o disponibilidad de recursos). μ_{cs} y ε_{ics} son los errores a nivel de escuela o clase y los individuales, respectivamente. λ, ϕ y $\frac{\gamma+\lambda\beta}{(1-\beta)}$ son los parámetros asociados al *background* familiar, el entorno escolar y los atributos de los compañeros de clase, respectivamente.

La ventaja de esta especificación respecto a la ecuación de interacción social explicada previamente es que elimina el problema de simultaneidad. Tal dificultad se generaría al incluir directamente los *scores* promedio ($\bar{R}_{(-i),cs}$) de los compañeros de clase y por tanto $\text{cov}(\bar{R}_{(-i),cs}, \varepsilon_{ics}) \neq 0$. De igual manera se elimina el problema de reflejo señalado por Manski (1993, 1995), salvo que se presentara un problema extremo de colinealidad perfecta en los regresores utilizados o por la manera como se haya diseñado la muestra.

Adicionalmente, es necesario tener en cuenta que el error, μ_{cs}, captura efectos correlacionados que condensan la influencia común que reciben los estudiantes a nivel de escuela.

No obstante, si se supone que los estudiantes se asignan aleatoriamente a nivel de clase en el interior de cada escuela, pero no entre escuelas, entonces el componente sistemático del *background* medio opera entre estas últimas. Así, al existir covarianza no nula a nivel de escuela (s), $\text{cov}(\bar{F}_{(-i),cs}, \mu_{cs}) \neq 0$, se garantizaría una estimación consistente de los parámetros, siempre y cuando la estimación incluya los efectos grupales pertinentes y logre eliminar o al menos modelar el componente omitido. Ver McEwan (2003) y Ammermueller y Pischke (2006).

Tal como se discutió en varios de los artículos, si las escuelas en Colombia claramente atraen a estudiantes con determinados perfiles socioeconómicos o provenientes de determinados estratos sociales es evidente entonces que las características no observadas que condensa el error μ_{cs} presenten correlación sistemática con $\bar{F}_{(-i),cs}$ al comparar escuelas. En tal sentido la hipótesis que subyace a esta última afirmación es que prevalece una modalidad de ***sorting socioeconómico*** entre los planteles educativos y por tal razón la especificación anterior podría reescribirse como sigue:

$$R_{ics} = \alpha_s^* + \lambda F_{ics} + \phi Sch_{cs} + \rho \bar{F}_{(-i)cs} + \mu_{cs} + \varepsilon_{ics}$$

En donde $\alpha_s^* = \frac{\alpha}{(1-\beta)}$ captura los efectos específicos de cada escuela y $\rho = \frac{\gamma + \lambda\beta}{(1-\beta)}$ la incidencia que tiene el *background* familiar de los compañeros de clase. Como μ_{cs} varía entre escuelas, entonces el intercepto compuesto es aleatorio.

Se aprecia de inmediato que la última ecuación muestra una estructura multinivel en la que los estudiantes están anidados en escuelas. Esto implica que la estimación de algún modelo que no considere esta estructura de agrupamiento sería incorrecta en la medida que supondría que los individuos se distribuyen de manera independiente. Por tales razones, la investigación se centró en la utilización de estructuras multinivel con sus mecanismos de correlación intra-clase.

Es bien conocido en la literatura de modelos jerárquicos y mixtos[2] generalizados que al no modelar la variación simultánea del intercepto con el componente de error asociado se generan sesgos de estimación que inflan la significancia de los predictores y violan uno de los supuestos básicos del teorema de Gauss-Markov. De acuerdo con la última especificación planteada, si $\alpha_s^* = \alpha_0 + \mu_s$, en donde α_0 es un efecto común a todas las escuelas y μ_s el error asociado, al no modelar la correlación existente se genera el sesgo mencionado. Por tal razón la especificación incluye de manera explícita los efectos de grupo, capturados a través de $\bar{F}_{(-i)cs}$ (3).

De manera general esta configuración jerárquica implica que la proficiencia de los individuos dependa de su vector de características específicas en una primera fase y que, en una segunda fase capture los efectos específicos relacionados con su entorno escolar o con el *background* de los compañeros de clase.

La técnica econométrica apropiada en este caso es entonces la de los modelos multinivel[4] ampliamente desarrollados en los trabajos de Goldstein (1999, 1993), Hox (2002), Rowe, Hill y Holmes-Smith (1995), Brik y Raudenbush (1992) y en los algoritmos recientes de Rabe-Hesketh, Skrondal y Pickles (2004)[5]

Así mismo, la utilización de técnicas multinivel aporta información de gran relevancia sobre la variabilidad total e intra-grupal de la muestra, tal y como se pudo apreciar en los diversos ejercicios realizados durante el desarrollo de la investigación y que se resumen en la siguiente sección.

3.4 Resultados concomitantes con los objetivos del estudio

De acuerdo con los objetivos trazados inicialmente en el estudio, el avance fue satisfactorio y no se solicitó a Colciencias ninguna prórroga o ampliación del plazo estipulado inicialmente para su realización, cumpliendo las metas establecidas y con una generación de productos parciales y finales que se detallan en los apartados 4 y 5 del presente informe.

A continuación se enumeran de manera sucinta los principales hallazgos, los cuales se encuentran desarrollados de manera amplia y con todos los protocolos de investigación en los diversos artículos enviados a revistas científicas, ponencias y documentos de divulgación que hemos publicado durante el desarrollo del estudio.

---

[2] Aquellos que combinan efectos fijos y aleatorios de manera simultánea.

[3] Se trata de un procedimiento remedial a un problema de variables omitidas.

[4] La bondad de este tratamiento reside en que los puntajes individuales no pueden ser tratados como observaciones aleatorias independientes, sino que dependen del cluster en el que se anidan, en este caso las escuelas y clases.

[5] Estos últimos creadores del módulo Glamm (Generalized Linear Latent and Mixed Models) recientemente incorporado en el Stata y que tiene la ventaja de explorar diferentes configuraciones de la estructura de covarianzas.

- Entre los hallazgos más relevantes se corrobora la hipótesis de que los factores de los entornos familiares y escolares inciden de manera importante en las diferencias de proficiencia. Para el caso de las pruebas en matemática, competencia lectora y educación para la ciudadanía, a partir de las pruebas TIMSS, PIRLS, ICCS y SABER, se identifican diferencias importantes en los puntajes obtenidos por género, localización de la escuela (urbano/ rural), naturaleza jurídica del establecimiento educativo (oficial/privada).

En relación con las competencias en matemática se destacan los siguientes hallazgos (ver A1 en los anexos del resumen técnico):

- Impacto fuerte de la variable sexo del estudiante sobre los resultados obtenidos en las pruebas TIMSS (2007), este efecto produce un mayor rendimiento en el alcance de logros en matemática para los niños en comparación al rendimiento logrado por las niñas.

- El hecho de poseer computador en los hogares de los estudiantes representa un efecto positivo en el alcance de logros en matemática.

- Se pudo verificar que unas mejores características de la escuela y del aula tienen un efecto positivo y significativo sobre los rendimientos en matemática de los estudiantes. De igual forma, una escuela que se encuentre ubicada en zona urbana y que sea de tipo privado tiene mayores posibilidades de tener mejores calificaciones promedio de sus estudiantes en las pruebas TIMSS.

- De todas las variables incluidas en el estudio las de mayor impacto sobre el alcance de logros en matemática fueron en su orden: el tipo de escuela, gusto por la matemática, zona y gusto por la escuela. Estos resultados sugieren la necesidad de generar mayor innovación en estrategias pedagógicas que busquen aumentar la motivación y el gusto por la matemática y la escuela en los estudiantes y así lograr mejores resultados, como por ejemplo los semilleros de matemática a nivel de escuela primaria y olimpiadas de matemática por zonas, entre otros.

- En este bloque de ejercicios en matemática se pudo verificar la fortaleza del uso de los modelos multinivel para el análisis de datos con estructura jerárquica ya que ofrecen distintas ventajas respecto a los modelos de regresión lineal tradicionales, como por ejemplo brindan resultados más realistas ya que modelan cada nivel de jerarquía, no requieren la hipótesis de independencia entre las medidas de la variable resultado y también dan estimaciones más precisas.

- Los resultados de los modelos mostraron la importancia del nivel de escuela al encontrar que los logros alcanzados por los estudiantes en gran medida se explican por las características propias de la institución como también por las características

de los docentes. Las estimaciones señalan que los mejores rendimientos corresponden a instituciones privadas en comparación con las instituciones públicas. De igual manera la zona donde se encuentra la escuela influye de manera positiva sobre el alcance de logros en las escuelas urbanas, estableciendo cierta desventaja frente a las escuelas de las zonas rurales

Competencias en lecto-escritura

- En el caso particular de las competencias lectoras se corroboró que los desarrollos o dificultades observados en la proficiencia lingüística y en las competencias en lectura en los primeros años guardan estrecha relación con los resultados obtenidos en las pruebas. De acuerdo con los resultados de un modelo incondicional de dos niveles con las pruebas Saber 11 la correlación intraclase llega a 27.3% que indica la variabilidad explicada de las diferencias en lenguaje por la diferencia entre colegios y que en nuestro caso se puede interpretar como un índice de segregación educativa. Al introducir un modelo incondicional al primer nivel y condicional al segundo nivel con un indicador socioeconómico de colegio proveniente de las pruebas Saber 11, el índice de correlación intra-clase cae a un valor de 11.58%, aportando evidencia fuerte de la presencia de *sorting* socioeconómico en los resultados de proficiencia. (Ver A2)

- Estos resultados se habían corroborado en los ejercicios previos exploratorios en ámbitos internacionales con las pruebas PIRLS de la IEA. De acuerdo con nuestras estimaciones, Colombia se ubicaba en el lugar 30 de un total de 35 países. El promedio general fue de 422 con una desviación estándar de 4.4. Las niñas registraron puntajes superiores a los de los niños. Las diferencias fueron significativas después de realizar pruebas de diferencia de medias de las dos distribuciones. El país que obtuvo los mayores *scores* fue Suecia con un promedio de 561 y una desviación de 2.2, seguido por Holanda (554) e Inglaterra (553). (Ver A3)

- En esta misma línea se detectó dificultades en la comprensión de textos literarios (con un puntaje medio en todos los ítems de 425 y desviación 4.2) y en los textos informativos (con un puntaje medio de 424 y una desviación de 4.3). En los procesos inferenciales y de extracción de información el valor plausible promedio fue 429 y en la interpretación, integración y evaluación de textos solamente llegó a 417.

Diferencias por regiones y países

- A partir de las pruebas SABER 2009, las estimaciones de los coeficientes de correlación intra-clase por departamentos –como medida de segregación escolar que se desprende de los ejercicios de descomposición de varianza—aportan evidencia de las fuertes disparidades en el rendimiento de los estudiantes. Un ejemplo de ello se muestra en el anexo a este apartado A4, donde se registra, con

base en el cálculo para Colombia, que los departamentos con mayor índice de correlación intra-clase fueron Atlántico, Valle, Nariño y Santander.

- Así mismo, los ingresos familiares, la dotación de ordenadores en el hogar y los factores asociados a la infraestructura y calidad de las escuelas, entre otros, inciden en los logros académicos alcanzados por los estudiantes.

- Cabe destacar que el nivel de segregación no se mantiene constante en las diferentes jornadas escolares, siendo la jornada completa la que mayor índice de segregación presenta, como se puede observar en el anexo A5. Este hallazgo, desagregado para Bogotá, Antioquia y Valle refleja que dicho resultado se encuentra estrechamente relacionado con el estrato socioeconómico del estudiante (ver A6); para el caso específico de la ciudad de Cali, se encontró que los colegios no oficiales son más heterogéneos que los oficiales, hallazgo que se presenta en los resultados de las pruebas de ciencias sociales, lenguaje y matemática (ver A7).

- En las comparaciones internacionales se destaca también el nexo entre los logros académicos y el Índice de Desarrollo Humano (IDH) o un indicador de PIB per cápita en paridades de poder de compra. Así mismo, se proporcionaron indicios de que el *background* familiar y otras variables del entorno escolar son significativas, pero con diferentes niveles de incidencia por países, tal como se evidencia en el anexo A8 para el caso de la prueba ICCS.

<u>Competencias ciudadanas</u>

- Entre los factores ligados al entorno escolar y asociados al logro académico, cabe anotar que las prácticas pedagógicas juegan un importante rol a la hora de explicar el rendimiento de los estudiantes. A partir del estudio ICCS de educación cívica se evidenció el estrecho vínculo entre las condiciones del entorno escolar y las prácticas pedagógicas de los docentes con los logros académicos de los estudiantes, lo que es, a nuestro juicio, uno de los puntos clave para un diseño adecuado de la política educativa del país.

- El trabajo de investigación con las pruebas internacionales ICCS, particularmente en lo que atañe a las dimensiones actitudinales de los estudiantes y profesores, sugiere que la formación para la ciudadanía constituye el núcleo central de un proceso formativo que coadyuve a la construcción de una sociedad más tolerante y con un manejo adecuado de los conflictos.

Creemos que un concepto integral de calidad educativa trasciende los mejoramientos en áreas temáticas como lenguaje, matemática, ciencias sociales, etc., y debe considerar en profundidad la formación de ciudadanos. Por tal razón el componente de competencias cívicas aparece como uno de los retos de mayor relevancia en el diseño de una política educativa que configure un sistema educativo de calidad.

3.5 Anexos – Resumen técnico

A1. Modelos multinivel con las pruebas TIMSS

La aplicación práctica y la interpretación de los distintos modelos multinivel de matemática se realizaron con el software MLWin, el cual es editado por el Centro de Modelamiento Multinivel del Instituto de Educación de Londres (GOLDSTEIN et al, 2000) y se constituye actualmente como una de las herramientas más importantes para trabajar modelos con estructura jerárquica. En primer lugar, la representación del modelo nulo realizado muestra que el modelo asume una variable de respuesta con distribución normal. El coeficiente beta se encuentra multiplicado por un vector constante que se introduce con la intención de dar el mismo tratamiento al punto de corte que a las pendientes del modelo. El software MLWin muestra también cómo la variación se encuentra estructurada en dos niveles, donde una parte le corresponde al nivel alumnos *e0ij* y al nivel escuela *μoj*. El modelo 1A muestra el modelo considerando solamente la variación del nivel alumnos; en el modelo 1B se incluyó solamente la variación del nivel escuelas. La existencia de varianza dentro de las escuelas indica que los sujetos se diferencian entre ellos dentro de las escuelas. La existencia de varianza en el segundo nivel indica que las escuelas también difieren en el rendimiento medio de cada una de ellas.

En el modelo 1A se tiene un $\beta_0 = 389{,}279$ que significa que la calificación promedio que obtuvieron los estudiantes es de 389,28 puntos. Por su parte el β_0 en el modelo 1B indica que la calificación promedio de las escuelas es de 377,74 puntos.

MODELO 1. COMPONENTES DEL MODELO NULO

Modelo	Parte fija		Parte aleatoria			-
	β_{0i}	β_{0j}	β_{0ij}	e_{0ij}	μ_{0j}	2*loglikelihood
A. Considerando únicamente la variación del nivel alumno	**389,279** (1,506)	No aplica	No aplica	**6962,891** (177,691)	No aplica	35864,030
B. Considerando únicamente la variación del nivel escuela	No aplica	**377,739** (7,565)	No aplica	No aplica	**7153,697** (899,876)	
C. Modelo nulo	No aplica	No aplica	**373,745** (5,371)	**4357,420** (113,547)	**3234,879** (452,774)	34757,672

Fuente: Cálculos propios. Entre paréntesis el error típico

En el modelo 1C (modelo nulo) β_0 es el rendimiento medio de todas las escuelas. $\square_{0j}$ es lo que la media de la escuela j se aparta de la media general y e_{ij} es lo que el rendimiento del alumno i de la escuela j se aparta de la media de su escuela. La varianza $\sigma^2_{\square 0}$ es la varianza entre las medias de las escuelas, y σ^2_e es la varianza entre los alumnos dentro de cada escuela.

El valor 373.745 es el punto de corte del modelo nulo, es decir, la calificación promedio en matemática obtenida en las pruebas TIMSS, con un error típico de 5.371. Si $\sigma^2_{\mu 0}$ no fuese estadísticamente significativa, querría decir que las escuelas no difieren entre sí. Sin embargo, en este caso se encontró varianza significativa tanto en el primer nivel como en el segundo. Por lo tanto, todos los esfuerzos se orientan en empezar a introducir variables en el modelo que disminuyan la varianza no explicada.

La comparación entre el estadístico –2*loglikelihood del modelo nulo y los modelos siguientes permite evaluar el aporte de cada variable, de igual manera la varianza de cada nivel permite analizar la manera como la inclusión de cada variable afecta la variabilidad debida a las escuelas y a los estudiantes. Por otro lado la correlación intraclase del modelo nulo muestra que el 42.6% de la variación total es explicado por el nivel escuela y el 57.4% lo explica el nivel alumno.

El siguiente paso consiste en ir incorporando predictores tanto del primer como del segundo nivel a la parte fija y aleatoria del modelo. En primer lugar se incluyó la variable sexo como predictora del primer nivel, esta variable ha sido definida como 0 para niñas y 1 para niños, obteniendo así el modelo 2. En el modelo 2 el parámetro fijo asociado a la constante (350,371(6,492)) es significativo, y se nota un descenso en su magnitud con respecto al modelo nulo. Esto es debido a que el grupo al que se refiere este punto de corte ya no es el mismo que en el modelo nulo. Ahora 350,371 es el rendimiento promedio en matemática de las niñas. El parámetro asociado a la variable sexo es positivo y significativo (15,618(2,455)) estos 15,6 puntos representan el incremento en el rendimiento medio en matemática entre niños y niñas. En este caso, los niños obtienen en promedio 15,6 puntos más que las niñas.

MODELO 2. APORTE DE LA VARIABLE SEXO

Variable	Parte fija		Parte aleatoria		-
	β_{0ij}	β_{1ij}	e_{0ij}	μ_{0j}	2*loglikelihood
Sexo del estudiante	350,371	15,618	4300,066	3216,042	34717,458
	(6,492)	(2,455)	(112,077)	(449,779)	

Fuente: Cálculos propios. Entre paréntesis el error típico

La diferencia entre el estadístico –2*loglikelihood del segundo modelo en comparación con el modelo nulo resulta estadísticamente significativo, lo que indica que esta variable aporta a la explicación del puntaje obtenido en la prueba TIMSS.

También se encuentran los valores de las variaciones residuales entre estudiantes (σ_e^2) y entre escuelas ($\sigma_{\square}^2$) donde se puede observar que la variación residual entre estudiantes es mayor que la variación residual entre escuelas. Adicionalmente con estas estimaciones se calcula la correlación intraclase (42.79%) que indica la proporción de variación residual que es explicada por cada uno de los niveles y el modelo en general. En el modelo 2 se tiene un valor de varianza total explicada del 0,992%.

Se prosigue con el modelo adicionando las otras variables del primer nivel, adicionando la variable “posee computador en el hogar” se obtuvo el modelo 3. En este modelo se observa una disminución en el punto de corte, ya que ahora se refiere al rendimiento medio de las niñas que no poseen computador en el hogar.

MODELO 3. APORTE DE TENER COMPUTADOR EN EL HOGAR

Variables	Parte fija		Parte aleatoria		-
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	2*loglikelihood
Sexo del estudiante		14,708			
	348,207	(2,464)	4291,125	3079,611	34704,830
Posee computador	(6,434)	9,100	(111,822)	(431,827)	
		(2,555)			

Fuente: Cálculos propios. Entre paréntesis el error típico

El modelo 4 resulta de incluir la variable que mide las deficiencias en informática y estudio. Tal como se esperaba el signo que acompaña al coeficiente de esta variable es negativo, esto indica que por cada incremento en la deficiencia en informática y estudio se disminuye la calificación obtenida en la prueba TIMSS en 9,175 puntos en promedio.

MODELO 4. APORTE DEL INDICADOR DE DEFICIENCIA EN INFORMÁTICA Y ESTUDIO

Variables	Parte fija		Parte aleatoria		-
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	2*loglikelihood
Sexo del estudiante		14,475			
		(2,456)			
Posee computador	362,648	7,837	4263,585	2902,969	34678,200
	(6,896)	(2,559)	(111,108)	(408,358)	
Deficiencia en		-9,175			
Informática y estudio		(1,772)			

Fuente: Cálculos propios. Entre paréntesis el error típico

Después de incluir las variables del nivel estudiante se prosigue con la inclusión de las variables del nivel escuela, inicialmente se incorpora el indicador de Materiales de instrucción en matemática (modelo 5). En este caso como se introduce una variable del nivel escuela se produce una reducción de la correlación intraclase de este nivel. El signo del coeficiente que acompaña a esta variable es positivo, lo que indica que un incremento en el indicador de materiales de instrucción aumenta la calificación promedio en 23,92 puntos. En otras palabras, tal y como se esperaba, una mejor dotación de materiales de instrucción de las escuelas ayuda a que los rendimientos de sus estudiantes sean mayores.

MODELO 5. APORTE DEL INDICADOR DE MATERIALES DE INSTRUCCIÓN EN MATEMÁTICA

Variables	Parte fija		Parte aleatoria		-
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	2*loglikelihood
Sexo del estudiante		14,347			
		(2,455)			
Posee computador		7,572			
	342,073	(2,560)	4263,658	2508,885	34661,142
Deficiencia en	(8,299)	-8,944	(111,102)	(357,684)	
Informática y estudio		(1,773)			
Materiales de instrucción		23,917			
		(5,645)			

Fuente: Cálculos propios. Entre paréntesis el error típico

En los modelos 6 a 11 se muestra el efecto de la adición de cada una de las variables del nivel escuela, se puede ver que todas tienen efecto positivo sobre la calificación TIMSS.

MODELO 6. APORTE DEL INDICADOR DE CARACTERÍSTICAS DEL AULA

Variables	Parte fija		Parte aleatoria		-
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	2*loglikelihood
Sexo del estudiante		14,345			
		(2,455)			
Posee computador		7,615			
		(2,559)			
Deficiencia en	326,848	-8,933	4264,066	2317,483	34651,772
Informática y estudio	(9,532)	(1,772)	(111,108)	(332,887)	
Materiales de instrucción		18,416			
		(5,726)			
Características del aula		17,945			
		(5,817)			

Fuente: Cálculos propios. Entre paréntesis el error típico

Al incluir la variable de características del aula, la variación en el nivel de escuela también presenta una reducción. El coeficiente de 17.945 implica que a medida que mejoran las características del aula el rendimiento del estudiante aumenta en ese valor.

En el modelo 7 se muestra que las características del docente presentaron un signo como el esperado, el parámetro asociado a esta variable fue de 12,705 lo que indica que a medida que se mejoran las características de los docentes se identifica un aumento en el rendimiento en matemática de los estudiantes.

MODELO 7. APORTE DEL INDICADOR DE CARACTERÍSTICAS DEL DOCENTE

Variables	Parte fija		Parte aleatoria		-
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	2*loglikelihood
Sexo del estudiante		14,373			
		(2,455)			
Posee computador		7,608			
		(2,559)			
Deficiencia en		-8,926			
Informática y estudio	308,011	(1,772)	4265,747	2222,667	34647,426
Materiales de instrucción	(13,112)	18,862	(111,158)	(319,767)	
		(5,625)			
Características del aula		17,614			
		(5,711)			
Características del docente		12,705			
		(6,118)			

Fuente: Cálculos propios. Entre paréntesis el error típico

Al incluir la variable tipo de escuelas también se obtuvo el resultado esperado, las escuelas privadas presentan mayores calificaciones en la prueba de matemática que las escuelas públicas. Aproximadamente 37,287 puntos más (ver modelo 8).

MODELO 8. APORTE DEL TIPO DE ESCUELA

Variables	Parte fija		Parte aleatoria		-2*loglikelihood
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	
Sexo del estudiante		14,370			
		(2,454)			
Posee computador		7,439			
		(2,560)			
Deficiencia en Informática y estudio		-8,818			
		(1,773)			
Materiales de instrucción	311,174	12,603		2105,763	4265,470
	(12,905)	(5,978)		(304,933)	(111,143)
Características del aula		14,225			
		(5,721)			
Características del docente		12,911			
		(5,975)			
Tipo de Escuela (Escuela Privada)		37,287			
		(14,146)			

Fuente: Cálculos propios. Entre paréntesis el error típico

Al igual que sucedió con el tipo de escuela, la zona donde se encuentra ubicada la escuela (modelo 9) también produce diferencias estadísticamente significativas en los resultados, siendo superiores las puntuaciones obtenidas por los estudiantes que pertenecen a escuelas de zona urbana. Los resultados de los estudiantes de estas escuelas presentan aproximadamente 22,515 puntos que los estudiantes de zona rural. En los modelos 10 y 11 se muestra el efecto de la inclusión de dos variables que tienen como característica principal que miden la actitud del estudiante frente a la escuela y frente a la matemática.

MODELO 9. APORTE DE LA ZONA DE UBICACIÓN DE LA ESCUELA

Variables	Parte fija		Parte aleatoria		-2*loglikelihood
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	
Sexo del estudiante		14,444			
		(2,454)			
Posee computador		7,196			
		(2,563)			
Deficiencia en Informática y estudio		-8,692			
		(1,774)			
Materiales de instrucción	296,035	10,384	4265,661	2039,727	34636,146
	(14,759)	(5,992)	(111,149)	(295,993)	
Características del aula		15,389			
		(5,670)			
Características del docente		12,124			
		(5,906)			
Tipo de Escuela (Escuela Privada)		32,894			
		(14,102)			
Zona de la Escuela (Escuela Urbana)		22,515			
		(11,023)			

Fuente: Cálculos propios. Entre paréntesis el error típico

La variable gusto por la escuela, como era de esperarse, presentó un parámetro positivo que indica que los estudiantes a los que les gusta ir a la escuela obtienen rendimientos más altos que aquellos a los que no les gusta. El gusto por la escuela produce aproximadamente 30,949 puntos adicionales en la calificación de matemática.

MODELO 10. APORTE DEL GUSTO POR LA ESCUELA

Variables	Parte fija		Parte aleatoria		-2*loglikelihood
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	
Sexo del estudiante		14,622 (2,441)			
Posee computador		6,751 (2,550)			
Deficiencia en Informática y estudio		-0,247 (0,050)			
Materiales de instrucción		0,271 (0,154)			
Características del aula	259,426 (15,824)	0,398 (0,145)	4217,504 (109,894)	2046,517 (296,589)	34601,871
Características del docente		0,473 (0,233)			
Escuela Privada		33,107 (14,118)			
Escuela Urbana		23,045 (11,024)			
Gusto por la escuela		30,949 (5,274)			

Fuente: Cálculos propios. Entre paréntesis el error típico

De igual manera la variable gusto por la matemática presentó un parámetro positivo que como en el caso del gusto por la escuela, está indicando que cuando a los estudiantes les gusta la matemática, obtienen calificaciones más altas en la prueba TIMSS. Aproximadamente 33,421 puntos más que aquellos estudiantes a los que no les gusta la matemática.

MODELO 11. APORTE DEL GUSTO POR LA MATEMÁTICA

Variables	Parte fija		Parte aleatoria		-2*loglikelihood
	β_{0ij}	$\beta_{.ij}$	e_{0ij}	μ_{0j}	
Sexo del estudiante		14,013 (2,405)			
Posee computador		5,521 (2,515)			
Deficiencia en Informática y estudio		-0,234 (0,049)			
Materiales de instrucción		0,262 (0,153)			
Características del aula	242,849 (15,742)	0,403 (0,144)	4092,287 (106,628)	2010,796 (291,379)	34509,709
Características del docente		0,478 (0,231)			
Tipo de Escuela (Escuela Privada)		34,435 (13,988)			
Zona (Escuela Urbana)		24,232 (10,914)			
Gusto por la escuela		18,973 (5,341)			
Gusto por la matemática		33,421 (3,460)			

Fuente: Cálculos propios. Entre paréntesis el error típico

El comportamiento de estas dos últimas variables pone de manifiesto que el aspecto motivacional de los estudiantes es uno de los que generan mayor impacto sobre su rendimiento académico.

A2. Modelos multinivel Saber 11 Lenguaje

Modelo incondicional de dos niveles

$$\text{TEMA_LENGUAJE}_{ij} \sim N(XB, \Omega)$$

$$\text{TEMA_LENGUAJE}_{ij} = \beta_{0ij}\text{constante}$$

$$\beta_{0ij} = 48.245(0.117) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 18.473(0.758) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 49.125(0.266) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 473288.606(69740 of 69740 cases in use)

Modelo incondicional al primer nivel y condicional al segundo

$$\text{TEMA_LENGUAJE}_{ij} \sim N(XB, \Omega)$$

$$\text{TEMA_LENGUAJE}_{ij} = \beta_{0ij}\text{constante} + 0.170(0.004)\text{INSE_Colegio_091}_j$$

$$\beta_{0ij} = 41.918(0.159) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 6.437(0.306) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 49.148(0.266) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 471996.881(69740 of 69740 cases in use)

INSE: índice socioeconómico de colegio Saber 11

A3. Modelos PIRLS

Diferencias en los Scores explicados por:						
Modelos Multinivel	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**
Entre Escuelas	53%	53%	57%	58%	64%	66%
Intra-escuela	47%	43%	39%	37%	31%	29%
Intra-clase	0%	5%	4%	5%	5%	5%
Número de individuos N	5131	5131	4035	4626	3981	3677
Número de grupos nivel de escuela	147	147	144	144	141	131
Número de grupos nivel de clase	---	196	193	193	190	174
Tamaño medio del grupo nivel escuela	34,9	34,9	28	32,1	28,2	28,8
Tamaño medio del grupo nivel de clase		26,2	20,9	24	21	21,1

[1]: Línea de base en 2 niveles

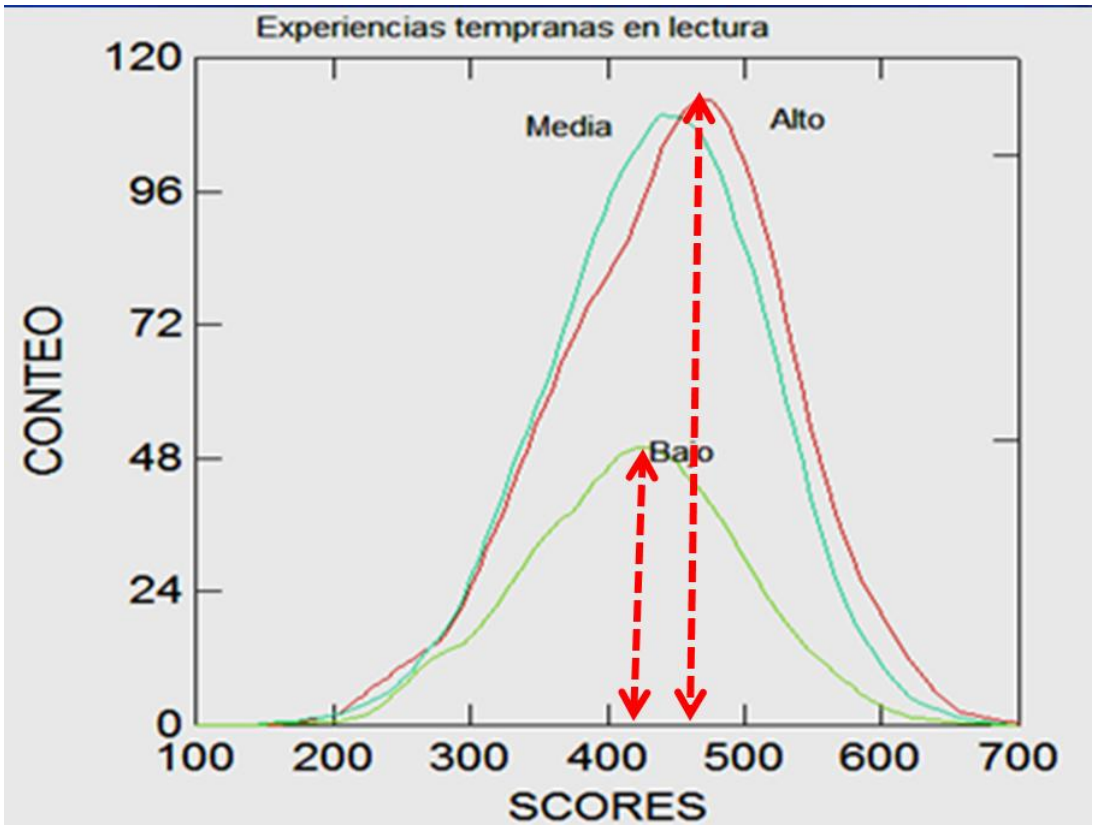
[2]: Línea de base en 3 niveles

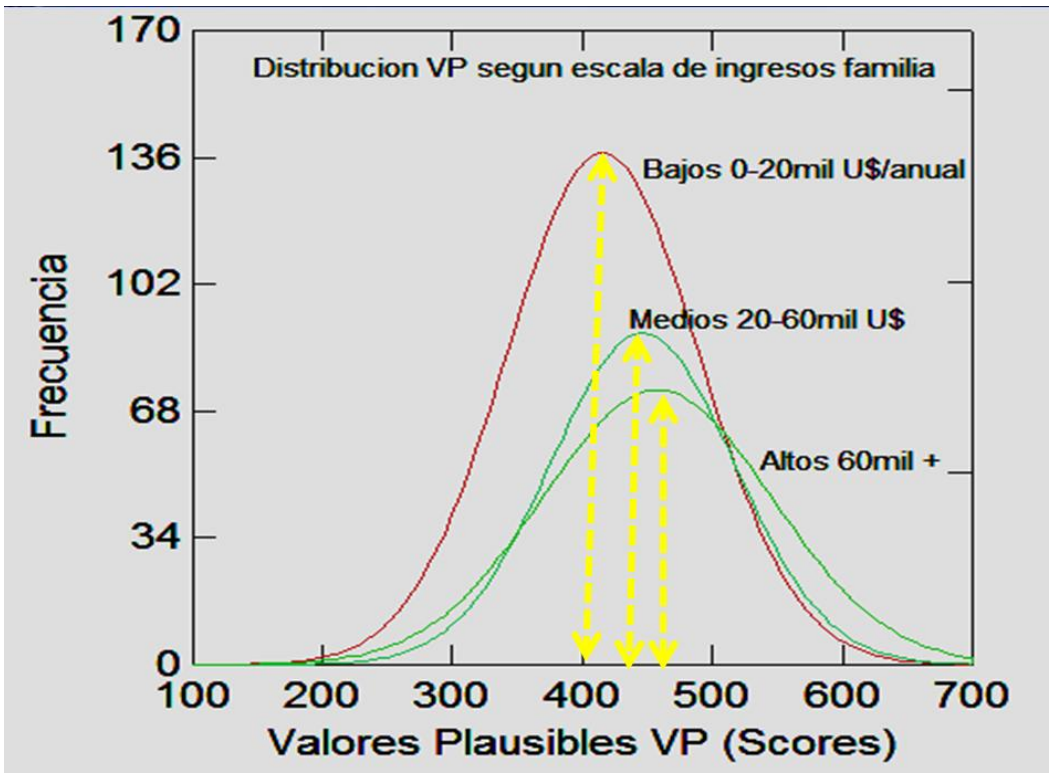
[3]: [2] + Estimulación en lectura+Educación de los padres

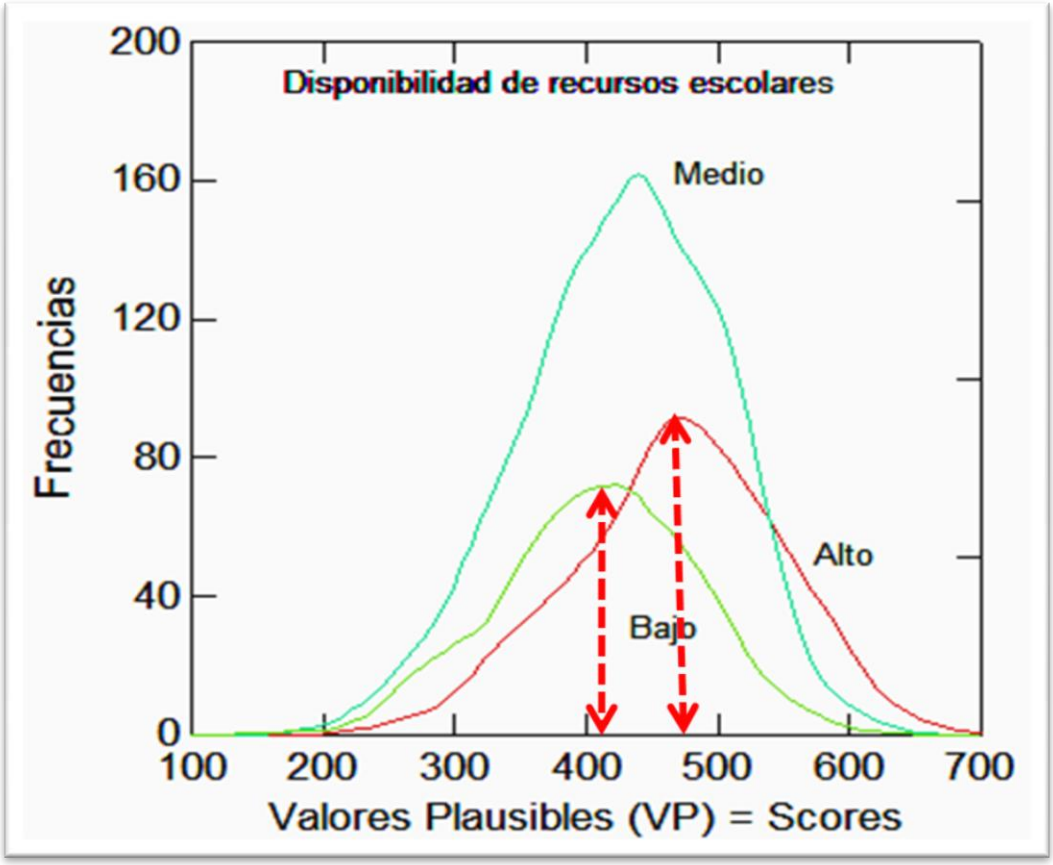
[4]: Posesión de libros en el hogar (en lugar de educ padres)

[5]: Estimulación en lectura +Posesión libros +Educación de los padres de los compañeros de clase

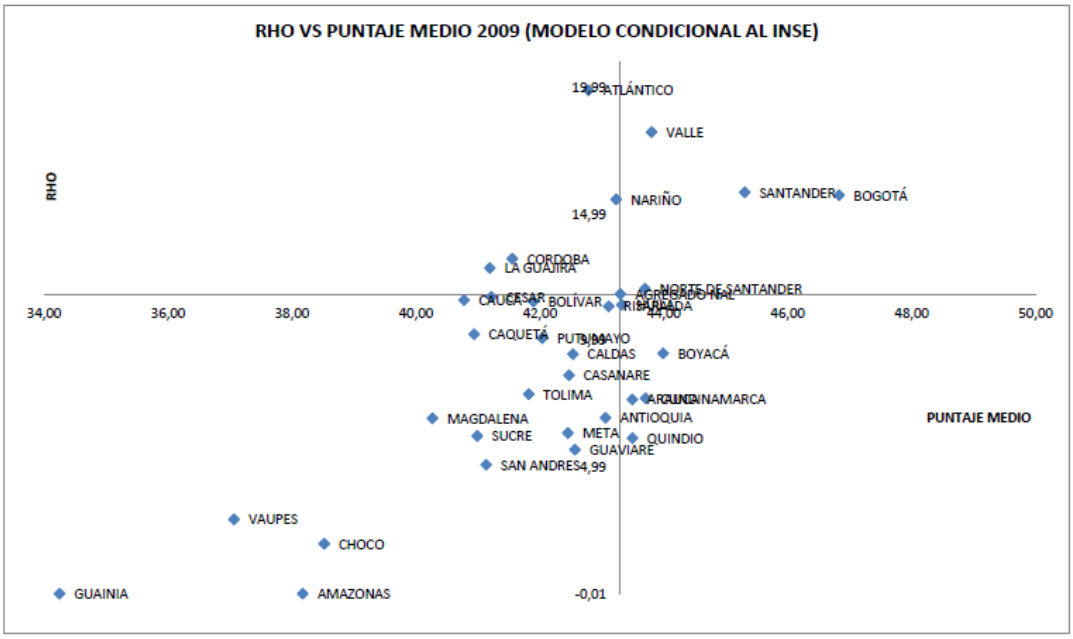
[6]: Estimulación en lectura+Posesión libros + Educación de los padres de los compañeros de clase + zona localiz escuela + Educación de los docentes




Distribucion VP segun escala de ingresos familia
Bajos 0-20mil U$/anual
Medios 20-60mil U$
Altos 60mil +
Frecuencia
170
136
102
68
34
0
100
200
300
400
500
600
700
Valores Plausibles VP (Scores)


Disponibilidad de recursos escolares
Medio
Alto
Bajo
Frecuencias
200
160
120
80
40
0
100
200
300
400
500
600
700
Valores Plausibles (VP) = Scores

A4. Correlación intra-clase (Rho) vs puntaje medio 2009 (Modelo condicional al INSE)



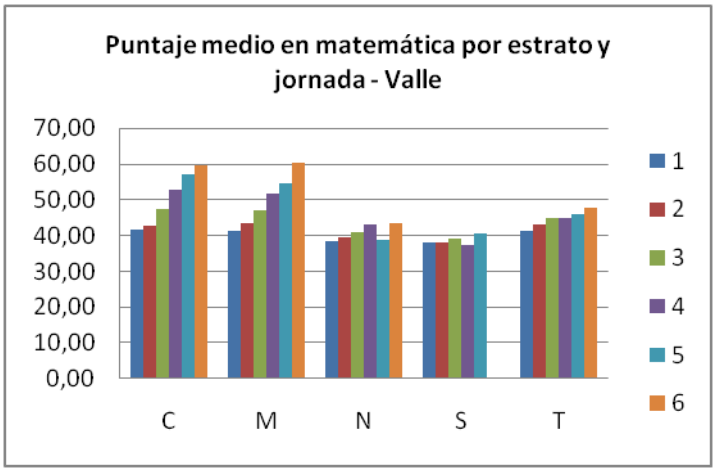
Fuente: Elaboración propia con base en las pruebas SABER 2009.

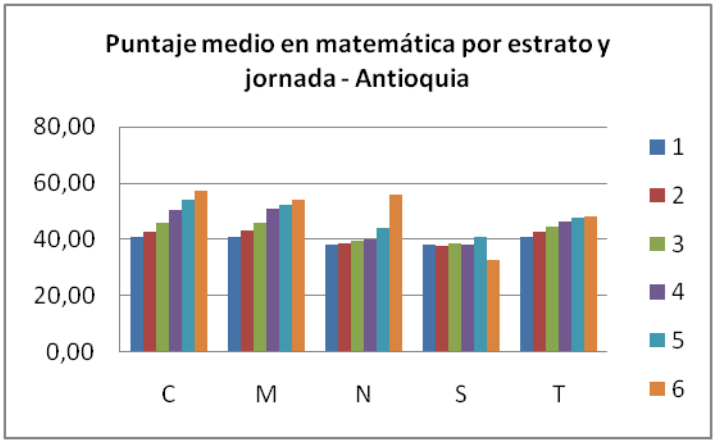
A5. Nivel de segregación según jornada escolar

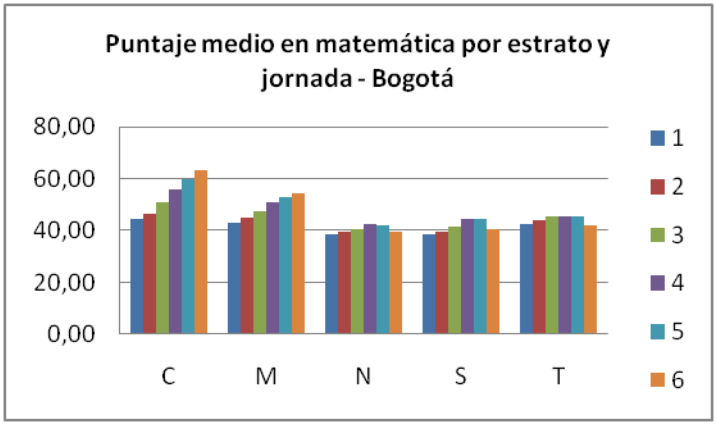
JORNADA	NULO		INSE		RHO NULO	RHO INSE
Completa	Media	45,02	Media	45,02	32,6844638	17,0926092
	desv col	6,55	desv col	4,20		
	desvest	9,40	desvest	9,25		
Mañana	Media	43,39	Media	43,39	18,3155182	10,2984872
	desv col	4,38	desv col	3,08		
	desvest	9,26	desvest	9,09		
Nocturna	Media	38,95	Media	38,95	2,35044746	5,87068993
	desv col	1,58	desv col	2,37		
	desvest	10,16	desvest	9,49		
Sabatina	Media	38,54	Media	38,54	2,13041544	7,14649546
	desv col	1,49	desv col	2,63		
	desvest	10,11	desvest	9,48		
Tarde	Media	42,79	Media	42,79	10,400842	7,64147027
	desv col	3,08	desv col	2,56		
	desvest	9,04	desvest	8,90		

Fuente: Elaboración propia con base en las pruebas SABER 2009.

A6. Nivel de segregación según jornada escolar por estrato socioeconómico

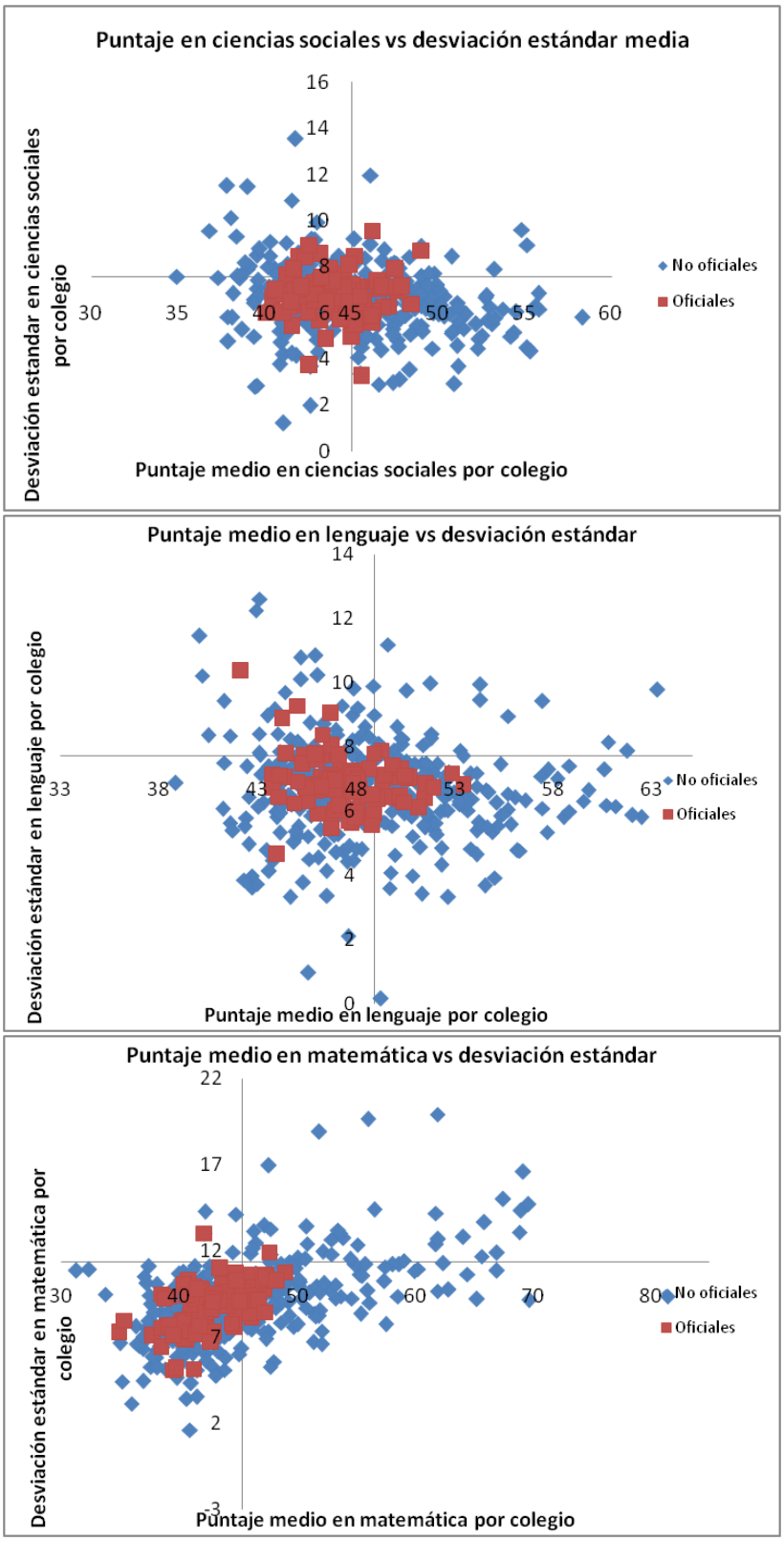






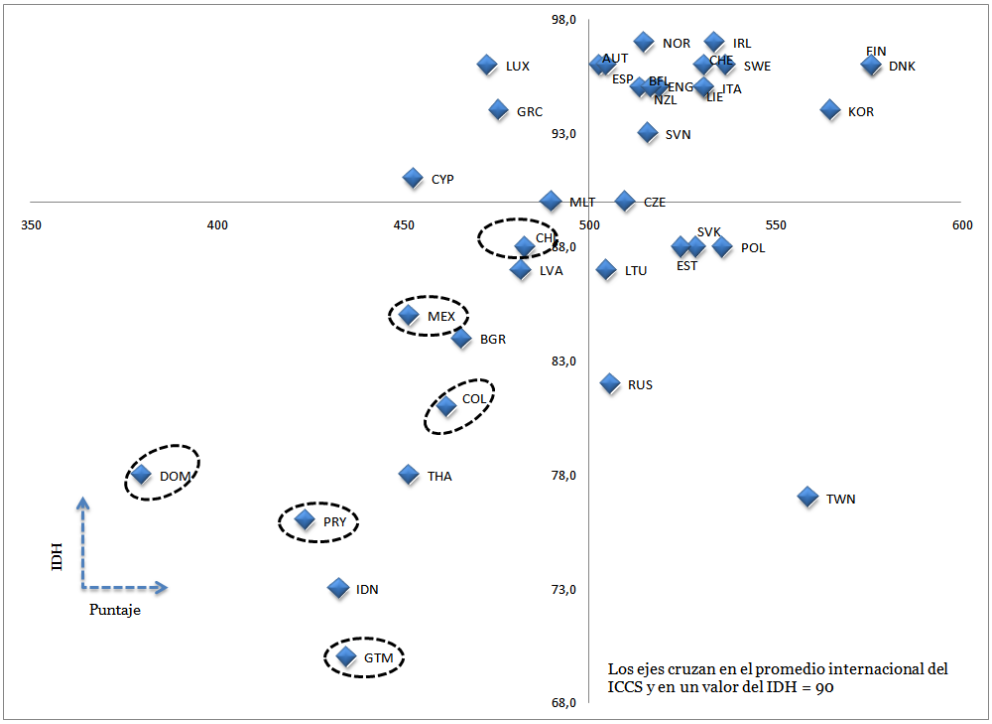
Fuente: Elaboración propia con base en las pruebas SABER 2009. C: Completa; M: Mañana; N: Nocturna; S: Sabatina; T: Tarde

A7. Nivel de segregación según naturaleza jurídica - Cali



Fuente: Elaboración propia con base en las pruebas SABER 2009.

A8. Promedio internacional de resultados del ICCS e Índice de Desarrollo Humano (IDH), 2009



Fuente: Elaboración propia a partir de la base de datos de ICCS - IEA, 2009. Los puntos con elipse corresponden a los países de América Latina incluidos en la muestra.

4. Cuadro de resultados de generación de conocimiento comprometidos contractualmente

OBJETIVOS [6] *(del proyecto aprobado)*	**RESULTADO ESPERADO**[7] *(según proyecto aprobado)*	**RESULTADO OBTENIDO**[8]	**INDICADOR VERIFICABLE DEL RESULTADO**[9]	**No. DE ANEXO SOPORTE**[10]	**OBSERVACIONES**[11]
Objetivo general Dilucidar los factores familiares y del entorno escolar que inciden en mayor proporción en la calidad educativa, medida a través de los indicadores de proficiencia de las pruebas de competencias nacionales e internacionales.					
Específico 1: Identificar a través de métodos de análisis multinivel los factores de los entornos familiares y escolares que inciden en mayor proporción en las diferencias de proficiencia.	RE1: De acuerdo con la literatura sobre los efectos del background familiar, de los pares (*peer effects*) y de los entornos escolares en los que se desarrollan los niños y jóvenes, se espera que estos factores	El artículo propone una estrategia empírica de estimación del potencial individual de logro educativo para Colombia (PCE) a través de modelos de ecuaciones estructurales (SEM). Estos resultados aportan evidencia de	Vivas Pacheco, Harvy, Correa, Juan Byron y Domínguez, Jorge Andrés. *"Potencial de logro educativo, entorno socioeconómico y familiar: una aplicación*	Anexo A	Artículo publicado en la revista Sociedad y economía, nº 21, julio – diciembre de 2011, pág. 99 – 124, Universidad del Valle. ISSN 1657-6357

[6] Se debe indicar el objetivo planteado de acuerdo con la ficha aprobada del proyecto.
[7] Se debe especificar el resultado esperado comprometido, correspondiente al objetivo planteado
[8] Elaborar una breve reseña del resultado obtenido
[9] Especificar el indicador de producto con el cual se puede verificar el logro de los resultados (artículo o libro publicado, manuscrito de artículo o libro sometido para publicación, nombre de patente presentada u homologada, norma establecida, software registrado, prototipo desarrollado, etc.).
[10] Relacionar el número del anexo que soporta o contiene el indicador del producto obtenido (copia de la publicación, patente, registro, norma, etc. o de la fuente de certificación o verificación respectiva).
[11] Incluir aquella información adicional que el investigador considere importante o necesario que Colciencias conozca, con relación al cumplimiento de los compromisos adquiridos contractualmente con el proyecto

	expliquen de manera significativa los logros académicos y los resultados en las pruebas de proficiencia.	que la calidad de los entornos locales (efectos de vecindad) presentan los efectos más grandes sobre el PCE. Las estimaciones fueron consistentes por los dos métodos utilizados (estimación bayesiana y máxima verosimilitud, M.L.).	*empírica con variables latentes para Colombia"*		
Específico 3: Comparar los resultados obtenidos en las exploraciones multinivel para Colombia con una muestra de países de América Latina incluidos en los estudios de competencias.	RE3: Se espera un fuerte nexo entre los resultados en las pruebas y los niveles de desarrollo de los países.	El artículo explora la relación entre la experiencia, las prácticas pedagógicas de los docentes y los logros académicos de los estudiantes colombianos en las pruebas de proficiencia en educación cívica y ciudadanía (ICCS, 2009) aplicada por la IEA . Los resultados de la exploración se comparan con los obtenidos por los demás países de América Latina	Vivas Pacheco, Harvy y Valencia Serna, Angélica. *"Logros académicos, experiencia y prácticas pedagógicas de los docentes en educación para la ciudadanía"*	Anexo B	Publicado en http://www.porunaeducaciondecalidad.org/Congreso/Html/IV%20Memorias.html ISBN 978-958-57047-2-5

		incluidos en la muestra.			
Específico 4: Simular, a partir de las estimaciones econométricas con modelos lineales jerárquicos, el impacto que podrían tener algunas políticas orientadas al mejoramiento de los entornos escolares en los que se forman los estudiantes.	Los ejercicios de estática comparativa que se desprendan de los diferentes modelos econométricas aportarán evidencia de la fuerte incidencia del background familiar y los entornos escolares.	El artículo propone una estrategia empírica de estimación del potencial individual de logro educativo para Colombia (PCE) a través de modelos de ecuaciones estructurales (SEM). Estos resultados aportan evidencia de que la calidad de los entornos locales (efectos de vecindad) presentan los efectos más grandes sobre el PCE. Las estimaciones fueron consistentes por los dos métodos utilizados (estimación bayesiana y máxima verosimilitud, M.L.).	Vivas Pacheco, Harvy, Correa, Juan Byron y Domínguez, Jorge Andrés. *"Potencial de logro educativo, entorno socioeconómico y familiar: una aplicación empírica con variables latentes para Colombia"*	Anexo A	Artículo publicado en la revista Sociedad y economía, nº 21, julio – diciembre de 2011, pág. 99 – 124, Universidad del Valle. ISSN 1657-6357

4.1 Otros resultados de generación de conocimiento no comprometidos contractualmente

OBJETIVOS *(del proyecto aprobado)*	**RESULTADO ESPERADO** *(según proyecto aprobado)*	**RESULTADO OBTENIDO**	**INDICADOR VERIFICABLE DEL RESULTADO**	**No. DE ANEXO SOPORTE**	**OBSERVACIONES**
Objetivo general Dilucidar los factores familiares y del entorno escolar que inciden en mayor proporción en la calidad educativa, medida a través de los indicadores de proficiencia de las pruebas de competencias nacionales e internacionales.					
Específico 3: Comparar los resultados obtenidos en las exploraciones multinivel para Colombia con una muestra de países de América Latina incluidos en los estudios de competencias.	RE3: Se espera un fuerte nexo entre los resultados en las pruebas y los niveles de desarrollo de los países.	El artículo analiza los resultados de los estudiantes de Colombia frente a los del resto de países latinoamericanos en la prueba de Educación cívica de la IEA (ICCS, 2009), en relación con sus creencias y actitudes, así como su participación en contextos relacionados con el ejercicio de la ciudadanía.	Valencia Serna, Angélica. *"Colombia en las pruebas de Educación Cívica de la IEA"*.	Anexo C	Publicado en Memorias del X Congreso Nacional de Sociología: Herencia y Ruptura en la Sociología Colombiana y Contemporánea ISBN 978-958-8357-54-6

5. Cuadro de otros resultados comprometidos contractualmente

OTROS RESULTADOS *(comprometidos contractualmente)*	**COMPROMISO ADQUIRIDO**	**LOGROS**	**ANEXO SOPORTE**
Formación de recurso humano (trabajo de grado o maestría o tesis de doctorado)	2 trabajos de maestría	**1- Logro 100%** (El trabajo obtuvo calificación meritoria) Maestría en Economía Aplicada, Universidad del Valle, Facultad de Ciencias Sociales y Económicas. Estudiante: Juan Carlos Zambrano. Tutor: Harvy Vivas Pacheco "Análisis multinivel del rendimiento escolar en matemática para grado cuarto de educación básica primaria en Colombia". Cali, 2012 **2- Logro 95%** (El trabajo se encuentra en revisión de la versión final) Maestría en Economía Aplicada, Universidad del Valle, Facultad de Ciencias Sociales y Económicas. Estudiante: Harold Calvache. Tutor: Harvy Vivas Pacheco "Incidencia del entorno escolar en los logros educativos	Anexo D

		de matemática en alumnos de secundaria: un estudio comparativo para Colombia".	
Capacitación (entrenamientos en investigación)	3 estudiantes	**Logro 100%** Entrenamiento de dos (2) monitores y dos (2) asistentes en el manejo de bases de micro-datos del ICFES y de las pruebas internacionales TIMSS, PISA e ICCS y en el tratamiento de datos anidados, especificación, estimación, análisis y validación de modelos lineales jerárquicos. Participaron tres estudiantes de la Universidad: dos estudiantes de pregrado en Economía y una estudiante de la maestría en Sociología.	Anexo E
Cursos organizados por el grupo, relacionados con el proyecto	1 curso de Economía de la educación	**Logro 100%** Se realizó el trámite de asignación de código de registro institucional y la tabla de contenido. El curso se ofrecerá en el segundo semestre del año 2012.	Anexo F
Publicaciones divulgativas (especificar)	___	___	___
Participación en eventos científicos	2 ponencias en eventos académicos	**1- Ponencia en evento internacional** Vivas Pacheco, Harvy y Valencia Serna, Angélica. *"Logros académicos, experiencia y prácticas pedagógicas de los docentes en educación para la ciudadanía"*	Anexo G

		Ponencia presentada en el I Congreso Iberoamericano / IV Nacional por una Educación de Calidad y III Encuentro de Docentes Iberoamericanos. Cartagena, octubre 24, 25 y 26 de 2011. **2- Ponencia en evento nacional** Valencia Serna, Angélica. *"Colombia en las pruebas de Educación Cívica de la IEA"*. Ponencia presentada en el X Congreso Nacional de Sociología. Cali, noviembre 2, 3 y 4, 2011.	
Organización de eventos científicos	___	___	___
Otros (especificar)	___	___	___

5.1 Cuadro de otros resultados no comprometidos contractualmente

OTROS RESULTADOS	LOGROS	ANEXO SOPORTE
Formación de recurso humano (trabajo de grado o maestría o tesis de doctorado)	Dos trabajos de grado en maestría (uno en Economía Aplicada y otro en Sociología) y dos en pregrado (Economía). Es preciso aclarar que cada uno de estos estudiantes recibió entrenamiento en el manejo de las bases de datos nacionales e internacionales, así como en el tratamiento estadístico de modelos lineales jerárquicos y en la operación de software econométrico especializado. **1- Avance 40%** Maestría en Economía Aplicada, Universidad del Valle, Facultad de Ciencias Sociales y Económicas. Estudiante: Edward Pacheco. Tutor: Harvy Vivas Pacheco *"Calidad educativa a partir de la incidencia docente y el entorno social con TIMSS 2007"*. El trabajo prosigue su avance normal en revisión de literatura y ejercicios exploratorios de modelamiento. **2- Avance 50%** Maestría en Sociología, Universidad del Valle, Facultad de Ciencias Sociales y Económicas.	Anexo H

	Estudiante: Angélica Valencia Serna. Tutor: Adolfo Adrián Álvarez *"La educación para la ciudadanía en Colombia: un análisis de las pruebas internacionales sobre educación cívica de la IEA".* El trabajo prosigue su avance normal en ejercicios exploratorios de la base de datos. La estudiante presentó una versión ampliada de esta propuesta a la convocatoria del ICFES para estudiantes de posgrado siendo seleccionado el trabajo para financiamiento. Tutor del proyecto ante el ICFES: Harvy Vivas Pacheco **3- Avance 90%** (El trabajo se encuentra en revisión de la versión final). Pregrado en Economía Aplicada, Universidad del Valle, Facultad de Ciencias Sociales y Económicas. Estudiante: Mónica F. Sánchez Tutor: Juan Byron Correa Fonnegra *"Calidad de la educación y rendimiento académico en Cali-2009"* **4- Avance 40%** Pregrado en Economía Aplicada, Universidad del Valle, Facultad de Ciencias Sociales y Económicas.	

	Estudiante: Eimmy Lara. Tutor: Harvy Vivas Pacheco *“Diferencias en la calidad educativa por regiones en Colombia”.* El trabajo prosigue su avance normal en revisión de literatura y en estadísticas descriptivas.	
Publicaciones divulgativas (especificar)	Un artículo en publicación seriada Vivas Pacheco, Harvy. *"Gran acuerdo nacional para una educación de calidad"* Artículo en la publicación periódica El Observador Regional de la Facultad de Ciencias Sociales y Económicas de la Universidad del Valle Cali, septiembre – diciembre de 2011 ISSN 2011-3420	Anexo I
Participación en eventos científicos	Una ponencia internacional y tres conferencias **1- Ponencia en evento internacional** Vivas Pacheco, Harvy y Valencia Serna, Angélica. *“Influencia del clima democrático en el salón de clase sobre los resultados en conocimiento cívico en América Latina”* Ponencia presentada en 10th Annual Conference: Education Across the Americas 2012. *Association of Latin American Students (ALAS). Teachers College, Columbia University*.	Anexo J

	Nueva York, marzo 30 de 2012. **2- Conferencia en evento internacional** Vivas Pacheco, Harvy. *"Competencias en lecto-escritura y calidad educativa en Colombia"* Conferencista invitado a la Mesa de Economía de la Educación en el marco del I Congreso Iberoamericano / IV Nacional por una Educación de Calidad y III Encuentro de Docentes Iberoamericanos. Cartagena, octubre 24, 25 y 26 de 2011. **3- Conferencia en espacio académico local** Vivas Pacheco, Harvy y Valencia Serna, Angélica. *"Clima democrático en el salón de clase y resultados en conocimiento cívico en América Latina: un estudio con las pruebas ICCS (2009) de la IEA"*. Conferencia presentada en el espacio *Charlas de los viernes* de la Facultad de Ciencias Sociales y Económicas de la Universidad del Valle Cali, abril 13 de 2012. **4- Conferencia en evento nacional** Vivas Pacheco, Harvy. *"Competencias en lecto-escritura, entorno escolar y familiar: un análisis con técnicas multinivel en Colombia"*.	

	Conferencia presentada en el Pre-congreso de estudiantes de Economía, preparatorio al Congreso de Estudiantes de Economía de FENADECO. Cali, septiembre de 2011.	
Capacitación (entrenamientos en investigación)	1- Asistencia del director del proyecto al curso de formación en manejo de bases de datos orientado por el ICFES. Bogotá, junio 20 – 24 de 2011. 2- Entrenamiento un estudiante en trámite de ceremonia de grado de pregrado en Economía en el manejo de bases de micro-datos del ICFES y de las pruebas internacionales TIMSS, PISA e ICCS y en el tratamiento de datos anidados, especificación, estimación, análisis y validación de modelos lineales jerárquicos.	Anexo K

6. Descripción del impacto actual o potencial de los resultados

Uno de los principales impactos del estudio se enmarca en su *contribución a los procesos evaluativos y en el diseño de políticas educativas nacionales, regionales y locales*. La estimación precisa de los diferentes efectos que inciden en el rendimiento de los estudiantes colombianos brinda una perspectiva más amplia a los *policy makers* para asignar los recursos del Estado orientados a la financiación de la educación pública con mayor responsabilidad, eficacia y eficiencia.

La identificación, estimación y análisis de los factores predominantes en la determinación y condicionamiento de los logros escolares de los estudiantes, aportan elementos de juicio científicamente elaborados para el diseño de políticas educativas, resaltando, como se comentó en un apartado anterior, que en un concepto integral de calidad educativa no es suficiente establecer programas de mejoramiento orientados específicamente a una u otra área de conocimiento sino que es imperativo considerar la formación ciudadana, de tal manera que las competencias cívicas entran a jugar un rol preponderante en un sistema educativo que apela por la pertinencia y calidad.

De otra parte, el desarrollo del proyecto impactó de manera positiva al Grupo GERA en términos del *fortalecimiento de la capacidad científica*, pues ofreció varios resultados que prometen futuras realizaciones académicas e investigativas, tales como:

- Amplió el margen de conocimiento en el estudio de la segregación escolar, para lo que se hizo uso de los coeficientes de correlación intra-clase, cuyos hallazgos con las pruebas SABER enriquecieron el modelo teórico de interacción social utilizado en el proyecto.

- Permitió innovar, en términos metodológicos, en el tratamiento estadístico de valores plausibles y variables actitudinales provenientes de las bases de datos internacionales, así como su utilización en modelos de ecuaciones estructurales (SEM).

- Ofreció la posibilidad de integrar un equipo de trabajo con participación de economistas, estadísticos, matemáticos y una investigadora en el campo de la pedagogía, lo que constituyó una de las principales virtudes de este estudio, permitiendo un trabajo interdisciplinario que retroalimentó a unos y otros participantes.

- Favoreció el entrenamiento de estudiantes de pregrado y de maestría en los procesos de investigación científica y, de manera particular, en el manejo de métodos y técnicas de modelamiento a partir de diseños teóricos desde el campo de la economía y la sociología de la educación, lo que representó un reto de gran trascendencia que se puede apreciar no solo en los productos del proyecto sino, a largo plazo, en el mayor número de profesionales que podrían contribuir a este tipo de estudios en el ámbito educativo.

- Facilitó la construcción de nuevas relaciones académicas a través de la participación en los eventos y el interés general que han suscitado los resultados divulgados hasta el momento. En este sentido, después de nuestra participación en el *Congreso Iberoamericano de Calidad Educativa* (Cartagena, octubre 2011) iniciamos contactos para conformar la Red de Economía de la Educación en la seccional del Occidente Colombiano. Así mismo, nuestra reciente participación como ponentes en la *Conferencia ALAS: Education Across the Americas en el Teachers College, Columbia University (Nueva York,* marzo 2012), permitió establecer vínculos tanto con investigadores nacionales como extranjeros que podrían rendir frutos a mediano plazo.

ANEXOS