



Efectos de la infraestructura sobre el fracaso escolar: evidencia empírica para Colombia

Hernando Bayona Rodríguez

Universidad de los Andes, Bogotá, Colombia (h.bayona28@uniandes.edu.co)

Recibido: agosto 01 de 2016 | Aceptado: noviembre 11 de 2016 | Publicado en línea: diciembre 30 de 2016

DOI: <http://dx.doi.org/10.18175/VyS7.2.2016.03>

INTRODUCCIÓN

Actualmente hay un amplio acuerdo acerca de los efectos positivos que tiene la educación sobre variables tales como la productividad, la salud y el crecimiento económico. Esto ha venido acompañado a su vez de un aumento importante de la inversión pública en educación, lo cual se evidencia en mayor medida en los países en desarrollo desde la década de los ochenta. Por ejemplo, América Latina y África Subsahariana han duplicado la inversión, Medio Oriente la ha triplicado, Asia Oriental la quintuplicó, y Asia del Sur casi la ha multiplicado por ocho (Glewwe *et al.*, 2011). Entender cómo la inversión en educación impacta su calidad es relevante en términos de política pública, pues permite priorizar los recursos y hacer más efectiva la inversión.

Por su parte, la literatura sobre calidad de la educación se ha concentrado en mayor proporción en su análisis vía resultados en pruebas estandarizadas (Hanushek, 1997; Glewwe *et al.*, 2011). Si bien esto es un resultado clave, variables de eficiencia interna como repetición, deserción y aprobación también son medidas importantes de la calidad educativa (Lee & Barro, 2001; Pritchett, 2004). Un marco conceptual sencillo apunta a que escuelas más efectivas o de mayor calidad son aquellas capaces de garantizar el éxito académico de sus estudiantes. Lo cual se debe observar no solo en mayores aprendizajes, sino en la reducción de las tasas de repetición y deserción. Este documento explora si el mejoramiento de la infraestructura tiene efectos sobre la repetición de los estudiantes, y adicionalmente analiza algunos de los posibles canales por los cuales esto podría ocurrir.

Entender los mecanismos por los cuales se puede reducir la repetición es importante por varias razones. En primer lugar, la reprobación de un año escolar tiene efectos negativos sobre el desarrollo emocional de los niños al disminuir la autoestima y reducir el interés por el estudio. En segundo lugar, la repetición es uno de los más importantes predictores de la deserción

escolar. Repetir por primera vez un grado aumenta la probabilidad de deserción entre el 40% y 50%, y repetir por segunda vez la aumenta al 90% (Unesco, 1998; Unesco, 2012). En tercer lugar, la repetición es mayor en la población más pobre, lo que genera un aumento de la inequidad (Unicef, 2007). En cuarto lugar, la repetición produce un impacto directo sobre la extraedad escolar, que afecta de manera directa la calidad de la educación. Además, extraedad y repetición en conjunto son factores asociados a la deserción escolar (García, Fernández & Sánchez, 2010; Rodríguez, 2014). En quinto lugar, las tasas de repetición son altas para países en desarrollo. Por ejemplo, América Latina presenta una tasa de repetición de 29% para toda la primaria y 42% para el primer grado. Siendo estas las más altas entre los países en desarrollo. Finalmente, además de los altos costos emocionales para los niños, la repetición tiene elevados costos económicos. Por ejemplo, América Latina gasta 4.200 millones de dólares cada año para atender los niños que repiten tan solo en primaria (Schiefelbein & Wolff, 2013).

Por otro lado, la literatura acerca de los efectos de la infraestructura sobre la calidad de la educación presenta resultados mixtos o heterogéneos. Parte de la literatura no encuentra efectos de la infraestructura sobre la calidad de la educación, mientras que otra ha encontrado efectos positivos y significativos. Esta ambigüedad en los resultados puede deberse a diferentes factores, tales como diversidad de los programas, contextos en el que estos se implementaron o el nivel de agregación de la información analizada (Hanushek, 1997; Glewwe *et al.*, 2011; Krishnaratne *et al.*, 2013). La presente investigación aporta a esta discusión empleando una rica y novedosa fuente de datos administrativos, donde la unidad de observación es el estudiante en cada año, y encuentra que la infraestructura reduce las tasas de repetición.

Esta investigación contribuye a la literatura en diferentes vías. Primero, presenta evidencia del efecto causal de la infraestructura sobre la tasa de repetición. Segundo, se analizan los canales por los cuales esto opera, por ejemplo, se estudian los efectos heterogéneos o diferenciales de los componentes de la infraestructura (área deportiva, biblioteca, laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora, salas especializadas, restaurante, sala para áreas artísticas, auditorio o coliseo y laboratorio de idiomas) sobre la repetición. Tercero, se emplea una gran fuente de información que permite asegurar la precisión de las estimaciones.

Una simple comparación entre colegios con acceso a una buena infraestructura y otros sin acceso podría conducir a estimaciones sesgadas, debido a que los efectos encontrados podrían deberse no necesariamente a la variación en infraestructura, sino a otra clase de insumos educativos o a las características de los estudiantes que asisten a cada una de ellas. Para resolver este problema, este trabajo explota la exogeneidad derivada de la construcción de 37 colegios en la ciudad de Bogotá entre 2006 y 2011, en el marco del Programa de Equidad en Educación en Bogotá. Estos colegios, que por sus características se han denominado megacolegios, fueron ampliamente dotados de diversos recursos locativos, tales como laboratorios, coliseos, restaurantes, entre otros. Los megacolegios fueron construidos en zonas donde se tenían carencias en oferta educativa e infraestructura. Esto indica que su construcción responde a necesidades propias del contexto. En algunos casos, los megacolegios reemplazaron colegios ya existentes con problemas de infraestructura.

Este trabajo emplea como estrategia de identificación la metodología de diferencias en diferencias (DID) a nivel del estudiante. Esta permite comparar el *mismo* niño antes y después de

estar en un megacolegio. Como controles, se emplean niños que estudian en colegios muy cercanos a los *megacolegios*, pero que no asistieron a un *megacolegio*. Las estimaciones se hicieron con efectos fijos de niño, para controlar por variables no observables, tales como las habilidades innatas o la preocupación de los padres. Se incluyen los efectos fijos de año, lo cual controla por inobservables constantes en el tiempo. Dado que la construcción de los megacolegios dependía de las condiciones de la zona en la que se construyeron (UPZ), las especificaciones restringen los datos a niños que asisten a colegios en las misma UPZ donde se construyeron los megacolegios o a UPZ vecinas.

Paralelamente al programa de megacolegios, la SED implementó el programa de reforzamiento de infraestructura y colegios en concesión. El primero consistía en adecuar las edificaciones que no cumplieran con la norma de sismorresistencia. Por su parte, el programa de concesión entregó veinticinco colegios públicos a particulares para su administración. Las estimaciones controlan por estos dos programas, que podrían sesgar los resultados si no se tuvieran en cuenta. Finalmente, es posible que la implementación del programa de megacolegios esté acompañada de cambios importantes en la planta de profesores o de directivos docentes. Esto podría confundir el efecto del programa de megacolegios. Por ejemplo, si al megacolegio llega un nuevo grupo de profesores y directivos docentes es difícil establecer si los efectos encontrados se debieron al mejoramiento de la infraestructura o al cambio de la planta docente o a una nueva administración. Para corregir esto, se incorporan dos controles adicionales: la tasa de rotación de profesores y la tasa de rotación de directivos docentes.

Este trabajo encuentra que colegios con mejor infraestructura tienen un efecto negativo y significativo sobre la tasa de repetición de 0,51 puntos porcentuales. Este es un efecto importante, si se tiene en cuenta que la tasa de repetición promedio para el periodo 2004-2012 fue de 6,35%, es decir, el impacto representa el 8,03% sobre la tasa promedio de repetición en dicho periodo. La literatura sugiere que mejores ambientes escolares generan mejores procesos educativos, que puede ser el canal por el cual opera este resultado (Durán-Narucki, 2008). Para someter a prueba esta hipótesis se estimaron los efectores parciales sobre once componentes de la infraestructura. Los efectos diferenciales entre los componentes de la infraestructura permiten concluir que los canales por los cuales la infraestructura opera son diversos. Se encuentra que los que generan impacto adicional en la reducción de la tasa de repetición son los laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora y laboratorio de idiomas.

Como ejercicio de robustez se hicieron las estimaciones empleando como placebo los datos justamente antes que el programa entrara en funcionamiento. La intuición de este ejercicio es que si el cambio de las tasas de repetición está asociado a la infraestructura, antes de la construcción no debería verse un efecto diferencial. Los resultados de estas estimaciones corroboran dicha intuición y sugieren que los efectos encontrados son causales y verdaderos.

El presente documento está dividido en ocho secciones; la sección dos presenta una revisión de la literatura que da cuenta de estudios que analizan el efecto de la infraestructura sobre calidad de la educación. La sección tres muestra las principales cifras de repetición. La sección cuatro explica de manera breve el programa que se está evaluado y la forma en que los individuos son elegidos y asignados a las instituciones educativas. La sección cinco explica la estrategia de identificación. La sección seis describe los datos empleados y sus fuentes, la sección

siete muestra y discute los resultados de las estimaciones y una serie de ejercicios de robustez. Finalmente, la sección ocho presenta las conclusiones.

REVISIÓN DE LA LITERATURA

La literatura acerca de los efectos de la infraestructura sobre la calidad de la educación presenta resultados mixtos. Parte de la literatura no encuentra efectos de la infraestructura sobre la calidad de la educación, mientras que otra ha encontrado efectos positivos y significativos. Poco se ha investigado el efecto de la infraestructura sobre variables como repetición o abandono. La presente investigación aporta a este debate al estudiar el efecto de la infraestructura sobre la repetición.

Hanushek (1997) hace una extensa revisión de la literatura sobre producción educativa para el periodo 1968-1994. Los cerca de cuatrocientos estudios analizados muestran que no existe consenso en la dirección de los efectos de la infraestructura sobre variables de resultado de los estudiantes, es decir, los resultados son mixtos. Adicionalmente, encuentra que tan solo el 25% de las investigaciones analizadas estudian variables de resultado diferentes a pruebas estandarizadas. Ninguno de los estudios analizados por Hanushek es a nivel de estudiante, y el 55,8% son a nivel de escuela. Por su parte, Glewwe *et al.* (2011) hacen una revisión de la literatura para el periodo 1990-2010. Su estudio, al igual que el de Hanushek (1997), muestra que los resultados de los efectos de la infraestructura sobre variables de logro académico, medidos como pruebas estandarizadas, son mixtos. Sus análisis sugieren que parte de la ambigüedad en los resultados puede deberse a que los tratamientos son heterogéneos, y a las diferencias de las circunstancias locales donde los programas se implementaron. Glewwe concluye que hay poca evidencia empírica sobre el efecto que podría tener una amplia variedad de características de la escuela y el maestro sobre la calidad. Los artículos analizados en esa revisión emplean como variable resultado las pruebas estandarizadas y no dan cuenta de efectos sobre variables de eficiencia interna como repetición y abandono, que son medidas importantes de la calidad educativa (Lee & Barro, 2001).

Krishnaratne *et al.* (2013) analiza una serie de estudios que abordan diferentes tipos de intervenciones sobre educación. En cuanto al tema particular de infraestructura, encuentra que la construcción de nuevos colegios impacta de manera positiva la matrícula, principalmente en el nivel de primaria y en las niñas. Esto puede deberse a que la matrícula de las niñas es más sensible a la distancia de la escuela al hogar. Estos resultados son consistentes con la evidencia presentada por Duflo (2001) en un estudio para Indonesia. Duflo (2001) encuentra que el aumento en el número de escuelas incrementó entre 0,12 y 0,19 años de estudio para niños entre 2 y 6 años de edad. Esto indica que aumentar la disponibilidad de infraestructura educativa, esto es, aumentar la oferta, tiene efectos positivos sobre la matrícula.

Mejorar las edificaciones existentes tiene un efecto positivo en asistencia escolar, en matrícula y en las pruebas estandarizadas de lenguaje y matemáticas. Por ejemplo, Earthman (2002) encuentra que el ambiente escolar influye de manera importante en los rendimientos escolares, y que una adecuada temperatura y calidad acústica del salón permiten una mayor concentración y motivación de los estudiantes hacia la clase. Por su parte, Durán-Narucki (2008),

en un estudio para 95 escuelas en Nueva York, encuentra que edificaciones menos deterioradas reducen la inasistencia escolar, mejoran la socialización de los estudiantes y el ambiente de aprendizaje, lo cual afecta positivamente las pruebas estandarizadas.

Por su parte, Glewwe & Jacoby (1994) presentan evidencia sobre el impacto del mejoramiento de las características de la escuela en el rendimiento estudiantil utilizando información para Ghana. Esta investigación presenta dos ejercicios, el primero emplea un modelo probit para identificar qué características de los colegios aumentan la probabilidad de que los padres inscriban a sus hijos. Las estimaciones muestran que los padres se sienten atraídos por escuelas con infraestructura y recursos didácticos mejores. El segundo ejercicio emplea mínimos cuadrados ordinarios en dos etapas para encontrar los determinantes de los resultados en pruebas estandarizadas de matemáticas y lenguaje. Aunque los resultados no son estadísticamente muy fuertes, estos sugieren que es más efectiva en cuanto a costos la adecuación o mejora de las instituciones educativas que dotación de bibliotecas, reducción del tiempo de desplazamiento de los estudiantes a la escuela y capacitación a maestros. Sin embargo, no se analiza el efecto sobre la repetición o deserción.

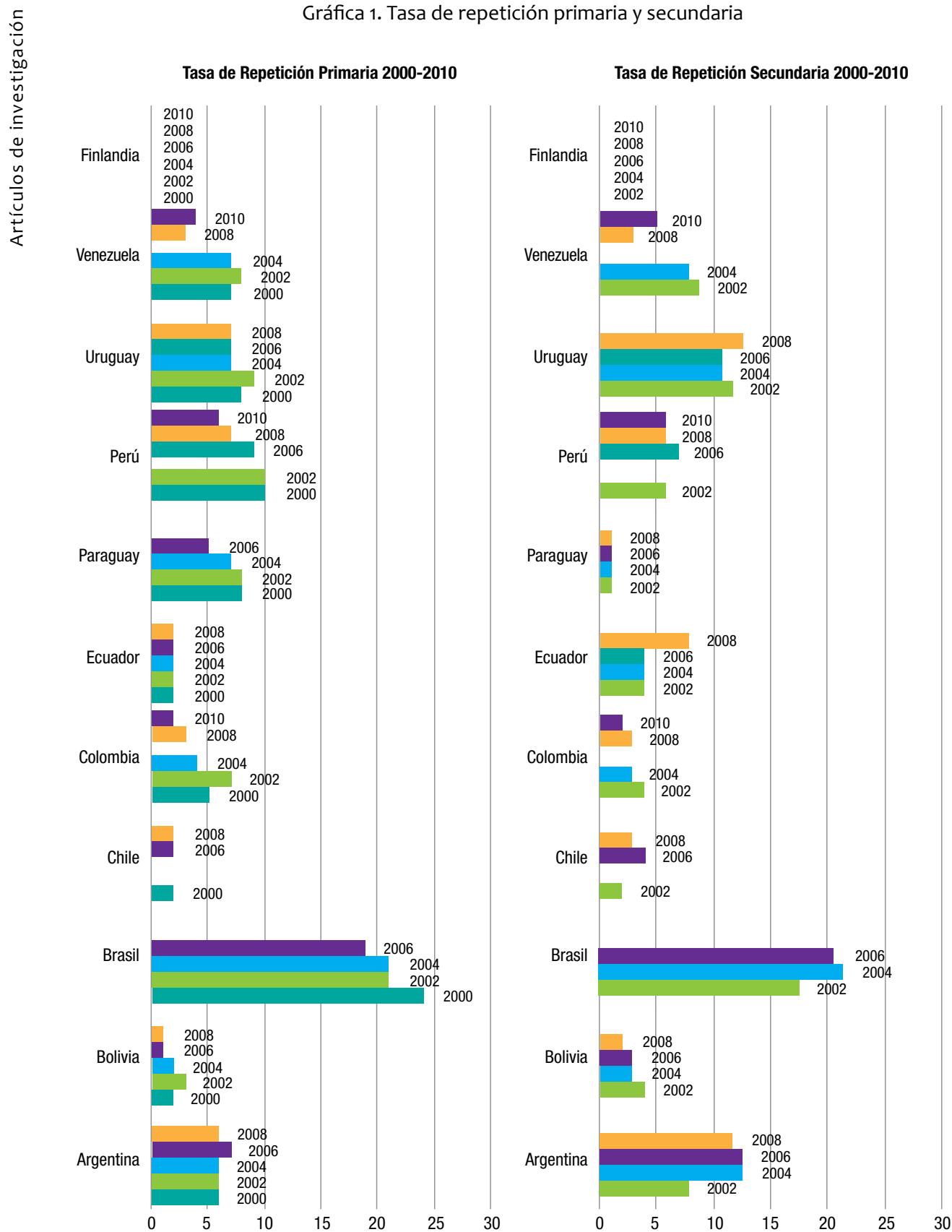
La presente investigación contribuye a la literatura en diferentes vías. Primero, presenta evidencia del efecto causal de la infraestructura sobre la tasa de repetición. Segundo, se analizan los canales por los cuales esto opera. Para esto se estudian los efectos parciales de once componentes de la infraestructura sobre la repetición. Los componentes son: área deportiva, biblioteca, laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora, salas especializadas, restaurante, sala para áreas artísticas, auditorio o coliseo y laboratorio de idiomas. Tercero, se emplea una gran fuente de información longitudinal de los niños que permite asegurar la precisión de las estimaciones; esta estructura de datos panel por niño permite controlar características no observadas que son constantes en el tiempo.

CIFRAS SOBRE REPETICIÓN

Las cifras sobre repetición en el mundo presentan una tendencia a la baja. Antes de 1985, alrededor del 15% de los niños repetían por lo menos un grado en la escuela primaria; en el 2000, este porcentaje llegó casi al 6%, y para 2010 estaba ligeramente por debajo del 5%. Las diferencias entre regiones son grandes. África Subsahariana, por ejemplo, redujo su tasa de repetición del 17% en 1985 al 9% en 2010, mientras que América Latina la redujo del 14% al 8%, siendo estas dos regiones las de tasas de repetición más altas. La historia en otras regiones es diferente. Por ejemplo, América del Norte y Europa Occidental han tenido tasas de repetición baja, las cuales entre 2000 y 2010 fueron alrededor del 1%. Por su parte, Asia Central cuenta con las tasas de repetición más bajas del mundo estando muy cerca del 0% entre 2000 y 2010 (Unesco, 2012).

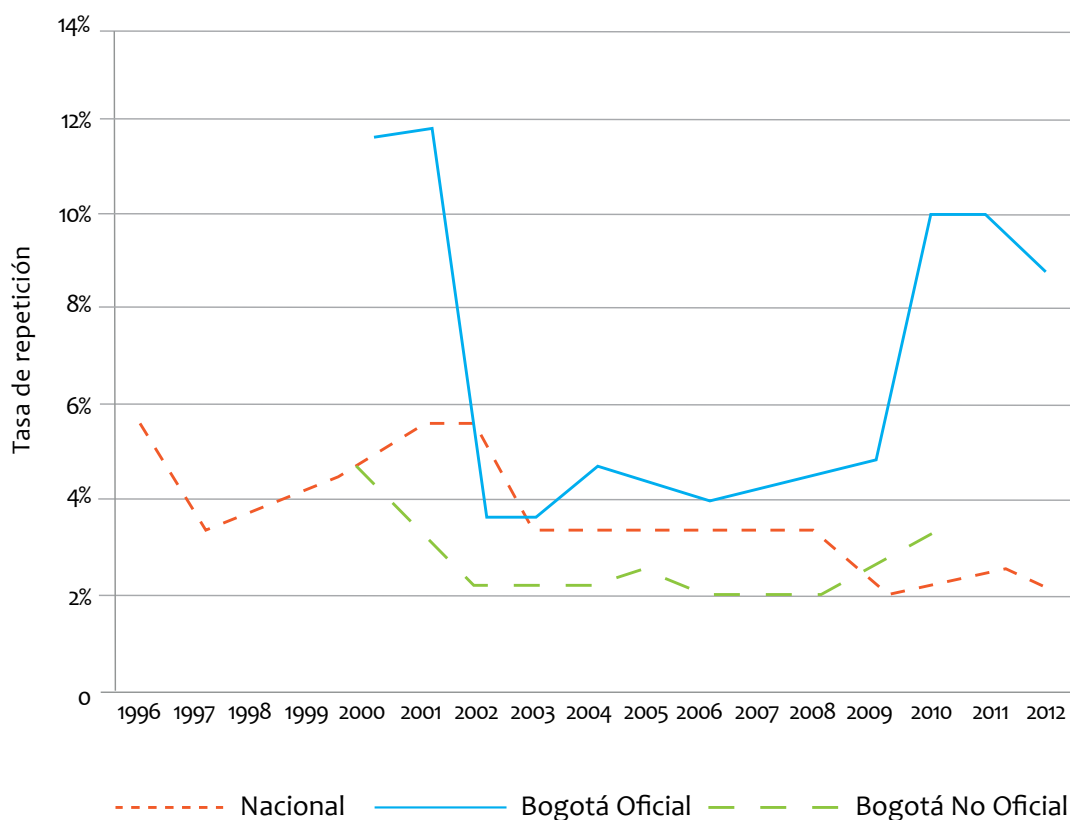
Para el caso particular de América Latina se observa una gran variabilidad entre países. La gráfica 1 muestra la tasa de repetición para diez países de la región. Brasil presenta la mayor tasa de repetición para el periodo del 2000 al 2006. Por ejemplo, Brasil en 2006 presentó una tasa de repetición de 19% en primaria y 21% en secundaria. Por su parte, el país con menores tasa de repetición en primaria para el 2006 es Bolivia con un 1% le sigue Chile y Ecuador, ambos con una tasa de repetición del 2% en el mismo año.

Gráfica 1. Tasa de repetición primaria y secundaria



Para el caso colombiano, la gráfica 2 muestra la evolución de la tasa de repetición nacional entre 1996 y 2012, y para Bogotá, entre 2000 y 2012, discriminada entre instituciones oficiales y no oficiales. La repetición en Colombia presentó un pico en 2002 de 5,47%, que cayó casi dos puntos porcentuales en 2003. A partir de 2009, la tasa de repetición se estabiliza alrededor del 2,28%, la más baja de los últimos veinte años. La reducción de la tasa de repetición, para el 2003, podría ser explicada por el cambio en la normativa sobre promoción que fue aprobada en 2002, mediante el Decreto 230, y derogada en 2009. Esta normativa dio instrucciones a las instituciones educativas para que la repetición no superara el 5% por grado. Esta norma también explica las dos grandes fluctuaciones en la tasa de repetición para Bogotá, la cual pasó del 11,8% en 2002 al 3,6% en 2003, y del 4,8% en 2009 al 10% en 2010. Aunque en Bogotá es evidente el aumento de la repetición debido al cambio de la norma, a nivel nacional la tasa agregada de repetición para los últimos tres años no subió significativamente.

Gráfica 2. Tasa de repetición en Colombia

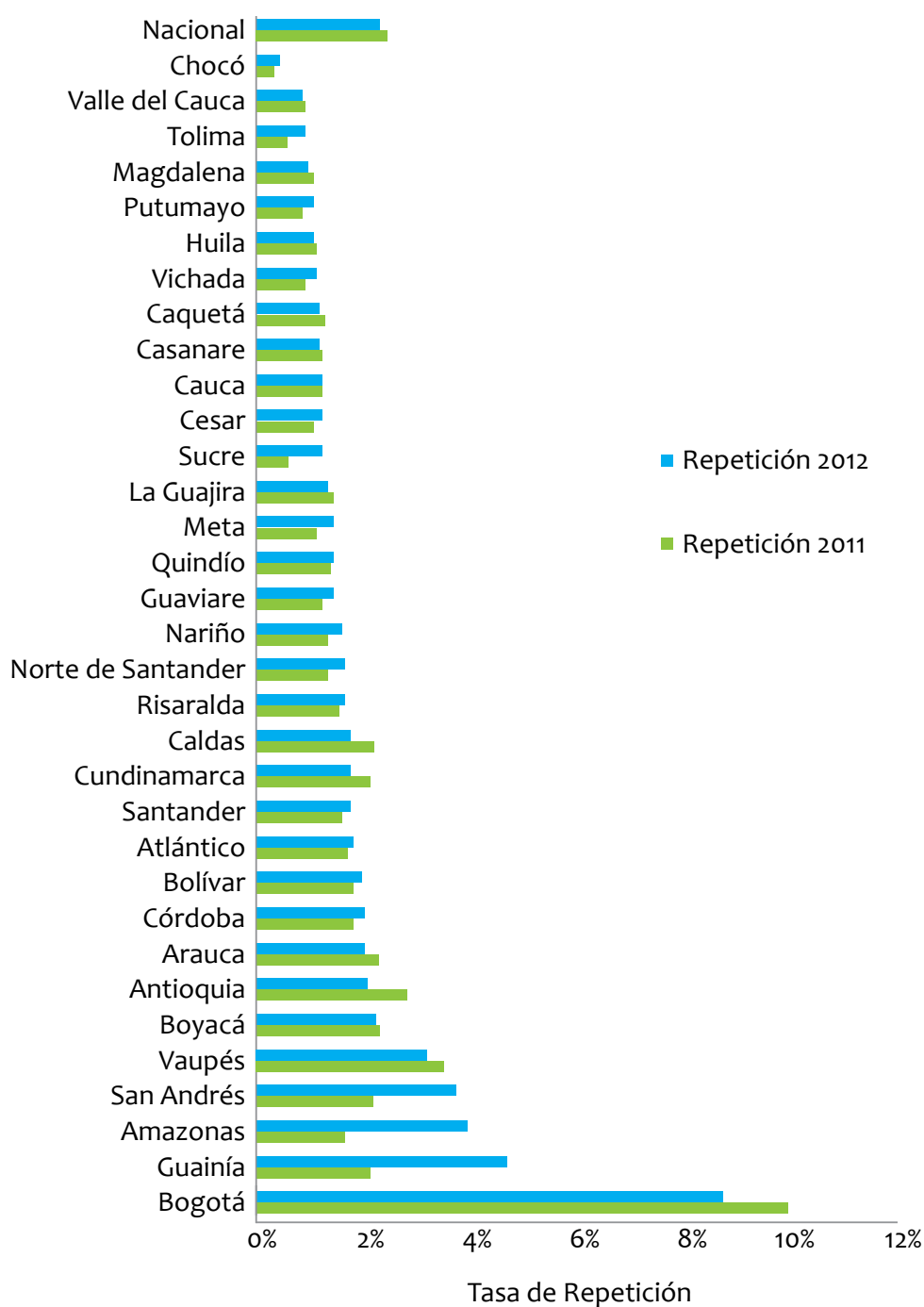


Fuente: MEN y SED

La disparidad entre regiones en Colombia es importante. La gráfica 3 muestra la tasa de repetición por departamento en Colombia en 2011 y 2012. La tasa más alta de repetición la tiene Bogotá: 9,91% en 2011 y 8,69% en 2012. Le siguen en la lista Guainía, Amazonas y San Andrés, departamentos que atienden a menos de 20.000 estudiantes (a diferencia de

Bogotá, que atiende a más de 800.000). Por otro lado, Antioquia atiende a un poco más de un millón de estudiantes; sin embargo, tuvo una tasa de repetición de 2,1% para el 2012. Esto muestra una gran heterogeneidad en cuanto a número de personas atendidas por región y tasa de repetición.

Gráfica 3. Tasa de repetición en Colombia, por departamentos



Fuente: MEN-SIMAT.

Adicionalmente, del total de niños que reprueban en Colombia, Bogotá puso el 41,3% en 2011 y el 38,6% en 2012, A pesar de que Bogotá solo tiene una participación en la matrícula del país cercana al 10%. La gráfica 2 muestra que Bogotá, particularmente el sector oficial, ha tenido tasas altas de repetición en los años en los que no operó la norma sobre topes máximos de reprobación.

PROGRAMA DE EQUIDAD EN EDUCACIÓN EN BOGOTÁ

En el marco del Plan de Desarrollo para el periodo 2004-2008, denominado Bogotá sin Indiferencia, la Alcaldía estableció políticas y metas para el desarrollo del sector educativo que responden a su preocupación de aumentar y mejorar la cobertura, aumentar los niveles de retención y graduación de los jóvenes y, en general, aumentar la calidad del servicio educativo. Para cumplir este propósito diseñó e inició en 2006 el Programa de Equidad en Educación en Bogotá.¹ Aunque la meta inicial del plan sectorial de educación era construir cincuenta nuevos colegios (SED, 2004), a 2012 se habían construido 37 colegios en Bogotá, que por sus dimensiones se denominaron *megacolegios*. Este programa cuenta, para la construcción de los últimos veinte megacolegios, con el apoyo financiero del Banco Interamericano de Desarrollo (BID) y una contrapartida importante por parte de la Secretaría de Educación de Bogotá (SED). La inversión total para este programa fue de 371.562 millones de pesos, lo que equivale a cerca del 30% del presupuesto para educación en 2012.

Para seleccionar los lugares en donde se debían construir los megacolegios se tuvieron en cuenta zonas donde la oferta de infraestructura educativa era precaria. Adicionalmente se analizaron factores socioeconómicos de la población por UPZ² que dieran cuenta de la vulnerabilidad de la población. Esto indica que los megacolegios fueron construidos en las zonas más necesitadas de Bogotá. Teniendo en cuenta que una de las condiciones para la construcción de un megacolegio era la precaria infraestructura, gran parte de los megacolegios se hicieron con el fin de reemplazar o consolidar en una sede varias sedes educativas ya existentes que estaban en malas condiciones o que eran sedes improvisadas (salones en zinc o prefabricados). Esto implicó que un gran número de estudiantes beneficiarios fueran estudiantes regulares de sedes ya existentes que fueron absorbidas por el megacolegio más cercano.

Asignación a colegios en Bogotá

La Secretaría de Educación Distrital recoge vía web las solicitudes de cupos de alumnos nuevos o de quienes desean traslado al final de cada año. El formulario de inscripción solicita un orden de preferencias sobre las instituciones en las cuales desea estudiar el niño, variables socioeconómicas y datos de ubicación. Con esta información se hace la asignación a los colegios en Bogotá de manera

1 Este programa articuló varias estrategias del Plan Sectorial de Educación Bogotá una Gran Escuela, para el período 2004-2008, en el marco del plan de desarrollo Bogotá sin Indiferencia.

2 Las Unidades de Planeamiento Zonal (UPZ) en Bogotá son aquellas zonas urbanas que reúnen una serie de barrios homogéneos entre sí y hacen parte de una localidad. Sirven como unidad para planificar el desarrollo urbano en el nivel zonal. Bogotá cuenta con 117 UPZ, distribuidas en veinte localidades.

centralizada, la cual responde a varios criterios, tales como disponibilidad de cupo, preferencias del estudiante, menor distancia del hogar al colegio y vulnerabilidad del hogar. Esto indica que la asignación a los colegios es difícilmente influenciable, por ejemplo, por padres más preocupados por sus hijos. Sin embargo, la asignación depende de las condiciones socioeconómicas de los hogares, haciendo que instituciones en zonas determinadas tengan estudiantes similares entre sí, pero diferentes entre instituciones. En cuanto a la asignación de estudiantes a los megacolegios, esta tuvo una dinámica diferente. Como los megacolegios, en su mayoría, se construyeron para reemplazar sedes antiguas y dispersas, los estudiantes de las sedes viejas fueron asignados a la nueva sede de manera automática. En el caso en que los megacolegios contaran con cupos adicionales, estos se llenaron empleando el método de asignación habitual y de manera centralizada. Esto indica que la elección de pertenecer a un megacolegio es en gran medida exógena a las variables no observables de los estudiantes, pero endógena a sus condiciones, y especialmente al lugar de residencia del hogar.

DATOS

El presente trabajo emplea principalmente dos fuentes de datos. El primer conjunto de datos es el panel censal de estudiantes de la Resolución 166 del Ministerio de Educación Nacional para el periodo 2004-2012, el cual cuenta con información del estado de la matrícula por niño en escuelas públicas. Esta base identifica el colegio en el que está matriculado el niño, el grado, la situación académica del año anterior (aprobó o reprobó), y si está repitiendo en el año actual; información sociodemográfica como fecha de nacimiento, género, si proviene de otro municipio, proviene del sector privado, estrato, víctima de conflicto, y si la madre es cabeza de familia. Una segunda fuente de datos empleados es la que contiene información sobre los proyectos de mejoramiento de la infraestructura de colegios en Bogotá suministrada por la SED. Esta base de datos contiene información sobre las fechas en las cuales iniciaron las obras, la fecha de terminación y entrega de las obras y los costos por sede educativa y las características de la infraestructura, particularmente la de los megacolegios. Esta última información se validó a través de comunicación directa con cada uno de los megacolegios. Es clave anotar que no todos los megacolegios fueron dotados exactamente con la misma infraestructura. Por ejemplo, como se puede ver en la tabla 1, el 78,4% de los megacolegios fueron dotados con salas de artes (sala de música, artística o danzas), 73% con bibliotecas, 70% con laboratorios de ciencias o químicas, 67,6% con salas especializadas (audiovisuales, salas polivalentes o ludotecas), 24,3 % con restaurante, 21,6% con laboratorio de idiomas y 18,9 con emisora. Esta variación se explota en el presente trabajo.

Tabla 1. Componentes de la infraestructura de los megacolegios

| | |
|-------------------------------|-------|
| Artes | 78,4% |
| Biblioteca | 73,0% |
| Ciencias o Química | 70,3% |
| Salas especializadas | 67,6% |
| Auditorio o coliseo | 48,6% |
| Restaurante | 24,3% |
| Laboratorio de idiomas | 21,6% |
| Emisora | 18,9% |

La tabla 2 presenta en el panel A las estadísticas descriptivas de todo el panel disponible, compuesto por todos los niños matriculados en alguna institución educativa pública en Bogotá observados entre 2004 y 2012. En promedio, el panel observa 919.179 individuos por año. Por su parte, el panel B de la misma tabla presenta las estadísticas descriptivas restringidas a 2006 para los grupos de estudiantes tratados (*estar en un megacolegio*) y los controles (*no estar en un megacolegio*). Se elige este año pues es a partir del 2007 que se construye la mayoría de los megacolegios. El panel formado por las observaciones de estos niños entre 2004 y 2012 corresponde a las empleadas en las estimaciones. Se tiene que el grupo tratado está formado por 55.685 individuos, y el control, por 341.769. El criterio para estar en este panel es tener observaciones antes y después de iniciado el tratamiento; además para el grupo control es necesario que el niño esté en un colegio en la misma UPZ de un megacolegio o en la UPZ vecina.

Tabla 2. Estadísticas descriptivas

| Variable | Panel A | | | Panel B: Año 2006 | | | | | | | |
|--------------------------|-------------|-------|------|-------------------|-------|------|---------------|-------|------|-------|-----|
| | Total Panel | | | Grupo Control | | | Grupo Tratado | | | Dif. | |
| | Obs. | Media | D.E. | Obs. | Media | D.E. | Obs. | Media | D.E. | | |
| Repetición | 8272612 | 0,06 | 0,24 | 341769 | 0,03 | 0,17 | 55685 | 0,03 | 0,16 | 0,00 | |
| Edad | 8272612 | 11,42 | 4,02 | 341769 | 8,72 | 2,41 | 55685 | 8,70 | 2,30 | 0,01 | |
| Mujer | 8272612 | 0,50 | 0,50 | 341769 | 0,50 | 0,50 | 55685 | 0,49 | 0,50 | 0,01 | *** |
| Pertenece a alguna etnia | 8272612 | 0,00 | 0,06 | 341769 | 0,01 | 0,07 | 55685 | 0,01 | 0,07 | 0,00 | |
| Víctima de Conflicto | 8272612 | 0,03 | 0,16 | 341769 | 0,01 | 0,12 | 55685 | 0,02 | 0,15 | -0,01 | *** |
| Discapacidad | 8272612 | 0,01 | 0,09 | 341769 | 0,01 | 0,10 | 55685 | 0,00 | 0,07 | 0,00 | *** |
| Reprobó año anterior | 8272612 | 0,05 | 0,21 | 341769 | 0,01 | 0,11 | 55685 | 0,01 | 0,11 | 0,00 | |
| Jornada completa | 8272612 | 0,16 | 0,37 | 341769 | 0,22 | 0,42 | 55685 | 0,09 | 0,28 | 0,13 | *** |
| Jornada mañana | 8272612 | 0,42 | 0,49 | 341769 | 0,42 | 0,49 | 55685 | 0,49 | 0,50 | -0,07 | *** |
| Jornada tarde | 8272612 | 0,35 | 0,48 | 341769 | 0,35 | 0,48 | 55685 | 0,42 | 0,49 | -0,07 | *** |
| Colegio técnico | 8272612 | 0,01 | 0,11 | 341769 | 0,00 | 0,00 | 55685 | 0,00 | 0,00 | 0,00 | |
| Preescolar | 8272612 | 0,06 | 0,23 | 341769 | 0,11 | 0,32 | 55685 | 0,11 | 0,31 | 0,00 | |
| Primaria | 8272612 | 0,41 | 0,49 | 341769 | 0,67 | 0,47 | 55685 | 0,68 | 0,47 | -0,01 | *** |
| Bachillerato | 8272612 | 0,48 | 0,50 | 341769 | 0,18 | 0,39 | 55685 | 0,18 | 0,38 | 0,01 | *** |
| Grado | 8180389 | 6,27 | 6,90 | 338755 | 3,61 | 5,25 | 55197 | 3,57 | 5,31 | 0,05 | * |
| Subsidiado | 8272612 | 0,18 | 0,38 | 341769 | 0,25 | 0,43 | 55685 | 0,07 | 0,26 | 0,17 | *** |
| Estrato | 8272612 | 1,97 | 0,75 | 341769 | 1,99 | 0,74 | 55685 | 1,81 | 0,70 | 0,18 | *** |
| Proviene sector privado | 8272612 | 0,03 | 0,18 | 341769 | 0,07 | 0,25 | 55685 | 0,07 | 0,25 | 0,00 | |
| Proviene otro municipio | 8272612 | 0,01 | 0,12 | 341769 | 0,01 | 0,11 | 55685 | 0,02 | 0,15 | -0,01 | *** |

D.E.= Desviación estándar.

* p<0.05, ** p<0.01, *** p<0.001.

En el panel A se observa que la tasa de repetición para el periodo de estudio es de 6,3%, alta si se compara con el 3,57% del promedio nacional. Por su parte, la proporción de hombres y mujeres es muy similar, lo cual se conserva en la población estudiada (panel B). El 42,1% de los niños estudian en la mañana, mientras que el 34,6% en la tarde, y tan solo el 16,2% lo hacen en jornada completa. En promedio, la población es de estrato dos.

El panel B, por su parte, muestra que la diferencia entre la repetición de los niños tratados y los controles, antes del tratamiento, no es estadísticamente significativa; esto indica que antes del tratamiento, la tasa de repetición entre los dos grupos es la misma. Aunque en general las diferencias de medias entre el grupo tratado y el de control son pequeñas, gran parte de estas son estadísticamente significativas. Esto indica que, en promedio, los niños que fueron a un megacolegio son sistemáticamente diferentes a los que no fueron. Sin embargo, la metodología de diferencias en diferencias reduce ese problema de sesgo, en especial para la especificación propuesta en el presente trabajo, pues esta compara el mismo niño antes y después de estar en un megacolegio. Lo anterior se puede hacer, debido a la rica fuente de datos con que se cuenta, ya que permite observar el mismo individuo a lo largo de varios periodos.

ESTRATEGIA DE IDENTIFICACIÓN

La repetición, al igual que la deserción y los resultados en pruebas estandarizadas, da cuenta de la calidad educativa. Escuelas más efectivas detectan mejor las debilidades de sus estudiantes y pueden establecer mecanismos que permitan evitar el fracaso académico. El presente trabajo muestra evidencia empírica sobre el efecto que tiene la infraestructura escolar sobre la repetición. En contraste con la literatura que emplea datos agregados por países, distritos escolares o escuelas, la presente investigación emplea datos longitudinales a nivel del estudiante. Esta estructura de datos permite controlar por un gran número de variables no observadas, y en particular tiene dos ventajas insuperables. Primero, es posible controlar por características no observables, íntimamente relacionadas con los resultados académicos como su habilidad innata o la importancia que recibe la educación en el hogar. Segundo, al observar cada estudiante es posible analizar cómo varía su probabilidad de repetir un grado académico cuando cambia de un colegio normal a un megacolegio.

Los megacolegios se construyeron en zonas donde la oferta educativa era escasa, y precaria la infraestructura, particularmente en zonas de la ciudad más pobres. Esto sugiere que los estudiantes que asisten a los megacolegios son potencialmente diferentes a los que asisten a otros colegios. Esto puede generar problemas de sesgo en las estimaciones. Para resolver este problema, la presente investigación emplea la metodología de diferencias en diferencias (DID), la cual aprovecha el hecho de que existen niños tratados y no tratados observados antes y después de tener la intervención. Esto permite comparar cada niño consigo mismo antes y después de estar en un colegio con mejor infraestructura.

Además, para que el efecto estimado se pueda considerar realmente causal, los niños que fueron a colegios con mejor infraestructura deben ser comparados con niños que

no fueron (control), pero que cuenten con características muy similares. En especial, se requiere que tengan la misma probabilidad de haber sido tratados. Dado que el criterio más importante de focalización del programa de megacolectivos fue que estos se ubicaran en las zonas más necesitadas, se eligieron como grupo control los niños que, por sus condiciones geográficas relacionadas con el colegio donde estudian, tuvieran condiciones similares a los niños tratados. Para la identificación del grupo control se buscaron niños que estudiaran en las mismas UPZ donde se construyeron los megacolectivos, o en las UPZ más próximas. Esto se hizo teniendo en cuenta que las condiciones socioeconómicas de la población entre distintas UPZ pueden ser diferentes entre sí; sin embargo, por la forma en que se diseñaron las UPZ, se asume que dentro de cada UPZ la población es similar en términos de vulnerabilidad, necesidades básicas y nivel socioeconómico.

El modelo DID que permite identificar el efecto de la infraestructura sobre la repetición se especifica en la siguiente ecuación:

$$Repetición_{i,t} = \beta_0 + \beta_1 Tratado_{i,t} + \psi X_{i,t} + \gamma_i + \gamma_t + \mu_{i,t} \quad (1)$$

donde $Repetición_{i,t}$ es una variable que indica que es 1 si el individuo i en el año t repite el año escolar, y cero en caso contrario; $Tratado_{i,t}$ es una variable que indica que es 1 si el individuo i en el año t está en un megacolectivo, y cero en otro caso. Para este estudio consideramos que el tratamiento es haber estado en un megacolectivo. Por su parte, $X_{i,t}$ es una matriz que representa las variables socioeconómicas o demográficas del estudiante y las variables de la escuela que varían en el tiempo. Los γ son los efectos fijos de individuo (i) y de año (t). Los efectos fijos permiten controlar, en el caso de los individuos, por características de los niños que son constantes en el tiempo y que no son observables, por ejemplo, el gusto por el estudio, la motivación, su capacidad intelectual, entre otras. Por su parte, los efectos fijos de año permiten controlar por aquellas cosas que cambian con la tendencia del temporal y que podrían afectar la variable resultado. Finalmente, $\mu_{i,t}$ es el término de error.

En la presente investigación, el coeficiente de interés es β_1 , el cual indica el efecto de estar en un megacolectivo sobre la repetición. Un signo negativo de β_1 indica que estar expuesto a una mejor infraestructura, ir a un megacolectivo, reduce la tasa de repetición.

Entender cuáles son los canales por los cuales la infraestructura puede operar es importante en términos de política pública. Esto permite entender qué tipo de infraestructura es más eficiente y, de esta manera, tomar decisiones más acertadas sobre la inversión. Se espera que escuelas con mejor infraestructura, que cuentan con aulas especializadas, laboratorios y restaurantes, generan ambientes de enseñanza y aprendizaje más adecuados que instituciones educativas con peor infraestructura (Earthman, 2002). Por otro lado, las instituciones con mejores instalaciones proveen un mejor ambiente que las casas de los estudiantes, promoviendo una mayor motivación frente a la permanencia de los estudiantes en los colegios (Durán-Narucki, 2008).

Con el fin de establecer los canales por los cuales actúa la infraestructura, la presente investigación estima los efectos parciales de los diferentes componentes de la infraestructura sobre la repetición. Para esto se agregó a la ecuación (1) la variable $CI_{i,t}$, la cual toma el valor de

1 si el individuo i en el año t está en un megacolegio que posee componente de infraestructura CI , y cero, de lo contrario. Adicionalmente, se incorpora a la ecuación la interacción entre esta nueva variable y la variable $T_{i,t}$, donde $T_{i,t}$ es una variable indicadora que es 1 si el individuo i en el año t está en un megacolegio, y cero en otro caso. La especificación es la siguiente:

$$Repetición_{i,t} = \beta_0 + \beta_1 T_{i,t} + \beta_2 CI_{i,t} + \beta_3 T_{i,t} CI_{i,t} + \psi X_{i,t} + \gamma_{UPZ} + \gamma_i + \gamma_t + \mu_{i,t} \quad (2)$$

De acuerdo con esta especificación, el efecto parcial del componente de infraestructura CI sobre repetición, con todas las demás variables fijas, es

$$\frac{\Delta Repetición}{\Delta CI} = \beta_2 + \beta_3 T \quad (3)$$

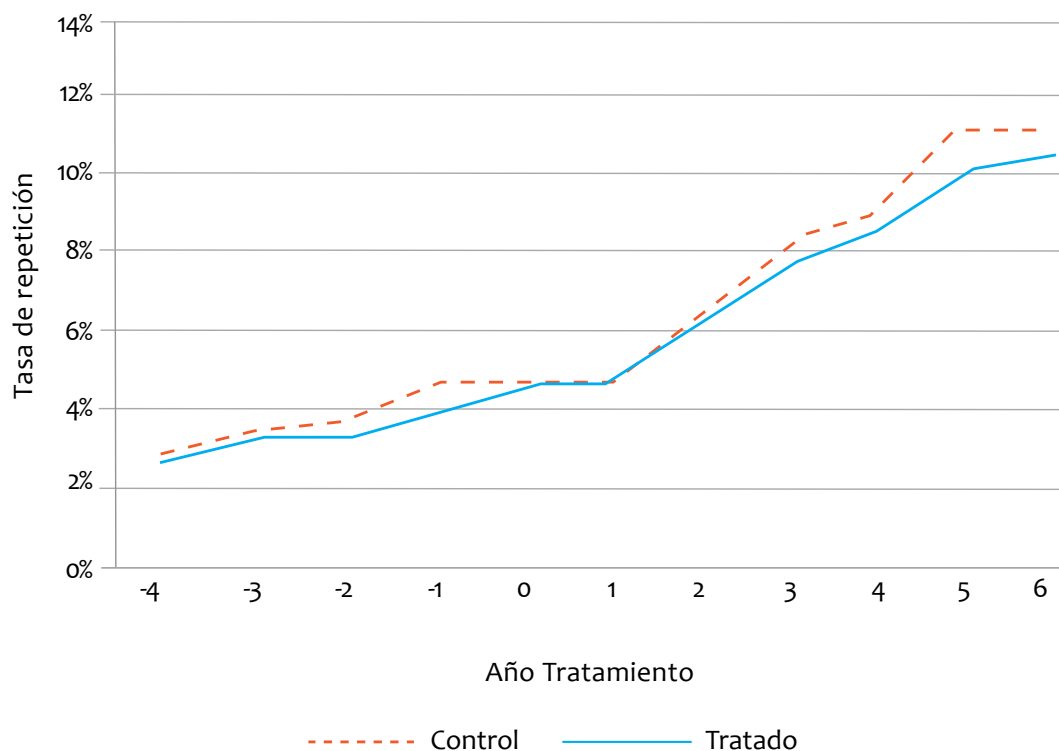
Si se hace $T=1$ en la ecuación (3), esto permite establecer el efecto parcial del componente de infraestructura en la repetición sobre los individuos tratados. En este caso, $\beta_2 + \beta_3 < 0$ implica que un megacolegio que cuenta con el componente de infraestructura CI produce una reducción adicional en repetición para los individuos tratados en comparación con los colegios que no tiene ese mismo CI .

RESULTADOS

Validez del modelo

El modelo DID tiene como supuesto que la pendiente de la variable de resultado, antes del tratamiento, sea la misma para el grupo tratado y el grupo control. Esto se denomina *supuesto de tendencias paralelas*. La gráfica 4 muestra la tasa de repetición para el grupo tratado y el grupo control reescalando los años de tal manera que el cero coincida con el año en que inicio el programa para cada niño tratado. La gráfica muestra que la pendiente de la tasa de repetición para estudiantes tratados y no tratados es similar para antes de iniciar el tratamiento ($t=0$). Para corroborar esta inspección visual se hizo una regresión entre la tasa de repetición de tratados y no tratados usando los datos para antes del inicio del tratamiento. En el caso de tendencias perfectamente paralelas se espera que la pendiente de la regresión sea 1. Los resultados muestran una pendiente de 0,98 y R cuadrado de 0,995 (anexo 1), lo cual indica que antes del tratamiento, las tasas entre tratados y no tratados tienen el mismo comportamiento. Tanto la gráfica como los resultados de la regresión sugieren que se cumple el supuesto de tendencias paralelas.

Gráfica 4. Tasa de repetición grupo tratado y control



Fuente: SED cálculos propios.

Repetición

La tabla 3 presenta los resultados para diferentes especificaciones de la ecuación 1. La estimación de la columna (1) usa efectos fijos de niño; esto permite controlar por variables no observables del niño, tales como sus habilidades innatas o la preocupación de sus padres, que pueden afectar la variable resultado. Además, incluye los efectos fijos de año, lo cual controla por las diferencias entre periodos, independientemente de si el individuo fue tratado o no. Dado que la construcción de los megacolegios dependía de las condiciones de la UPZ, en la columna (2) se emplean efectos fijos de UPZ; esto controla la estimación por variables omitidas que sean constantes entre las UPZ.

La columna (3) agrega condiciones del individuo que pueden cambiar en el tiempo, tales como edad, edad al cuadrado, un indicador de si proviene de otro municipio, un indicador de si proviene del sector privado, jornada académica y nivel educativo (primario o bachillerato). Esto con el fin de controlar por características observables que permiten mejorar la comparabilidad de los individuos tratados y no tratados.

Tabla 3. Efecto de la infraestructura sobre la repetición

| | (1) | (2) | (3) | (4) | (5) |
|--|------------|------------|------------|------------|------------|
| Tratamiento | -0,0051*** | -0,0045*** | -0,0066*** | -0,0039*** | -0,0124*** |
| | (0,001) | (0,001) | (0,001) | (0,001) | (0,001) |
| Controles | | | | | |
| Socioeconómicas y académicas | No | No | Sí | Sí | Sí |
| Otros programas | No | No | No | Sí | Sí |
| Tasa rotación profesores y directivos docentes | No | No | No | No | Sí |
| Constante | -0,017*** | 0,0004 | 0,458*** | 0,459*** | 0,484*** |
| | (0,000) | (0,012) | (0,013) | (0,013) | (0,013) |
| Observaciones | 3,064,317 | 2,986,609 | 2,986,609 | 2,986,609 | 2,986,609 |
| Efectos Fijos de Niño | Sí | Sí | Sí | Sí | Sí |
| Efectos fijos año | Sí | Sí | Sí | Sí | Sí |
| Efectos fijos UPZ | No | Sí | Sí | Sí | Sí |
| Errores estándar robustos en paréntesis | | | | | |
| *** p<0.01, ** p<0.05, * p<0.1 | | | | | |

Paralelamente al programa de construcción de megacolegios, la SED realizó un plan de reforzamiento de infraestructura de colegios que consistía en adecuar las edificaciones que no cumplieran con la norma de sismorresistencia. Además, la SED entregó en esta época veinticinco colegios en concesiones, esto es, el estado entregó colegios públicos a particulares para su administración. Para controlar por estos otros programas se construyeron dos variables indicadoras: concesión y reforzamiento. La variable *concesión* le asigna 1 al individuo en el año que estuvo matriculado en un colegio en concesión, y cero, de lo contrario. Por su parte, la variable *reforzamiento* le asigna 1 al individuo en el año que estuvo matriculado en un colegio que fue reforzado, y cero, de lo contrario. La columna (4) incluye los controles de estos otros programas.

Es posible que la implementación del programa de megacolegios esté acompañada de cambios importantes en la planta docente o administrativa. Esto podría confundir el efecto del programa de megacolegios. Por ejemplo, si al megacolegio llega un nuevo grupo de profesores y directivos docentes es difícil establecer si los efectos encontrados se debieron al mejoramiento de la infraestructura o al cambio de la planta docente o a una nueva administración. Para controlar por los posibles efectos que podrían derivarse del cambio de maestros o directivos docentes, la columna (5) incorpora dos controles adicionales: la tasa de rotación de profesores y la tasa de rotación de directivos docentes. La tasa de rotación de profesores se mide como el número de profesores nuevos que llegan a la institución educativa sobre el total de profesores. . Por su parte, la tasa de rotación de directivos docentes es el porcentaje de directivos docentes nuevos que llegan a cada institución.

Teniendo en cuenta que el Gobierno nacional emitió un decreto que estableció en 2002 una tasa de repetición máxima del 5%, y que este decreto fue derogado en 2010, se construyó una variable indicadora, que es 1 si el año es igual o superior a 2010, y cero, de lo contrario. Además, esta variable interactuó con el tratamiento, para controlar por los efectos diferenciales que pudiera tener la norma sobre el grupo tratado y el grupo control. Todas las especificaciones tiene este control.

Todas las estimaciones muestran consistentemente un efecto negativo y estadísticamente significativo del tratamiento. Esto sugiere la existencia de un efecto negativo de la infraestructura sobre la repetición. En la columna (1), el coeficiente de la variable *Tratado* es negativo y significativo, lo cual indica que el impacto del tratamiento reduce la repetición en 0,51 puntos porcentuales. Este es un efecto importante, si se tiene en cuenta que la tasa de repetición promedio para el periodo 2004-2012 fue de 6,35%, es decir, el impacto de la infraestructura sobre la repetición equivale a una reducción del 8,03% de la tasa promedio de repetición para el periodo analizado.

Canales

Los resultados muestran que una mejor infraestructura tiene un efecto sobre la reducción de la tasa de repetición de los estudiantes. Sin embargo, estas estimaciones no permiten ver qué canales de la infraestructura están operando. Para superar esta dificultad se aprovecha que no todos los megacolegios fueron dotados con exactamente la misma infraestructura. La presente investigación clasificó la infraestructura en once componentes y estimó para cada componente sus efectos parciales sobre la repetición, ecuación (3). Las componentes de la infraestructura son: área construida, área deportiva, biblioteca, laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora, salas especializadas, restaurante, sala para áreas artísticas, auditorio o coliseo y laboratorio de idiomas.

Cada columna de la tabla 4 muestra las estimaciones realizadas para uno de los once componentes de la infraestructura. Cada estimación presenta el coeficiente del tratamiento, el coeficiente del componente de la infraestructura, CI, el coeficiente de la interacción entre CI y el tratamiento y el efecto parcial del componente de infraestructura CI sobre repetición con $T=1$ en la ecuación (3). Para ejemplificar la forma en que se interpretan los resultados se analiza el componente de infraestructura laboratorios de ciencias o química, columna (4). El efecto parcial de tener un laboratorio de ciencias o química sobre la repetición es -0,002. Esto indica que dotar un megacolegio con laboratorios de ciencias o química, bajo el supuesto de que todos lo demás permanece constante, tiene un efecto adicional de reducción de la tasa de repetición de 0,2 puntos porcentuales. Los demás resultados se interpretan de manera similar.

Tabla 4. Efectos parciales de los componentes de la infraestructura sobre la repetición

| | Área construida | Área deportiva | Biblioteca | Laboratorio ciencias o química | Laboratorio tecnología o informática | Emisora | Salas especializadas | Restaurante | Sala áreas artísticas | Auditorio o coliseo | Laboratorio idiomas |
|--|-----------------|----------------|------------|--------------------------------|--------------------------------------|------------|----------------------|-------------|-----------------------|---------------------|---------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| Tratamiento | -0,0078*** | -0,0080*** | -0,0034 | -0,0112*** | -0,0055* | -0,0041** | 0,0003 | -0,0079*** | -0,0072*** | -0,0103*** | -0,0074*** |
| | (0,003) | (0,002) | (0,003) | (0,002) | (0,003) | (0,002) | (0,002) | (0,002) | (0,002) | (0,002) | (0,002) |
| CI (coeficiente del componente de infraestructura) | 0,0005*** | 0,0002* | 0,0098** | -0,0116*** | -0,0545*** | 0,0115** | 0,0079** | 0,0018 | 0,0312*** | 0,0248*** | -0,0144** |
| | (0,000) | (0,000) | (0,004) | (0,004) | (0,006) | (0,005) | (0,004) | (0,004) | (0,004) | (0,004) | (0,006) |
| Tratamiento x CI | 0,0003 | 0,0002 | -0,0042 | 0,0096*** | -0,0011 | -0,0128*** | -0,0144*** | 0,0083** | 0,0024 | 0,0081*** | 0,0132** |
| | (0,000) | (0,000) | (0,003) | (0,003) | (0,004) | (0,004) | (0,003) | (0,003) | (0,003) | (0,003) | (0,005) |
| Efecto Parcial con T=1 | 0,0008 | 0,0004 | 0,0056 | -0,002 | -0,0556 | -0,0013 | -0,0065 | 0,0101 | 0,0336 | 0,0329 | -0,0012 |
| Constante | 0,4896*** | 0,4899*** | 0,4894*** | 0,4894*** | 0,4894*** | 0,4895*** | 0,4896*** | 0,4893*** | 0,4894*** | 0,4895*** | 0,4894*** |
| | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) | (0,013) |
| Efectos fijos año | Si | Si | Si | Si | Si | Si | Si | Si | Si | Si | Si |
| Efectos fijos UPZ | Si | Si | Si | Si | Si | Si | Si | Si | Si | Si | Si |
| Observaciones | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 | 2.752.917 |
| Estadístico F | 629,15 | 629,15 | 629,1 | 629,15 | 629,1 | 629,18 | 629,12 | 629,13 | 629,11 | 629,13 | 629,12 |
| Errores estándar robustos en paréntesis | | | | | | | | | | | |
| *** p<0.01, ** p<0.05, * p<0.1 | | | | | | | | | | | |

Los resultados muestran una gran variabilidad entre los efectos parciales de cada uno de los componentes de la infraestructura. Mientras que para área construida, área deportiva, biblioteca, restaurante, sala para áreas artísticas y auditorio o coliseo, los resultados sugieren que los efectos parciales aumentan la tasa de repetición, para laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora, salas especializadas y laboratorio

de idiomas, los resultados sugieren que los efectos parciales disminuyen la tasa de repetición, aunque no hay evidencia empírica sobre por qué estos canales se comportan de esta manera. Una posible explicación podría ser que parte de la nueva infraestructura es completamente novedosa para los estudiantes, por ejemplo, laboratorio de idiomas, mientras que otro tipo de infraestructura, aunque nueva, podría ser menos novedosa. Por ejemplo, los estudiantes antes de moverse al megacolegio también tenían áreas deportivas. Otra potencial explicación es que para ciertos procesos pedagógicos o proyectos educativos de la institución (PEI³), algunos componentes de la infraestructura pueden ser más útiles que otros. Por ejemplo, si el PEI de la institución tiene como énfasis las ciencias naturales, un laboratorio de ciencias puede potenciar el desarrollo del PEI.

Los efectos diferenciales entre los distintos componentes de la infraestructura sugieren que los canales por los cuales la infraestructura opera son diversos. Los que generan impacto adicional en la reducción de la tasa de repetición son los laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora y laboratorio de idiomas.

Ejercicio de robustez

En esta sección se llevan a cabo diferentes ejercicios para mostrar que el resultado es robusto respecto a distintas especificaciones y posibles problemas de endogeneidad. La tabla 5, panel A, muestra las estimaciones para diferentes restricciones de datos. La columna (1) incluye todos los registros sin restricción de UPZ. La columna (2) incluye registros de estudiantes de control que solamente estudian en la misma UPZ donde fueron construidos los megacolegios. Las columnas (3), (4) y (5), además de tener solo personas de la misma UPZ o vecinas, restringen la muestra por el número de observaciones que cada individuo tiene en la estimación, esto es, más de 6, 7 y 8, respectivamente. En todas las especificaciones se mantuvieron el signo y la significancia.

Tabla 5. Efecto de la infraestructura sobre la repetición: ejercicios de robustez

| Panel A: Diferentes restricciones de datos | | | | | |
|--|-----------|------------|-----------|-----------|-----------|
| | (1) | (2) | (3) | (4) | (5) |
| Tratamiento | -0,011*** | -0,0116*** | -0,010*** | -0,008*** | -0,008*** |
| | (0,001) | (0,001) | (0,002) | (0,002) | (0,002) |
| Controles | | | | | |
| Socioeconómicas y académicas | Sí | Sí | Sí | Sí | Sí |
| Otros programas | Sí | Sí | Sí | Sí | Sí |
| Tasa rotación profesores y directivos docentes | Sí | Sí | Sí | Sí | Sí |

(continuación)

3 PEI: Proyecto Educativo Institucional, es un documento que orienta el quehacer pedagógico en cada institución. Estos proyectos tienen diversos énfasis, por ejemplo, énfasis en valores, humanidades, ciencias, etc.

| Panel A: Diferentes restricciones de datos | | | | | |
|--|-----------|-----------|-----------|-----------|-----------|
| | (1) | (2) | (3) | (4) | (5) |
| Constante | 0,469*** | 0,4735*** | 0,452*** | 0,448*** | 0,446*** |
| | (0,020) | (0,012) | (0,014) | (0,016) | (0,017) |
| Efectos fijos de niño | Sí | Sí | Sí | Sí | Sí |
| Efectos fijos año | Sí | Sí | Sí | Sí | Sí |
| Efectos fijos UPZ | Sí | Sí | Sí | Sí | Sí |
| Observaciones | 2.886.192 | 3.230.646 | 2.146.750 | 1.702.595 | 1.673.691 |
| Panel B: Falsificación | | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Tratamiento | -0,000 | -0,001 | -0,001 | -0,001 | -0,002 |
| | (0,001) | (0,001) | (0,001) | (0,001) | (0,001) |
| Controles | | | | | |
| Socioeconómicas y académicas | No | No | Sí | Sí | Sí |
| Otros programas | No | No | No | Sí | Sí |
| Tasa rotación profesores y directivos docentes | No | No | No | No | Sí |
| Constante | 0,004*** | 0,020 | 0,198*** | 0,198*** | 0,208*** |
| | (0,000) | (0,018) | (0,020) | (0,020) | (0,020) |
| Efectos fijos de niño | Sí | Sí | Sí | Sí | Sí |
| Efectos fijos año | Sí | Sí | Sí | Sí | Sí |
| Efectos fijos UPZ | No | Sí | Sí | Sí | Sí |
| Observaciones | 1.574.037 | 1.512.752 | 1.512.752 | 1.512.752 | 1.512.752 |
| Errores estándar robustos en paréntesis | | | | | |
| *** p<0,01, ** p<0,05, * p<0,1 | | | | | |

Un segundo ejercicio de robustez es la estimación de la ecuación (1) empleando como placebo los datos justamente antes que el programa entrara en funcionamiento. Para esto, a los estudiantes que en el año t pasan a un megacolegio se les marcó como tratados en el año $t-1$. Para los años anteriores a $t-1$, estos mismos estudiantes se marcaron como no tratados. La intuición de este ejercicio es que si el cambio de las tasas de repetición está asociado al tratamiento, antes del tratamiento no debería verse un efecto diferencial, más aún si la condición de tendencia paralela se cumple. La tabla 5, panel B, presenta estos resultados para diferentes especificaciones. Los resultados indican consistentemente que el tratamiento placebo no tiene ningún efecto sobre la repetición. Esto sugiere que es el tratamiento el que genera el cambio en la tendencia de la tasa de repetición entre el grupo tratado y el grupo control.

CONCLUSIONES

Este trabajo encuentra que colegios con mejor infraestructura reducen las tasas de repetición de sus estudiantes. Estos resultados son robustos en diferentes especificaciones. Esto es consistente con la evidencia que sugiere que mejores ambientes escolares generan mejores procesos educativos (Durán-Narucki, 2008).

Este trabajo explora once componentes de la infraestructura como posibles canales por los cuales puede operar la disminución de la repetición. Se encuentran efectos diferenciales entre componentes de la infraestructura, lo que sugiere que el tipo de infraestructura puede afectar de manera diferencial las variables de resultado de los estudiantes, en este caso, la repetición. Los tipos de infraestructura que generan un impacto adicional sobre la repetición son los laboratorios de ciencias o química, laboratorio de tecnología o informática, emisora y laboratorio de idiomas.

REFERENCIAS

- Duflo, E. (2001). Schooling and labor market consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91 (4), 795-813.
- Durán-Narucki, V. (2008). School building condition, school attendance, and academic achievement in New York City public schools: A mediation model. *Journal of Environmental Psychology*, 28, 278-286.
- Earthman, G. I. (2002). School facility conditions and student academic achievement. UCLA's Institute for Democracy, Education, & Access. <http://repositories.cdlib.org/idea/wwws/wwws-rr008-1002>
- García, S., Fernández, C. & Sánchez, F. (2010). *Deserción y repetición en los primeros grados de la básica primaria: Factores de riesgo y alternativas de política pública*. Bogotá: Gente Nueva Editorial.
- Glewwe, P. & Jacoby, H. (1994). Student achievement and schooling choice in low-income countries: Evidence from Ghana. *Journal of Human Resources*, 29 (3), 843-864.
- Glewwe, P., Hanushek, E., Humpage, S. & Ravina, R. (2011). School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. Working Paper 17554.
- Hanushek, E. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19 (2), 141-164.
- Krishnaratne, S., White, H., & Carpenter, E. (2013). Quality education for all children? What works in education in developing countries. *New Delhi: International Initiative for Impact Evaluation (3ie), Working Paper*, (20).
- Lee, J. W. & Barro, R. J. (2001). Schooling quality in a cross-section of countries. *Economica*, 68 (272), 465-488.
- Pritchett, L. (2004). Towards a new consensus for addressing the global challenge of the lack of education. Center for Global Development Working Paper, 43.
- Rodríguez, C. (2014). Niños y jóvenes en Colombia: Su evolución en el periodo 2010-2013. En: *Colombia en movimiento: los cambios en la vida de los hogares a través de la Encuesta Longitudinal Colombiana de la Universidad de los Andes ELCA* (pp. 99-120). Bogotá: Universidad de los Andes - Facultad de Economía.

Schiefelbein, E. & Wolff, L. (2013). Repetition and inadequate achievement in Latin America's primary school: A review of magnitudes causes relationships and strategies. *Estudos em avaliação educacional*, 07, 45-87.

SED (2004). *Plan Sectorial de Educación, Bogotá: una Gran Escuela*. http://www.educacionbogota.edu.co/Centro_Documentacion/anexos/publicaciones_2004_2008/plan_sectorial_2004_08.pdf

Unesco. (1998). *Oportunidades perdidas: Cuando la escuela no cumple su misión*.

Unesco. (2012). *Oportunidades perdidas: El impacto de la repetición y de la salida prematura de la escuela*.

Unicef. (2007). *Todos pueden aprender. Propuestas para superar el fracaso escolar*. Buenos Aires.

ANEXO 1.

Tasa de repetición tratados vs. control antes del tratamiento

