

El efecto del desarrollo municipal sobre la distribución de los resultados en matemáticas para Colombia: una aproximación a partir de modelos jerárquicos lineales por cuartiles

Por

Jhon James Mora y Daniela Estrada¹

Resumen

En este artículo nosotros analizamos el efecto del entorno de desarrollo económico y la tipología del municipio sobre el desempeño que ésta tiene sobre el componente de matemáticas de la Prueba Estatal Colombiana Saber 11. A partir de modelos lineales jerárquicos con regresión cuantílicas nuestros resultados sugieren que tanto el entorno de desarrollo como la tipología municipal tiene un efecto significativo en el logro académico en matemáticas de los estudiantes para los cuartiles estudiados. De esta forma, los hacedores de política pública deberán tener en cuenta las condiciones del lugar donde viven los estudiantes a la hora de formular políticas educativas encaminadas a mejorar la calidad educativa. Especialmente, deberían enfocarse en atender las necesidades de los municipios de tipología de desarrollo incipiente, así como estudiar las prácticas pedagógicas y administrativas que se imparten en los municipios de desarrollo económico robusto. Finalmente, al comparar los resultados de la regresión multinivel cuantílica con los que se obtendrían utilizando solo regresiones cuantílicas y modelos multinivel para el estudiante promedio, se encuentra que hay diferencias en la magnitud y significancia de la asociación, lo que evidencian la importancia de emplear esta metodología.

JEL: A22, C31

Palabras Clave: Pruebas Saber 11, Regresiones multinivel Cuantílicas, Desempeño en Matemáticas

1. Introducción

Resulta preocupante que en Colombia la mayoría de los estudiantes próximos a graduarse de secundaria presenten bajos niveles de rendimiento en la competencia de matemáticas. Como se puede observar en la Figura 1, el componente de matemáticas de las pruebas del Programa Internacional para la Evaluación de los Estudiantes (o prueba PISA, por sus siglas en inglés)² es donde peor les va a los estudiantes colombianos. Para 2012, año en que ésta fue el área de conocimiento en que se profundizó, Colombia se ubicó en el puesto 62 entre 65 países que participaron en la prueba; cabe observar que tres de cada cuatro estudiantes que la tomaron fueron calificados en la categoría de bajo rendimiento (OECD, 2014). Hanushek (2013) ha encontrado que aquellos que están en la categoría de bajo rendimiento, tienen más dificultades para participar en el mercado laboral y tienen una mayor probabilidad de tener dificultades para beneficiarse de nuevas oportunidades

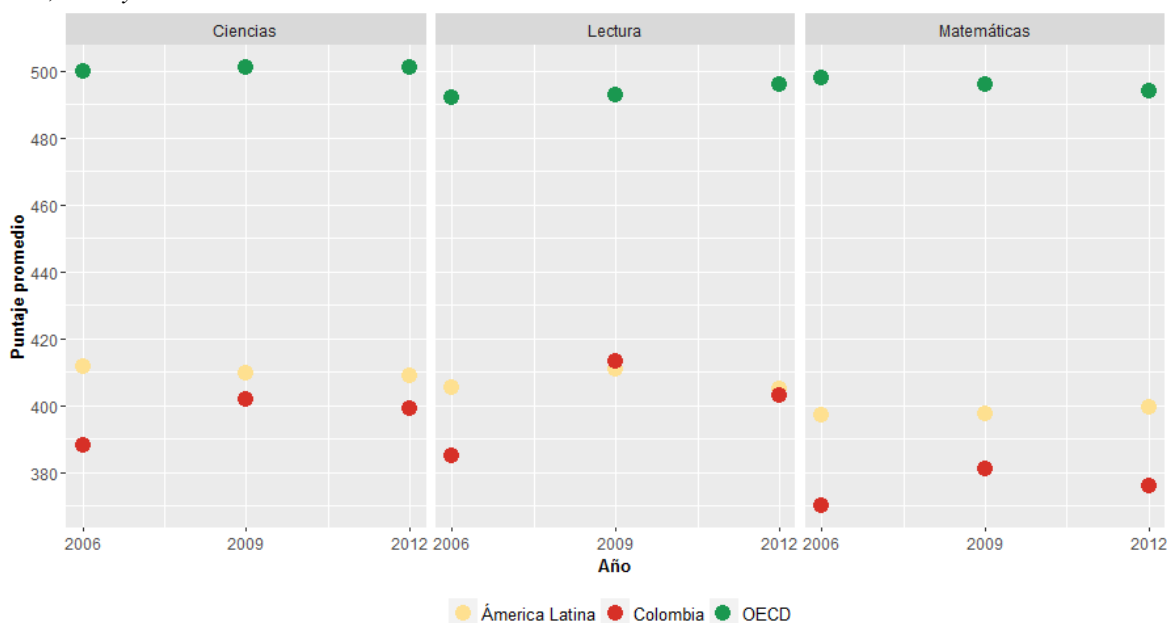
¹ Jhon James Mora, Profesor Titular del Departamento de Economía, Miembro de la Academia Colombiana de Ciencias Económicas y Director del Doctorado en Economía de los negocios. Daniela Estrada Nates, Master en Economía de la Universidad Icesi e Investigadora del CIENFI-ICESI. E.mail: jjmora@icesi.edu.co.

² La prueba PISA es una prueba estandarizada realizada cada tres años y que evalúa a estudiantes de 15 años (muestras representativas, entre 4000 y 10000 estudiantes por país) de diferentes países del mundo en tres competencias: ciencias, lenguaje y matemáticas. (OCDE, 2007)

educativas formales porque carecen de las habilidades para entender problemas complejos, utilizar fórmulas matemáticas y aplicar esos conocimientos en su vida diaria.

Además, una buena competencia matemática impacta otras dimensiones de la vida del ser humano. En la dimensión laboral, se ha encontrado que los individuos con habilidades pobres en matemáticas tienen menos probabilidades de emplearse, y si son empleados es menos probable que los promuevan (Smith, 2004). En la dimensión financiera, Litster (2013) reporta que los ingleses con bajo rendimiento en matemáticas tienden a obtener ingresos inferiores que aquellos con un buen nivel de matemáticas. Por su parte, Melzer et al. (2006) al analizar la base de datos del Estudio Longitudinal Inglés sobre Envejecimiento (ELSA por sus siglas en inglés) encontraron una fuerte correlación entre tener un alto nivel en esta competencia y la riqueza financiera acumulada a lo largo de la vida.

Figura 1: Puntaje promedio para América Latina, Colombia y países que componen OECD en pruebas PISA 2006, 2009 y 2012



Fuente: OECD PISA resultados 2006, 2009 y 2012

Nota: Los países de América Latina que toman estas pruebas son Argentina, Brasil, Uruguay, Chile y Perú; este último se tiene en cuenta a partir de 2009

Reyna y Brainerd (2007) encuentran que existe una correlación positiva entre una buena competencia matemática y la prevención y tratamiento exitosa de diferentes enfermedades; lo anterior, explican los autores, se debe a que los individuos con esta competencia pueden participar activamente en la toma de decisiones médicas, son más conscientes de la importancia de tomarse sus medicamentos y comprenden mejor la información relacionada con sus enfermedades. De igual forma, Huizinga et al. (2008) encuentra una correlación positiva y significativa entre esta competencia y un peso saludable, medido a través del índice de masa corporal.

Por su parte, DfEE (1999) halla que una buena competencia en matemáticas tiene un efecto positivo en la productividad promedio de la fuerza laboral y que ésta explica

una proporción significativa de las diferencias en el desempeño económico que hay entre los países.

Por otro lado, la estructura de la prueba PISA permite comparar el rendimiento de los alumnos a lo largo del tiempo dentro del país y con otros países; lo anterior brinda una oportunidad para generar indicadores de calidad de los sistemas educativos de los países participantes, señalar la necesidad de realizar modificaciones a estos, y evaluar si los cambios en los resultados se relacionan con la implementación de reformas en las políticas educativas y prácticas de enseñanza; en últimas, es una forma de evaluar la calidad de la educación de una región geográfica (OCDE, 2007). Esto es importante puesto que la calidad de la educación tiene efectos a nivel microeconómico y macroeconómico: impacta las probabilidades de ingresar a la educación superior (Jimenez et al., 1991), es uno de los determinantes de los ingresos y su distribución (Hanushek y Woessmann, 2008); además que tiene una correlación significativa con el crecimiento económico y la disminución de la pobreza ((Delgado et al., 2014); (Hanushek y Woessmann, 2008); Cooray (2009); (van der Berg, 2008)).

Para Colombia, Mora (2003) utilizando la prueba de Hungerford y Solon (1987) y una regresión cuantílica encuentra que las ganancias adicionales (en términos de ingresos) de la educación secundaria son del 13% para el año 2000. Las ganancias adicionales a lo largo de la distribución de ingresos varían de acuerdo al percentil analizado, de 7.6% para el percentil 20, al 17.1% para el percentil 80.

Es de observar, que la Figura 1 también envía una señal a los diferentes actores del sector educativo de que los jóvenes colombianos están recibiendo una educación que presenta deficiencias en la calidad³ puesto que el puntaje promedio de los estudiantes colombianos en las tres competencias evaluadas se ha ubicado por debajo del promedio de los países que conforman la Organización para la Cooperación y el Desarrollo Económico (OECD) y los países de la región de América Latina (Argentina, Brasil, Chile, Uruguay y Perú se incluye desde 2009) que también participan de esta prueba. El Gobierno no ha sido ajeno a esta situación y ha empezado a implementar diferentes proyectos enmarcados en el Plan Sectorial 2010-2014 y el Plan Nacional de Desarrollo 2014-2018 para solventarla. Algunos de estos proyectos son la implementación de guías y materiales didácticos para el desarrollo de competencias (Ministerio de Educación Nacional, 2010); acompañamiento y formación a docentes y directivos docentes bajo el programa Excelencia docente, Jornada Única⁴ y Educación media para todos⁵ (Nacional, 2015).

No obstante, uno de los problemas que existen para guiar la política pública a nivel municipal es que la categorización de las entidades territoriales⁶ no responde a la

³ La calidad de la educación es un concepto amplio; Rodríguez (1994) propone 4 categorías para interpretar la calidad: cobertura, gestión pedagógica, resultados y gestión administrativa. Este artículo se concentra en la calidad de la educación medida a través de los resultados de pruebas estandarizadas

⁴ Este programa consiste en extender el tiempo que permanecen los niños que asisten a colegios del sector oficial. Entre los objetivos que tiene está fortalecer las competencias básicas (como matemáticas) de los estudiantes al aumentar el tiempo dedicado a actividades pedagógicas. Se espera que su implementación se dé de forma gradual y finalice en 2025 en áreas urbanas y en 2030 en las zonas rurales (Nacional, 2015).

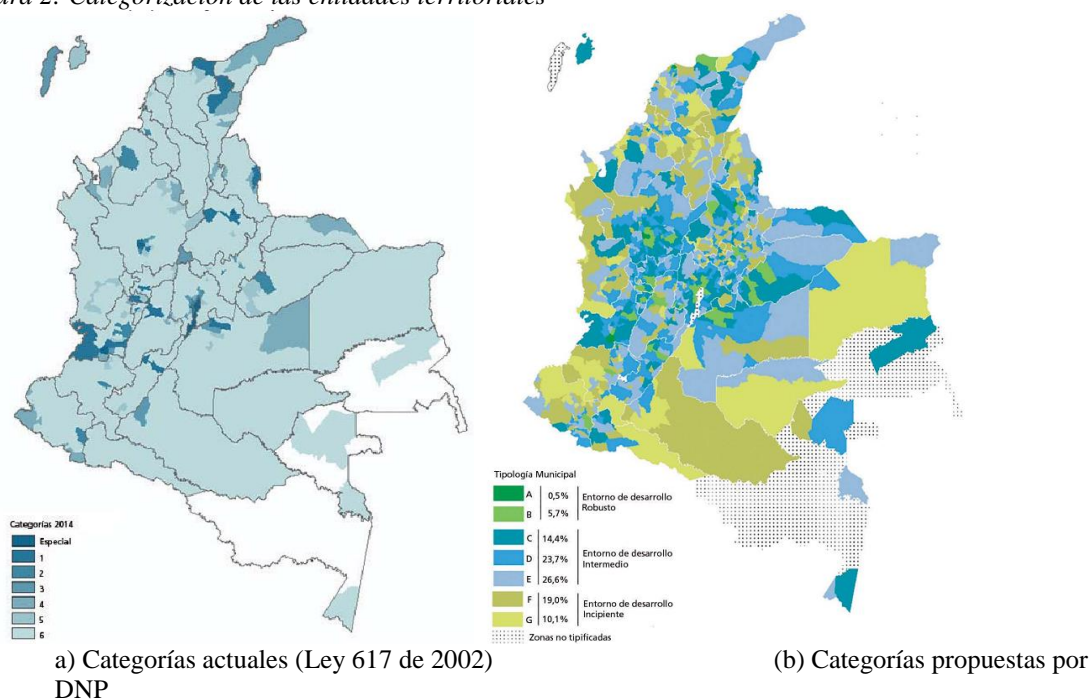
⁵ Este programa busca ampliar la cobertura en educación media del país, la cual era del 77.31% en 2015, al volverla obligatoria. Dicho plan se está implementando de manera gradual con las entidades territoriales certificadas (Nacional, 2015).

⁶ Actualmente se basa en dos criterios: tamaño de la población y los ingresos corrientes anuales (Ley 617 de 2002). Lo anterior implica que el 90% de los municipios se encuentran clasificados en la misma categoría, la sexta. Es decir en municipios con bajos ingresos corrientes anuales y una población pequeña.

heterogeneidad municipal ni refleja las brechas fuertes que existen entre las regiones en el país. Éstas contribuyen a la desigualdad (Cortés y Vargas, 2012), la cual ha persistido en el tiempo (Meisel Rosas y Romero, 2007) y que en términos educativos se evidencia en la inequidad espacial en la distribución de las dotaciones educativas (número de colegios y número de maestros) a nivel regional (Bonet, 2005) y la concentración regional de maestros con altos niveles de calificación (Galvis y Bonilla Mejía, 2011).

Como respuesta a dicha carencia, el Departamento Nacional de Planeación (DNP) propuso el diseño de una tipología municipal y departamental que permita identificar grupos más homogéneos (ver Figura 2). Se espera que ésta se convierta en una herramienta para que la política pública puede responder de manera más satisfactoria a las necesidades y características particulares de las diferentes entidades territoriales, de tal forma que la focalización del gasto público entre entidades territoriales ayude en la ejecución de programas que permitan la reducción de las brechas regionales (Sánchez et al., 2014).

Figura 2: Categorización de las entidades territoriales



Fuente: Tomado de Aguilar Londoño et al. (2015)

Hasta el momento dicha categorización no se ha empleado en algún estudio, luego partiendo tanto de la importancia de las matemáticas y de esta nueva categorización, el presente artículo espera contribuir a llenar un vacío actual en la literatura colombiana al responder: ¿Cómo afecta el entorno de desarrollo del municipio donde vive el estudiante a la distribución de los resultados del componente de matemáticas de la prueba Saber 11 cuando se controla por las características del estudiante y de la institución educativa (específicamente la naturaleza jurídica y el tipo de jornada diurna)?

Para lograr lo anterior se emplean los resultados del componente de matemáticas Saber 11 del 2014-2, y las categorías de entorno de desarrollo y tipología municipal propuestas por el DNP mediante la implementación de un modelo lineal jerárquico por cuartiles. Con lo anterior se busca analizar si el entorno de desarrollo municipal está

relacionado con los resultados en matemáticas (Saber 11) por cuartiles (0.25, 0.5 y 0.75) cuyo se controla por las características individuales y familiares del estudiante, y del colegio. Asimismo, evaluar si las tipologías municipales (que conforma los entornos de desarrollo) ayudan a explicar las variaciones en el puntaje de matemáticas para cada uno de los cuartiles estudiados. Por último, comparar estos resultados con los que se obtendría empleando una regresión cuantílica y un modelo jerárquico lineal, de manera independiente.

Este artículo contiene además de la introducción, 5 secciones adicionales. En la siguiente se describen los factores que pueden afectar el desempeño en matemáticas, y la evidencia empírica a nivel internacional y nacional que se ha encontrado al respecto. En la tercera parte se explica la aproximación empírica. Posteriormente, se caracterizan los datos empleados y se presentan las estadísticas descriptivas. En la quinta sección se encuentran los resultados de estimar el modelo jerárquico lineal por cuartiles y su respectivo análisis. En la última sección se encuentran las conclusiones.

2. Revisión de la literatura

Esta sección presenta los diferentes modelos que se pueden emplear para responder la pregunta planteada, y ejemplos de los hallazgos que se han encontrado en el ámbito internacional y nacional sobre los factores que pueden afectar el rendimiento en matemáticas de los estudiantes. Siguiendo a Restrepo y Alviar (2005) esos insumos de educación se pueden agrupar en las siguientes categorías:

- Individuales
- Familiares
- Colegio
- Factores del lugar de residencia

Generalmente para estimar cuáles factores están relacionados con los rendimientos en pruebas estandarizadas se emplean funciones de producción de educación (Hoeryner y Lemke (2006); Hanushek (1986)). Los resultados pueden variar dependiendo del contexto que se analice, de la riqueza de la base de datos empleada para caracterizar a los colegios e individuos y si el estudio se enfoca en los resultados del estudiante promedio o en la distribución de puntajes obtenidos por los estudiantes.

2.1. Modelos

Es importante resaltar, que el desempeño en una prueba en particular puede analizarse tanto a nivel de estudiante, a nivel de colegio o utilizando modelos multiniveles (es decir modelos lineales jerárquicos que incluyen la estructura jerarquizada natural de los datos donde los estudiantes se agrupan en salones que a su vez se agrupan en colegios).

De acuerdo con Hox (2002) este último tipo de modelos es ampliamente utilizado en educación pues cubre dos de los inconvenientes de los modelos anteriores. El primero es que corrige el problema estadístico derivado de la falta de independencia entre los factores asociados al logro dentro de un mismo colegio, el cual se presenta a nivel estudiante cuyo se estima a través de mínimos cuadrados ordinarios. El segundo, evita la pérdida de

información que se da en los modelos a nivel de colegio debido a la agregación, la cual puede llevar a una pérdida de poder estadístico. Por otro lado, evita un problema conceptual que se puede dar en el momento de interpretar los resultados para un nivel de la jerarquía incorrecto. Uno de esos problemas es el de la falacia ecológica donde se infiere información que ha sido agregada (por ejemplo, a nivel colegio) como si fuera un resultado a nivel de estudiante. El otro es el problema de falacia atomista que se da cuyo a partir de las asociaciones encontradas a nivel individual se infiere sobre la variabilidad grupal (Hox,2002).

Por otras parte, cuyo se quiere analizar al estudiante, ya sea empleando un modelo a nivel de estudiante o multinivel, se puede buscar entender cómo los factores afectan al estudiante promedio (por ejemplo, empleando mínimos cuadrados ordinarios) o cómo varía el efecto a lo largo de la distribución de los estudiantes al escoger diferentes percentiles de la distribución y observar si los factores elegidos afectan de igual forma a todos los estudiantes (ver Sección 3 para un detalle de estos métodos de estimación). Este último método de estimación se llama regresión cuantílica y es un método que ha cobrado importancia puesto que diferentes estudios (Eide y Showalter (1998); Koenker y Bassett (1978)) muestran que las variables que son significativas para el estudiante promedio no lo son necesariamente para los estudiantes a lo largo de la distribución de los resultados. Lo anterior puede resultar crucial para fines de política pública ya que al utilizar esta metodología no sólo se responde si un factor es importante, también para quién lo es.

2.2. La literatura Internacional

A continuación, para cada una de las categorías expuestas anteriormente, se expondrán algunos de los resultados encontrados en la amplia literatura que hay sobre el tema.

Con respecto a los resultados individuales, algunos estudios han encontrado que los hombres tienden a obtener mejores resultados que las mujeres (Demir et al., 2010). Sin embargo, diferentes estudios de meta análisis encuentran que la brecha es pequeña y no es significativa. Por otra parte, hay estudios que incluyen la raza del individuo; por ejemplo, Haile y Nguyen (2008) concluye que los individuos afroamericanos e hispanos tienen menores rendimientos en matemáticas, ciencias y lectura que los individuos blancos, lo anterior aplicando una regresión cuantílica a los datos para el 2000 del Estudio Longitudinal Nacional de Educación (NELS por sus siglas en inglés).

Con respecto a los resultados familiares, se ha encontrado que el nivel socioeconómico es uno de los determinantes del desempeño en matemáticas. Bassett et al.(2002) encuentra que los estudiantes de las familias más pobres tienen un peor desempeño en la Prueba Universitaria Americana (ATC por sus siglas en inglés) que aquellos que provienen de una familia acaudalada. Chiu y Xihua (2008) encuentra algo similar para 41 países, pero empleando un índice creado a partir de los años de educación de la madre y el padre, y el estatus laboral más importante entre los dos. De igual forma hay estudios que controlan por el nivel educativo de los padres (Bassett et al. (2002);Haile y Nguyen (2008)), ocupación de los padres (Haile y Nguyen, 2008), número de hermanos (Chiu y Xihua, 2008), entre otras.

Con respecto al colegio, Bassett et al. (2002) encontró que el porcentaje de maestros con maestría y la razón profesor estudiante son significativas y tienen una relación positiva con los puntajes en la prueba ACT lo largo de la distribución para una escuela pública de Illinois, Estados Unidos. El efecto es positivo para la cola inferior de la distribución y negativo en la cola superior para el caso del porcentaje de profesores con maestría.

Mientras que Eide y Showalter (1998) no encuentra que estas variables sean significativas empleando la base de datos nacional y longitudinal HS&B (High School y Beyond).

Por otro lado, Hanushek (2002) investiga la relación entre el tamaño de las clases y las variables de rendimiento de los estudiantes. Los resultados son mixtos. De los 277 estudios que revisa, informa que el 14% muestran que la relación entre el tamaño de la clase y rendimiento de los estudiantes es significativo y positivo, mientras que el 14% muestran que la relación es significativa, pero negativa, y el 72% de los estudios encuentra una relación insignificante.

Asimismo, hay estudios que incluyen la naturaleza jurídica de los colegios (si son privados o públicos) y los resultados también son mixtos. Por ejemplo, Lubienski y Lubienski (2006) al analizar los resultados de matemáticas de la Evaluación Nacional del Progreso Educativo (NAEP, por sus siglas en inglés) obtienen como principal resultado que, si no se controla por las características personales y familiares de los individuos, los estudiantes que asisten a colegios privados obtienen en promedio puntajes mayores que aquellos que asisten a colegios públicos. Sin embargo, una vez se controla por esas características el efecto desaparece. Por el contrario, Jiménez et al. (1991) hallan para Colombia, República Dominicana, Filipinas, Tanzania y Tailandia que incluso si se controla por el ingreso del jefe del hogar, el nivel educativo de los padres y la ocupación de estos, encuentran que los estudiantes que asisten a colegios privados obtienen mejor desempeño que los que asisten a colegios públicos. Por otro lado, hay documentos que han analizado cómo la duración de la jornada escolar afecta el rendimiento escolar. Lavy (2015) utiliza los datos de PISA para 2006 (50 países) y encuentra evidencia de que el tiempo de instrucción está positiva y significativamente relacionada con las calificaciones de los exámenes, tanto en los países desarrollados como los países en desarrollo. Rivkin y Schiman (2015) realizan un estudio similar empleando los resultados de la prueba PISA 2009 (72 países) y encuentran que entre mayor sea la duración de la jornada escolar mejor es el rendimiento de los estudiantes.

Con respecto al lugar de residencia, existen análisis desde una perspectiva de país, de departamento, municipio, ciudad o incluso de barrio. Desde la perspectiva del país de residencia se encuentran los artículos de Chiu y Xihua (2008) y Fryer Roly G. y Levitt (2010). El primero incluye PIB per cápita, gini, y grado en que la cultura es individualista para caracterizar a 41 países. Los resultados evidencian que los valores culturales no son significativo para explicar el desempeño en matemáticas; sin embargo, los países más inequitativos en términos de ingresos obtienen puntajes menores, de igual forma aquellos que son más pobres. Por su parte, el segundo artículo incluye un índice de inequidad por género y encuentra que las brechas son casi nulas en países con grandes grados de inequidad por género.

En la revisión de la literatura internacional en el campo educativo sólo se encontraron dos estudios que estiman un modelo multinivel empleando un método para tener en cuenta la distribución de los puntajes de los estudiantes. El primero es de Tian y Chen (2006), quien estudia la relación entre el desempeño en matemáticas de alumnos de secundaria (de décimo a doceavo grado) de 35 colegios en Alberta (Canadá) con otros factores sociales, económicos y culturales. Su principal conclusión es que el tamaño de los colegios tiene poco efecto en los resultados en matemáticas, mientras que el suburbio en el que está localizado el colegio explica en una medida significativa las variaciones en el desempeño. El segundo es el de Costanzo (2015), quien responde cómo y a quiénes afecta un programa enfocado en mejorar los resultados en las pruebas de matemáticas mediante el entrenamiento de profesores en la aplicación de las matemáticas a la vida diaria. Utiliza un

modelo de dos niveles donde los estudiantes se agrupan en salones, y lo compara los resultados de estimar el modelo multinivel solo. Encuentra que el programa tiene un efecto significativo y positivo en el desempeño de los estudiantes (tanto el estudiante promedio como a lo largo de la distribución), además que éste beneficia más a los estudiantes que se encuentran en la cola inferior de la distribución que a los que se encuentran en la cola superior.

2.3. La literatura nacional

A continuación, se presentan los resultados de varios estudios realizados para el contexto colombiano por las categorías de factores que se han encontrado que están correlacionados con el rendimiento en matemáticas. En este ámbito hay estudios tanto a nivel de estudiante, como a nivel del colegio y multiniveles. La mayoría emplean los resultados de las pruebas Saber 11 ejecutadas por el Instituto Colombiano para el Fomento de la Educación Superior (ICFES).

Independientemente de los datos empleados y el modelo, el género y la edad son las variables más utilizadas dentro del grupo de factores personales. Estas variables tienden a presentar el mismo signo, el cual indica que los hombres en promedio obtienen puntajes en matemáticas superiores a las mujeres, y que entre mayor es la edad del estudiante menor es su puntaje (ver Cuadro 1).

Cuadro 1: Evidencia nacional de factores personales y familiares asociados al puntaje de matemáticas.

Autor	Datos	Modelo	Factores			
			Hombre	Edad	Condiciones económicas	Nivel educativo padres
Piñeros, L. y Rodríguez, A. (1998)	Saber 11 - 1997 II	Multinivel	(+)	N.I	NSE: +	N.I
Correa, J.J. (2004)	Puntaje total: Saber 11 2001-I. Cali	Multinivel	(+)	(-)	Ingresos familiares: +	Padre con el mayor nivel educativo (+)
Barrientos, J. (2008)	Saber 11 - 1999-2004-2006	Estudiante	(+)	(-)	Estrato: (+/-)	N.I
Bonilla, L. (2011)	Puntaje total: Saber 11 - 2009	Estudiante	(+)	(-)	Ingresos familiares: +	Madre (+)
Barón (2012)	Saber 11 2009: Barranquilla	Estudiante	NA	N.I	Ingresos familiares: +	Ambos padres (+) para mujeres y mixto para hombres
	Saber 11 2009: Bogotá	Estudiante	NA	N.I	Ingresos familiares: +	Ambos padres (+) para mujeres y hombres

Montes, I., Garcés,
J., Chica, S., y
Jaramillo, A.
(2014)

Saber 11 - 2010 Multinivel (+) (-) NSE: + N.I

Nota: N.I Variable no se incluyó y NA: no aplica * Sólo se incluyeron los años de la base de datos de Saber 11 utilizada y la variable dependiente en caso de que no sea el puntaje en matemáticas † Sólo se incluyeron aquellas variables que fueron significativas al menos al 95 % de confianza ‡ Regresión cuantílica.

Fuente: Elaboración propia.

En cuanto a las condiciones socioeconómicas del hogar se encuentra que entre mejores sean éstas (independientemente de cómo se midan, nivel socioeconómico (Piñeros y Rodríguez, 1998), estrato (Marín, 2008) o los ingresos familiares medidos por medio de una variable categórica y ordinal (Barón, 2012) tienen un efecto positivo en el puntaje en matemáticas. No obstante, (Barrientos, 2007) encuentra que para los estudiantes que se encuentran en el percentil más bajo (0.05) pertenecer a un estrato más alto no es significativo, mientras que para los demás percentiles sí lo es.

Otra variable que se incluye bajo este grupo de factores es el nivel educativo de los padres (de la madre, el padre o ambos). En términos generales aquellos estudiantes cuyos padres han alcanzado un mayor nivel educativo obtienen en promedio mejores puntajes. No obstante, eso puede diferir por regiones; por ejemplo, Barón (2012) halla que para los hombres de Barranquilla sólo el nivel educativo de la madre es significativo, mientras que para los hombres de Bogotá ambos son significativos.

Con respecto al colegio, las variables más utilizadas son la naturaleza jurídica de la institución (si es oficial o no) y la jornada escolar. Como se puede observar en el Cuadro 2, en dos de tres estudios que incluyen esta variable se encuentra que aquellos que pertenecen al sector oficial obtienen en promedio mejores resultados que aquellos que estudian en instituciones no oficiales. En cuanto a la jornada escolar, se observa que los estudiantes que asisten a jornada completa (es decir que reciben 2 horas adicionales de clase respecto a las otras jornadas diurnas) obtienen mejores resultados que aquellos que asisten a la jornada de la tarde o mañana; mientras que, el signo que acompaña a las otras jornadas puede variar dependiendo de cuál sea la jornada base.

Cuadro 2: Evidencia nacional de factores del colegio asociados al puntaje de matemáticas.

Autor	Datos	Modelo	Factores colegio	
			Oficial	Jornada escolar
Piñeros, L. y Rodríguez, A. (1998)	Saber 11 - 1997 II	Multinivel	N.I.	Mañana (+)
				Tarde (+)
				Completa (+)
Correa, J.J. (2004)	Puntaje total: Saber 11 2001-I. Cali	Multinivel	(+)	Mañana (+) Tarde (-) Noche (-)
Barrientos, J. (2008)	Saber 11 1999- 2004-2006	Estudiante	N.I.	Tarde (+) : 0.25, 0.5 y 0.75
Bonilla, L. (2011)	Puntaje total / Saber 11 - 2009	Estudiante	(+)	Completa (+)
Barón (2012)	Saber 11 2009: Barranquilla y Bogotá	Estudiante	N.I.	Completa o mañana (+)

Montes, I., Garcés, J., Chica, S., y Jaramillo, A. (2014)	Saber 11 2010	Multinivel	(-)	Mañana (+)
				Tarde (+)
				Completa (+)
Hincapié, D. (2016)	Saber 5 y 9 - 2002, 2005, y 2009.	Estudiante	NA	Completa (+)

Nota: N.I Variable no se incluyó y NA: no aplica * Sólo se incluyeron los años de la base de datos de Saber 11 utilizada y la variable dependiente en caso de que no sea el puntaje en matemáticas † Sólo se incluyeron aquellas variables que fueron significativas al menos al 95 % de confianza ‡ Regresión cuantílica.

Fuente: Elaboración propia.

De igual forma se pueden incluir otras variables como modalidad (técnico, académico, o ambos) (Piñeros y Rodríguez, 1998); género de la institución (masculino, femenino o mixto) (Bonilla Mejía, 2011), valor de la pensión (Barón, 2012), características de los docentes y estrategia de enseñanza (Montes et al., 2014), entre otras. La cantidad de variables que se incluyen depende del objetivo del investigador y de la calidad de las bases de datos con la que éste cuenta.

Con respecto al lugar de residencia del estudiante existen pocos estudios que contemplen variables de este grupo. Por ejemplo, Restrepo y Alviar (2005) buscan determinar cuál es el efecto colegio sobre la variabilidad de los resultados de Saber 11 en el departamento de Antioquia encuentran que a nivel departamental la naturaleza jurídica y la clase de jornada no son significativas. Sin embargo, cuando analizan por subregiones encuentra que asistir a la jornada de la mañana está correlacionada con menores puntajes en las subregiones diferente al Valle de Aburrá. De igual forma, los resultados evidencian que los estudiantes que atienden colegios públicos tienen un desempeño menor a los estudiantes que asisten a colegios privados en esas subregiones.

Por otro lado, Bonilla Mejía (2011) incluye características de diferente índole (sociales, educativas, económicas, entre otras) de las 10 ciudades principales del país y aunque encuentra que éstas son significativas no discute las implicaciones de estos factores en términos de política pública. Lo mismo sucede con Martínez y Turriago (2015), quienes realizan un análisis gráfico y estadístico (empleando el coeficiente de correlación y el coeficiente de determinación) y evidencian que menores índices de necesidades básicas insatisfechas y un mayor nivel de preparación de los docentes se asocian directamente con el puntaje total de los estudiantes en las pruebas.

Cuadro 3: Evidencia nacional de factores del lugar de residencia asociados al puntaje de matemáticas.

Autor	Datos	Modelo	Municipio	Hallazgos
Restrepo, P.P., y Alviar, M. (2004)	Saber 11- 1999	Multinivel	Subregiones Antioquia	Oficial (-) Jornada mañana (-)
Bonilla, L. (2011)	Puntaje total Saber 11 - 2009	Estudiante	10 ciudades principales	% Urbano (-) % NBI (-) % analfabetismo (-) % homicidios (-) Desempeño fiscal (+) NBI (-)
Martínez, W.H., y Turriago, A. (2014)	Puntaje total: Saber 11 - 2005-2012, semestre II	NA	Todos	% Alumnos del sector privado (+) % Profesores con educación superior (+)

Nota: N.I Variable no se incluyó y NA: no aplica * Sólo se incluyeron los años de la base de datos de Saber 11 utilizada y la variable dependiente en caso de que no sea el puntaje en matemáticas † Sólo se incluyeron aquellas variables que fueron significativas al menos al 95 % de confianza.

Fuente: Elaboración propia.

En síntesis, dentro de la revisión de la literatura colombiana se encontraron dos vacíos que este estudio pretende empezar a cubrir al emplear un modelo multinivel por cuantiles, el cual no ha sido utilizado con anterioridad; y por otro lado, reconocer la heterogeneidad municipal, la cual ha sido ignorada en la mayoría de los estudios puesto que estos emplean los puntajes en matemáticas desde la perspectiva nacional o de las ciudades principales como Bogotá, Cali, Medellín y Barranquilla.

3. Aproximación empírica

Las aproximaciones empíricas utilizadas para analizar el efecto que tienen diferentes factores en el desempeño escolar son: regresiones por cuantiles (QR por sus siglas en inglés) y regresiones de modelos lineales jerárquicos. Sin embargo, la regresión por cuantil no puede lidiar con datos cuya naturaleza es jerárquica, como en el caso educativo, donde los estudiantes se agrupan en salones y estos a su vez en colegios. Parte de la solución a ese inconveniente se ha dado en años recientes con el desarrollo de modelos y algoritmos que permiten aplicar regresiones por cuantiles a datos con una estructura jerárquica (e.g. Tzavidis y Brown (2010) Geraci y Bottai (2007), Smith et al. (2015)). En este caso se estimará el modelo empleado el método de regresión por cuantiles con efectos mixtos (LQMM) de Geraci y Bottai (2014) en el software R (2013), el cual permite estimar el

cuantil condicional de la variable dependiente en función de efectos fijos y aleatorios. A continuación, se presentan las principales características de los tres métodos.

3.1. Regresión por cuartiles

La regresión por cuartiles fue propuesta por Koenker y Bassett (1978) y consiste en minimizar las desviaciones estándar de los residuos (ver Ecuación 1). Entre sus principales ventajas están que permite analizar toda la distribución condicional de la variable dependiente y no sólo la media condicional como lo hace el método de mínimos cuadrados ordinarios. Este resultado es importante porque al estudiar diferentes cuartiles se puede entender no sólo si un factor es importante sino para quién lo es (Costanzo, 2015). Asimismo, este modelo es robusto a los supuestos que se realicen sobre la distribución que sigue el residuo; lo anterior, facilita el cálculo y la interpretación del modelo (Koenker, 2005).

$$\text{Min}_{\beta \in R^k} \sum_{i: y_i \geq x_i \beta} \theta |y_i - x_i \beta| + \sum_{i: y_i \leq x_i \beta} (1 - \theta) |y_i - x_i \beta| \quad (1)$$

Donde y_i es la variable dependiente, x_i es vector de k por 1 que contiene las variables explicativas, θ es el cuartil a ser estimado, y β son los coeficientes, los cuales pueden variar dependiendo del cuartil que se analice.

3.2 Regresión de modelos lineales jerárquicos

Los modelos lineales jerárquicos también conocidos como modelos multinivel o de efectos aleatorios se caracterizan por permitir la correlación entre las observaciones que pertenecen a una misma unidad, como es el caso de estudiantes que pertenecen a un mismo colegio (Raudenbush y Bryk, 1986). También permiten estimar la variabilidad al interior de los grupos (colegios) al incluir efectos aleatorios (ver Ecuación 2) y estudiar de manera simultánea la relación dentro de un nivel jerárquico dado, y entre los niveles. Además, se puede estimar el coeficiente de correlación intraclase ($\text{ICC} = \frac{\Psi_u^2}{\Psi_u^2 + \Psi^2}$), el cual es una medida de cuánto de la varianza total se puede explicar por la clusterización de los datos (Raudenbush y Bryk, 1986).

$$y_{ij} = \theta_{00} + \sum_{k=1} \theta_{0k} X_{ijk} + \sum_{p=1} \theta_{p0} Z_{pj} + \sum_{p=1} \sum_{k=1} \theta_{pk} X_{ijk} Z_{pj} + U_{j0} + \sum_{k=1} U_{jk} X_{ijk} + R_{ij} \quad (2)$$

Donde y_{ij} es la variable dependiente; X_{ij} es la matriz de variables independientes que caracterizan al individuo i que hace parte del colegio j ; Z_{pj} es la matriz de variables independientes que caracterizan al colegio j ; $X_{ij} Z_{pj}$ corresponde a las interacciones entre ambos niveles de la jerarquía; U_{j0} y U_{jk} son los efectos aleatorios y R_{ij} es el término de error del individuo i asociado al colegio j .

3.3 Regresión por cuartiles de modelos lineales jerárquicos

El modelo LQMM (GeracyBottai, 2014) combina las características anteriormente mencionadas. Las covariables pueden ejercer diferentes efectos dependiendo del cuantil que se analice, tal como sucede en la regresión por cuantil estándar. El grado de heterogeneidad no observada, como en los modelos lineales jerárquicos, se caracteriza con

los parámetros específicos de la varianza, la cual depende del percentil estudiado y que se identifica por θ . La ecuación básica que se estima es:

$$y_{ij} = x_{ij}^T \beta_{\theta} + z_{ij} u_j + \varepsilon_{ij} \quad (3)$$

Donde y_{ij} es la variable dependiente, x es el vector de efectos fijos, z es en un vector de efectos aleatorios, ε es el término de error que sigue una distribución de Laplace (Sánchez y Lachos, 2013; Geraci y Bottai, 2014; Marino y Farcomeni, 2015).

En términos generales, la estimación se basa en la relación existente entre la minimización de las desviaciones estándar de la regresión cuantílica y la maximización de una distribución Laplace. La estimación de parámetros se lleva a cabo mediante el uso de la función de distribución asimétrica de Laplace que se ha introducido en el contexto de regresiones por cuantiles por Yu et al. (2003) y Koenker y Machado (1999). Para un detalle del algoritmo de estimación ver Geraci y Bottai (2007) y Geraci y Bottai (2014).

3.2. Aproximación a seguir

1. Estimar el modelo vacío o de efectos fijo (Ecuación 4) y calcular el coeficiente de correlación interclase (Ecuación 5):

$$y_{ij} = 1 * \theta^{\tau} + 1 * u_j + \varepsilon_{ij}^{\tau} \quad (4)$$

$$ICC = p^{\tau} = \frac{\Psi_{u_{\tau}}^2}{\Psi_{u_{\tau}}^2 + \Psi_{\varepsilon_{\tau}}^2} \quad (5)$$

Donde $\Psi_{\varepsilon_{\tau}}^2$ representa la varianza al interior del colegio (entre los alumnos que se encuentran en el cuartil τ , y $\Psi_{u_{\tau}}^2$ representa la varianza entre colegios en el cuartil τ . Luego $\Psi_{u_{\tau}}^2 + \Psi_{\varepsilon_{\tau}}^2$ indica la varianza total.

2. Estimar el modelo incluyendo la variable independiente del entorno de desarrollo:

$$y_{ij} = x_{ij}^T \theta^{\tau} + z_{ij}^T u_j + \varepsilon_{ij}^{\tau} \quad (6)$$

Donde

- τ : son los percentiles 0.25, 0.5 y 0.75
- y_{ij} : es el puntaje en matemáticas del estudiante i que asiste al colegio j
- $x_{ij}^T = (1, \text{munintermedia}_{ij}, \text{munincipiente}_{ij}, \text{edad}_{ij}, \text{mujer}_{ij}, \text{msecundaria}_{ij}, \text{mterciaria}_{ij}, \text{oficial}_{ij}, \text{jmanana}_{ij}, \text{jtarde}_{ij})$ (Ver Anexos para una descripción de cada una de las variables.)
- $\theta^{\tau} = (\theta_0^{\tau}, \theta_1^{\tau}, \theta_2^{\tau}, \theta_3^{\tau}, \theta_4^{\tau}, \theta_5^{\tau}, \theta_6^{\tau}, \theta_7^{\tau}, \theta_8^{\tau}, \theta_9^{\tau})^T$
- z_{ij} : 1
- u_j : son los efectos aleatorios en el intercepto $u \perp \varepsilon_{ij}$ y $u \sim N(0, \Psi^{\tau})$ por lo tanto las estimaciones se calcularán mediante integración numérica utilizando el método de cuadratura Gauss- Hermite (Geraci y Bottai, 2014).
- ε : término de error $\varepsilon \sim \text{Al}(0, \sigma, \tau)$.

3. Estimar el modelo incluyendo la variable independiente de las tipologías municipales que conforman los entornos de desarrollo:

$$y_{ij} = x_{ij}^T \delta^\tau + z_{ij}^T u_j + \varepsilon_{ij}^\tau \quad (6)$$

Donde

- $x_{ij}^T = (1, B_{ij}, C_{ij}, D_{ij}, E_{ij}, F_{ij}, G_{ij}, \text{mujer}_{ij}, \text{msecundaria}_{ij}, \text{mterciaria}_{ij}, \text{edad}_{ij}, \text{oficial}_{ij}, \text{jmanana}_{ij}, \text{jtarde}_{ij})$
- $\delta^\tau = (\delta_0^\tau, \delta_1^\tau, \delta_2^\tau, \delta_3^\tau, \delta_4^\tau, \delta_5^\tau, \delta_6^\tau, \delta_7^\tau, \delta_8^\tau, \delta_9^\tau, \delta_{10}^\tau, \delta_{11}^\tau, \delta_{12}^\tau, \delta_{13}^\tau)^T$

4. Estimar el modelo de la ecuación 6 y 7 empleando una regresión cuantílica y un modelo lineal jerárquico.

Los errores estándar para determinar la significancia individual de los coeficientes estimados se obtienen por el método de bootstrapping (se emplean 50 repeticiones).

4. Los Datos

Los datos que utilizaremos provienen de dos fuentes. En primer lugar, la información sobre el entorno de desarrollo municipal y tipología municipal del Departamento Nacional de Planeación (DNP). Esta información está disponible sólo para el 2014 (ver Cuadro 10). En segundo lugar, los datos sobre los puntajes en matemáticas y los factores individuales, familiares y de los colegios provienen de Saber 11 del segundo semestre del 2014.

4.1. DNP

El Entorno de Desarrollo y tipologías municipales Las tipologías y clasificaciones de entorno de desarrollo municipal⁷ provienen de la propuesta elaborada por el DNP. El aporte principal de la propuesta, según los autores, consiste en facilitar el diseño de políticas públicas, de tal forma que se pueda responder de manera asertiva e integral a las necesidades de las diferentes entidades territoriales.

La metodología empleada para construir las tipologías se basó en el análisis de componentes principales (ver Aguilar Londoño et al. (2015) para un detalle de la metodología y las variables utilizadas en la construcción del índice). De lo anterior resultan tres entornos de desarrollo: i) Entorno de Desarrollo Robusto que corresponde a los municipios pertenecientes a la tipología municipal A y B ii) Entorno de Desarrollo Intermedio corresponde a los municipios que pertenecen a la tipología C, D y E; iii) Entorno Incipiente, son los municipios pertenecientes a la tipología F y G. Como se observa en el Cuadro 4, si bien la tipología A agrupa la menor cantidad de municipios del país (0.5%), contiene casi un tercio de los habitantes del país para el 2014. Las otras tipologías contienen un menor número de habitantes, siendo la tipología G la que contiene la menor cantidad (3.6% del total de habitantes). Por otro lado, entre todas las tipologías, es la E la que agrupa la mayor cantidad de municipios, seguida de la D, F y C.

Cuadro 4: Número de municipios y habitantes por Tipología municipal

Entorno de Desarrollo	Tipología	Número de municipios	Participación (%)	Población	% Población	Ejemplo
Robusto	A	6	0,5%	15 293 275	32,17%	Cali
	B	63	5,7%	9 179 477	19,31%	Manizales
Intermedio	C	158	14,4%	7 250 124	15,25%	Tunja

⁷ Bogotá D.C no se incluyó en la Tipología Municipal sino en la Tipología Departamental, por lo que de ésta se extrajo la información correspondiente.

Incipiente	D	261	23,7%	5 812 304	12,23%	Florencia
	E	293	26,6%	5 126 952	10,79%	El Cocuy
	F	209	19,0%	3 173 444	6,68%	Moñitos
	G	111	10,1%	1 700 315	3,58%	Currillo
Total general		1101	100%	47535891	100%	

Fuente: Adaptado de Carmona, Supelano y Osejo (2015).

Las tipologías reflejan las características propias de cada municipio, en relación con seis dimensiones: Urbano regional, condiciones sociales, económica, ambiental, institucional y de seguridad (Ver Anexo 10 para una descripción de detallada de estas dimensiones). Cada una de éstas se mide por medio de un índice que va de 0 a 1, donde entre más cerca esté el índice de la entidad territorial a 1 mejor se encuentra en esa dimensión (ver Cuadro 5).

Cuadro 5: Puntaje promedio en las dimensiones por Entorno de Desarrollo

Entorno de desarrollo	Dimensiones					
	Urbana	Económica	Calidad de Vida	Ambiental	Seguridad	Institucional
Robusto	0.28	0.74	0.78	0.58	0.64	0.74
Intermedio	0.08	0.42	0.52	0.54	0.73	0.52
Incipiente	0.06	0.29	0.36	0.45	0.71	0.34

Fuente: Cálculos propios y DNP.

4.2. Saber 11

La prueba Saber 11 es efectuada por el Instituto Colombiano para la Evaluación de la Educación (ICFES). Esta prueba es importante porque todos los estudiantes que se deseen graduar del bachillerato deben tomarla (es decir, es de carácter censal y obligatoria), además las instituciones de educación superior emplean los resultados como un mecanismo para seleccionar a los estudiantes y otorgar becas educativas⁸ (Ministerio de Educación Nacional, 2015).

De acuerdo con el ICFES (2015) la prueba Saber 11 evalúa 5 competencias: matemáticas, lectura crítica, sociales y competencias ciudadanas, ciencias naturales e inglés. Los puntajes de estos componentes se miden en una escala de 0 a 100, con media 50 y desviación estándar 10. La distribución de estos puntajes se debe a la aplicación de la metodología de respuesta al ítem a los puntajes iniciales que obtienen los estudiantes (Abadía y Bernal, 2016). Por lo tanto, el puntaje no indica el número de respuestas correctas, y por ahora no tiene asociado una caracterización de los niveles de desempeño para la competencia, excepto para inglés.

Aquí utilizamos la base del segundo semestre de 2014 por dos razones. En primer lugar, el indicador de tipología municipal está construido sólo para el 2014. En segundo lugar, en ese semestre se introdujeron cambios a la forma y estructura de la prueba por lo que los resultados del primer semestre no son estrictamente comparables con los del segundo semestre (Ministerio de Educación Nacional, 2014).

⁸ Uno de los principales criterios utilizado por el Gobierno en el programa “Ser Pilo Paga” consiste en otorgar becas 100 % condonables para acceder a las instituciones de educación superior acreditadas en Alta Calidad a partir del puntaje total que los estudiantes obtienen en Saber 11.

La base de datos de 2014 se caracteriza en términos generales por incluir:

- Estudiantes que tienen entre 14 y 28 años.
- Estudiantes que asisten a colegios públicos y privados, en jornada diurna (mañana, tarde, y completa).
- Viven en cualquier municipio de Colombia excepto en San Yrés Islas, Belén de Bajira, El Encanto, La Chorrera, La Pedrera, La Victoria, Mirití - Parana, Puerto Santyer, Tarapaca, Barranco Minas, Maripipana. Lo anterior debido a que para estos municipios no existe una tipología definida por el DNP.
- Pertenecen a alguna institución educativa.
- Toman la prueba por primera vez.
- Libres, es decir no están privados de su libertad.
- Reportaron información para todas las variables de interés.
- Pertenecen a instituciones educativas que presentaron como mínimo 10 estudiantes.

De acuerdo con lo anterior, se obtuvieron 410.608 observaciones de 533.133; es decir el 77,0% de las observaciones originales.

4.3. Estadísticas descriptivas

El Cuadro 6 muestra las estadísticas descriptivas de todas las variables usadas en la estimación a nivel nacional y por entorno de desarrollo (Ver Anexos para las estadísticas descriptivas por tipología municipal).

Cuadro 6: Estadísticas descriptivas a nivel nacional y por entorno de desarrollo

Variable	Entorno de desarrollo			
	Nacional	Incipiente	Intermedio	Robusto
No. De estudiantes	410709	35111	150928	224670
Personales				
Promedio puntaje Matemáticas	50.88	46.58	49.32	52.60
Desviación Estándar puntaje Matemáticas	9.68	8.29	9.04	9.90
Primer cuartil puntaje Matemáticas	45	41	44	46
Segundo cuartil puntaje Matemáticas	50	46	49	52
Tercer cuartil puntaje Matemáticas	56	52	55	58
Mujer (%)	54.89	53.30	55.61	54.65
Edad (promedio años)	16.82	17.14	16.87	16.74
Familiares (%)				
Nivel educativo madre: ninguno o primaria	33.21	58.14	43.11	22.67
Nivel educativo madre: secundaria	45.18	33.34	41.45	49.53
Nivel educativo madre: terciaria	21.61	8.52	15.44	27.80
Colegio (%)				
Oficial	77.47	97.82	90.51	65.54
Jornada Mañana	59.63	77.23	64.86	53.37

Jornada Tarde	18.36	8.40	14.45	22.55
Jornada Completa	21.98	14.36	20.64	24.07

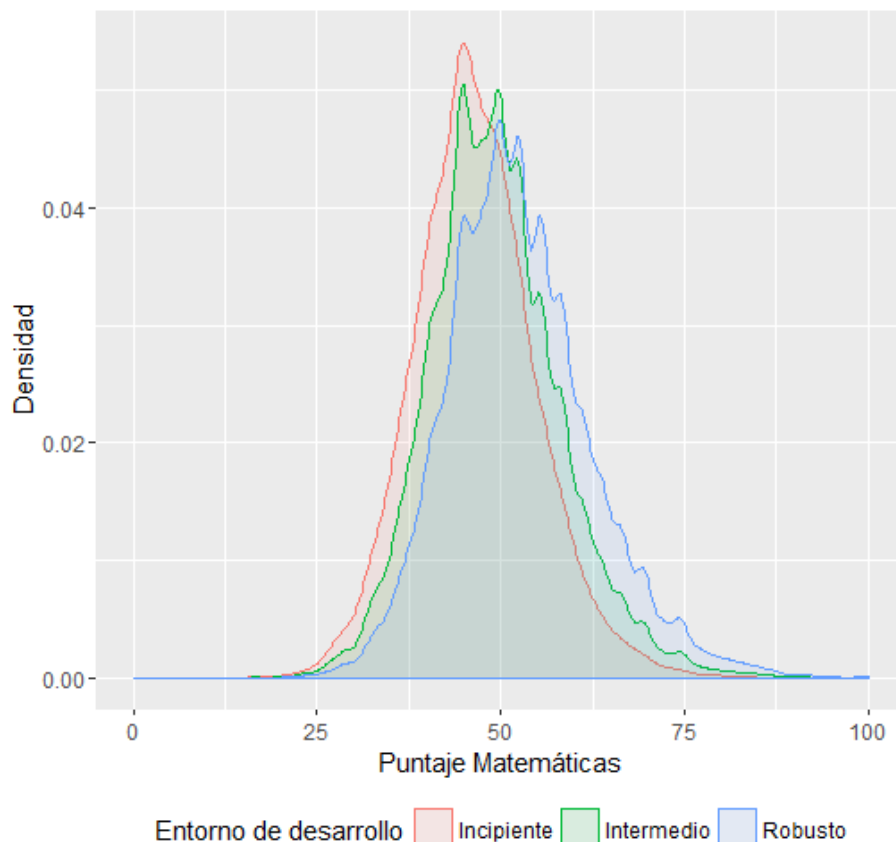
Fuente: Cálculos propios.

En cuanto a las características personales, se puede observar que el puntaje promedio en matemáticas de los estudiantes que viven en los entornos de desarrollo incipiente e intermedio está por debajo del puntaje promedio nacional; mientras que, el promedio en el entorno robusto está por encima. Este comportamiento se mantiene si se comparan los puntajes en matemáticas obtenidos por cuartil, tanto por entornos de desarrollo con en el ámbito nacional.

De igual forma hay diferencias al interior de los entornos, por ejemplo, se observa que la diferencia entre la cola inferior (primer cuartil) y la cola superior (tercer cuartil) de la distribución del puntaje aumenta ligeramente por entorno de desarrollo: 9 puntos para entorno de desarrollo incipiente; 11 puntos para entorno de desarrollo intermedio; y 12 para entorno de desarrollo robusto.

Asimismo, hay diferencias marcadas entre los entornos de desarrollo para un mismo cuartil lo que indica que el lugar de residencia puede ser un factor relevante. Por ejemplo, el puntaje del primer cuartil para el estudiante que vive en un municipio incipiente es de 41 puntos, y para un estudiante que vive en un entorno de desarrollo robusto es de 46; mientras que, para el tercer cuartil el puntaje es de 52 y 58, respectivamente. Lo anterior se hace más evidente en la Figura 3 muestra el kernel de la distribución del puntaje de matemáticas para cada uno de los entornos de desarrollo. Se observa que en la parte media y en la cola superior, la distribución de los puntajes de los estudiantes que viven en un entorno de desarrollo incipiente o intermedio está a la izquierda de aquellos que viven en municipios con un entorno de desarrollo robusto. La diferencia de la brecha a lo largo del kernel entre los entornos de desarrollo evidencia la importancia de tener en cuenta este factor en la estimación y de emplear un modelo que tenga en cuenta la distribución de los puntajes y no solo al estudiante promedio.

Figura 3 Distribución kernel del puntaje de matemáticas por entorno de desarrollo



Fuente: Elaboración propia

Por otra parte, el porcentaje de mujeres es relativamente similar y superior al 50 % para todos los casos. La edad promedio del estudiante es de 16.82 años a nivel nacional, para el municipio incipiente e intermedio es ligeramente mayor y para los estudiantes de municipios robustos es menor.

En cuanto a las características familiares, en el ámbito nacional el 45.18 % de los estudiantes reportaron que el máximo nivel educativo alcanzado por su madre había sido secundaria (incompleta o completa), en este mismo nivel educativo se concentra la mayoría de respuestas de los estudiantes que viven en un entorno de desarrollo robusto. Por el contrario, tanto en el entorno incipiente como en el intermedio el mayor porcentaje de respuesta se concentró en ninguno o primaria (incompleta o completa).

Por último, se encuentra más del 90 % de los estudiantes que viven en un entorno de municipio incipiente o intermedio asisten a colegios oficiales, mientras que en los municipios robustos ese porcentaje es del 65.5 %. En los tres entornos, más del 50 % de los estudiantes asiste a jornada mañana, seguido de jornada completa y por último jornada de la tarde.

Cuadro 7: Coeficiente de correlación interclase del modelo vacío

Cuartil	CCI (%)
0.25	64
0.50	76
0.75	68

Fuente: Elaboración propia

5. Resultados

En esta sección se presentan los resultados de la aplicación de la aproximación empírica tanto a nivel de entorno de desarrollo como de tipología municipal. Previo a eso se estima el modelo vacío de la ecuación 4.

Los resultados de iniciales indican que el colegio explica el 64 % de la variación en los resultados del puntaje de matemáticas en 2014 II de los estudiantes que se encuentran en la cola inferior de la distribución. Mientras que, ese porcentaje aumenta en 12 puntos porcentuales si el estudiante se encuentra en la mediana de la distribución, y en 4 puntos porcentuales para el estudiante del tercer cuartil. Lo anterior apoya la hipótesis de que es adecuado emplear modelos lineales jerárquicos.

El hecho de que factores del colegio expliquen un alto porcentaje de las variaciones en los puntajes de matemáticas indica que hay algunos colegios que alcanzan unos altos niveles de desempeño, mientras que otros tienen desempeños muy bajos. Lo que podría estar indicio que en términos de equidad en el área educativa todavía hay un largo camino que recorrer Foy (2004); además, estos resultados estarían acordes con los estudios que revelan que la asignación de recursos físicos y humanos en el área educativa no están distribuidos equitativamente entre las entidades territoriales (Bonet (2005); Bonilla Mejía (2011)).

Por otro lado, hay que tener en cuenta que estos porcentajes se encuentran por encima de lo reportado en otros estudios (donde también aplican modelos lineales jerárquico de dos niveles pero para el estudiante promedio) para el caso de matemáticas: 28.6 % empleo datos de Saber 93 (Castaño, 1998); 35 % en colegios no oficiales y 23 % en colegios oficiales con datos de Saber 1997-II (Piñeros y Rodríguez, 1998); 46.64 % para la ciudad de Medellín al utilizar los puntajes de Saber 11-2010 (Montes et al., 2014). Luego, esos cambios en los porcentajes del ICC se pueden deber a los cambios (forma y conceptualización) que la prueba Saber 11 ha experimentado en los últimos años; asimismo, al hecho de que la aproximación empírica empleada difiere de las anteriores, entre otras razones.

5.1. Entorno de desarrollo

En el Cuadro 8 se encuentran los resultados de la estimación del modelo jerárquico lineal por cuartiles (LQMM), del modelo lineal jerárquico (HML) y de la regresión cuantílica (RQ). El intercepto, el cual es significativo y positivo para todos los cuartiles, se puede interpretar como el puntaje en matemáticas promedio respectivo a cada cuartil que obtendría un estudiante hombre con 16.82 años en promedio, cuya madre no ha alcanzado un nivel educativo o sólo alcanzó la primaria (completa o incompleta), que vive en un entorno de desarrollo robusto, asiste a un colegio no oficial en la jornada de completa, respectivamente para cada cuartil. Como era de esperarse la magnitud del coeficiente aumenta a medida que se pasa de cuartil de estudio. El intercepto del modelo HML si bien es significativo, por su misma naturaleza, oculta las diferencias en el desempeño que existen a lo largo de la distribución.

Los resultados muestran que vivir en un municipio catalogado con un entorno de desarrollo incipiente o intermedio tiene un efecto negativo en el desempeño en matemáticas

comparado con vivir en un municipio de entorno de desarrollo robusto. Lo anterior también se evidencia con la aplicación del modelo HML y RQ. Vivir en un municipio intermedio para un estudiante de bajo desempeño (primer cuartil) no tiene un efecto significativo sobre su desempeño; en cambio para un estudiante que se encuentra en la mediana o en la cola superior de la distribución el efecto es significativo con el 99 % de confianza. Por otra parte, un estudiante de la cola inferior que viva en un entorno de desarrollo incipiente obtiene en promedio dos puntos menos que aquel que vive en un entorno de desarrollo robusto, esa diferencia es mayor para aquellos que se encuentran en el segundo o tercer cuartil.

Por otro lado, los coeficientes que acompañan los controles utilizados también son relevantes y evidencian los signos esperados de acuerdo a la revisión de la literatura realizada. El coeficiente de regresión para la edad y el sexo son aproximadamente constantes a lo largo de la distribución y negativos para todos los cuartiles. Aquellos individuos con más edad tienen un desempeño en matemáticas menor a sus pares más jóvenes. Las mujeres tienen un desempeño menor al de los hombres. De hecho, se puede observar que el ser mujer afecta más a los estudiantes de alto desempeño (tercer cuartil). El nivel educativo alcanzado por la madre también es positivo y estadísticamente significativo. Un estudiante cuya madre haya alcanzado un nivel terciario de educación obtiene en promedio 3 puntos adicionales en promedio si se compara con un estudiante cuya madre no alcanzó algún nivel educativo o sólo alcanzó primaria.

Cuadro 8: Estimación modelo jerárquico lineal, modelo jerárquico lineal por cuantiles y regresión cuantílica por entornos de desarrollo: Puntaje Matemáticas

Variable	HML	Puntaje en matemáticas					
		LQMM			Regresión cuantílica		
		Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 1	Cuartil 2	Cuartil 3
Efectos fijos							
Intercepto	54.91*** (0.11)	49.19*** (0.58)	54.25*** (0.46)	60.15*** (0.37)	47.51*** (0.09)	53.48*** (0.08)	59.67*** (0.19)
Entorno de desarrollo Intermedio	-1.85*** (0.09)	-1.00 (0.69)	-2.00*** (0.44)	-2.02*** (0.33)	-1.95*** (0.05)	-2.00*** (0.04)	-2.06*** (0.05)
Entorno de desarrollo Incipiente	-3.12*** (0.14)	-2.00* (1.04)	-3.00*** (0.52)	-3.69*** (0.43)	-3.48*** (0.09)	-3.66*** (0.06)	-3.98*** (0.06)
Efectos fijos: Controles							
Mujer	-2.74*** (0.03)	-2.00*** (0.01)	-2.06*** (0.42)	-3.04*** (0.04)	-1.76*** (0.05)	-2.45*** (0.05)	-2.96*** (0.05)
Edad (centrada)	-1.19*** (0.01)	-1.00*** (0.01)	-1.02*** (0.05)	-1.18*** (0.04)	-1.28*** (0.02)	-1.39*** (0.02)	-1.47*** (0.03)

Nivel educativo madre: secundaria	0.87*** (0.03)	0.99*** (0.01)	0.98*** (0.06)	0.92*** (0.07)	1.00*** (0.05)	1.22*** (0.05)	1.51*** (0.09)
Nivel educativo madre: terciaria	3.04*** (0.04)	2.99*** (0.03)	2.98*** (0.03)	3.35*** (0.11)	4.45*** (0.07)	5.22*** (0.05)	6.53*** (0.10)
Oficial	-1.67*** (0.12)	-1.00 (0.72)	-1.01 (0.70)	-1.60*** (0.25)	-0.29*** (0.05)	-0.63*** (0.05)	-1.14*** (0.06)
Jornada mañana	-2.38*** (0.12)	-3.00*** (0.66)	-3.01*** (0.49)	-2.40*** (0.38)	-2.01*** (0.05)	-2.21*** (0.06)	-2.31*** (0.08)
Jornada tarde	-2.81*** (0.15)	-3.00*** (0.99)	-4.01*** (0.72)	-2.83*** (0.39)	-2.50*** (0.08)	-2.76*** (0.08)	-3.25*** (0.10)
Efectos aleatorios							
Colegio (sd)	3.94	3.23	3.22	3.69			
CCI	20.00 %	50.62 %	54.34 %	56.18 %			

Error estándar entre paréntesis

*** Nivel de confianza 99 % ** Nivel de confianza 95 % * Nivel de confianza 90 %

Fuente: Elaboración propia

Con respecto a las variables del colegio, se encuentra que asistir a un colegio oficial está correlacionado negativa y significativamente sólo con aquellos estudiantes que se encuentran en la cola superior, lo anterior no es capturado por el modelo HML o RQ. Se ha encontrado para el caso colombiano que esto puede obedecer no solamente a unas mejores condiciones socioeconómicas de los estudiantes que asisten a estos colegios, también a la forma en que son administrados y estructura de incentivos que se maneja al interior de estos (Gaviria y Barrientos, 2001).

Nuestros resultados también muestran que los estudiantes que asisten a la jornada de mañana o de la tarde tienen un menor desempeño que aquellos que asisten a la jornada completa. Ese efecto negativo es mayor para aquellos que asisten a la jornada de la tarde que para aquellos que asisten a la jornada de la mañana. Por último, se puede observar que los coeficientes de correlación interclase para el modelo LQMM disminuyeron respecto a los obtenidos con el modelo vacío. Estos indican que aproximadamente el 50 % de la variabilidad en los resultados se debe a factores del colegio, una vez se controla por ciertas variables de la dimensión personal, familiar y del lugar de residencia del individuo.

5.2. Tipología Municipal

Al igual que en el caso del entorno de desarrollo, se encuentra que el intercepto es significativo y positivo para todos los cuartiles, se puede interpretar como el puntaje en matemáticas promedio que obtendría un estudiante hombre con 16.82 años en promedio, que vive en un municipio de tipología A, cuya madre no ha alcanzado un nivel educativo o sólo de primaria (completa o incompleta), asiste a un colegio no oficial a la jornada de completa. Al igual que en la sección anterior el intercepto para el estudiante promedio (HML) y de la regresión cuantílica son significativos y positivos (Ver Cuadro 9. Al analizar por tipologías se encuentra que aquellos estudiantes que viven en un municipio de tipología B obtienen un puntaje levemente mayor a aquellos que estudian en un municipio de tipología A; excepto, para el estudiante que se encuentra en el segundo cuartil.

Resultados similares se obtienen con la regresión cuantílica, excepto para el estudiante de bajo desempeño para el cual vivir en un municipio de tipología B no resulta significativo; mientras que al aplicar el modelo LQMM se encuentra que vivir en el municipio de tipología B para un estudiante del primer cuartil está asociado significativamente, con 95 % de confianza, con un aumento de 1 punto, respecto al puntaje del estudiante que vive en un municipio de tipología A, en el resultado de las pruebas de matemáticas.

Cuadro 9: Estimación modelo jerárquico lineal, modelo jerárquico lineal por cuantiles y regresión cuantílica por tipología municipal: Puntaje Matemáticas

Variable	Puntaje en matemáticas						
	HML	LQMM			Regresión cuantílica		
		Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 1	Cuartil 2	Cuartil 3
Efectos fijos							
Intercepto	54.81*** (0.11)	49.20*** (0.70)	54.23*** (0.43)	60.13*** (0.30)	47.58*** (0.08)	53.52*** (0.07)	59.39*** (0.15)
Entorno Robusto							
Tipología B	0.22** (0.09)	0.99** (0.50)	0.00 (0.33)	0.49** (0.24)	-0.08 (0.06)	0.00 (0.04)	0.25*** (0.08)
Tipología C	-1.08*** (0.12)	-1.00 (0.74)	-1.01** (0.38)	-0.86 (0.39)	-1.25*** (0.08)	-1.19*** (0.05)	-0.92*** (0.08)
Entorno Intermedio							
Tipología D	-1.92*** (0.13)	-2.01** (0.80)	-2.01*** (0.45)	-1.95*** (0.34)	-2.08*** (0.06)	-2.08*** (0.06)	-2.01*** (0.06)
Tipología E	-2.59*** (0.14)	-2.01** (0.82)	-2.03*** (0.44)	-2.93*** (0.18)	-3.08*** (0.08)	-3.24*** (0.07)	-3.34*** (0.09)
Entorno Incipiente							
Tipología F	-3.07*** (0.16)	-3.01*** (0.99)	-3.01*** (0.31)	-3.61*** (0.29)	-3.43*** (0.11)	-3.58*** (0.06)	-3.62*** (0.10)
Tipología G	-3.19*** (0.21)	-3.00** (1.36)	-3.02*** (0.33)	-3.88*** (0.36)	-3.96*** (0.10)	-4.08*** (0.08)	-4.12*** (0.11)
Efectos fijos: Controles							
Mujer	-2.74*** (0.03)	-2.00*** (0.01)	-2.04*** (0.37)	-3.07*** (0.04)	-1.78*** (0.05)	-2.43*** (0.04)	-2.99*** (0.02)
Edad (centrada)	-1.19*** (0.01)	-1.01*** (0.01)	-1.01*** (0.06)	-1.17*** (0.03)	-1.29*** (0.02)	-1.38*** (0.02)	-1.41*** (0.03)
Nivel educativo madre: secundaria	0.86*** (0.03)	0.98*** (0.02)	0.99*** (0.07)	0.92*** (0.06)	0.97*** (0.05)	1.16*** (0.04)	1.5*** (0.05)
Nivel educativo madre: terciaria	3.03*** (0.04)	2.98*** (0.04)	2.98*** (0.07)	3.35*** (0.09)	4.40*** (0.07)	5.16*** (0.05)	6.58*** (0.08)
Oficial	-1.53*** (0.12)	-1.01* (0.59)	-2.00*** (0.42)	-1.63*** (0.27)	-0.17*** (0.06)	-0.51*** (0.05)	-0.97*** (0.07)

Jornada mañana	-2.45*** (0.12)	-2.99*** (0.45)	-2.01*** (0.39)	-2.35*** (0.27)	-2.07*** (0.06)	-2.30*** (0.05)	-2.41*** (0.07)
Jornada tarde	-2.91*** (0.15)	-3.98*** (0.79)	-3.01*** (0.51)	-2.82*** (0.32)	-2.57*** (0.07)	-2.90*** (0.06)	-3.16*** (0.11)
Efectos aleatorios	3.91	3.21	3.22	3.69			
Colegio (sd)	20.00 %	50.44 %	54.41 %	56.15%			

Error estándar entre paréntesis

*** Nivel de confianza 99 % ** Nivel de confianza 95 % * Nivel de confianza 90 %

Fuente: Elaboración propia

Vivir en un municipio de Tipología D, E, F o G está asociado con tener un menor resultado que aquellos que viven en un municipio de tipología A, independientemente del cuartil en el que se encuentren y de la metodología aplicada. Las diferencias entre el modelo LQMM y la regresión cuantílica radica en la magnitud de los coeficientes. Etre el modelo LQMM y el modelo HML consiste en que éste no evidencia las variaciones que existen entre los cuartiles. Por ejemplo, un estudiante de bajo desempeño que vive en un municipio de Tipología G (entidades territoriales con el menor avance en la mayoría de las dimensiones que contempla el índice) obtiene en promedio 3 puntos menos en la prueba de matemáticas que aquellos estudiantes de bajo desempeño que viven en un municipio de Tipología A; mientras que para el estudiante de alto desempeño ese valor es de casi 4 puntos menos respecto al que obtendría su semejante que vive en un municipio con las mejores condiciones de desarrollo posibles en el país.

Al analizar las tipologías por el entorno de desarrollo al que pertenecen, se encuentra dentro de las tipologías que conforman el entorno de desarrollo intermedio (C, D y E) sí hay diferencias entre estos. Para un estudiante de bajo desempeño vivir en un municipio de tipología C no está asociado con tener un menor desempeño que aquel que vive en municipio de tipología A; por el contrario, para ese mismo estudiante que vive en un municipio de tipología D o E, éste obtiene dos puntos menos en la prueba respecto al estudiante que vive en la tipología A con un 95 % de confianza.

Para el estudiante del tercer cuartil sucede algo similar, excepto porque sí hay una diferencia entre pertenecer a un municipio de tipología D y E de casi un punto, se tendría que confirmar si esta diferencia es estadísticamente significativa. Por último, para el estudiante que se encuentra en el segundo cuartil y vive en un municipio de tipología C se halla que éste obtiene 1 punto menos en la prueba de matemáticas que el que vive en una tipología A; mientras que aquellos que viven en un municipio de tipología D o E obtienen dos puntos menos, esto con un 99 % de confianza. Para las tipologías que conforman el entorno de desarrollo incipiente (F y G) no se observa que haya diferencias a lo largo de la distribución entre vivir en un municipio de tipología F ó G respecto a vivir en el municipio de tipología A, puesto que en ambos casos se obtiene en promedio tres puntos menos en la prueba de matemáticas para el estudiante del primer y segundo cuartil, y de aproximadamente 4 puntos menos para el estudiante de alto desempeño. Por últimos los coeficientes que acompañan las variables de control (individuales, familiares y del colegio) son en su mayoría significativas y tienen los signos esperados.

6. Conclusiones

Nuestros resultados muestran una relación estadística significativa entre el entorno de desarrollo y el desempeño en matemáticas, luego los hacedores de políticas deben tener en cuenta las condiciones del lugar donde viven los estudiantes a la hora de formular políticas educativas encaminadas a mejorar la calidad educativa. Especialmente, deberían enfocarse en atender las necesidades de los municipios de entorno de desarrollo incipiente (tipología F y G), y en estudiar las prácticas pedagógicas y administrativas que se imparten en los municipios del entorno de desarrollo robusto (tipología A y B). De igual forma, la política educativa debe ser diferenciada entre los estudiantes, ya que el entorno de desarrollo y las características de los colegios no afectan en la misma magnitud a todos los estudiantes.

Lo anterior cobra aún más importancia si se tiene en cuenta los hallazgos del estudio de Cendales et al. (2015). En éste los autores establecen que el sistema de educación media en la zona industrial de Colombia (Medellín, Cali y Bogotá, las tres hacen parte del entorno de desarrollo robusto y la tipología A) incentiva la reproducción de una estructura social inequitativa. Lo anterior puesto que los estudiantes de bajos recursos tienen menores oportunidades de acceder a la educación superior, por ende, tienen a repetir la baja productividad que caracteriza a sus padres; lo contrario sucede para aquellos con mejores condiciones socioeconómicas. Los autores concluyen que si se tiene en cuenta que en estas ciudades la mayor cantidad de la mano obra industrial (91 % en 2012 y 2013) corresponde a estudiantes de bajo recursos entonces estos resultados tienen implicaciones negativas en el desarrollo de la industria moderna del país.

El problema de tener una estructura social inequitativa que se regenera a sí misma se vuelve mayúsculo cuyo se tiene en cuenta que en los 1117 municipios restantes, donde vive el 73.3 % de la población de estudiantes de educación media según las pruebas Saber 11 2014-2, el 96.2 % de los estudiantes son de bajos recursos. Cuantificar la magnitud de ese potencial obstáculo en el desarrollo de la industria moderna del país requiere de su estudio respectivo, pero estas cifras ponen de manifiesto la importancia de incluir en estos análisis a los municipios de los entornos de desarrollo intermedio e incipiente.

Por otro lado, entre las limitaciones del estudio se encuentra que hay varias variables a nivel de colegio y de la familia que no se tuvieron en cuenta por restricciones de los datos, por ejemplo, los años de experiencia de los profesores, tamaño de la clase, salarios de los profesores o la mensualidad del colegio, el estrato socioeconómico o el ingreso de la familia. Una segunda limitación consiste en que no diferenció a los colegios privados que prestan su servicio al sector educativo público, de los que no, mediante contrato. Lo anterior podría ser importante puesto que se ha encontrado que hay una diferencia entre los estudiantes que asisten a estos colegios de los que asisten a colegios oficiales Godoy (2013).

Una tercera limitación consiste en que, aunque aquí se incluye una conceptualización del entorno de desarrollo, se desconoce cuál o cuáles dimensiones tienen una asociación más fuerte con el desempeño y por lo tanto en qué condición o característica del lugar donde reside el estudiante se deberían focalizar la inversión pública. De igual forma el índice construido por el DNP carece de una dimensión importante que es la educativa. En esta se podría incluir elementos como si la secretaría de educación municipal está certificada, nivel educativo de los profesores, características de la infraestructura educativa, entre otras. Luego, un próximo estudio podría construir un índice de desempeño educativo que permita analizar cómo esos factores educativos, que se derivan de decisiones institucionales, afectan el desempeño de los estudiantes en matemáticas, y otras competencias básicas, como lectura crítica e inglés.

Referencias

Abadía, L. K y Bernal, G. (2016). Brechas de género en el rendimiento escolar a lo largo de la distribución de puntajes: evidencia pruebas Saber 11.

Aguilar Londoño, C. M., Osejo Villamil, I., Carmona Sanchez, C. O., Supelano González, D., Gaitán Alvarez, J., & Taborda, J. C. (2015). Tipologías municipales de Colombia: herramienta para la consolidación de regiones desarrolladas

Barón, J. D. (2012). Diferencias en las características de los estudiantes y la brecha de rendimiento académico entre Barranquilla y Bogotá: Una descomposición semiparamétrica. *Ensayos sobre Política Económica*, 30(68), 164-215.

Barrientos, J. H., & Rios, P. (2007). Evaluación de la gestión privada del servicio público educativo en Medellín. *Lecturas de Economía*, 66, 147.

Barrientos, M. (2012). Diferencias regionales en la distribución del ingreso en Colombia. *Sociedad y Economía*, 108, 43–68.

Bassett Jr, G. W., Tam, M. Y. S., & Knight, K. (2002). Quantile models and estimators for data analysis. *Metrika*, 55(1-2), 17-26.

Bonet, J., & Bonett, J. (2005). *Inequidad espacial en la dotación educativa regional en Colombia*. Banco de la República.

Castaño, E. (2010). El efecto colegio sobre la variabilidad del rendimiento en matemáticas. *Lecturas de economía*, 49(49), 47-57.

Cendales, A., Mora, J.J., y Campo, J. (2015). Educación media y mano de obra industrial en Colombia: Inequidad y pobreza, En VI Seminario Internacional de Investigación sobre la Calidad Educativa Bogotá.

Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction*, 18(4), 321-336.

Cooray, A. (2009), The role of education in economic growth, en 38 Australian Conference of Economists University of Wollongong, 1–27.

Cortés, D. y Vargas J.F (2012). Inequidad Regional en Colombia, *Serie Documentos Cede*, 34.

Costanzo, A. (2015). The Effect of M@ tabel on Italian Students' Performances: A Quantile Regression Approach. *Procedia-Social and Behavioral Sciences*, 197, 236-244. Educación Nacional, Ministerio, Colombia, la mejor educada en el 2025, Technical Report, Ministerio Nacional de Educación, Bogotá 2015.

Delgado, M. S., Henderson, D. J., & Parmeter, C. F. (2014). Does education matter for economic growth?. *Oxford Bulletin of Economics and Statistics*, 76(3), 334-359.

Demir, İ., Kılıç, S., & Ünal, H. (2010). Effects of students' and schools' characteristics on mathematics achievement: findings from PISA 2006. *Procedia-Social and Behavioral Sciences*, 2(2), 3099-3103.

DfEE (1999). A, Moser Report.

Eide, E., & Showalter, M. H. (1998). The effect of school quality on student performance: A quantile regression approach. *Economics letters*, 58(3), 345-350.

Foy, P. (2004). Intraclass correlation y variance components as population attributes y measures of sampling efficiency in PIRLS 2001, in The 1st IEA International Research Conference, 1–13.

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2), 210-240.

Galvis, L. A., & Bonilla, L. (2012). Desigualdades regionales en el nivel educativo de los profesores en Colombia. *Revista de Economía Institucional*, 14(26), 223-240.

Gaviria, A., & Barrientos, J. (2001). Características del plantel y calidad de la educación en Bogotá. *Coyuntura Social*, 25, 81-98.

Geraci, M., & Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, 8(1), 140-154.

Geraci, M., & Bottai, M. (2014). Linear quantile mixed models. *Statistics and Computing*, 24(3), 461-479.

Godoy, S. (2013), Análisis del programa de ampliación de cobertura en Cali con respecto a la calidad de la educación, medido a través de las pruebas de Estado.

Haile, G. A., & Nguyen, A. N. (2008). Determinants of academic attainment in the United States: A quantile regression analysis of test scores. *Education Economics*, 16(1), 29-57.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3), 1141-1177.

Hanushek, E. A. (2002). Evidence, politics, y the class size debate, in Lawrence Mishel y Richard Rothstein, eds., *The Class Size Debate*, Washington, DC: Economic Policy Institute, chapter 2, 37–66.

Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37, 204-212.

- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, 46(3), 607-668.
- Hoerandner, C. M., & Lemke, R. J. (2006). Can No Child Left Behind Close the Gaps in Pass Rates on Standardized Tests?. *Contemporary Economic Policy*, 24(1), 1-17.
- Hox, J. (2002). *Multilevel Analysis: Techniques y Applications.*, New Jersey: Lawrence Erlbaum Associates.
- Huizinga, M. M., Beech, B. M., Cavanaugh, K. L., Elasy, T. A., & Rothman, R. L. (2008). Low numeracy skills are associated with higher BMI. *Obesity*, 16(8), 1966-1968.
- ICFES (2015), *Guía de Interpretación y Uso de Resultados de las pruebas SABER 11*, 57.
- Jimenez, E., Lockheed, M. E., & Paqueo, V. (1991). The relative efficiency of private and public schools in developing countries. *The World Bank Research Observer*, 6(2), 205-218.
- Koenker, R., & Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, 94(448), 1296-1310.
- Koenker, R. (2005). *Quantile regression* (No. 38). Cambridge university press.
- Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33-50.
- Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal*, 125(588), F397-F424.
- Litster, J. (2013). The impact of poor numeracy skills on adults, Technical Report , University of London, London.
- Lubienski, C., & Lubienski, S. T. (2006). *Charter, private, public schools and academic achievement: New evidence from NAEP mathematics data* (Vol. 16). New York: National Center for the Study of Privatization in Education, Teachers College, Columbia University.
- Marín, J. B. (2008). Calidad de la educación pública y logro académico en Medellín 2004-2006. Una aproximación por regresión intercuartil. *Lecturas de Economía*, 68(68), 121-144.
- Marino, M. F., & Farcomeni, A. (2015). Linear quantile regression models for longitudinal experiments: an overview. *Metron*, 73(2), 229-247.
- Mateus, W. A. M. (2016). Análisis de distribución geográfica y espacial de los resultados de las Pruebas Saber 11 del Instituto Colombiano para el Fomento de la Educación Superior-ICFES-. 2005-2012. Colombia. *Cuadernos Latinoamericanos de Administración*, 11(21), 39-50.

Meisel, R. y Romero, J. (2007). Igualdad de oportunidades para todas las regiones, en Fernández, M., Guerra W., y Meisel A., eds., Políticas para reducir las desigualdades regionales en Colombia, primera ed., Cartagena: Banco de la República, 1–415.

Melzer, David, E Gardener, Iain Lang, B McWilliams, y JM Guralnik, Retirement , health y relationships of the older population in Engly: The 2004 English Longitudinal Study of Ageing number July 2006.

Ministerio de Educación Nacional (2010). Plan sectorial de educación 2010 - 2014, Technical Report.

Ministerio de Educación Nacional (2014). El ICFES cambia la clasificación de resultados, jul 2014.

Ministerio de Educación Nacional (2015), Información de la prueba Saber 11.

Montes, I., Garcés, J. D., Chica, S. M., & Jaramillo, A. (2014). Rendimiento académico: ¿qué papel juegan los factores institucionales? Documentos de trabajo Economía y Finanzas, (14), 1–39.

Mora, J. J. (2003). Sheepskin effects and screening in Colombia. *Colombian Economic Journal*, 1(1), 95-108.

OCDE (2007). El programa PISA de la OCDE Qué es y para qué sirve, Technical Report, Organización para la cooperación y el desarrollo económico, París.

OECD (2014). PISA 2012 Results in Focus, Technical Report.

Jiménez, L. J. P., & Pinzón, A. R. (1998). Los insumos escolares en la educación secundaria y su efecto sobre el rendimiento académico de los estudiantes: un estudio en Colombia. *Washington, DC: The World Bank/Latin America and the Caribbean Region/Department of Human Development*.

Raudenbush, S., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of education*, 1-17.

Restrepo, P. P., & Alviar, M. (2009). El logro académico y el efecto colegio en las pruebas ICFES en Antioquia. *Lecturas de Economía*, 60(60), 67-95.

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences*, 17(2), 147-159.

Rivkin, Steven G y Jeffrey C Schiman, Instruction time, classroom quality, y academic achievement, *The Economic Journal*, 2015, 125 (588), F425–F448.

Sánchez, C., Supelano, D., y Osejo, I. (2014). Tipologías Departamentales y Municipales: una propuesta para comprender las entidades territoriales colombianas, Technical Report, Departamento Nacional de Planeación - Dirección de Desarrollo Territorial Sostenible.

Sánchez, B. L., Lachos, H. V., & Labra, V. F. (2013). Likelihood based inference for quantile regression using the asymmetric Laplace distribution. *Journal of Statistical Computation and Simulation*, 81, 1565-1578.

Smith, A. (2004). Making mathematics count, Report Commissioned by the UK Government, 2004.

Smith, L. B., Fuentes, M., Gordon-Larsen, P., & Reich, B. J. (2015). Quantile regression for mixed models with an application to examine blood pressure trends in China. *The Annals of Applied Statistics*, 9(3), 1226-1246.

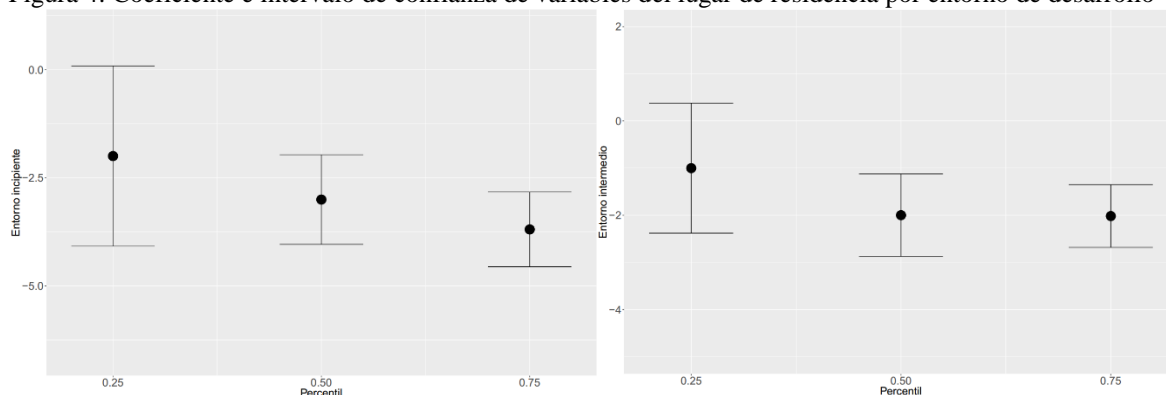
Tian, M., & Chen, G. (2006). Hierarchical linear regression models for conditional quantiles. *Science in China Series A: Mathematics*, 49(12), 1800-1815.

Tzavidis, N., & Brown, J. (2010). Using M-quantile models as an alternative to random effects to model the contextual value-added of schools in London.

Yu, K., Lu, Z., & Stander, J. (2003). Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3), 331-3

Anexos

Figura 4: Coeficiente e intervalo de confianza de variables del lugar de residencia por entorno de desarrollo



Cuadro 10: Conceptualización de las dimensiones para agrupar las entidades territoriales por entornos de desarrollo

Dimensión	Descripción
Funcionalidad Urbana	Se identifica el soporte físico territorial de los asentamientos poblacionales y de las actividades sociales y económicas en la base natural. Se expresa en la estructura espacial de nodos y flujos urbanos, rurales y regionales a través de las redes de infraestructuras, equipamientos, vivienda y servicios necesarios para sustentar el crecimiento y desarrollo. La caracterización de estos elementos y su comprensión de manera interrelacionada como sistema, permite establecer las condiciones del entorno territorial construido como soporte o limitante del desarrollo regional. Este enfoque, en línea con la Nueva Geografía Económica, busca explicar la formación de una amplia variedad de aglomeraciones económicas en el espacio geográfico, cada una como prestadora de diferentes servicios funcionales en razón de su densidad, distancia y división (Fujita & Krugman, 1999).
Condiciones de vida	Describe las carencias, déficits y necesidades básicas insatisfechas, y plantea las posibles soluciones. Esto obedece a que los desarrollos más recientes de políticas y estrategias para el mejoramiento de la calidad de vida hacen énfasis en la importancia de tener presente las nociones de necesidades y satisfactores (Manfred A., 1998).
Económica	Analiza la estructura económico-productiva de la entidad territorial. Examina elementos relacionados con el tamaño de la economía, las tendencias de crecimiento, la comparación con entornos externos de otras entidades territoriales, las infraestructuras y capacidades en Ciencia Tecnología e Innovación y algunos elementos relacionados con la disponibilidad de recursos que potencian la economía local. Se parte de que la economía real requiere la interrelación con el territorio para ser un factor de desarrollo que posibilita la construcción de cadenas de valor y sistemas productivos territoriales (Esser et al., 1996).

Ambiental Identifica las estructuras básicas eco-sistémicas del territorio que califican la riqueza ambiental y las capacidades de gobernanza ambiental de los territorios. Esta temática está en línea con los principios contenidos en la Convención Ramsar que plantea el uso racional de los ecosistemas, como el mantenimiento de sus características ecológicas, logrado mediante la implementación de enfoques por ecosistemas dentro de un concepto de desarrollo sostenible (Ramsar, 1971).

Institucional Se examina tres conceptos: i) Gobernar, como ejercer una delegación para liderar, gestionar y representar el conjunto de lo público con idoneidad y transparencia, de forma participativa y democrática buscando el bien común en el territorio (TorresMelo, 2011). ii) Gobernabilidad, como la capacidad que tiene la sociedad para construir y decidir participativa y legítimamente su propio desarrollo y bienestar, gestionándolo con equidad, transparencia y democracia (DNP-DDTS, 2010). Y iii) Gobernanza, como el proceso a través del cual se organizan y coordinan las múltiples relaciones e interacciones establecidas entre actores y niveles de gobierno diversos que están presentes en un territorio (Abad Aragón, 2010).

Seguridad La temática de seguridad ciudadana protege pues un núcleo importante, central, del desarrollo de las personas contra las amenazas que atentan contra su vida, su integridad o su patrimonio. La seguridad ciudadana hace referencia a amenazas refiriéndose únicamente a aquellas que derivan de las acciones violentas de los grupos delictivos comunes u organizados; y, por ello mismo, puede concebirse como la protección de determinadas opciones u oportunidades de las personas respecto del delito (PNUD, 1994, pág. 31).

Fuente: Tomado de (Aguilar Londoño et al., 2015).

Cuadro 11: Puntaje promedio en las dimensiones por Tipología Municipal

Entorno de desarrollo	Tipología	Dimensión					
		Institucional	Seguridad	Ambiental	Condiciones de vida	Económico	Urbano Regional
Robusto	A	0,70	0,48	0,64	0,81	0,95	0,78
	B	0,75	0,65	0,58	0,78	0,72	0,23
	C	0,63	0,71	0,57	0,62	0,54	0,11
Intermedio	D	0,54	0,74	0,54	0,53	0,42	0,08
	E	0,44	0,75	0,51	0,45	0,36	0,06
Incipiente	F	0,38	0,73	0,47	0,39	0,31	0,06
	G	0,28	0,67	0,43	0,31	0,27	0,06

Cuadro 12: Diccionario de Variables

Variable	Descripción	Categoría Base
$munintermedia_{ij}$	Variable dummy, toma el valor de 1 si el estudiante i que asiste al colegio j vive en un municipio de entorno de desarrollo intermedio, 0 en caso contrario.	
$munincipiente_{ij}$	Variable dummy, toma el valor de 1 si el estudiante i que asiste al colegio j vive en un municipio de entorno de desarrollo incipiente, 0 en caso contrario.	$munrobusto_{ij}$
$edad_{ij}$	La edad del individuo se encuentra centrada en 16.81971 años, por lo que esta variable indica el número de años por debajo o por encima de esa cifra que el individuo i que asiste al colegio j tiene.	
$mujer_{ij}$	Variable dummy, toma el valor de 1 si el estudiante i que asiste al colegio j es mujer, 0 en caso contrario.	$hombre_{ij}$
$msecundaria_{ij}$	Variable dummy, toma el valor de 1 si nivel educativo más alto alcanzado por la madre del estudiante i que asiste al colegio j es de secundaria incompleta o completa, 0 en caso contrario.	
$mterciaria_{ij}$	Variable dummy, toma el valor de 1 si nivel educativo más alto alcanzado por la madre del estudiante i que asiste al colegio j es de técnica, tecnológica, universitaria o de postgrado.	Ninguno o primaria
$oficial_{ij}$	Variable dummy, toma el valor de 1 si el colegio j al que asiste el estudiante i es oficial, 0 en caso contrario.	$nooficial_{ij}$
$jmanana_{ij}$	Variable dummy, toma el valor de 1 si la jornada a la que asiste el estudiante i del colegio j es la de la mañana, 0 en caso contrario.	
$jtarde_{ij}$	Variable dummy, toma el valor de 1 si la jornada a la que asiste el estudiante i del colegio j es la de la tarde, 0 en caso contrario.	$jcompleta_{ij}$
B_{ij}	Variable dummy, toma el valor de 1 si el estudiante i del colegio j vive en un municipio de tipología B, 0 en caso contrario.	
C_{ij}	Variable dummy, toma el valor de 1 si el estudiante i del colegio j vive en un municipio de tipología C, 0 en caso contrario.	
D_{ij}	Variable dummy, toma el valor de 1 si el estudiante i del colegio j vive en un municipio de tipología D, 0 en caso contrario.	
E_{ij}	Variable dummy, toma el valor de 1 si el estudiante i del colegio j vive en un municipio de tipología E, 0 en caso contrario.	
F_{ij}	Variable dummy, toma el valor de 1 si el estudiante i del colegio j vive en un municipio de tipología F, 0 en caso contrario.	A_{ij}
G_{ij}	Variable dummy, toma el valor de 1 si el estudiante i del colegio j vive en un municipio de tipología G, 0 en caso contrario.	

Cuadro 13: Estadísticas descriptivas por tipología municipal

Variable	Tipología Municipal						
	A	B	C	D	E	F	G
N	137692	86978	62224	49235	39469	23353	11758
Personales							
Promedio puntaje Matemáticas	52.80	52.30	50.47	49.25	47.61	46.90	46.0
Desviación Estándar puntaje Matemáticas	10.05	9.66	9.42	8.87	8.35	8.40	8.05
Primer cuartil puntaje Matemáticas	46	45	45	44	42	41	41
Segundo cuartil puntaje Matemáticas	52	52	50	49	47	47	45
Tercer cuartil puntaje Matemáticas	58	58	56	55	53	52	51
Mujer (%)	54.54	54.81	55.88	55.53	55.29	53.82	52.27
Edad (promedio años)	16.75	16.73	16.83	16.85	16.95	17.06	17.29
Familiares (%)							
Nivel educativo madre: ninguno o primaria	21.02	25.28	36.48	45.43	50.68	57.19	60.03
Nivel educativo madre: secundaria	49.17	50.10	44.75	40.50	37.42	33.59	32.84
Nivel educativo madre: terciaria	29.81	24.62	18.77	14.07	11.90	9.22	7.14
Colegio (%)							
Oficial	60.15	74.07	85.34	92.08	96.69	97.71	98.03
Jornada Mañana	48.47	61.12	66.30	60.89	67.54	75.35	80.97
Jornada Tarde	23.92	20.39	16.00	15.07	11.21	9.94	5.35
Jornada Completa	27.61	18.47	17.69	23.87	21.25	14.71	13.68

Figura 5: Coeficiente e intervalo de confianza de variables del lugar de residencia por tipología municipal

