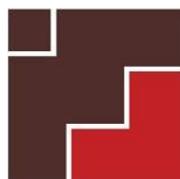


**Evaluación de Impacto del Bono de
Incentivo al Desempeño Escolar o
“Bono Escuela”**

JUAN LEÓN



FORGE

FORTALECIMIENTO
DE LA GESTIÓN DE LA
EDUCACIÓN EN EL PERÚ

Canada¹⁰¹



El presente documento se realizó por encargo del Proyecto Fortalecimiento de la Gestión de la Educación en el Perú (FORGE) que es implementado por el Grupo de Análisis para el Desarrollo – GRADE con el apoyo técnico y financiero del Gobierno de Canadá a través de Global Affairs Canada. (Proyecto N° A-034597)

Evaluación de Impacto del Bono de Incentivo al Desempeño Escolar o “*Bono Escuela*”

Informe final: Proyecto FORGE

Lima, noviembre 2016

Autor: Juan León, investigador principal del Grupo de Análisis para el Desarrollo (GRADE)

Cuidado de estilo: María Fernanda Torres

<p>Las opiniones y recomendaciones vertidas en este documento son responsabilidad de sus autores y no representan necesariamente los puntos de vista de GRADE ni de las instituciones auspiciadoras.</p>
--

Resumen

El presente documento tiene como objetivo principal presentar los principales resultados de la evaluación de impacto cuasi-experimental del *Bono Escuela* que viene implementando el Ministerio de Educación, mediante el uso de bases de datos administrativas que recolecta anualmente el Ministerio de Educación. Los aspectos que se evalúan en el presente estudio son dos: i) el efecto que genera la competencia por el *bono escuela*, estimado mediante análisis de emparejamiento comparando instituciones educativas urbanas estatales y no estatales con similares características, y ii) el efecto de continuidad en los buenos resultados educativos por parte de las instituciones educativas que recibieron el *bono escuela*, estimado mediante el uso de regresiones discontinuas alrededor del punto de corte fijado para la entrega del *bono escuela* a las instituciones educativas.

Entre los principales resultados del presente estudio están:

- *Efecto de la competencia*: los resultados del análisis de emparejamiento muestran que las instituciones educativas estatales urbanas tienen en promedio un mayor rendimiento que sus pares de instituciones educativas no estatales en el rendimiento promedio y niveles de desempeño en matemática, mientras en el caso de comprensión lectora no se encuentran diferencias entre ambos grupos de instituciones educativas. Así mismo, se realizó los análisis usando el diferencial del rendimiento y los niveles de desempeño entre el 2014 y 2015, encontrándose solamente efectos consistentes y robustos en la reducción de estudiantes que no dominan los contenidos para su grado en comprensión de lectura, a favor de las instituciones educativas tratadas.
- *Continuidad de los efectos*: existe un impacto positivo y significativo del *bono escuela* en los resultados educativos (estimación paramétrica y no paramétrica). Los análisis muestran que aquellas instituciones educativas que recibieron el incentivo monetario en el 2014, obtienen un mejor rendimiento promedio en Matemática y Comprensión lectora en el 2015, es decir, sus estudiantes mantienen el buen performance a pesar de que son cohortes de estudiantes diferentes. En relación a cambios en los niveles de desempeño, sólo en el caso de Matemática, se aprecia que los estudiantes de 2do de primaria de las instituciones educativas intervenidas muestran mejoras en su desempeño dado que un menor porcentaje de sus estudiantes no logran dominar las habilidades iniciales matemáticas. Por otro lado, los efectos en los puntajes promedios de matemática, a favor de las instituciones educativas que recibieron el incentivo, son robustos dado que son estadísticamente significativos tanto mediante el uso de estrategias paramétricas y no paramétricas.
- *Asistencia docente*: se encontró que existe un efecto positivo y significativo en la asistencia de los docentes y directores. En otras palabras, aquellas instituciones educativas que han recibido el *bono escuela*, son aquellas donde se aprecia una menor inasistencia de los docentes y directores, a diferencia de aquellas instituciones educativas donde no recibieron el bono. Este aspecto resulta importante dado que la mayor asistencia de los docentes permite que se pierdan menos días de clase y los docentes puedan desarrollar más contenidos del currículo, aspecto que favorece el aprendizaje de los estudiantes (Das et al, 2005; Suryadarma et ál., 2006; Cueto et al, 2008).
- *Rendimiento en Matemática*: existe un efecto robusto en matemática y no en comprensión lectora. Estos resultados pueden estar relacionados al nivel de dominio de los contenidos a enseñar por parte de los docentes. En el caso de Matemática, de acuerdo a la UMC (2015), el 31% de los estudiantes de 2do grado se ubica en el nivel en inicio, mientras en comprensión de lectura se ubica en este mismo nivel el 13% de los estudiantes. Este hecho podría estar haciendo que los docentes estén trabajando más contenidos matemáticos dado que lograr mejoras en esta área a los estudiantes es mucho más factible dado que se requiere simplemente que dominen de forma parcial los contenidos para el grado que están siendo evaluados. Mientras en comprensión de lectora, si requeriría que los docentes dominen mejor las estrategias de enseñanza para ubicar a sus estudiantes en un nivel satisfactorio dado el bajo porcentaje de estudiantes en el nivel en inicio.

- *Tamaño del efecto*: en cuanto al tamaño del efecto encontrado, si usamos los criterios o *benchmarks* planteados por Cohen (1988), se considerarían los efectos encontrados como pequeños debido a que están por debajo de 0.50 desviaciones estándar¹. Sin embargo, el problema que presenta considerar este criterio para evaluar el effect size radica en que Cohen no toma en consideración el contexto en el cual se desarrollan las intervenciones. Motivo por el cual, Hill et al (2007) plantean que para considerar un tamaño de efecto como pequeño, mediano o grande, se pueden usar tres diferentes criterios: i) cambios esperados de acuerdo al grupo de población estudiado, ii) metas políticamente relevantes (reducción de brechas), iii) tamaños de efecto en estudios similares. Así, para el presente estudio aplica considerar el último criterio y ver qué tan diferentes son los efectos en relación a estudios similares. Así, al comparar los efectos del *bono escuela* sobre el rendimiento con los resultados en estudios similares (McEwan, 2015; Imberman, 2015), se puede apreciar que los tamaños de efecto son similares tanto para Matemática como para Comprensión de lectura. Por otro lado, en el caso de la asistencia docente, los tamaños de efecto encontrados (0.23 desviaciones estándar) son similares a aquellos encontrados en el meta-análisis desarrollado por Guerrero et al (2013) para países en desarrollo, donde los efectos oscilaban entre 0.15 a 0.30 desviaciones estándar.

¹ De acuerdo a Cohen (1988), un tamaño de efecto pequeño es de 0.20 desviaciones estándar, un efecto mediano es de 0.50 desviaciones estándar y un efecto grande es de 0.80 desviaciones estándar.

Índice

1. Introducción	7
2. Revisión de Literatura	8
2.1 Experiencias internacionales y locales	8
2.2 Balance de la literatura	13
3. La herramienta <i>Bono Escuela</i>	14
3.1 Revisión de la herramienta a la luz de las Experiencias Internacionales	18
4. Objetivos de la Evaluación de Impacto	19
5. Estrategia de Evaluación de Impacto	20
5.1 Unidad de Análisis	22
5.2 Bases de datos y tamaño de muestra	22
5.3 Variables de estudio	23
5.4 Análisis de poder y tamaño de la muestra	25
5.5 Especificación de los Modelos Estadísticos	26
6. Resultados de los análisis estadísticos desarrollados	30
6.1 Efecto de la Competencia del Bono Escuela	30
6.2 Efectos de continuidad del Bono Escuela	38
7. Conclusiones y Recomendaciones de Política	48
8. Referencias	52
Anexos	
Anexo A. Programas de Incentivos en países en vías de desarrollo	56
Anexo B. Tamaño de Muestra para los análisis a nivel de estudiantes	62
Anexo C. Análisis de verificación del modelo de Emparejamiento	63
Anexo D. Resultados adicionales para los modelos de Emparejamiento	66
Anexo E. Gráficos de verificación del modelo de Regresión Discontinua	68
Anexo F. Análisis adicionales de la estimación Paramétrica del Modelo de Regresión Discontinua	78
Anexo G. Análisis adicionales de la estimación no paramétrica del Modelo de Regresión Discontinua	86
Lista de Cuadros	
Cuadro 1. Criterios para identificar a las instituciones educativas participantes	15
Cuadro 2. Niveles de agrupación de las instituciones educativas participantes del bono escuela	16
Cuadro 3. Indicadores para la generación del índice de desempeño escolar	16
Cuadro 4. Fórmulas para la generación del índice de desempeño escolar en base a los indicadores de cada grupo	16
Cuadro 5. Montos de las bonificaciones por grupos de instituciones educativas	18
Cuadro 6. Tamaño de muestra por grupo para un diseño RCT y RD	26
Cuadro 7. Tipos de función de enlace entre la variable de selección y resultado de los estudiantes	29
Cuadro 8. Rendimiento y niveles de desempeño (2015) promedio sin ajustar en Matemática y Comprensión de lectura por grupo de estudio	31
Cuadro 9. Diferencias en Diferencias en el rendimiento y niveles de desempeño promedio (2015-2014) sin ajustar en Matemática y Comprensión de lectura por grupo de estudio	32
Cuadro 10. Diferencias en el rendimiento y niveles de desempeño (2015) ajustados en Matemática y Comprensión de lectura por grupo de estudio (coeficientes no estandarizados)	33
Cuadro 11. Diferencias en Diferencias en el rendimiento y niveles de desempeño ajustados en Matemática y Comprensión de lectura por grupo de estudio (coeficientes no estandarizados)	34
Cuadro 12. Resultados de la Regresión Discontinua Paramétrica en Matemática	42
Cuadro 13. Resultados de la Regresión Discontinua Paramétrica en Comprensión de Lectura	42
Cuadro 14. Resultados de la Regresión Discontinua Paramétrica para las variables Intermedias	44
Cuadro 15. Anchos de banda óptimos para las variables de resultado de acuerdo al método de Imbens y Kalyanaraman (2012)	45
Cuadro 16. Resultados de la Regresión Discontinua No Paramétrica en Matemática	46

Cuadro 17. Resultados de la Regresión Discontinua No Paramétrica en Comprensión de lectura	46
Cuadro 18. Resultados de la Regresión Discontinua No Paramétrica en Resultados intermedios	47

Lista de Figuras

Figura 1. Teoría de cambio de los incentivos monetarios a los docentes	13
Figura 2. Probabilidad de participación en el bono	28
Figura 3. Distribución de los puntajes en Comprensión de lectura y niveles de desempeño (2015) para las escuelas de tratamiento y control en el soporte común	35
Figura 4. Distribución de los puntajes en Matemática y niveles de desempeño (2015) para las escuelas de tratamiento y control en el soporte común	36
Figura 5. Distribución de las variaciones en Comprensión de lectura y niveles de desempeño (2015 - 2014) para las escuelas de tratamiento y control en el soporte común	37
Figura 6. Distribución de las variaciones en Matemática y niveles de desempeño (2015 - 2014) para las escuelas de tratamiento y control en el soporte común	37
Figura 7. Discontinuidad de las variables de resultado final (2015) alrededor del punto de corte de la variable de asignación.	39
Figura 8. Discontinuidad de las variables intermedias (2015) alrededor del punto de corte de la variable de asignación.	40
Figura 9. Distribución de los puntajes en Matemáticas para las escuelas por encima y debajo del punto de corte (ancho de banda 0.25)	47
Figura 10. Distribución de los puntajes en Comprensión de lectura para las escuelas por encima y debajo del punto de corte (ancho de banda 0.25)	48

1. Introducción

En las últimas décadas en el Perú se han visto grandes cambios en materia educativa, siendo el más importante el incremento en la cobertura en educación básica a nivel nacional. Entre el 2003 y 2015, la tasa neta de matrícula en el nivel inicial y secundario mostró incrementos significativos pasando de 53% a 83% y de 70% a 84% respectivamente, mientras a nivel primario se mantuvo alrededor del 92%². Por otro lado, se pueden apreciar que las brechas de acceso a la educación por sexo, lengua materna y área geográfica vienen reduciéndose en el tiempo. Sin embargo, las mejoras en el acceso y brechas de acceso a la educación, no han ido de la mano con una mejora en la calidad de los aprendizajes de los estudiantes peruanos. Diferentes evaluaciones nacionales de rendimiento que se han venido llevando desde hace tres décadas muestran que los estudiantes peruanos en los diferentes niveles educativos tienen un bajo rendimiento en las áreas que han sido evaluadas del currículo (MECEP, 2002; UMC, 2004; UMC, 2006; Miranda, 2008; UMC, 2013). Así mismo, evaluaciones internacionales (PERCE, SERCE, TERCE, PRIDI y PISA) donde ha participado el Perú muestran la misma tendencia con respecto al bajo rendimiento de los niños peruanos, e incluso en la última evaluación de PISA (2012), los estudiantes peruanos se ubicaron en el último lugar de todos los países que participaron de dicho estudio (OECD, 2014).

La investigación en educación, a raíz de los bajos resultados educativos de los estudiantes peruanos, buscó identificar qué factores están asociados con el rendimiento de los estudiantes peruanos (MECEP, 2002; UMC, 2004; UMC, 2006; Cueto, 2007; Miranda, 2008; UMC, 2013; Balarin, 2016). Los estudios de investigación y balances de investigación desarrollados en el Perú han podido identificar diferentes factores a nivel individual, familiar o escolar que están asociados con el rendimiento de los estudiantes. Sin embargo, a nivel escolar, un factor que cumple un rol importante para explicar el resultado de los estudiantes es la labor que realiza el docente. Así, el docente es una herramienta clave para la mejora de los aprendizajes a través de diferentes formas o medios (Guerrero y León, 2015; Cueto et al, 2013; Guadalupe et al, 2013; UMC, 2013; Miranda, 2008; Cueto, 2007; UMC, 2004; MECEP, 2002), como: i) prácticas pedagógicas que utiliza en el aula para la enseñanza, ii) clima al interior del aula de clase, iii) las buenas relaciones que mantiene con otros actores educativos (estudiantes, padres de familia, director, entre otros), y iv) desempeño del docente (asistencia del docente, conocimiento de la materia, entre otros). Así, los docentes se convierten en pieza fundamental para mejorar los resultados educativos e influyen de diferentes formas en sus aprendizajes.

Dado lo anterior, en los últimos años, el Ministerio de Educación viene implementando una política educativa orientada a revalorizar la carrera docente, para lo cual se han diseñado una serie de herramientas que buscan trabajar diferentes aspectos asociados con la revalorización de la carrera de los docentes, siendo uno de estos aspectos la implementación de reconocimientos e incentivos al personal docente en función del desempeño de sus estudiantes. Así, en el 2014, el Ministerio de Educación a través del Decreto de Urgencia 002-2014 crea el Bono de Incentivo al Desempeño Escolar o *Bono escuela* que surge con la finalidad de premiar el esfuerzo que realizan los docentes y directores de instituciones educativas estatales por lograr buenos resultados educativos con sus estudiantes, de tal forma que al ver reconocido su esfuerzo se vean motivados para seguir haciéndolo continuamente.

El presente documento tiene como objetivo principal presentar los principales resultados de la evaluación de impacto del *Bono Escuela* que viene implementando el Ministerio de Educación, para lo cual, se evalúan dos aspectos fundamentales de su implementación. En primer lugar se evalúa el efecto de la competencia que genera la implementación del incentivo monetario anual, y en segundo lugar, se evalúan los efectos a un año en los resultados educativos de la entrega del incentivo, apreciándose de esta forma si el bono genera los incentivos suficientes para que los docentes sigan esforzándose continuamente. El documento está dividido en siete secciones, siendo la primera la presente introducción. La segunda da un alcance de la literatura internacional y local sobre el tema, la tercera describe la herramienta y brinda un balance de la misma a la luz de la revisión de la literatura, la cuarta presenta los objetivos del presente estudio, la quinta da un alcance del diseño de investigación y cómo este responde a los objetivos planteados, la sexta presenta los resultados de los análisis de datos secundarios realizados y la séptima expone las principales conclusiones del estudio.

² Los datos estadísticos son obtenidos de la web ESCALE de la Unidad de Estadística e Informática del Ministerio de Educación (escale.minedu.gob.pe)

2. Revisión de Literatura

Los incentivos a los docentes pueden ser analizados desde el punto de vista económico usando el enfoque de la agencia o el problema del principal y agente. En el cual se asume que hay un contrato entre ambos donde el principal (empleador) señala las actividades que tiene que realizar el agente (empleado) acordándose cierta suma de dinero para que realice su trabajo. Sin embargo, un problema que se presenta es que el principal no puede observar las acciones que realiza el agente y solo puede observar el producto final que resulta de las acciones tomada por el agente, lo cual origina que se dé una asimetría de información dado que el principal no ve las acciones y el agente sabe qué acciones realizó. Así, para resolver el problema de asimetría de información y el agente realice acciones en favor de los intereses del principal, la solución es la implementación de esquemas de incentivos a los agentes (Fama, 1980; Baker, 1992; Baker, 2000; Levacic, 2009).

Existen dos formas de otorgar los incentivos, una basada en insumos donde se premia al agente en función a su esfuerzo o trabajo que realiza y la segunda basada en resultados, donde se premia a los agentes (docentes) en función a los intereses del principal (Estado). Sin embargo, existe un *trade-off* al usar uno u otro enfoque para la entrega de los incentivos. Al usar incentivos basados en insumos, se premia al docente por sus capacidades o conocimientos que no necesariamente están alineadas con los intereses del Estado; mientras al usar incentivos basados en resultados, si bien los intereses están alineados a los del Estado, debido a la asimetría de información, los docentes pueden actuar de acuerdo a sus propios intereses para lograr el resultado fijado por el Estado (Prendergast, 2000).

Así, en el caso educativo, se tiene que el Estado (el principal) no puede observar fácilmente la calidad de los docentes y directores (agentes), lo cual origina problemas de asimetría de información para que cumplan con los intereses del gobierno en materia educativa. Una manera de solucionar el problema es mediante un adecuado esquema o sistema de incentivos (monetarios o no monetarios) que estimulen a los directores y docentes (agentes) a dar un mayor esfuerzo y realizar su trabajo de forma eficaz que ayude a mejorar el rendimiento de sus estudiantes. La literatura sobre la implementación de esquema de incentivos a docentes y directores es diversa siendo la mayoría de ella en países desarrollados que en vías de desarrollo; asimismo, la mayoría de los esquemas implementados usan como enfoque para dar incentivos basados en resultados educativos antes que insumos dado que existe bastante incertidumbre dado que no pueden observar las acciones que toman los docentes. A continuación se brinda un alcance de los principales esquemas de incentivos implementados a nivel internacional y local.

2.1 Experiencias internacionales y locales

En el caso de países desarrollados, como los Estados Unidos, Figlio y Kenny (2006) usando las bases de datos de la encuesta longitudinal NELS y una encuesta escolar elaborada por los autores, encuentran una relación positiva y significativa entre los incentivos individuales a los docentes y el rendimiento de los estudiantes e incluso esta relación es mucho más fuerte en escuelas donde los padres tienen un mayor nivel de participación. Como medida de logro educativo utilizan el promedio de las notas en las pruebas de matemática, lectura, ciencias e historia que se encuentran en la encuesta NELS. Como medidas de incentivo monetario emplearon información sobre el aumento salarial de los docentes de acuerdo al mérito (experiencia y estudios), el rango del salario y si los docentes reciben bonos. Esta información se recogió a través de la encuesta hecha a los directores de las escuelas. Sus resultados, dependiendo de la especificación que utiliza, para la muestra completa con la que cuentan los autores, varían de 1.1 a 2.1 puntos que se consiguen por tener algún sistema de incentivo como los citados en las escuelas. Este resultado es comparable con el de tres años más de educación de las madres, según los autores.

Springer et al (2010) realizan una evaluación de impacto experimental del sistema de incentivos docentes de 5to a 8vo grado en Nashville durante el periodo 2006 a 2009. El sistema de incentivos era individual basado en el progreso educativo de sus estudiantes, el cual se mide como el incremento en el rendimiento promedio de los estudiantes año a año en las pruebas estandarizadas estatales de matemática. Los montos de los incentivos eran de 5 000, 10 000 o 15 000 dólares americanos, los cuales se entregaban si sus alumnos se ubicaban en el percentil 80, 85 o 95 respectivamente. Los autores encuentran que no existen efectos de este sistema de incentivos en el

rendimiento de los estudiantes, así como en variables relacionadas con la labor docente como las prácticas pedagógicas en el aula de clase. Sin embargo, si encuentran efectos en los estudiantes de 5to grado, indicando que la posible explicación para este efecto es que los estudiantes de este grado aun cuentan con un solo docente para todas las materias que se dictan en clase.

Goodman & Turner (2013) realizan una evaluación de impacto del sistema de incentivos grupales a docentes en escuelas de Nueva York. El incentivo o bono otorgado a los docentes era de aproximadamente \$ 3 000 dólares americanos que representaba entre el 3% y 7% del salario anual de los docentes. Este bono era entregado a los docentes si en promedio los estudiantes de la escuela superaban la valla fijada por el programa en cuanto al rendimiento de los estudiantes. Los autores utilizan una muestra de 200 escuelas con población escolar de bajos recursos para realizar su evaluación, encontrando efectos positivos pero pequeños (0.08 desviaciones estándar) de la implementación del sistema de incentivos en el rendimiento de matemáticas promedio de las escuelas, pero al segundo año de implementación del programa. Por otro lado, evaluaron si había efectos en variables relacionadas al desempeño docente como el ausentismo, pero no encontraron resultados significativos.

Imberman & Lovenheim (2013) estudian el efecto del programa de incentivos Accelerating Student Progress, Increasing Results and Expectations (ASPIRE) en el rendimiento de los estudiantes del distrito de Houston en los Estados Unidos durante el año escolar 2007-2008. Este sistema de incentivos era grupal y se premiaba a los docentes que logren incrementar los puntajes de sus estudiantes por encima de la media del distrito en las áreas de matemática, inglés, ciencias y personal social, y los incentivos monetarios podían llegar a representar hasta el 14% del salario anual de los docentes. Los resultados del estudio sugieren que mientras más pequeño es el grupo de docentes, el efecto del esquema de incentivos sobre los estudiantes es mayor.

Por otro lado, Fryer (2013) realizó una evaluación de impacto experimental de un esquema de incentivos en escuelas de Nueva York, donde se premia a las escuelas si se ubican en el percentil 75% o más del índice de progreso escolar que está compuesto por tres indicadores que son: i) el rendimiento de los estudiantes (25%), ii) la calidad del ambiente escolar (15%), y iii) el progreso de los estudiantes (60%). El pago se realiza a toda la escuela, la cual reparte el premio según sus propios criterios. Así, los docentes pueden llegar a ganar hasta 3,000 dólares americanos al interior de cada escuela premiada. Los resultados de este experimento muestran que en promedio, este esquema de incentivos no tuvo efecto en el rendimiento de los estudiantes e incluso tuvo un efecto negativo en el logro de los estudiantes, más que nada en escuelas grandes o con un gran número de estudiantes matriculados.

Glazerman & Seifullah (2012) hacen una evaluación del Teacher Advancement Program desarrollado en Chicago en los Estados Unidos. Dicho programa tiene como objetivo brindar incentivos a los docentes en función al rendimiento de sus estudiantes y de su desempeño en las aulas de clase. Así mismo, este programa promovía el desarrollo profesional de los docentes mediante el otorgamiento de mayores responsabilidades como el ser docentes tutores o acompañantes de otros docentes. Los incentivos monetarios oscilaban entre 1 100 y 1 400 dólares americanos, mientras los docentes acompañantes o tutores podían recibir hasta 20 000 dólares americanos al año adicionales a su sueldo. Los autores realizaron una evaluación de impacto experimental usando la información disponible de este esquema de incentivos entre los años 2007 y 2010, a la vez que hicieron análisis de robustez usando un diseño cuasi-experimental. Los resultados no son contundentes y más bien los autores plantean que no existen efectos significativos de la implementación del esquema de incentivos implementado en Chicago debido a que los resultados dependen mucho del área, año o método usado para ver el impacto del esquema.

En el caso de países en vías de desarrollo, Lavy (2004) analiza el impacto de un esquema de incentivos individuales a docentes en escuelas en Israel. El autor encuentra un efecto positivo y significativo en el rendimiento de los estudiantes, lo cual lo explica por la ampliación de la jornada escolar en las escuelas que son parte del esquema de incentivos y un mayor esfuerzo por parte de los docentes en darles retroalimentación a sus estudiantes o el uso de diferentes prácticas pedagógicas para el trabajo en grupo en clase. El rendimiento de los estudiantes se mide en las asignaturas de matemática, inglés y hebreo (opcional para las escuelas). Los docentes fueron premiados con bonos en efectivo por el incremento del puntaje (créditos por examen que aprueban, cada examen vale de 1 a 5 créditos) que obtienen en promedio los estudiantes a los que cada docente enseña. Los premios que podían recibir los docentes estaban en el rango de 6% a 25% del salario anual, dependiendo del

desempeño de sus estudiantes según el ranking, es más un docente podía ganar más de un premio si tenía a su cargo más de una clase, pues el concurso era por cursos. El programa de incentivos tuvo un impacto significativo de 0.24 créditos más en matemática y 0.18 en inglés en el promedio de los alumnos cuyos docentes estaban en el programa de incentivos.

En la India, Duflo & Hanna (2011) realizan un experimento aleatorio en 120 escuelas rurales para ver el impacto de un sistema de incentivos individuales a los docentes en el ausentismo y el rendimiento académico de sus estudiantes. El esquema de incentivos consiste en un aumento salarial por cada día extra de asistencia a la escuela y una reducción por cada día de ausencia, con respecto a una jornada mensual de 20 días. Entre sus principales resultados a un año de implementación están que el ausentismo docente es 20 puntos porcentuales menos en el grupo de tratamiento que el de contraste y la diferencia en el rendimiento de los estudiantes entre el grupo de tratamiento y contraste fue de 0.15 desviaciones estándar para Matemática y 0.16 desviaciones estándar para Lenguaje, siendo en ambos casos las diferencias estadísticamente significativas. Sin embargo, al final de la evaluación (después de un año), las diferencias en el rendimiento no fueron significativas y solo se mantuvo el efecto en el ausentismo de los docentes.

Otro estudio realizado en la India es el experimento aleatorio realizado por Muralidharan & Sundararaman (2009) en 300 escuelas rurales, para ver impacto de dos sistemas de incentivos a los docentes. El primer esquema otorga incentivos individuales a los docentes por el incremento en por lo menos 5% del rendimiento de sus estudiantes en matemática y lenguaje, mientras el segundo esquema otorga incentivos a todos los docentes de la escuela en base al incremento promedio en el rendimiento académico de la escuela. Los resultados indican que existen diferencias a favor de las escuelas que reciben incentivos (individuales o grupales) en matemática y comprensión de lectura al año y dos años de implementado el programa, y no hubo una diferencia significativa estadísticamente en el rendimiento entre ambos esquemas de incentivos en el primer año, pero sí en el segundo, donde los incentivos individuales tuvieron un mayor impacto. Se midió el rendimiento en pruebas estandarizadas en razonamiento matemático y razonamiento verbal, que medían las competencias del año escolar previo y el año en curso. Las pruebas medían dos componentes, denominados mecánicos y conceptuales. El efecto después del primer año de la intervención los estudiantes de las escuelas tratadas tenían un efecto promedio en matemática y razonamiento verbal de 0.15 desviaciones estándar más, sin que existan diferencias significativas entre el esquema de incentivos individual y mixto. A finales del segundo año, las diferencias entre los alumnos cuyos docentes recibieron los incentivos fueron de 0.27 desviaciones estándar en matemáticas y 0.17 en verbal, en promedio. A finales del segundo año la diferencia de rendimiento en los estudiantes entre el mecanismo de incentivo individual y el grupo de referencia fue de 0.28 desviaciones estándar en promedio, mientras que el mecanismo de incentivo grupal fue 0.15. Finalmente, en cuanto a qué variables podrían ser los posibles canales a través de los cuales los docentes lograron incrementar el rendimiento de sus estudiantes, los autores encontraron que en las escuelas del grupo de tratamiento los docentes daban tareas extra a los estudiantes, realizaban clases adicionales y daban una atención especializada a los estudiantes de bajo rendimiento.

En Kenia, Glewwe et al (2003) analizan el impacto de un esquema de incentivos grupales a docentes en escuelas rurales mediante una evaluación experimental. Los docentes reciben un premio en forma de bono en base al rendimiento de los estudiantes en la prueba que mide las competencias en el distrito cada año. Se premian dos tipos de escuelas, las cuales sus estudiantes tienen los puntajes más altos y las que mejoraron en mayor medida en función de un año base. Todos los docentes de la escuela reciben el mismo premio si esta logra obtener algún premio. El rango de premio va de un 21% a 43% del salario mensual de un docente, siendo aproximadamente 24 de 50 las escuelas premiadas. Los autores encuentran un efecto positivo y significativo del esquema de incentivos a los docentes; sin embargo, indican que estos efectos se dan en el corto plazo y más bien se disipan una vez que se termina el esquema de incentivos a las escuelas, lo cual puede explicarse por el esfuerzo que realizan los docentes durante el periodo del esquema de incentivos dado el incremento de sus remuneraciones. El resultado que hallan es que en promedio, el programa de incentivos tiene un efecto no significativo al comienzo del programa, un efecto de 0.14 desviaciones estándar en la nota total de la prueba que toma el distrito escolar en el segundo año y para el tercer y último año del programa, el efecto es casi nulo. Además, hallan que el programa incentiva que los docentes en las escuelas tratadas den clases de refuerzo (tutorías) para preparar a sus alumnos para los exámenes.

Loyalka et al (2015) realizan una evaluación de impacto aleatoria en el este de China evalúa tres esquemas de incentivos basados en resultados académicos de los estudiantes. Los tres métodos están basados en un ranking construido a partir del resultado en matemática de los alumnos en la prueba estandarizada de cada distrito educativo. El primero de ellos es un esquema de premios según el percentil que alcanza cada estudiante en el ranking formado a partir de la evaluación, el segundo se basa en el promedio del resultado de los alumnos de cada docente al final del año y el tercero en el incremento promedio de la calificación de los estudiantes. Los autores definen dos esquemas de pago diferenciados dentro de cada sistema de ranking. Los resultados muestran que el sistema de incentivos según percentiles tiene un efecto de entre 0.13 y 0.15 desviaciones estándar sobre el rendimiento académico de los estudiantes, mientras que los otros esquemas de incentivos no tienen un efecto significativo sobre el rendimiento de los estudiantes. Además los autores encuentran que cualquiera de los tres esquemas de incentivos tiene un efecto significativo en las actitudes y prácticas de los docentes, específicamente en la cobertura del currículo del curso, según encuestas aplicadas a los estudiantes.

Imberman (2015) hace una revisión sistemática de la efectividad de los programas de incentivos monetarios a docentes en países desarrollados y en vías de desarrollo. Dicho estudio encuentra que la evidencia es mixta en cuanto a los efectos de los incentivos monetarios a los docentes en países desarrollados, mientras en el caso de los países en vía de desarrollo encuentra evidencia que en la mayoría de los casos es positiva y se logran mejorar los resultados educativos de los estudiantes. Por otro lado, encuentran que los sistemas de incentivos que cuentan con un buen diseño son aquellos que logran resultados positivos con sus estudiantes. Finalmente, encuentra que los sistemas de incentivos grupales (a la escuela) son más eficientes para incrementar los resultados educativos de los estudiantes que aquellos sistemas que premian de manera individual a los docentes.

En América Latina, McEwan & Santibañez (2005) realizan un estudio sobre el impacto del programa Carrera Magisterial en México. Este sistema de incentivos se implementó en 1993 y era voluntario. Los docentes que participan de este programa, tienen un esquema de estímulos salariales que se basan en el desempeño individual de cada docente que está en función de su preparación profesional, formación continua y aprovechamiento escolar de sus estudiantes. El premio se basa en un aumento del salario mensual en 20% durante el resto de su vida laboral. Los autores encuentran efectos heterogéneos de este esquema de incentivos a docentes sobre el rendimiento de los estudiantes, siendo los resultados no significativos para el caso de primaria y positivos y significativos en el caso de educación secundaria, e incluso solo se apreciaron en aquellos docentes que estaban en un nivel mayor del escalafón salarial. Finalmente, el efecto promedio de este incentivo salarial es menor a 0.10 desviaciones estándar.

Otro estudio realizado en México es el de Vivanco (2013), dicho estudio busco evaluar el efecto del Programa del Estímulo para la Calidad Docente (PECD) mexicano que brinda bonos salariales a docentes de escuelas públicas de primaria y secundarias en México, de acuerdo al nivel de las calificaciones de sus alumnos y la incremento de estas. Para la evaluación, usa las bases de datos del PECD del 2010 y los resultados de rendimiento de la Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE) del 2010 y 2011. Dado que el bono es entregado casi de manera universal a todas las escuelas públicas de primaria y secundaria, el autor hace uso de los modelos de regresión discontinua (RD) para evaluar la efectividad de la implementación del esquema de bonos salariales a los docentes. Los resultados indican que no existe efecto del programa de incentivos sobre el rendimiento de los estudiantes; sin embargo, se aprecia un efecto negativo y significativo (0.25 desviaciones estándar) para las escuelas primarias indígenas.

En Chile, Mizala & Romaguera (2005) realizan un estudio donde dan un alcance del efecto del esquema de incentivos implementado por el gobierno central en 1996 denominado Sistema Nacional de Evaluación del Desempeño de los Establecimientos Educacionales Subvencionados (SNED). Este esquema, otorga incentivos a los docentes de escuelas consideradas con un desempeño escolar excelente. Así, para determinar que una escuela ha tenido un desempeño excelente, se utiliza un índice que está compuesto por seis factores que son ponderados para el establecimiento del índice global. Los factores considerados para el índice son: i) resultados educativos de las escuelas, ii) incremento de los puntajes en el tiempo, iii) logro de innovaciones educativas y el compromiso de los actores educativos con la escuela, iv) mejora en las condiciones de trabajo, vi) igualdad de oportunidades en las escuelas, vii) participación de los padres de familia en la escuela. El premio se otorga a las escuelas todas las escuelas según el índice de desempeño hasta completar el 25% de la matrícula del grupo de escuelas creado para el concurso. El 90% del bono se divide por igual, mientras el 10% restante lo reparten

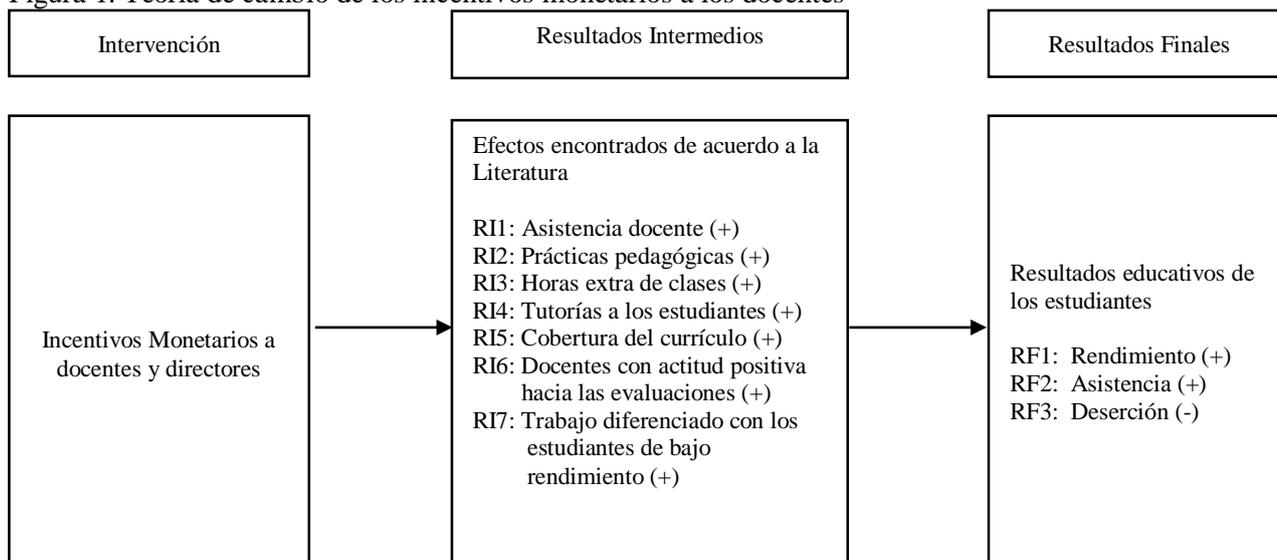
dentro de las escuelas a los docentes que consideran que hayan contribuido más. Este bono que se entrega una vez al año representa en promedio el 40% del salario mensual de un docente. Los resultados muestran que el SNED ha tenido un efecto positivo en mejorar el resultado de los estudiantes, a la vez que ha hecho que los docentes estén más abiertos a considerar evaluaciones de desempeño basadas en resultados educativos de sus estudiantes.

Otro estudio que explora la implementación de incentivos monetarios a los docentes en Chile es el realizado por Rau & Contreras (2009). Los autores utilizan técnicas cuasi-experimentales para evaluar los efectos del esquema de incentivos (SNED) en forma de bono en el rendimiento de los estudiantes. Un primer efecto que estiman es el de participar o no en el programa, motivo por el cual, hacen uso de la metodología de emparejamiento o *propensity score matching*. Así, emparejan escuelas que recibieron el incentivo con aquellas que no formaron parte del esquema de incentivos como son las escuelas privadas. Entre los principales resultados, encuentran que existe un efecto positivo y significativo en la implementación del esquema de incentivos a los docentes, e incluso este efecto es mayor conforme se restringe más la muestra a aquellas escuelas que no son seguros ganadores o seguros perdedores del incentivo. Un segundo efecto que exploran es la sostenibilidad de la mejora del rendimiento de los estudiantes en escuelas que recibieron el incentivo, comparándolas con aquellas escuelas que si bien pudieron recibir el incentivo, no lo recibieron. Para tal fin, hacen uso de los modelos de regresión discontinua dado que utilizan a las escuelas cerca al punto de corte para recibir el incentivo como grupos de tratamiento y contraste debido a que sus diferencias serían aleatorias. Los resultados de este análisis muestran que si bien existe un efecto positivo de haber recibido el incentivo en el rendimiento de los estudiantes (entre 0.2 y 0.3 desviaciones estándar), esta diferencia no es diferente a la mejora que tienen las escuelas que no participaron o recibieron el incentivo.

En el caso peruano, hay dos etapas de los sistemas de incentivos. En el año 2003, se implementó el programa piloto (cuatro regiones del país) de incentivos monetarios a docentes denominado Mejor Educación a través de más tiempo en el aula (META). Este programa tenía como finalidad brindar incentivos monetarios a los docentes en zonas rurales por el tiempo de asistencia en la escuela y no por el rendimiento de los estudiantes, el programa duró solo tres años para luego ser desactivado. Cueto y otros (2008) realizaron un estudio para evaluar el impacto del programa en la asistencia de los docentes y en el rendimiento de los estudiantes. Los autores, usando técnicas cuasi-experimentales, encuentran que existe un efecto positivo y significativo de los incentivos a los docentes en la asistencia (alrededor de 17 días); sin embargo, se encontraron efectos heterogéneos en el rendimiento de los estudiantes dado que estos dependían del área evaluada o del grado.

Finalmente, el siguiente mapa conceptual sintetiza la revisión de literatura realizada en relación a como son los mecanismos a través de los cuales la entrega de incentivos monetarios a los docentes influye en los resultados educativos de los estudiantes.

Figura 1. Teoría de cambio de los incentivos monetarios a los docentes



2.2 Balance de la literatura

Se ha podido apreciar que existen diferentes experiencias internacionales y locales sobre el uso de incentivos monetarios a los docentes, ya sean individuales o colectivos, a nivel primario o secundario, y en países desarrollados o en vías de desarrollo³. Sin embargo, los aspectos que quedan claros de esta revisión de literatura son:

- Existen efectos mixtos en la implementación de los incentivos, es decir, los efectos en la mejora del rendimiento de los estudiantes por su implementación no son contundentes encontrándose en algunos estudios no efectos del sistema de incentivos a docentes (Imberman, 2015; Vivanco, 2013; Mendoza del Solar 2008), efectos dependiendo del tipo de metodología utilizada (Rau & Contreras, 2009), efectos en determinados grupos de estudiantes y/o escuelas (Cueto et al, 2008), efectos positivos y significativos (Glewwe et al, 2010; Muralidharan & Sundararaman, 2009; McEwan & Santibañez, 2005; Duflo & Hanna, 2005; Lavy 2004).
- En el caso de los estudios con efectos positivos, estos se han dado tanto en evaluaciones de impacto experimentales (Muralidharan & Sundararaman, 2009; Duflo & Hanna, 2005), como cuasi-experimentales (Rau & Contreras, 2009; Cueto et al, 2008; McEwan & College, 2005; Lavy, 2004).
- La mayoría de estudios que encuentran un efecto positivo y significativo de usar un esquema de incentivos han sido desarrollados en países en vías de desarrollo como son el caso de Perú (Cueto et al, 2008), Chile (Rau & Contreras, 2009; McEwan & Santibañez, 2005), Kenya (Duflo & Hanna, 2005), India (Muralidharan & Sundararaman, 2009) e Israel (Lavy, 2004). Mientras que, en países desarrollados, no se ha encontrado mayores efectos contundentes y más bien se han encontrado efectos negativos de la implementación de este tipo de esquemas.
- Diversos estudios, principalmente en países en vías de desarrollo, han encontrado efectos de los esquemas de incentivos en variables relacionadas al desempeño docente como son la asistencia (Cueto et al, 2008; Duflo & Hanna, 2011), mejoras en las prácticas pedagógicas (Lavy, 2004; Loyalka et al, 2015), horas o días extra de clases a los estudiantes (Lavy, 2004; Glewwe et al, 2010), y actitudes positivas hacia la evaluación docente (Mizala & Romaguera, 2005; Loyalka et al, 2015). Esto estaría indicando que los incentivos estarían motivando a los docentes a mejorar su desempeño y lograr mejorar los resultados de los estudiantes. Sin embargo, se ha podido apreciar en la literatura que no hay estudios que midan el efecto de la implementación

³ En el Anexo A, se presenta una matriz con la descripción detallada de los programas incluidos en la revisión de literatura.

de los esquemas de incentivos en el desempeño de los Directores, aspecto que debería ser explorado dado que varios de los esquemas son grupales.

- Se han encontrado efectos de los incentivos a docentes tanto en esquemas grupales como individuales (Muralidharan & Sundararaman, 2009), siendo los esquemas grupales aquellos que mejores resultados educativos logran en los estudiantes de acuerdo a la revisión de literatura realizada por Imberman (2015). Sin embargo, se ha encontrado evidencia que los incentivos grupales funcionan en escuelas chicas o con un menor número de docentes, dado que en las escuelas grandes se presenta el riesgo que varios docentes reciban los incentivos si haber realizado ninguna acción para merecerlo, es decir se presenta el problema del *free-rider* (Imberman, 2015; Goodman & Turner, 2013).

3. La herramienta *Bono Escuela*

Bono Escuela se crea mediante el Decreto de Urgencia 002-2014, que de acuerdo a su artículo 10 indica:

“Autorízase al Ministerio de Educación y los Gobiernos Regionales a otorgar, de manera excepcional, un Bono de Incentivo al Desempeño Escolar como reconocimiento de la mejora en el aprendizaje de los estudiantes de las instituciones educativas públicas durante el ejercicio anterior, a favor del personal directivo, personal jerárquico y personal docente nombrado y/o contratado de las instituciones educativas públicas de Educación Básica Regular con mejor desempeño.”

Posteriormente, el Decreto Supremo 287-2014-EF proporciona los criterios, requisitos y condiciones para su adjudicación. El objetivo de esta herramienta es brindar un reconocimiento monetario a los docentes y directores de las instituciones educativas públicas por el logro de resultados educativos de sus estudiantes. Está dirigido al personal directivo y docentes nombrados y/o contratados de instituciones educativas públicas de Educación Básica Regular.

En el año 2014, las instituciones educativas participantes en el *Bono escuela* fueron aquellas instituciones educativas públicas de educación básica regular del nivel primario registradas como activas en el padrón de instituciones educativas correspondiente al año previo del otorgamiento del *bono*, publicado por la Unidad de Estadística Educativa del MINEDU.

Para el otorgamiento del *bono escuela*, se divide a las instituciones educativas en tres grupos, y se rigen bajo los siguientes criterios para participar:

Cuadro 1. Criterios para identificar a las instituciones educativas participantes

Criterios	Grupo A	Grupo B	Grupo C ^{1/}
Clasificación			
Clasificación de grupos	IIEE que participaron de la Evaluación Censal de Estudiantes (ECE) en segundo grado de primaria en Comprensión de Lectura y Matemáticas, en los dos años previos, según los registros de la UMC.	IIEE que implementan la Educación Intercultural Bilingüe (EIB) y que participaron de la ECE en cuarto grado de primaria en Comprensión de Lectura en castellano como segunda lengua, en los dos años previos, según los registros de la UMC.	IIEE donde no se ha programado y/o aplicado la ECE en alguno de los dos años previos (o ambos) por motivos externos a esta, según los registros de la UMC ^{2/} .
Elegibilidad			
Porcentaje de estudiantes y número de secciones programados que rindieron la última ECE	i) si la IIEE cuenta con una sola sección: 80% ii) si la IIEE cuenta con dos o más secciones: 90%	i) si la IIEE cuenta con una sola sección: 80% ii) si la IIEE cuenta con dos o más secciones: 90% iii) Si la IIEE cuenta con tres o más secciones: 95% (a partir del 2015).	-
Registro de matrícula en el SIAGIE ^{2/}	IIEE que tienen estudiantes matriculados y registrados en el SIAGIE ^{3/} en el último año previo.	IIEE que tienen estudiantes matriculados y registrados en el SIAGIE ^{3/} en el último año previo.	IIEE que tienen estudiantes matriculados y registrados en el SIAGIE ^{3/} en el último año previo.

1/ En este grupo también se encuentran instituciones educativas que son parte del programa EIB pero que no hayan participado de las Evaluaciones Censales en Lengua indígena.

2/Si la ECE no ha sido aplicada en la IE por motivos atribuibles a esta, no es tomada en cuenta para la selección de Bono escuela.

3/ Sistema de Información de Apoyo a la Gestión de la Institución Educativa

Para que una institución educativa sea considerada en el proceso de selección de las instituciones educativas con mejor desempeño debe cumplir dichos criterios o requisitos de manera simultánea. Cabe señalar, que a partir del 2015, a diferencia del años anterior, aquellas instituciones educativas que cumplan con todos los requisitos anteriormente indicados y que durante dos años consecutivos haya logrado tener al 80% o más de sus estudiantes en el nivel satisfactorio, se les otorgará automáticamente el *bono escuela*, con la finalidad de poder premiar su excelencia académica (DS 203-2015-EF).

Una vez identificadas las instituciones educativas que van a participar del *bono escuela*, se procede a agruparlas de acuerdo a diferentes aspectos como se detalla en el siguiente cuadro.

Cuadro 2. Niveles de agrupación de las instituciones educativas participantes del bono escuela

Grupo A	Grupo B	Grupo C^{1/}
1) De acuerdo a la Dirección Regional de Educación (DRE) a la que pertenece. 2) De acuerdo a la Unidad de Gestión Educativa Local (UGEL) a la que pertenece. 3) Característica de la IIEE (polidocente, multigrado o unidocente).	1) De acuerdo a la Dirección Regional de Educación (DRE) a la que pertenece. 2) Característica de la IIEE (polidocente, multigrado o unidocente).	1) De acuerdo a la Dirección Regional de Educación (DRE) a la que pertenece ^{2/} .

1/ No se incorpora la característica de la institución educativa debido a que sólo el 1% de las instituciones educativas son polidocentes completas en este grupo.

2/ A partir del año 2015 se tomará como nivel de agrupación la UGEL a la que pertenece la IE.

Una vez agrupadas las instituciones educativas participantes se procede a la selección de aquellas instituciones educativas con mejor desempeño. Para tal fin se hace uso del *índice de desempeño escolar* que está conformado por cinco indicadores para las instituciones educativas del grupo A y B, y por tres indicadores para las instituciones educativas del grupo C. A continuación se dan detalles de los indicadores:

Cuadro 3. Indicadores para la generación del índice de desempeño escolar^{1/}

Nº	Indicadores	Grupo A	Grupo B	Grupo C
1	Superación Promedio de la variación anual de los puntajes de las pruebas aplicadas en las instituciones educativas.	Indicador_1A= [variación comprensión lectora + variación matemáticas]/2	Indicador_1B= [variación comprensión lectora en castellano (2da lengua)]	-
2	Efectividad Puntaje promedio de los estudiantes evaluados en las instituciones educativas.	Indicador_2A=[puntaje matemáticas último año + puntaje comprensión lectora último año]/2	Indicador_2B= [puntaje comprensión lectora en castellano (2da lengua)]	-
3	Eficiencia Tasa de retención de estudiantes de toda las instituciones educativas en el año previo al otorgamiento del bono.	Indicador_3= [(Nº de estudiantes aprobados + Nº de estudiantes desaprobadados)/(Nº de estudiantes aprobados + Nº de estudiantes desaprobadados+ + Nº de estudiantes retirados)]		
4	Oportunidad en el registro en el SIAGIE^{2/} del acta de evaluación y nómina de matrícula para todas las secciones de la institución educativa correspondientes al año en que se otorga el bono.	Indicador_4= [Nº de estudiantes de las secciones que cumplen las tres condiciones ^{4/} / Nº de estudiantes total por institución educativa]		
5	Registro en el SIAGIE^{2, 3/} de grado o grados de enseñanza de los docentes	-	-	Indicador_5= [Nº de docentes de la institución educativa]

Nº	Indicadores	Grupo A	Grupo B	Grupo C
	que laboran en la institución educativa en el año que se otorga el bono.			con registro de grado de enseñanza en el SIAGIE/ total de docentes de la institución registrados en NEXUS]

1/ Los indicadores se re escalan al interior de cada subgrupo

2/ Sistema de Información de Apoyo a la Gestión de la Institución Educativa

3/ Para el *bono escuela* 2015 se cambiaría este indicador por “Oportunidad en la rendición de cuentas del Programa de Mantenimiento de la Infraestructura y Mobiliario de los Locales Escolares”

4/ La condición 1 es que la nómina de matrícula sea remitida o rectificada. Si el estado es “aprobado”, se considera que la nómina ha sido remitida o rectificada previamente. La condición 2 refiere a que el acta de evaluación sea remitida o rectificada. Si el estado es “aprobado”, se considera que el acta de evaluación ha sido remitida o rectificada previamente. Finalmente, la condición 3 refiere a que el acta de evaluación sea al menos remitida o rectificada antes del 1º marzo del año del otorgamiento del Bono. Si el estado es “aprobado” antes de esa fecha, se considera que el acta de evaluación ha sido remitida o rectificada previamente.

De ahí, para construir el *índice de desempeño*, se realiza una suma ponderada para cada grupo, como se muestra en el siguiente cuadro:

Cuadro 4. Fórmulas para la generación del índice de desempeño escolar en base a los indicadores de cada grupo

	Fórmula
Grupo A	$IME = 0.60 * \text{Superación} + 0.20 * \text{Efectividad} + 0.15 * \text{Eficiencia} + 0.05 * \text{Registro Estudiantes}$
Grupo B	$IME = 0.60 * \text{Superación} + 0.20 * \text{Efectividad} + 0.15 * \text{Eficiencia} + 0.05 * \text{Registro Estudiantes}$
Grupo C	$IME = 0.60 * \text{Eficiencia} + 0.30 * \text{Registro Estudiantes} + 0.10 * \text{Registro Docentes}$

Para fijar qué instituciones educativas reciben el *bono escuela*, se ordenan de forma descendente todas las escuelas participantes al interior de cada subgrupo o estrato definido⁴, según el índice obtenido.

Las instituciones educativas con mejor desempeño son aquellas que se encuentran en algunos de los dos siguientes grupos:

- Top 25: instituciones educativas ubicadas en el percentil 25 de la distribución del puntaje del índice de desempeño al interior de cada subgrupo o estrato, según el ordenamiento antes mencionado.
- Top 25-35: instituciones educativas ubicadas entre el percentil 25 y 35 de la distribución del puntaje del índice de desempeño al interior de cada subgrupo o estrato, según el ordenamiento antes mencionado.

Uno de los principales aspectos de este esquema de incentivos es su carácter colectivo. Así, si una institución educativa sale seleccionada por haber mejorado su desempeño, en principio, todos los docentes y director de la institución educativa reciben la bonificación, siempre y cuando cumplan con los requisitos establecidos en el Decreto Supremo 287-2014-EF. Así, aquellos docentes o directivos que no cumpla con algún requisito no recibirán el incentivo monetario.

Los requisitos que deben cumplir el personal docente, jerárquico o directivo de la institución educativa seleccionada para recibir la bonificación son:

- Cuentan con plaza de nombramiento o contrato asignado a la institución educativa seleccionada durante por lo menos 6 meses (continuos o discontinuos) del último año previo, según la información disponible del sistema de administración y control de plazas (NEXUS⁵).

⁴ Los estratos son: Dirección Regional, Unidad de Gestión Local, Característica de la Escuela.

⁵ Sistema para la Administración y Control de Plazas docentes, administrativos del sector educación.

- El personal directivo, jerárquico y docente nombrado y/o contratado de las instituciones educativas seleccionadas debe encontrarse registrado en el AIRHSP⁶ a cargo del Ministerio de Economía y Finanzas.
- Cuentan con plaza de nombramiento o contrato vigente según la información disponible del NEXUS al cierre del antepenúltimo mes anterior a la fecha del Informe de la Secretaría de Planificación Estratégica del Ministerio de Educación que presenta la lista de instituciones educativas y personal beneficiario del *bono escuela*.
- No cuentan con sanción administrativa disciplinaria de suspensión en el cargo, cese temporal en el cargo o destitución del servicio vigente al cierre del antepenúltimo mes anterior a la fecha del Informe de la Secretaría de Planificación Estratégica del Ministerio de Educación que presenta la lista de instituciones educativas y personal beneficiario del *bono escuela*, de acuerdo al Registro Nacional de Sanciones de Destitución y Despido.

Los beneficiarios del *bono escuela* son los docentes y directores (y demás personal jerárquico) del nivel primario, siendo los montos del bono diferenciados de acuerdo al grupo en que se encuentre la institución educativa que labora (incluye a todos los docentes, con aula a cargo o no, de las IE con EBR primaria). Cabe señalar, que en el caso de instituciones educativas integradas (cuenten con más de un nivel en la institución educativa) también reciben el *bono escuela* los directores que no sean del nivel primario (implementado para el año 2015).

Los montos de la bonificación varían de acuerdo al grupo en que se encuentra la institución educativa (A, B o C), el percentil del *índice de desempeño escolar* en el que se ubique y el cargo/grado de enseñanza. A continuación se detallan los montos de las bonificaciones que reciben los docentes y directores de acuerdo a los grupos de instituciones educativas.

Cuadro 5. Montos de las bonificaciones por grupos de instituciones educativas

	Grupo A ECE		Grupo B ECE - EIB		Grupo C No ECE	
	Top 25	Top 25-35	Top 25	Top 25-35	Top 25	Top 25-35
Director ^{1/}	3,000	2,000	3,000	2000	1,500	1,000
1°	2,500	1,667	2,500	1,667	1,500	1,000
2°	3,000	2,000	2,500	1,667	1,500	1,000
3°	1,500	1,000	2,500	1,667	1,500	1,000
4°	1,500	1,000	3,000	2,000	1,500	1,000
5°	1,500	1,000	1,500	1,000	1,500	1,000
6°	1,500	1,000	1,500	1,000	1,500	1,000
Sin grado	1,500	1,000	1,500	1,000	1,500	1,000

1/ Para el año 2015, a diferencia del año anterior, en el caso de instituciones educativas integradas, si el director era de otro nivel también recibe la bonificación por la mejora en el desempeño de su escuela primaria.

Finalmente, un cambio para el año 2016 es incorporar al nivel secundario en el *bono escuela* dado que se cuenta con información de rendimiento (Comprensión de Lectura y Matemática) para este nivel desde el 2015.

3.1 Revisión de la herramienta a la luz de las Experiencias Internacionales

Se puede apreciar que la herramienta implementada por el Ministerio de Educación ha considerado diferentes aspectos que la literatura sobre esquemas de incentivos indica se debe tomar en cuenta con la finalidad de minimizar los posibles problemas durante su implementación. A manera de resumen, estos aspectos son:

- *No usar solo rendimiento como indicador para dar el incentivo:* diferentes estudios e informes (Solmon y Podgursky, 2000; Baker et al, 2010; Fraser Institute, 2014) indican que es necesario incorporar otras

⁶ Aplicativo Informático para el Registro Centralizado de Planillas y de Datos de los Recursos Humanos del Sector Público

medidas de desempeño de la escuela en los esquemas de incentivos monetarios a docentes y no usar sólo rendimiento como indicador, de forma tal que estos puedan funcionar mejor y evitar posibles efectos no deseados como que los preparen solo para las pruebas. Un ejemplo claro es el caso Chileno donde su esquema de incentivos incorpora diferentes indicadores que reflejan el desempeño de una escuela.

- *Criterios de cobertura para que una institución educativa participe:* una de las consecuencias negativas detalladas en la literatura es que durante las evaluaciones el número de estudiantes evaluados es menor con relación a la matrícula regular en las escuelas debido a que se excluyen a cierto grupos de estudiantes de las evaluaciones (Solmon y Podgursky, 2000). Motivo por el cual, el *bono escuela* considera dentro de sus criterios para participar del bono contar con una cobertura del 80% de los estudiantes matriculados en instituciones educativas de una sección y de 90% de los estudiantes matriculados en el caso de instituciones educativas con dos o más secciones. De esta forma, se asegura que no se excluya de las evaluaciones a los estudiantes de bajo rendimiento y se tienda a inflar el puntaje de las instituciones educativas.
- *Hacer el incentivo colectivo en lugar de individual:* si bien estudios experimentales como el realizado por Muralidharan y Sundararaman (2009) muestran que podría haber diferencias entre esquemas individuales y grupales, los incentivos grupales permiten o generan espacios de colaboración entre los docentes para el logro de una meta común que sería mejorar el desempeño de los estudiantes, aspecto que hace importante el uso de incentivos colectivos (Bryk y Schneider, 2002; OECD, 2011). Sin embargo, es posible que haya incentivos para acciones oportunistas dentro de los esquemas de incentivos grupales cuando cada docente tiene menos responsabilidad a cargo (Goodman y Turner, 2012).
- *Usar no solo el incremento en el tiempo sino los puntajes simples también,* de forma tal de no generar desincentivos en las escuelas de altos rendimientos: en el caso del esquema de incentivos en Chile, un aspecto que se tomó en cuenta para evitar desincentivar a las escuelas que ya vienen obteniendo buenos resultados es incorporar como parte del indicador para otorgar el incentivo el rendimiento de las escuelas en el periodo que se otorga el bono (Mizala y Romaguera, 2005). Sin embargo, de forma adicional, el esquema de *bono escuela* en Perú incorpora un mecanismo automático (siempre y cuando cumplan los requisitos para ser participante del esquema) para la entrega de la bonificación para las instituciones educativas con rendimiento sobresalientes o como Rau y Contreras (2009) llaman los que siempre serán ganadores (*winner*s) y que se pueden ver perjudicados por incorporar el incremento en el rendimiento en la fórmula del *índice de desempeño escolar*.
- *Brindar incentivos monetarios que representen montos significativos con respecto a los salarios de los docentes:* estudios como los desarrollados por Kelley et al (2000), Lincove (2012) y Fraser Institute (2014) y otros anteriormente citados indican que el tamaño del incentivo es un factor crucial para el éxito de un esquema. Así, en el caso de la herramienta implementada por el Ministerio de Educación, la bonificación más baja es de S/. 1,000 que representa el 65% del salario mensual de un docente en el nivel I de la escala de la carrera pública magisterial y 33% para un docente en el nivel V por una jornada de 30 horas semanales, mientras la bonificación más alta de S/. 3,000 representa el 194% del salario mensual del nivel I y 98% del nivel V para la misma jornada. De esta forma, se puede apreciar que el incentivo representa montos significativos del sueldo docente.
- *Hacer los rankings específicos a cada contexto en que se encuentran las escuelas:* de acuerdo a la OECD (2011) un sistema de incentivos debe premiar la buena performance de la escuela y la mejora relativa; es decir, se debe de comparar la mejora de las escuelas con otras en contextos y características similares. De esta forma, al considerar la estratificación por UGEL y característica de la institución educativa, el *bono escuela* estaría siendo más efectivo para evaluar el progreso de las instituciones educativas.

4. Objetivos de la Evaluación de Impacto

La presente evaluación tiene como principal objetivo evaluar el impacto de la herramienta *Bono Escuela* en los resultados educativos de las instituciones educativas estatales. Se va explorar el impacto del programa en diferentes variables intermedias relacionadas con el desempeño de los docentes y/o directores en sus instituciones educativas (asistencia docente, asistencia del director, carga docente, horas de clase al día y contar

con documentos de gestión). El impacto del programa se va evaluar de dos formas, la primera está relacionada con el efecto de la competencia que genera el programa en las instituciones educativas estatales (sea que ganen o no el incentivo), es decir, se va explorar si los resultados educativos de los estudiantes mejoran solo por el hecho de saber que pueden ganar el *bono escuela*. La segunda guarda relación con la continuidad de los buenos resultados educativos en el siguiente periodo de recibir el *bono escuela*, este efecto busca ver si el hecho de recibir el incentivo monetario en el año t genera algún cambio en los resultados educativos en el año $t+1$.

En ambos casos, se va explorar el efecto del programa en diferentes variables de resultados educativos, buscándose responder las siguientes preguntas: *¿Cuál es el efecto del programa (competencia y un año después) en el rendimiento y niveles de desempeño de las instituciones educativas estatales en comprensión lectora?* y *¿Cuál es el efecto del programa (competencia y un año después) en el rendimiento y niveles de desempeño de las instituciones educativas estatales en matemática?*

En el caso del efecto de continuidad de los resultados educativos, se va buscar responder las siguientes preguntas adicionales⁷: *¿Cuál es el efecto del programa en la asistencia de los docentes/directores en las instituciones educativas estatales?* *¿Cuál es el efecto del programa en variables relacionadas a procesos escolares (días de clase al año, horas de clase al día, entre otras) y gestión escolar (contar con los documentos de gestión, contar con comité de aula) de las instituciones educativas estatales?*. Se espera que el otorgamiento de incentivos monetarios a los docentes y directores de las instituciones educativas, los motiven a desarrollar acciones o propiciar un clima que favorezca el aprendizaje de los estudiantes, a través del desarrollo de diferentes prácticas pedagógicas y/o gestión escolar al interior de la institución educativa o el aula de clase.

5. Estrategia de Evaluación de Impacto

Para poder responder a los diferentes objetivos de investigación planteados, es necesario tener un diseño de evaluación que permita medir los efectos causales del *bono escuela*. Sin embargo, debido al carácter universal del *bono escuela* (cualquier institución educativa estatal a nivel nacional puede recibir este incentivo siempre y cuando cumpla los criterios y requisitos para ser elegible) no se puede contar con un diseño de evaluación experimental o aleatorizado que sería el mejor método para medir el impacto de cualquier programa de intervención social. Así, se plantea usar diseños de evaluación cuasi-experimentales para poder responder a los diferentes objetivos del presente estudio.

Los diseños cuasi-experimentales se diferencian de los experimentales en cuanto a la forma de asignación de los individuos a un grupo o situación de estudio. En el caso de los diseños cuasi-experimentales, a diferencia de los experimentales, la asignación de los individuos a un grupo o estado no es de manera aleatoria pero sí se cuenta con un grupo de tratamiento y otro de contraste. Los diseños cuasi-experimentales son diversos, entre los cuales se tiene a los modelos de emparejamiento (Rosenbaun y Rubin 1983, Heckmak, Ichimura & Todd 1997), modelos de diferencias en diferencias (Ashenfelter & Card, 1985; Meyer, 1995; Abadie, 2005) y modelos de regresiones discontinuas (Trochin 1994; Lee & Lemieux 2010).

Para medir el efecto de la competencia, se hace uso de los modelos de emparejamiento. Los modelos de emparejamiento son comúnmente usados en evaluaciones de impacto de programas con diseños cuasi-experimentales. Estos tipos de modelos permiten al evaluador recurrir a un grupo de técnicas estadísticas para garantizar que los grupos que compara (tratamiento y contraste) sean lo más similar posible en términos de aspectos observables para la participación del programa. Así, dado el carácter universal del *bono escuela*, aspecto que origina que todas las instituciones educativas estatales puedan ser elegibles y ganar el bono; se optó por usar como grupo de comparación a las instituciones educativas privadas. La selección de este tipo de instituciones educativas como grupo de comparación responde a: i) no están afectas al *bono escuela*, y ii) dado su crecimiento en los últimos 20 años, atienden a grupos de estudiantes con similares características

⁷ No se incluyen el análisis de estos resultados educativos para el efecto de la competencia debido a que no se cuenta con información sobre procesos o gestión escolar para el grupo de comparación seleccionado que son las instituciones educativas privadas.

sociodemográficas que los estudiantes de instituciones educativas privadas, aspecto que favorece la comparación entre grupos⁸.

El usar esta metodología para medir el efecto de la competencia presenta sus ventajas y desventajas. Entre las principales ventajas está el poder contar con una muestra comparable de acuerdo a diferentes características de la composición demográfica de los estudiantes, características de las instituciones educativas y características del contexto donde se ubica la institución educativa. Otro aspecto está relacionado con la no linealidad de su diseño, es decir, el modelo de emparejamiento no asume que existe una forma funcional entre la variable de tratamiento y la variable de resultado⁹. Una última ventaja está relacionada con el tamaño de muestra, dado que al empatar por el puntaje de propensión entre ambos grupos, los requerimientos de tamaño de muestra para un buen emparejamiento son de al menos 1000 observaciones (Yang et al, 2007). En cuanto a las limitaciones, la más importante es que solo permite controlar por posibles sesgos en variables observables, mientras diseños experimentales o cuasi-experimentales como regresiones discontinuas permiten controlar también por variables no observables dado que este último evalúa individuos cercanos al punto de discontinuidad.

Para medir la continuidad del efecto (buenos resultados educativos) en las instituciones educativas que reciben el *bono escuela*, se plantea usar un modelo de regresión discontinua. Este tipo de modelación permite comparar a individuos que participaron y no participaron del programa en función a un punto de corte fijado por una variable de asignación. En el caso del *Bono Escuela*, la variable de asignación es el índice de desempeño escolar y el punto de corte es el percentil 65 donde toda institución educativa por encima de este percentil recibe el incentivo. Existen dos formas de estimación para ver el efecto del programa que se está evaluando, que son: i) estimación paramétrica y ii) estimación no paramétrica. En el caso de la estimación paramétrica, el énfasis está en identificar la mejor forma funcional (o polinomio) entre la variable de asignación y la variable de resultado. En el caso de la estimación no paramétrica, se fija un ancho de banda¹⁰ alrededor del punto de corte y se corre una estimación lineal para ver el efecto del programa. No es necesario incluir variables de control dado que los individuos alrededor de este punto de corte tendrían características similares y la única diferencia sería haber recibido el incentivo. La ventaja de este último tipo de estimación es que permite tener individuos con similares características, aspecto que elimina posibles sesgos por variables no observables, lo que hace que este diseño se asemeje más a una evaluación experimental¹¹.

Los modelos de RD están siendo usados cada vez más en la investigación educativa o evaluaciones de impacto de programas en educación. A manera de ejemplo, se tienen que han sido usados para evaluar la provisión de educación pre-escolar universal (Gormley, Gayer, Phillips y Dawson 2005; Barnett, Lamy y otros 2005; Wong, Cook, Barnett y Jung 2008), programas de *tracking* en las escuelas (Duflo, Dupas y Kremer 2001), educación bilingüe (Chin, Daysal, Imberman 2011), incentivo monetarios a los docentes en las escuelas (Lavy 2002; Jinnai 2012) y programas de transferencias condicionadas como PROGRESA en México y Bono de Desarrollo Humano en Ecuador (Buddelmeyer y Skoufias 2004; Ponce y Bedi 2008).

Por otro lado, el usar esta metodología tiene sus ventajas y desventajas. En cuanto a las ventajas se tiene: i) los estimados de una regresión discontinua (no paramétrica) son similares a los de una evaluación de tipo experimental debido a que se comparan individuos o sujetos cercanos al punto de discontinuidad, lo cual hace que sean similares y la única diferencia sea el hecho de recibir o no el tratamiento (Trochin 1994; Lee & Lemieux 2010; Jacob, Zhu, Somers & Blomm 2012), ii) poseen un buen nivel de validez interna debido a las diferentes formas funcionales que se prueban para poder capturar los efectos del programa (Imbens & Lemieux 2008, Jacob, Zhu, Somers & Blomm 2012). La principal desventaja de este tipo de diseños está en la pérdida de validez externa dado que se utiliza solo una parte de la muestra (alrededor del punto de corte) y no la totalidad de observaciones (Trochin 1994, Lee & Lemieux 2010).

⁸La selección de este grupo de comparación se basa en un análisis similar realizado por Rau y Contreras (2009) para evaluar el impacto de los incentivos monetarios en Chile, donde usan a las escuelas privadas como grupo de control para ver el efecto de la competencia que genera el sistema de incentivos chilenos.

⁹La mayoría de métodos cuasi-experimentales asumen que existen una forma funcional para la variable de resultado (lineal), mientras los modelos de emparejamiento se basa en encontrar grupos comparables con la finalidad de reducir los posibles sesgos.

¹⁰ Existen diferentes métodos para fijar el ancho de banda óptimo. Para mayores referencias ver Lee & Lemieux (2010).

¹¹ Dina Pomeranz (2011) realiza una revisión de los diferentes métodos de evaluación y concluye que el mejor método cuasi-experimental para hacer una evaluación de impacto es el de Regresiones Discontinuas.

5.1 Unidad de Análisis

El *bono escuela* tiene como objetivo reconocer el esfuerzo de los docentes y directores de las instituciones educativas estatales que logran mejorar los resultados educativos de sus estudiantes. Motivo por el cual, la unidad de análisis considerada para la presente evaluación de impacto es la institución educativa.

5.2 Bases de datos y tamaño de muestra

En cuanto a las bases de datos que se van a usar para la evaluación de impacto estas son:

- La evaluación Censal de Estudiantes (2014, 2015): esta base de datos se recoge anualmente y permite tener información de resultados de aprendizaje de los estudiantes en Matemática y Comprensión de Lectura de 2do grado de primaria en escuelas castellano hablantes con matrícula de 5 alumnos o más y 4to grado en escuelas bilingües con 5 alumnos o más matriculados. A partir del año 2015, se comenzó a evaluar también estudiantes del 2do año de educación secundaria. Las áreas de aprendizaje que mide son Comprensión de Lectura y Matemática en 2do grado de primaria y 2do año de secundaria, mientras en comprensión lectora en castellano y lengua indígena en 4to de primaria de escuelas EBI.
- Censo Escolar (2014-2015): esta base de datos contiene información sobre la infraestructura escolar, personal docente, resultados educativos y el contexto escolar de todas las escuelas a nivel inicial, primario y secundario a nivel nacional. Para el presente informe, se hace uso de la información disponible para el nivel primario.
- Semáforo Escuela (2015): esta base de datos contiene información sobre la infraestructura escolar, personal docente, procesos escolares y performance de los docentes para una muestra de escuelas estatales del nivel inicial, primario y secundario de todo el país. La representatividad es a nivel de cada Unidad de Gestión Local, al interior de cada región del Perú. Al igual que el Censo Escolar, sólo se hace uso de la información correspondiente al nivel primario. Así mismo, no se puede usar todas las variables del semáforo escuela debido a que durante los primeros meses (Abril-Julio) estuvo en proceso de consolidación de preguntas y temas a indagar, motivo por el cual algunos temas relevante como por ejemplo el acompañamiento o supervisión de los directores a los docentes, recién se incluye en los cuestionarios a partir de agosto, aspecto que hace que varios indicadores cuenten con un número de observaciones bastante reducido.
- Sistema de Focalización de Hogares (2012-2013): esta base de datos cuenta con información de los hogares a nivel nacional a nivel censal y está disponible en la sección de microdatos de la página web del Instituto Nacional de Estadística e Informática. Para el presente, se construyen diferentes indicadores a nivel de hogar que luego son agregados a nivel distrital para ser combinados con la información de las otras bases de datos a través del código de ubigeo de cada institución educativa.

De esta manera, se usó la información de las bases de datos antes mencionadas con la finalidad de poder estimar los efectos del *bono escuela*. La muestra final de instituciones educativas para ver los efectos de la competencia y los efectos de sostenibilidad del programa a un año de entregado el incentivo son diferentes debido a que los grupos de comparación son distintos.

El tamaño de muestra para evaluar el efecto de la competencia que genera el incentivo, fue de 8,855 instituciones educativas de las cuales 2,338 son estatales (o tratamiento) y 6,517 no estatales (o control). Para llegar a este tamaño de muestra, se partió de un total de 37,888 instituciones educativas primarias estatales y no estatales, muestra que se redujo debido a la aplicación de los siguientes filtros:

- Se excluyen instituciones educativas rurales debido a que la oferta de educación no estatal es primordialmente urbana (22,365 I.E. eliminadas).
- Se excluyen a las estatales que no eran elegibles para el bono (2,123 I.E. eliminadas).

- Se excluyen instituciones educativas que participan del programa soporte pedagógico (2,243 I.E. eliminadas).
- Se excluyen instituciones educativas que indican participar del programa de acompañamiento rural o ser parte del programa ASPI (433 I.E. eliminadas).
- Se excluyen instituciones educativas que no cuentan con datos de rendimiento de la ECE 2015¹² (1,768 I.E. eliminadas).
- Se excluyen instituciones educativas en regiones donde no se cuente con instituciones educativas que sean parte del tratamiento (51 I.E. eliminadas).

El tamaño de muestra para evaluar el efecto de continuidad del efecto a un año de la entrega del incentivo, fue de 5,858 instituciones educativas. Para llegar a este tamaño de muestra, se partió de un total de 13,633 primarias estatales que son parte del grupo A de instituciones educativas elegibles para recibir el *bono escuela* en el 2014. Sin embargo, esta muestra se redujo debido a la aplicación de los siguientes filtros:

- Se excluyen instituciones educativas que no tengan resultados de rendimiento en la ECE 2015¹³ (2,954 I.E. eliminadas).
- Se excluyen instituciones educativas de gestión privada pero financiamiento público (160 I.E. eliminadas).
- Se excluyen instituciones educativas que participaron del programa de Soporte Pedagógico (2,301 I.E. eliminadas).
- Se excluyen instituciones educativas que participan del programa de acompañamiento pedagógico rural en el 2014 o 2015 (2,149 I.E. eliminadas).
- Se excluyen instituciones educativas que participan del programa ASPI (211 I.E. eliminadas).

5.3 Variables de estudio

El impacto del programa se va medir sobre las siguientes variables a nivel de cada institución educativa. En primer lugar, se tienen las variables relacionadas a resultados finales, estas son:

- Rendimiento en Matemática (RF1¹⁴): puntaje promedio de los estudiantes por institución educativa para el 2014 y 2015¹⁵.
- Niveles de desempeño en Matemática (RF1): porcentaje de estudiantes que se ubican en el nivel de desempeño satisfactorio, en proceso y en inicio para el 2014 y 2015¹⁶.
- Rendimiento en Comprensión Lectora (RF1): puntaje promedio de los estudiantes en la institución educativa para el 2014 y 2015.

¹² Esto se debe a que el índice se construye en función de la información del puntaje del 2014 y 2013, es decir, hay escuelas que participaron esos años pero que no cuentan con datos de rendimiento en el 2015.

¹³ Se excluyen instituciones educativas que si bien son elegibles de acuerdo a la muestra del 2014, en el 2015 no se cuenta con información de rendimiento.

¹⁴ Códigos de acuerdo a la teoría de cambio presentado en una sección previa. RF hace referencia a Resultado Final y RI hace referencia a resultado intermedio.

¹⁵ Cabe señalar que los puntajes obtenidos por los estudiantes fueron calculados usando el modelo Rasch o Teoría de Respuesta al Ítem de 1 parámetro. Este modelo usa como parámetro para la estimación de los puntajes, el nivel de dificultad de los ítems. Así, estima la probabilidad de que un estudiante responda correctamente cada ítem ya sea que lo responda bien o no. La media se fijó en 500 puntos, mientras que la desviación estándar se fijó en 100 para el año 2007. Los años posteriores se equiparó los puntajes de forma tal que los puntajes del 2007 al 2015 sean comparables en el tiempo.

¹⁶ Estos niveles de desempeño fueron fijados en el 2007 por la Unidad de Medición de Calidad Educativa del Ministerio de Educación. El nivel suficiente indica que los estudiantes logran desarrollar las tareas para su grado, los del nivel en proceso están en proceso de lograr desarrollar las tareas para su grado y en inicio indican que los estudiantes no pueden ni resolver las tareas más básicas para el área que se evalúa.

- Niveles de desempeño en Comprensión Lectora (RF1): porcentaje de estudiantes que se ubican en el nivel de desempeño satisfactorio, en proceso y en inicio para el 2014 y 2015.
- Asistencia de los estudiantes (RF2): porcentaje de estudiantes que se encuentran presentes durante la visita del monitoreo de semáforo escuela en la institución educativa.

En segundo lugar se cuenta con las variables relacionadas a resultados intermedios. Se han considerado variables relacionadas al desempeño de los docentes y directores, así como variables relacionadas con el quehacer pedagógico y que guardan relación con los procesos escolares al interior de la institución educativa. Estas variables son:

- Asistencia docente (RI1): porcentaje de docentes en la institución educativa que se encontraban presentes al momento de la visita del monitor del sistema de monitoreo a instituciones educativas. Esta variable guarda relación con la performance de los docentes.
- Asistencia de los directores (RI1): variable cualitativa que toma el valor de 1 si el director se encontraba presente durante la visita del monitor del sistema de monitoreo a instituciones educativas. Esta variable guarda relación con la performance de los directores.
- Horas de Clase al día (RI2): número de horas cronológicas de clase al día que programan los directores durante el año escolar 2015, descontando el tiempo dedicado al desayuno/almuerzo/refrigerio y recreo durante el día. Esta variable es *proxy* para el número de horas de aprendizaje al día que tienen los estudiantes y que guarda relación con las oportunidades de aprender de los estudiantes (Carroll, 1969; Bloom, 1980).
- Días de clase (RI2): número de días de clase al año que programan los directores durante el año escolar 2015. Al igual que las horas de clase al día, esta variable es *proxy* para las oportunidades de aprender que tienen los estudiantes durante el año escolar (Carroll, 1969; Bloom, 1980). Si bien, el número de días de clase puede estar fijado desde el gobierno central, es una variable que puede ser modificada por los directores y docentes en cada institución educativa.
- Carga docente (RI2): número de estudiantes por docente a nivel primario para el 2015. El número de estudiantes por aula de clase guarda relación con el tiempo que los docentes les puede dedicar a cada uno de sus estudiantes y por ende con sus oportunidades de aprender en el aula (Carroll, 1969; Bloom, 1980).
- Comité de tutoría (RI2): variable cualitativa que toma el valor de 1 si la institución educativa cuenta con el comité de tutoría, y 0 en cualquier otro caso. El cual busca promover un clima de convivencia democrático al interior de las instituciones educativas aspecto que refleja los procesos escolares al interior de la misma.
- Documentos de gestión escolar (RI2): variable cualitativa que toma el valor de 1 si la institución educativa cuenta con el plan anual de trabajo o Existe una matriz con diagnóstico, objetivos, metas y actividades por cada indicador de los Compromisos de Gestión Escolar para el 2015, y 0 en cualquier otro caso. Este indicador refleja una adecuada gestión escolar en pro del aprendizaje de los estudiantes al interior de la institución educativa.

En cuanto a las variables independientes más importantes, se tiene a las siguientes variables:

- Variable de asignación (2014 y 2015): el índice de desempeño calculado por la Secretaria de Planificación Estratégica del Ministerio de Educación. Cabe señalar que debido a la presencia de múltiples rankings por el diseño mismo del *bono escuela*, se centró la variable alrededor del valor del índice en el percentil 65 de cada ranking, permitiendo de esta manera que todos los rankings estén centrados en 0. Esta variable fue usada solamente para los modelos de regresión discontinua.
- Tratamiento: variable cualitativa que toma el valor de 1 si la institución educativa recibió el *bono escuela* y 0 en cualquier otro caso. Esta variable se usó para la regresión discontinua. En el caso del modelo de

emparejamiento, la variable cualitativa toma el valor de 1 si la institución educativa es de gestión pública y elegible para recibir el Bono Escuela y 0 en caso sea una institución educativa privada.

En cuanto a las variables de control, se usaron el siguiente grupo de variables:

- Composición socio-demográfica de los estudiantes: porcentaje de estudiantes mujeres, porcentaje de estudiantes en extra-edad, y porcentaje de estudiantes indígenas.
- Características del personal docente: porcentaje de docentes mujeres, porcentaje de docentes con título pedagógico y porcentaje de docentes que estudiaron para ser docentes de educación primaria.
- Características de la infraestructura educativa: cuenta con los tres servicios básicos (luz, agua y desagüe), computadoras con acceso a internet y ambientes escolares (sala de cómputo, lona deportiva, sala de profesores, patio y oficinas administrativas).
- Características del contexto a nivel distrital: porcentaje de hogares con servicio de agua, porcentaje de hogares con servicio de desagüe, porcentaje de hogares con servicio electricidad, porcentaje de hogares que se encuentran hacinados (3 personas o más por habitación), porcentaje de la población de 18 o más que tiene educación secundaria completa o más, porcentaje de la población de 15 años o más que trabaja en alguna rama económica, porcentaje de población rural y efectos fijos por región.

5.4 Análisis de poder y tamaño de la muestra

Uno de los aspectos claves que se debe tomar en consideración para asegurar que tan bueno es el diseño del estudio para detectar los efectos de un programa sobre la variable de resultado, es el poder estadístico de la muestra utilizada. De acuerdo a Cohen (1988), en caso de querer comparar dos grupos¹⁷, el tamaño de muestra necesario para cada grupo, se calcula usando la siguiente formula¹⁸:

$$n = \frac{2\sigma^2(Z_\beta + Z_{\alpha/2})^2}{ES^2}$$

Donde:

n	:	Tamaño de muestra cada grupo de estudio
σ^2	:	Varianza de la variable de resultado
Z_β	:	Poder estadístico
$Z_{\alpha/2}$:	Nivel de significancia
ES	:	Es el tamaño del efecto del programa

Sin embargo, en modelos de Regresión Discontinua (RD), a comparación de diseños de evaluación experimentales, los tamaños de muestra son más grandes dado que para estimar efectos *insesgados* bajo este tipo de diseño requieren emplear modelos de regresión para ajustar por la variable de asignación al tratamiento (Shadish, Cook & Campbell, 2002; Deke & Dragoset, 2012). Así, diferentes estudios indican que el tamaño de muestra adecuado para un diseño de evaluación con Regresión Discontinua debe ser 3 a 5 veces más grande que un diseño experimental (Cappelleri, Darlington y Trochim, 1994; Shochet 2008).

A continuación, se presentan los tamaños de muestra necesarios por grupo para un diseño RCT y de RD, para los cálculos se asume un nivel de confianza del 95%.

¹⁷ Se asume igual proporción de instituciones educativas por grupo de estudio (tratamiento y contraste).

¹⁸ Cabe señalar, que debido a que la unidad de análisis es la institución educativa no se realiza el ajuste por correlación intra-cluster en la variable sobre la cual se quiere ver el efecto del programa. Sin embargo, en el Anexo B se da detalles del tamaño de muestra en caso se considere la posibilidad de cambiar la unidad de análisis a los estudiantes.

Cuadro 6. Tamaño de muestra por grupo para un diseño RCT y RD

Poder	RCT					RD con efecto de diseño de 3				
	TE: 0.10	TE: 0.15	TE: 0.20	TE: 0.25	TE: 0.30	TE: 0.10	TE: 0.15	TE: 0.20	TE: 0.25	TE: 0.30
70%	1,234	548	309	197	137	3,702	1645	926	592	411
80%	1,568	697	392	251	174	4,704	2091	1176	753	523
90%	2,100	933	525	336	233	6,299	2799	1575	1008	700
95%	2,592	1152	648	415	288	7,776	3456	1944	1244	864

TE: Tamaño del efecto en desviaciones estándar

En cuanto a que tamaño de efecto se debe de plantear para la presente evaluación, revisiones sistemáticas de programas de intervención orientados a mejorar el rendimiento de los estudiantes y/o desempeño de los docentes (p.ej.: ausentismo) encuentran que el tamaño del efecto oscila entre 0.10 y 0.30 desviaciones estándar (McEwan, 2014:19; Guerrero et al, 2013:482); motivo por el cual, se propone para la presente evaluación muestras que consideren un poder estadístico del 80% y tamaño de efecto de 0.20 desviaciones estándar o menos como adecuados para ver el impacto del *bono escuela*, es decir, muestras de 2,352 instituciones educativas o más (1176 escuelas por grupo de estudio según el diseño RD con efecto de diseño de 3).

5.5 Especificación de los Modelos Estadísticos

A continuación se da una breve especificación de cada uno de los modelos usados para responder a cada uno de los objetivos del presente estudio.

Efecto de la Competencia: Modelo de Emparejamiento

Los modelos de emparejamiento (Rosenbaun y Rubin 1983, Heckmak, Ichimura & Todd 1997) proponen en primer término, un modelo de elección discreta para obtener la “*probabilidad de participación*” (*propensity score*) de ser tratado o no que puede ser estimado usando una función normal o logística. Esta variable indica la probabilidad de que los individuos reciban o no el tratamiento dado una serie de características que pueden estar asociadas con su participación. Para el presente estudio se considera una función logística¹⁹, siendo la ecuación para tal fin la siguiente:

$$P(y_i = j | X) = \frac{\exp(X_i \beta_j)}{1 + \sum_{p=1}^J \exp(X_i \beta_p)}$$

La probabilidad de participación frente a la de no participar del programa (una para cada grupo de comparación) es obtenida a partir de características que hacen diferentes a ambos grupos que participan del programa (contenidos en X_i). Una vez estimado el puntaje, se procede a emparejar a los individuos usando diferentes técnicas estadísticas, entre las más comunes están las de emparejamiento uno a uno (*one to one matching*), por distribución de puntajes (*kernell matching*) y por el número de vecinos más cercanos (*nearest neighbor matching*). Para el presente estudio, se optó por usar el emparejamiento *Uno a Uno* y el método de vecinos más cercanos (*nearest neighbor matching*). Así mismo, en ambos casos, el emparejamiento se hizo sin reemplazos dado que se dispone con un número considerable de controles y se quiere evitar que las comparaciones sean siempre con el mismo grupo de controles que comúnmente sucede cuando se opta por un emparejamiento con reemplazo (Stuart & Rubin, 2008). Una vez seleccionado el método de emparejamiento de los puntajes estimados, se calcula la diferencia en la variable de impacto (p.ej.: rendimiento) del grupo de tratamiento y grupo de contraste para los individuos que se encuentren dentro del soporte común²⁰. Finalmente, para ver la robustez de los resultados, en el caso del *Uno a Uno*, se usa diferentes medidas de distancia (caliper) para hacer el emparejamiento, mientras en el caso del vecino más cercano (*nearest neighbor matching*), se va reduciendo

¹⁹ Para detalles de la especificación probabilística ver Heckmak, Ichimura & Todd (1997).

²⁰ Es el rango común de probabilidades entre el grupo de tratamiento y contraste.

el número de vecinos que se usa. Así mismo, en ambos casos, se hace uso del *bootstrapping* o remuestreo con la finalidad de verificar que los errores estándar estimados sean robustos.

Para este análisis, se considerarán los siguientes aspectos:

- Se van a considerar como grupo de contraste a las instituciones educativas privadas; es decir, se va a comparar a las instituciones educativas que son elegibles para el *bono escuela* 2015 con un grupo de instituciones educativas privadas que sean comparables. Sin embargo, el uso de este grupo de comparación tiene como principal limitación restringir la muestra sólo a contextos urbanos²¹ y por ende sólo podría llevarse a cabo para el grupo A de instituciones educativas que participan del *bono escuela*. Cabe señalar que no se consideran a las instituciones educativas públicas de gestión privada (p.ej.: Fe y Alegría) como parte del grupo de tratamiento dado que tienen un sistema de gestión distinto al de las escuelas públicas debido a que son de gestión privada. Sin embargo, su exclusión no presentaría mayores problemas dado que el número de instituciones educativas de este tipo es reducido a nivel nacional²².
- Emparejar a las instituciones educativas eliminando cualquier diferencia debido a características observables relacionadas con insumos escolares (p.ej.: infraestructura educativa), composición sociodemográfica del alumnado (p.ej.: porcentaje de mujeres) y características del contexto donde se ubican las instituciones educativas (p.ej.: efectos fijos por región). La forma funcional de la ecuación que se va a estimar es:

$$P(T_i=1|X_i) = p_i = F(\alpha_0 + \alpha_1(\text{composición demográfica del alumnado}) + \alpha_2(\text{insumos escolares}) + \alpha_3(\text{variables contextuales}) + \alpha_4(\text{efectos fijos por región})) = F(X_i'\alpha) \quad \text{para } i=1, \dots, n$$

Una vez estimado el puntaje de propensión para cada grupo, se va a proceder a comparar los resultados educativos de aquellas instituciones educativas que se ubiquen en el soporte común. En relación a las variables de resultado que se van a comparar entre ambos grupos de estudio, se tiene:

- El rendimiento promedio y los niveles de desempeño de los estudiantes en las instituciones educativas que participaron del *bono escuela* en el 2015 para ver el efecto de la competencia que genera el *Bono Escuela* en el sistema educativo.
- El diferencial del rendimiento promedio de los estudiantes en las instituciones educativas que participaron del bono escuela entre los años 2015 y 2014²³. Este indicador a diferencia del anterior, permite ver el efecto de la competencia pero con una medida que elimina cualquier sesgo por característica invariante en el tiempo²⁴.

Cabe señalar, como se mencionó anteriormente, que para ver la robustez de los resultados obtenidos, se van a usar dos estrategias: i) uso de diferentes metodologías de emparejamiento como son el *vecino más cercano* y *uno a uno*, y ii) realizar análisis de re-muestreo (*bootstrapping*) con 100 réplicas para cada diferencia estimada.

Continuidad del Efecto: Regresión Discontinua

El modelo de regresión discontinua consiste en comparar resultados entre tratados y no tratados alrededor del punto de corte (percentil 35) de una variable de selección continua (*índice de desempeño escolar*), el cual es tomado como referencia para definir el grupo de tratamiento. Por ejemplo, si una institución educativa obtiene un valor de 0.65 y otra 0.63 en el *índice de desempeño escolar* y el punto de corte es 0.64 o más, la primera

²¹ En áreas rurales la oferta de educación privada es menor o igual al 2% del total de escuelas (Cueto, León y Miranda, 2015).

²² De acuerdo al padrón de instituciones educativas del 2014, el 0.5% de escuelas primarias a nivel nacional son públicas de gestión privada, mientras el porcentaje de escuelas primarias públicas a nivel nacional es del 76%.

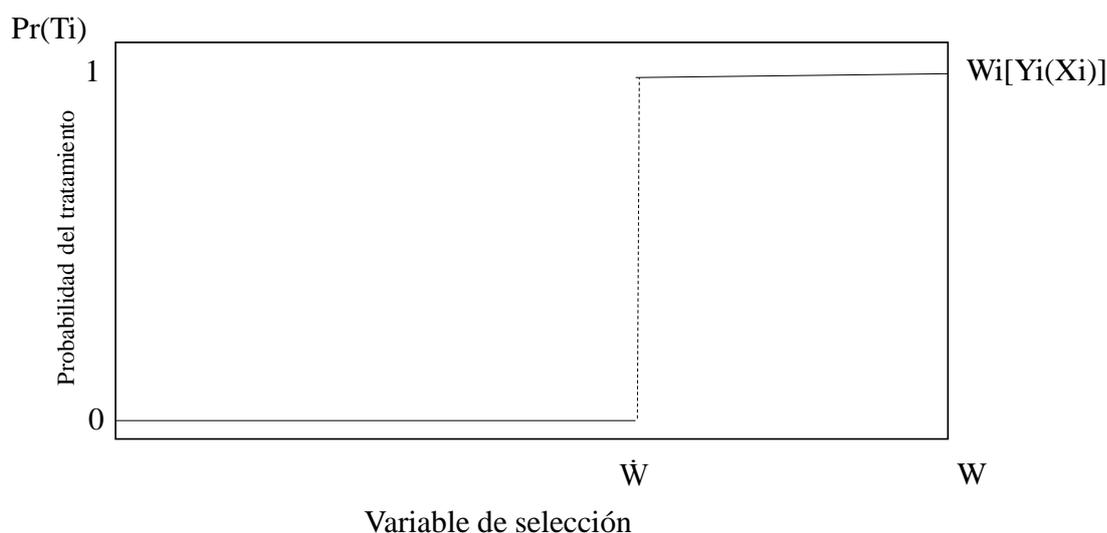
²³ Se utiliza el puntaje del 2014 como línea de base dado que la implementación del bono escuela en el 2014 se realizó a fines de año y no dio el tiempo suficiente a los docentes para influir en los resultados educativos considerados en el índice de desempeño escolar de acuerdo al Decreto Supremo 287 de Octubre del 2014.

²⁴ Este procedimiento sería más riguroso y sería similar al planteado por Blundell y Costa-Dias (2000) donde el estimador del efecto de tratamiento promedio en los tratados (ATT) sería: $\hat{\alpha}_{PSMDD} = \sum_{i \in T} (\Delta Y_i - \sum_{j \in NT} W_{ij} (\Delta Y_j)) w_i$

institución educativa recibirá el *bono escuela* y la segunda no. Sin embargo, dado que la segunda institución educativa está muy cercana al punto de corte fijado puede utilizarse como grupo de contraste de la primera dado que tendrían características similares.

De esta manera, el punto de corte que se fija determina si los que están a un lado o al otro reciben el *bono escuela*, la figura 2 muestra un ejemplo de cómo se fijaría los grupos de tratamiento y contraste, donde individuos a la derecha del punto de corte (\hat{W}) reciben el bono y a la izquierda no.

Figura 2. Probabilidad de participación en el bono



Donde:

T_i : Indicador de tratamiento (recibe el bono)

Y_i : Variable de resultado²⁰¹⁵ (p.ej.: logro educativo de los estudiantes)

X_i : Variables de control que influyen en el rendimiento promedio de las escuelas²⁰¹⁵

W_i : Variable de selección (*índice de desempeño escolar* elaborado por la SPE²⁰¹⁴)

\hat{W} : Punto de corte para la selección (percentil 35²⁰¹⁴)

Sin embargo, un problema que se puede presentar es que haya instituciones educativas por encima del punto de corte que no reciban el *bono escuela* o escuelas por debajo del punto de corte y reciban el *bono escuela*. Dado este escenario, la variable de selección (*índice de desempeño escolar*) no identifica perfectamente al tratamiento por lo que una mejor aproximación para la estimación del efecto sería un diseño de *Regresión Discontinua Difuso*. En el caso específico del *bono escuela*, se han presentado casos de instituciones educativas que a pesar de haber obtenido puntajes por encima del percentil 35 en la variable de selección, no han recibido el incentivo debido a los criterios para la elegibilidad de los docentes, aspecto que origina que no necesariamente toda institución educativa por encima del percentil 35 haya recibido el incentivo o tratamiento. Este hecho origina que si bien la participación en un grupo de estudio u otro dependa de la variable de selección, esta no sea de manera determinística sino aleatoria. Dado este último punto, se va considerar un modelo de *Regresión Discontinua Difuso*.

Cabe señalar, que el estimador de *Regresión Discontinua* parte de cuatro supuestos claves que tienen que ser verificados antes de proceder a realizar la estimación del impacto del programa y estos son: i) debe poder observarse discontinuidad en la probabilidad del tratamiento alrededor del punto de corte, ii) no debe observarse manipulación de la variable de asignación del bono, para lo cual se debe de utilizar test estadísticos como los de McCrary (2008) para ver la distribución de las observaciones, iii) se debe observar discontinuidad alrededor del punto de corte de selección en las variables de resultado (logro, asistencia, entre otras), y iv) debemos

verificar que las características de las instituciones educativas alrededor del punto de corte no presenten alguna discontinuidad²⁵.

Una vez que verificados los supuestos mencionados anteriormente, se procede a realizar la estimación del efecto del programa, para lo cual se tienen dos estrategias (tipos de estimación) que se van emplear.

- **Estrategia 1 (estimación paramétrica):** Se utiliza toda la muestra de instituciones educativas participantes para poder estimar el efecto del programa, para lo cual el principal problema es encontrar la mejor forma funcional que se ajuste a los datos que se tienen. La ventaja principal de esta estrategia es que permite reducir el sesgo en la estimación del impacto de programa dado que utiliza todas las observaciones de la muestra; sin embargo, la desventaja está en que reduce la precisión en la estimación de los efectos dado que usa observaciones alejadas al punto de corte para la decisión de participar en el programa o no. Asimismo, con la finalidad de ver la robustez de los resultados se van a correr los análisis reduciendo el tamaño de la muestra utilizada (los extremos) en un 5%, 10% y 15%.
- **Estrategia 2 (estimación no paramétrica):** a diferencia de la estrategia anterior, no se va asumir ninguna forma funcional para la relación entre la variable de selección y de resultado, pero en su lugar, se realiza los análisis en una parte de la muestra que son parte de la evaluación o en un vecindario local. Para esto, es necesario fijar un segmento de banda (h) a la derecha e izquierda del punto de corte que será donde se hará los análisis de regresión para ver el impacto del programa. Sin embargo, surgen dos problemas al usar esta metodología. En primer lugar está el fijar el tamaño del segmento de banda para los análisis. De acuerdo a la literatura existe dos tipos de forma de fijar este valor: i) *fijado de manera arbitraria, es decir, el investigador fija los valores de acuerdo a los cuales se hacen los análisis*, y ii) *fijado a través de un método estadístico como lo hacen Imbens y Kalyanaraman (2012) o Imbens y Lemieux (2008)*. Para la evaluación de impacto, se van a utilizar los tres métodos con la finalidad de comparar la robustez de los resultados. El segundo problema está en que el usar simple diferencia de medias en el vecindario establecido puede llevar a posibles sesgos debido a la forma funcional de la relación entre la variable de selección y la de resultado, motivo por el cual se van a testear en el vecindario establecido las mismas formas funcionales testeadas con la estrategia anterior.

Así, el modelo de RD que se plantea es:

$$\text{Variable de resultado educativo}_i^{2015} = \alpha_0 + \alpha_1 T_i^{2014} + f(W_i^{2014}) + \alpha_2 \text{Características de las IE}^{2015}_i + \alpha_3 \text{Características del contexto}_i + \theta + \varepsilon_i \quad (1)$$

Donde:

- $V. \text{ resultado}_i^{2015}$: variables relacionadas a resultados escolares
- T_i^{2014} : variable dicotómica que indica que recibió el bono escuela o no
- $f(W_i^{2014})$: la función lineal, cuadrática o cúbica que describe la relación entre la variable de selección y la variable de resultado.
- θ : son efectos fijos por DRE/UGEL.

Las formas funcionales que se van a usar para ver la relación entre la variable de selección y la variable de resultado para ambas estrategias se describen en el siguiente cuadro:

Cuadro 7. Tipos de función de enlace entre la variable de selección y resultado de los estudiantes

Tipo	Forma funcional
Lineal	$\alpha_2 W_i$
Lineal + interacción	$\alpha_2 W_i + \alpha_3 W_i * T_i$
Cuadrática	$\alpha_2 W_i + \alpha_3 W_i^2$
Cuadrática + interacción	$\alpha_2 W_i + \alpha_3 W_i^2 + \alpha_4 W_i * T_i + \alpha_5 W_i * T_i^2$
Cúbica	$\alpha_2 W_i + \alpha_3 W_i^2 + \alpha_4 W_i^3$

²⁵ En el Anexo C se puede apreciar análisis gráficos de los datos.

Cúbica + interacción	$\alpha_2 W_i + \alpha_3 W_i^2 + \alpha_4 W_i^3 + \alpha_5 W_i * T_i + \alpha_6 W_i * T_i^2 + \alpha_6 W_i * T_i^3$
----------------------	---

Por otro lado, dado que vamos a considerar un modelo *RD Difuso* debido a las características del mismo programa, se va utilizar el método de variables instrumentales para poder obtener estimaciones insesgadas de los parámetros estimados en la ecuación 1. Así, la variable que se va usar como instrumento es la misma regla de asignación a pertenecer al grupo de tratamiento o contraste debido a que cumple las condiciones de estar altamente correlacionada con participar o no del programa y no guarda relación con los resultados educativos de las instituciones educativas en el 2015. A continuación se presenta la ecuación de la primera etapa:

$$T_i^{2014} = \beta_0 + \beta_1 f(W_i^{2014}) + \beta_2 \text{Características de las IE}^{2015}_i + \beta_3 \text{Características del contexto}_i + \mu_i \quad (2)$$

Así, sustituyendo (2) en (1) se obtiene:

$$\text{Variable de resultado educativo}_i^{2015} = \alpha_0 + \alpha_1 \check{T}_i^{2014} + f(W_i^{2014}) + \alpha_2 \text{Características de las IE}^{2015}_i + \alpha_3 \text{Características del contexto}_i + \theta + \varepsilon_i \quad (3)$$

Donde, \check{T}_i es la probabilidad predicha de ser parte del grupo de tratamiento que viene de la ecuación 2.

Un aspecto adicional que se debe tomar en cuenta dadas las características del *bono escuela*, es que se cuentan con diferentes rankings (uno por estrato en cada grupo) como se detalló en la descripción del programa. Este aspecto origina que el punto de corte para cada *ranking* sea distinto y que el número de observaciones alrededor del punto de corte no sea suficiente para tener un buen nivel de precisión en los estimados. Motivo por el cual, se plantea usar la metodología usada por Cerqua y Pellegrini (2014) para evaluar la efectividad de un programa de subsidios a empresas en Italia.

En dicho estudio, dado que se contaba con diferentes rankings para las empresas, se optó por utilizar dos metodologías para estimar el efecto global del programa. La primera consistía en ponderar los diferentes efectos del programa en cada ranking y así obtener un efecto global del programa. Para poder combinar los diferentes efectos de cada ranking, usaron dos tipos de ponderadores: i) número de firmas que se usaron para el cálculo del efecto del programa²⁶, ii) la inversa de la varianza del efecto estimado para cada efecto²⁷.

Finalmente, la segunda metodología consiste en usar toda la base en su conjunto y estimar un efecto global, para lo cual es necesario centrar la variable de asignación del incentivo de forma tal que en todos los rankings el valor del punto de corte sea 0²⁸, de esta forma todas las instituciones que tengan un valor por encima de ese punto de corte han recibido el *bono escuela* el 2014. Así, una vez centrados los rankings, se procede a realizar el análisis de regresión discontinua de acuerdo a las estrategias planteadas anteriormente. Para la evaluación de impacto del *bono escuela*, la metodología más adecuada es la segunda dado el poco número de observaciones en varios de los estratos donde se asigna el incentivo y que conllevaría a estimaciones poco precisas del *bono escuela* si se utiliza la primera metodología.

6. Resultados de los análisis estadísticos desarrollados

A continuación los resultados de los principales resultados de los análisis realizados para evaluar el impacto del bono escuela.

6.1 Efecto de la Competencia del Bono Escuela

El primer objetivo de la presente evaluación de impacto es medir el impacto de la competencia que genera la implementación del *bono escuela* en diferentes variables relacionadas a resultados educativos. Se usó como

²⁶ $\alpha^{MRDD} = \frac{\sum_{r=Rankings} N_r \alpha^{RDD}}{N}$, $\sigma^{MRDD} = \sqrt{\sum_{r=Rankings} N_r^2 \sigma_r^2 / N^2}$

²⁷ $\alpha^{MRDD} = (\sum_{r=Rankings} \alpha^{RDD} / \sigma_r^2) / (\sum_{r=Rankings} 1 / \sigma_r^2)$, $\sigma^{MRDD} = \sqrt{\frac{1}{(\sum_{r=Rankings} 1 / \sigma_r^2)}}$

²⁸ Para normalizar los puntajes en función al punto de corte de cada estrato se realiza lo siguiente: $W^N_r = W - \bar{W}_{ranking}$

grupo de comparación a las instituciones educativas no estatales dado que no son elegibles para el *bono escuela*. Como ya se mencionó previamente, el incremento en la oferta educativa no estatal a nivel nacional y predominantemente en zonas urbanas, ha originado que la población escolar a la que atiende la oferta no estatal sea de similares características a la población escolar que atiende las instituciones educativas estatales. Motivo por el cual, los resultados de la presente sección solo considera instituciones educativas urbanas.

Primer, es necesario tener una idea de cuáles son las diferencias sin ajustar en el rendimiento y niveles de desempeño de los estudiantes de acuerdo a cada grupo de estudio. Así, en el siguiente cuadro, se aprecia que sin realizar ningún tipo de ajuste, el rendimiento promedio y los niveles de desempeño en comprensión de lectura son mayores en el grupo de contraste (no estatales) que el grupo de tratamiento (estatales), siendo estas diferencias estadísticamente significativas tanto para el puntaje promedio como para cada nivel de desempeño. Mientras, el rendimiento promedio y niveles de desempeño en Matemática, se aprecia que es a favor del grupo de tratamiento, siendo estas diferencias estadísticamente significativas para el puntaje promedio y cada nivel de desempeño. Similar tendencia es reportada por la Unidad de Medición de la Calidad Educativa, donde se observa que son las instituciones educativas estatales urbanas las que obtienen los mejores resultados en Matemática y las No Estatales en Comprensión lectora (UMC, 2016).

Cuadro 8. Rendimiento y niveles de desempeño (2015) promedio sin ajustar en Matemática y Comprensión de lectura por grupo de estudio

	Tratamiento (Estatal)	Contraste (No estatal)	Diferencia (E-NE)	
Comprensión de lectura				
Puntaje	575.2 (1.208)	599.3 (0.586)	-24.1 (1.342)	***
Satisfactorio	0.41 (0.006)	0.53 (0.003)	-0.11 (0.007)	***
En proceso	0.51 (0.005)	0.44 (0.003)	0.08 (0.006)	***
En inicio	0.07 (0.003)	0.04 (0.001)	0.04 (0.003)	***
Matemática				
Puntaje	576.5 (1.691)	551.5 (0.819)	24.9 (1.879)	***
Satisfactorio	0.28 (0.005)	0.20 (0.003)	0.09 (0.006)	***
En proceso	0.41 (0.004)	0.40 (0.002)	0.01 (0.005)	+
En inicio	0.31 (0.006)	0.40 (0.003)	-0.10 (0.007)	***
Observaciones	2,338	6,517	8,855	

Errores robustos entre paréntesis

*** p<0.001, ** p<0.01, * p<0.05, + p<0.10

Además de las variables de rendimiento y los niveles de desempeño del 2015, se cuenta con el rendimiento promedio y los niveles de desempeño del 2014 para la mayoría de instituciones educativas de la muestra. Así, se procedió a calcular los incrementos en los puntajes de las cohortes de estudiantes de 2do grado de primaria y ver si las escuelas de tratamiento tienen un mayor incremento promedio tanto en el puntaje promedio como los niveles de desempeño en cada área evaluada. El siguiente cuadro presenta los resultados para cada grupo de estudio sin realizar ningún tipo de ajuste. Se puede apreciar, en primer lugar, que tanto en comprensión de lectura como matemática, son las instituciones educativas estatales quienes presentan mejores resultados; sin embargo, los incrementos en el puntaje promedio y niveles de desempeño son pequeño y no en todos los casos

estadísticamente significativos. Se aprecia que en el caso de comprensión de lectura, son las instituciones educativas estatales quienes en promedio tienen un mayor incremento en el puntaje promedio, al igual que en el porcentaje de estudiantes en los niveles de desempeños más bajos de la prueba (en proceso e inicio). En relación a matemáticas, si bien no se aprecian diferencias estadísticamente significativas en el puntaje promedio entre ambos grupos de estudio, si se aprecia que en promedio el incremento de estudiantes en los niveles de desempeño satisfactorio y en proceso es mayor en el grupo de instituciones educativas estatales que en las no estatales.

De esta manera, se aprecia que no solo las instituciones educativas estatales están obteniendo buenos resultados educativos en el 2015, sino también la mejora en los aprendizajes de los estudiantes año a año es mayor en este grupo de instituciones educativas también.

Cuadro 9. Diferencias en el rendimiento y niveles de desempeño promedio (2015-2014) sin ajustar en Matemática y Comprensión de lectura por grupo de estudio

	Tratamiento (Estatal)	Contraste (No estatal)	Diferencia (E-NE)	
Comprensión de lectura				
Puntaje	23.7 (1.111)	19.4 (0.529)	4.3 (1.230)	**
Satisfactorio	0.08 (0.006)	0.07 (0.003)	0.01 (0.007)	
En proceso	-0.02 (0.006)	-0.06 (0.003)	0.04 (0.007)	***
En inicio	-0.05 (0.004)	-0.01 (0.001)	-0.05 (0.004)	***
Matemática				
Puntaje	24.6 (1.775)	21.7 (0.827)	2.9 (1.958)	
Satisfactorio	0.03 (0.006)	0.01 (0.003)	0.03 (0.006)	***
En proceso	0.08 (0.005)	0.11 (0.003)	-0.03 (0.006)	***
En inicio	-0.11 (0.006)	-0.12 (0.004)	0.01 (0.007)	
Observaciones	2,338	5,873	8,211	

Errores robustos entre paréntesis

*** p<0.001, ** p<0.01, * p<0.05, + p<0.10

Los resultados de los cuadros anteriores no toman en consideración posibles sesgos en variables observadas y asociadas a diferentes características de las instituciones educativas, tales como tamaño de la escuela, composición demográfica de los estudiantes, entre otras. Motivo por el cual, se procedió a realizar el análisis de emparejamiento para eliminar cualquier tipo de sesgo debido a estas características de las instituciones educativas. Así, una vez que se elimina cualquier tipo de sesgo en variables observables relacionadas a la composición sociodemográfica de los estudiantes, insumos escolares y características del contexto donde se ubican las instituciones educativas²⁹, se aprecia que las diferencias en comprensión lectora entre ambos grupos dejan de ser estadísticamente significativas; mientras en el caso de matemática, las diferencias se mantienen a favor del grupo de tratamiento (ver cuadro 10)³⁰.

²⁹ En el anexo C se presentan el análisis del balance de las variables antes y después del emparejamiento, así como la distribución de los puntajes de propensión antes y después de realizar el emparejamiento.

³⁰ En el anexo D se presentan los resultados estandarizados.

De esta manera, se aprecia que una vez que se comparan instituciones educativas con similares características observables, la generación de la competencia en el grupo de instituciones educativas estatales tiene un efecto positivo y significativo en el rendimiento en matemática pero no en comprensión de lectura. Por otro lado, al realizar el análisis de robustez de los resultados, disminuyendo la distancia³¹ entre las instituciones que se están comparando o reduciendo el número de vecinos con los cuales se compara, se aprecia que se mantienen los resultados a favor de las instituciones educativas estatales urbanas en matemática, siendo la diferencia de 27 puntos (0.38 desviaciones estándar) y cuentan con un mayor porcentaje de estudiantes que domina los contenidos de manera satisfactoria para su grado o en proceso; y el porcentaje de estudiantes que domina de forma inicial los contenidos es menor en el tratamiento que en el contraste.

Cuadro 10. Diferencias en el rendimiento y niveles de desempeño (2015) ajustados en Matemática y Comprensión de lectura por grupo de estudio (coeficientes no estandarizados)

	Vecino más cercano ^{1/}				Uno a Uno	
	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia
	(E – NE) ^{2/} (n=4)	(E – NE) (n=3)	(E – NE) (n=2)	(E – NE) (caliper: 0.10)	(E – NE) (caliper:0.05)	(E – NE) (caliper:0.01)
Comprensión de lectura						
Puntaje	-5.2 (3.271)	-5.0 (3.292)	-5.1 (3.339)	-5.8 * (2.930)	-7.9 ** (2.717)	-4.2 (3.346)
Satisfactorio	-0.02 (0.016)	-0.02 (0.016)	-0.02 (0.016)	-0.02 (0.014)	-0.03 * (0.013)	-0.01 (0.016)
En proceso	0.02 (0.014)	0.02 (0.014)	0.02 (0.014)	0.01 (0.013)	0.02 (0.012)	0.00 (0.014)
En inicio	0.00 (0.007)	0.00 (0.007)	0.00 (0.007)	0.01 (0.006)	0.01 (0.006)	0.01 (0.007)
Matemática						
Puntaje	32.0 *** (4.605)	32.4 *** (4.712)	31.6 *** (4.851)	27.0 *** (3.935)	24.8 *** (4.121)	27.0 *** (4.703)
Satisfactorio	0.10 *** (0.015)	0.10 *** (0.016)	0.10 *** (0.016)	0.08 *** (0.012)	0.08 *** (0.014)	0.08 *** (0.015)
En proceso	0.02 + (0.013)	0.02 + (0.013)	0.02 + (0.013)	0.02 * (0.009)	0.02 ** (0.008)	0.03 *** (0.010)
En inicio	-0.12 *** (0.017)	-0.12 *** (0.017)	-0.12 *** (0.018)	-0.11 *** (0.014)	-0.10 *** (0.013)	-0.11 *** (0.015)
Observaciones						
Total	3,000	3,000	3,000	1,300	1,277	1,200
Estatal	1,129	1,129	1,129	637	633	601
No Estatal	1,871	1,871	1,871	663	644	599

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se usó como distancia (caliper) entre vecinos 0.01.

2/ E – NE = Estatal – No Estatal.

Errores estándar robustos entre paréntesis, los cuales fueron calculados mediante la metodología de *bootstrapping* con 100 réplicas.

Nota: Las variables usadas en el modelo de regresión *logit* para el emparejamiento son: porcentaje de estudiantes mujeres, porcentaje de estudiantes en extra-edad, porcentaje de estudiantes indígenas, alumnos por sección, porcentaje de docentes mujeres, porcentaje de docentes con título pedagógico, porcentaje de docentes que se han formado en educación primaria, I.E. multigrado, la I.E. cuenta con los tres servicios básicos, la I.E. cuenta con computadoras con conexión a internet, la I.E. tiene sala de computo, la I.E. tiene sala de profesores, la I.E. tiene losa deportiva, la I.E. tiene patio, la I.E. tiene oficinas administrativas, proporción de hogares con agua potable en el distrito, proporción de hogares que se encuentran hacinados en el distrito

³¹ De acuerdo a Austin (2011), la distancia óptima para el emparejamiento es la quinta parte (0.2) de la desviación estándar del puntaje de propensión del emparejamiento. Usando este criterio, la distancia óptima (caliper) para el emparejamiento resulta ser de 0.07. Sin embargo, para los análisis se prueban diferentes valores de distancia como son 0.10, 0.05 y 0.01.

(3 o más personas por habitación), proporción de la población de 18 años o más con educación secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito y efectos fijos por región.

De igual forma, se estimó las diferencias ajustadas por características observables de las instituciones educativas para el incremento de los puntajes y los niveles de desempeño. En el siguiente cuadro, se aprecia que son las escuelas estatales quienes entre el 2014 y 2015 han logrado reducir en mayor medida los estudiantes que dominan de manera inicial las habilidades de comprensión de lectura a diferencia de sus pares no estatales. Así mismo, se aprecia que son las instituciones educativas estatales quienes entre el 2014 y 2015 han logrado incrementar el porcentaje de estudiantes que logra dominar los contenidos para su grado o ubicarse en el nivel satisfactorio. En relación a los demás efectos, no se aprecian resultados muy claro sino más bien mixtos; sin embargo, un aspecto común en todos los análisis son los mayores promedios o mejores resultados que han obtenido las instituciones educativas estatales a comparación de sus pares no estatales.

Cuadro 11. Diferencias en Diferencias en el rendimiento y niveles de desempeño ajustados en Matemática y Comprensión de lectura por grupo de estudio (coeficientes no estandarizados)

	Vecino más cercano				Uno a Uno	
	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia
	(E – NE)* (n=4)	(E – NE) (n=3)	(E – NE) (n=2)	(E – NE) (caliper: 0.10)	(E – NE) (caliper:0.05)	(E – NE) (caliper:0.01)
Comprensión de lectura						
Puntaje	6.1 * (2.962)	5.9 + (3.037)	5.2 + (3.102)	5.8 * (2.696)	2.8 (2.466)	4.7 (3.112)
Satisfactorio	0.03 + (0.016)	0.03 + (0.017)	0.03 (0.017)	0.04 ** (0.014)	0.02 (0.013)	0.02 (0.016)
En proceso	0.00 (0.015)	-0.01 (0.015)	0.00 (0.016)	-0.02 (0.012)	0.00 (0.013)	0.00 (0.015)
En inicio	-0.03 ** (0.009)	-0.02 ** (0.009)	-0.02 ** (0.009)	-0.02 * (0.007)	-0.01 * (0.006)	-0.02 ** (0.007)
Matemática						
Puntaje	10.5 * (4.226)	9.6 * (4.306)	9.3 * (4.437)	7.6 * (3.662)	4.4 (3.576)	5.6 (4.587)
Satisfactorio	0.04 ** (0.014)	0.04 * (0.015)	0.03 * (0.015)	0.03 * (0.013)	0.03 * (0.012)	0.03 (0.016)
En proceso	-0.02 (0.014)	-0.02 (0.014)	-0.02 (0.014)	-0.02 (0.013)	-0.03 * (0.013)	-0.03 * (0.014)
En inicio	-0.02 (0.017)	-0.02 (0.017)	-0.02 (0.017)	-0.01 (0.014)	0.00 (0.014)	0.00 (0.018)
Observaciones						
Total	2,635	2,635	2,635	1,215	1,202	1,111
Estatal	984	984	984	596	595	549
No Estatal	1,651	1,651	1,651	619	607	562

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se usó como distancia (caliper) entre vecinos 0.01.

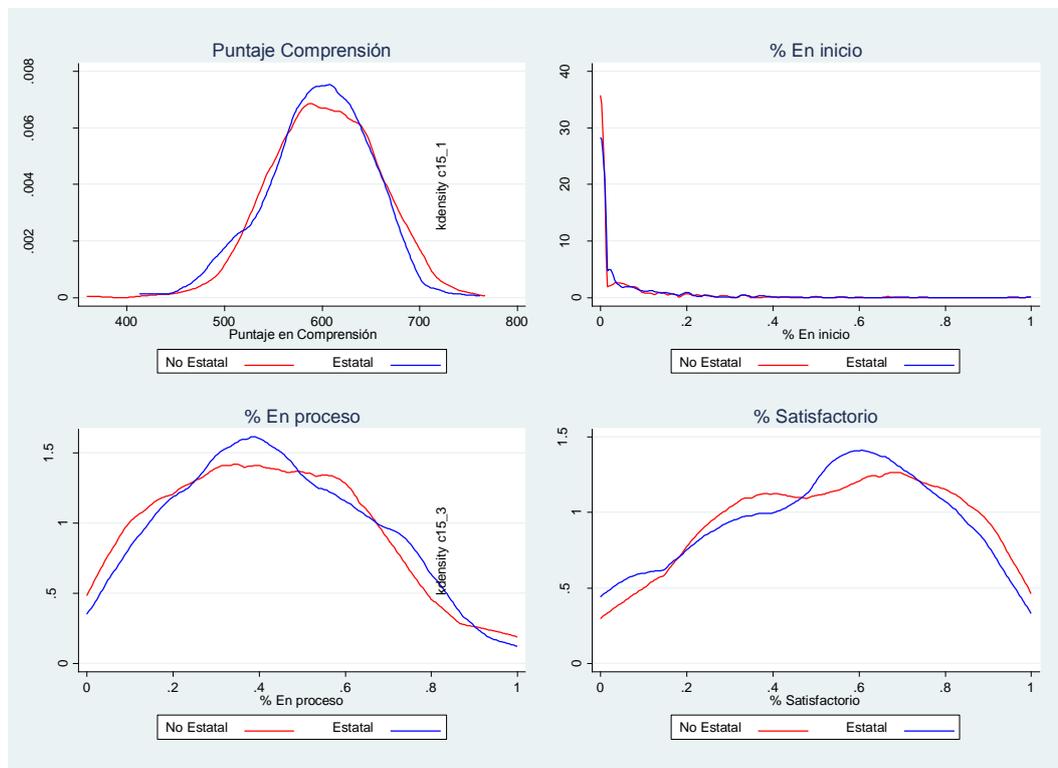
2/ E – NE = Estatal – No Estatal.

Errores estándar robustos entre paréntesis, los cuales fueron calculados mediante la metodología de *bootstrapping* con 100 réplicas.

Nota: Las variables usadas en el modelo de regresión *logit* para el emparejamiento son: porcentaje de estudiantes mujeres, porcentaje de estudiantes en extra-edad, porcentaje de estudiantes indígenas, alumnos por sección, porcentaje de docentes mujeres, porcentaje de docentes con título pedagógico, porcentaje de docentes que se han formado en educación primaria, I.E. multigrado, la I.E. cuenta con los tres servicios básicos, la I.E. cuenta con computadoras con conexión a internet, la I.E. tiene sala de computo, la I.E. tiene sala de profesores, la I.E. tiene losa deportiva, la I.E. tiene patio, la I.E. tiene oficinas administrativas, proporción de hogares con agua potable en el distrito, proporción de hogares que se encuentran hacinados en el distrito (3 o más personas por habitación), proporción de la población de 18 años o más con educación secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito y efectos fijos por región.

Finalmente, se realiza un análisis gráfico de la distribución de los puntajes y niveles de desempeño para cada grupo de estudio en el soporte común³². Se aprecia, en el caso de Comprensión de lectura, que no existen mayores diferencias entre ambos grupos de estudio tanto para el puntaje promedio de las escuelas como para el promedio del porcentaje de niños y niñas en cada nivel de desempeño.

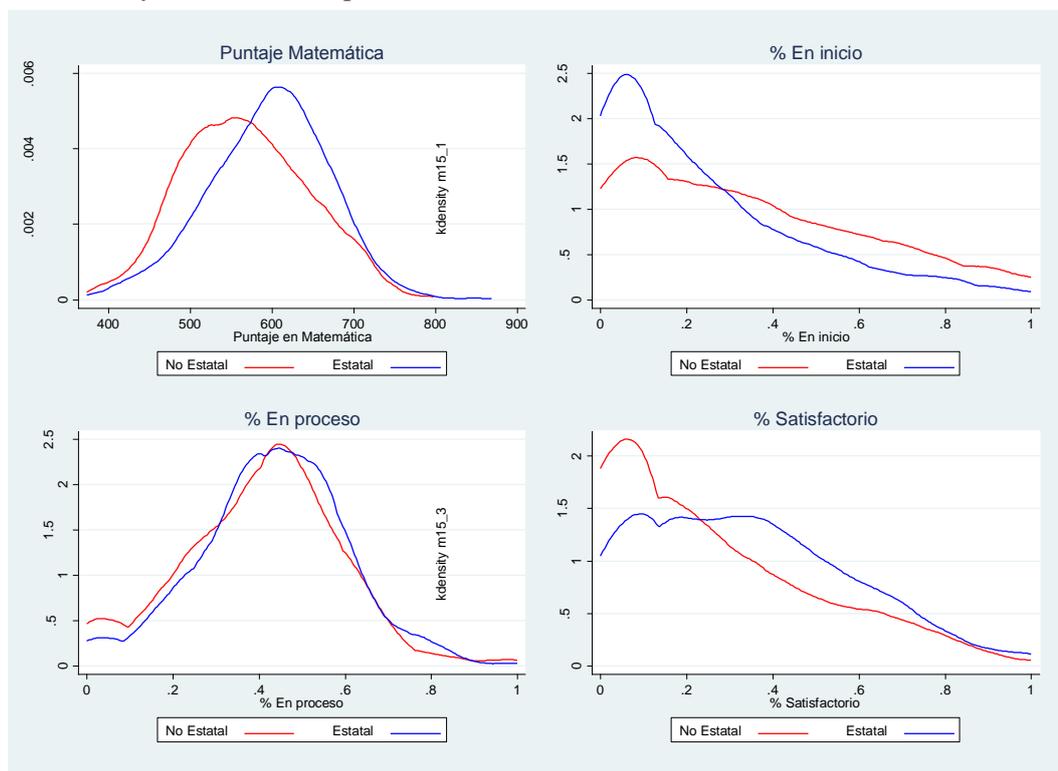
Figura 3. Distribución de los puntajes en Comprensión de lectura y niveles de desempeño (2015) para las escuelas de tratamiento y control en el soporte común



Un escenario completamente distinto se ve en Matemática. En la siguiente figura, al igual que en el análisis de emparejamiento, se aprecia que existen claras diferencias en la distribución de los puntajes y niveles de desempeño, a favor de las escuelas estatales. Tanto el puntaje promedio y el porcentaje de estudiantes en los niveles de “En inicio” y “Satisfactorio” son mayores en las escuelas estatales que sus pares no estatales. Sólo en el caso de la distribución del porcentaje de estudiantes ubicados en el nivel de desempeño “En proceso”, no se observan claras diferencias entre ambos grupos. De esta manera, se puede apreciar que uno de los logros que se tiene con esta estrategia es reducir de forma considerable los estudiantes que no dominan las habilidades básicas en Matemática, y por ende incrementar el puntaje promedio de los estudiantes.

³² Para este ejercicio se tomó el soporte común del emparejamiento Uno a Uno con una distancia de 0.05.

Figura 4. Distribución de los puntajes en Matemática y niveles de desempeño (2015) para las escuelas de tratamiento y control en el soporte común



Por último, se exploró la distribución de los diferenciales tanto del puntaje promedio por escuela como del porcentaje de estudiantes en cada nivel de desempeño por escuela. Las siguientes figuras muestran que no existen mayores diferencias en la distribución de las variaciones del puntaje o el porcentaje de estudiantes por cada nivel de desempeño en Comprensión de lectura o Matemática. Sólo en el caso de comprensión de lectura, se aprecia que existen diferencias en la distribución de las variaciones (2015-2014) para el porcentaje de estudiantes en el nivel de desempeño “En inicio”. Se aprecia que una cantidad considerable de escuelas del grupo no estatal, no muestran variaciones entre el 2014 y 2015, aspecto que se aprecia en la alta concentración de casos alrededor de 0 para este grupo de escuelas.

Figura 5. Distribución de las variaciones en Comprensión de lectura y niveles de desempeño (2015 - 2014) para las escuelas de tratamiento y control en el soporte común

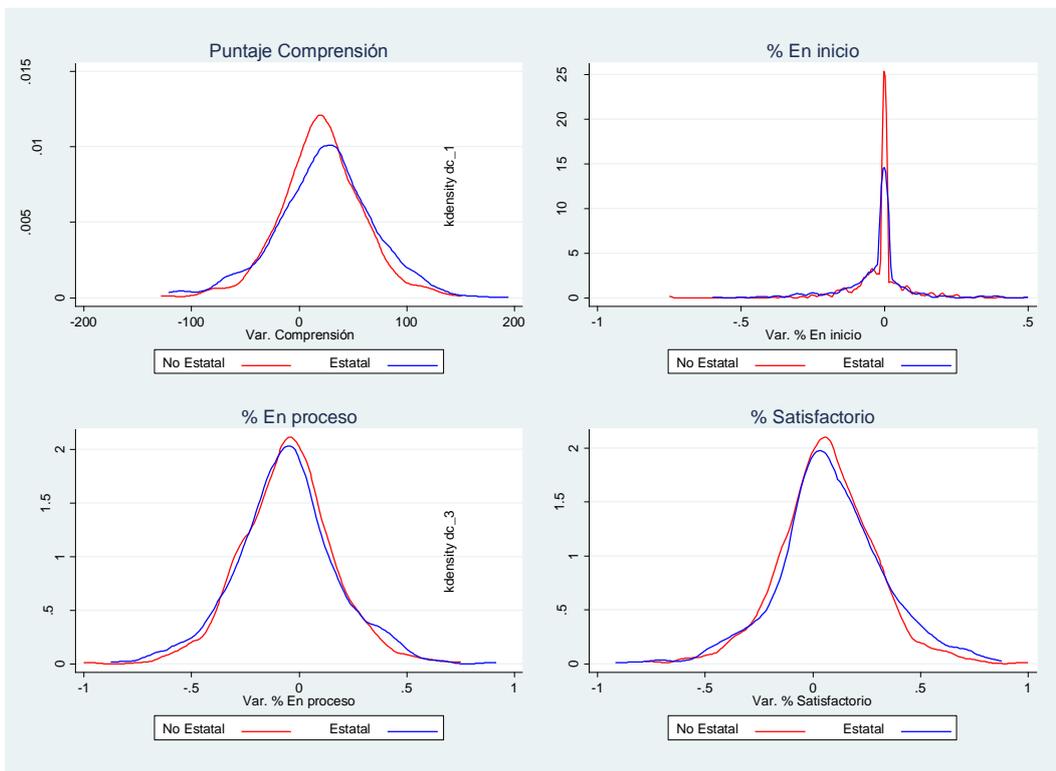
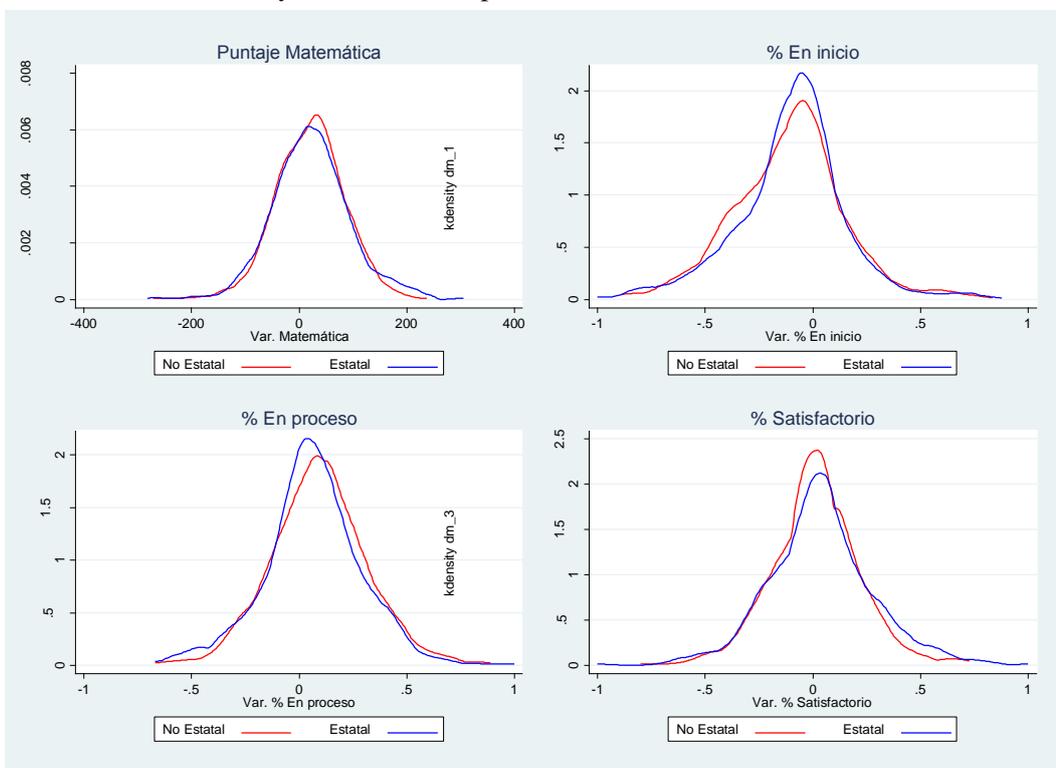


Figura 6. Distribución de las variaciones en Matemática y niveles de desempeño (2015 - 2014) para las escuelas de tratamiento y control en el soporte común



6.2 Efectos de continuidad del Bono Escuela

Un segundo objetivo del presente estudio es ver el efecto en los resultados educativos de las instituciones educativas a un año de entregado el incentivo monetario (continuidad de buenos resultados). Para poder ver este efecto, el mejor diseño es un modelo de regresión discontinua (RD) dado que se cuenta con una regla de asignación del incentivo monetario, es decir, instituciones educativas que se ubican por encima de cierto valor (percentil 65) reciben el incentivo monetario y las otras no. Sin embargo, la entrega del incentivo también depende de si los docentes cumplían con el criterio de elegibilidad, motivo por el cual existen instituciones educativas que si bien se ubicaban por encima del punto de corte pero no recibieron el incentivo monetario dado que sus docentes no cumplían con estos criterios. Este último aspecto origina que en lugar de usar un modelo de RD nítido, se opte por un modelo de RD difuso que permita tomar en consideración este aspecto mediante la estimación de un modelo de Mínimos Cuadrados en 2 Etapas, usando como instrumento la variable de asignación para la entrega del incentivo.

Uno de los primeros aspectos que se exploró dentro del análisis RD es que se cumplieran con las diferentes condiciones para validar el uso de esta metodología de evaluación de impacto. El primer paso está relacionado con la no manipulación de la regla de asignación. Motivo por el cual, se exploró este aspecto de dos formas: i) forma gráfica y ii) mediante el test de Manipulación propuesto por Cattaneo et al (2015) que usa estimaciones no paramétricas para ver la existencia de discontinuidad alrededor del punto de corte de la variable de asignación. Los resultados de ambos análisis muestran que no existe manipulación de la variable de asignación alrededor del punto de corte para la entrega del incentivo tanto de forma gráfica como mediante el test estadístico realizado³³.

En segundo lugar, se debe cumplir que la discontinuidad se dé solo en las variables de resultados (finales o intermedios), y no en otras variables o características de los individuos que se están estudiando (instituciones educativas) o del contexto donde se ubican. Las siguientes figuras presentan la discontinuidad alrededor del punto de corte para las variables de resultados (finales e intermedias) que se van a analizar con el modelo RD difuso. Se aprecia en el caso de las variables de resultado final que existe una discontinuidad alrededor del punto de corte tanto en el rendimiento como en los niveles de desempeño de Matemática y Comprensión de lectura. Sin embargo, la discontinuidad es más marcada en Matemáticas que en Comprensión. En relación a las variables intermedias, se aprecia que existe discontinuidad en la asistencia de los docentes y del director, mientras en el caso de las otras variables, se puede apreciar que no existe una discontinuidad clara alrededor del punto de corte.

En cuanto al análisis de discontinuidad en las variables observables de las instituciones educativas relacionadas con la composición sociodemográfica del alumnado (p.ej.: porcentaje de estudiantes indígenas), personal docente (p.ej.: porcentaje de docentes con título pedagógico), infraestructura educativa (p.ej.: material de los techos) y contexto (p.ej.: porcentaje de hogares hacinados en el distrito) alrededor del punto de corte. El análisis gráfico realizado muestra que no existe una discontinuidad o salto alrededor del punto de corte de la variable de asignación, es decir, esto nos asegura que las observaciones alrededor del punto de corte son similares y que no existirían mayores diferencias. Además del análisis gráfico, se realizaron también pruebas estadísticas de discontinuidad alrededor del punto de corte en las variables observadas seleccionadas para asegurar la validez del diseño³⁴. Los resultados de estos análisis³⁵ indican que para la mayoría de las variables no existe una discontinuidad sistemática alrededor del punto de corte en el ancho de banda seleccionado. Así mismo, se probó diferentes formas funcionales (lineal y cuadrática), encontrando de igual forma la no presencia de discontinuidad en la mayoría de las variables analizadas. Así, tanto el análisis gráfico como los test estadísticos realizados, nos aseguran la validez del uso del RD para la presente evaluación. Por último, se analizó la posible existencia de discontinuidad en diferentes puntos de corte³⁶, no encontrándose diferencias en esos otros puntos de corte.

³³ Para mayores detalles ver el Anexo E.

³⁴ Este análisis consiste en correr el análisis RD usando como variable dependiente a las variables observadas, en la literatura al respecto se le denomina “placebo test”. La idea es que no exista ningún salto o que la variable asociada a pertenecer al tratamiento no sea significativa en el modelo RD planteado (Imbens & Lemieux, 2007; de la Cuesta & Imai, 2016).

³⁵ Ver anexo E para los resultados del análisis gráfico y pruebas de discontinuidad de las variables observables consideradas.

³⁶ Ver Anexo E para los resultados del análisis de continuidad en otros puntos de corte.

Figura 1. Discontinuidad de las variables de resultado final (2015) alrededor del punto de corte de la variable de asignación.

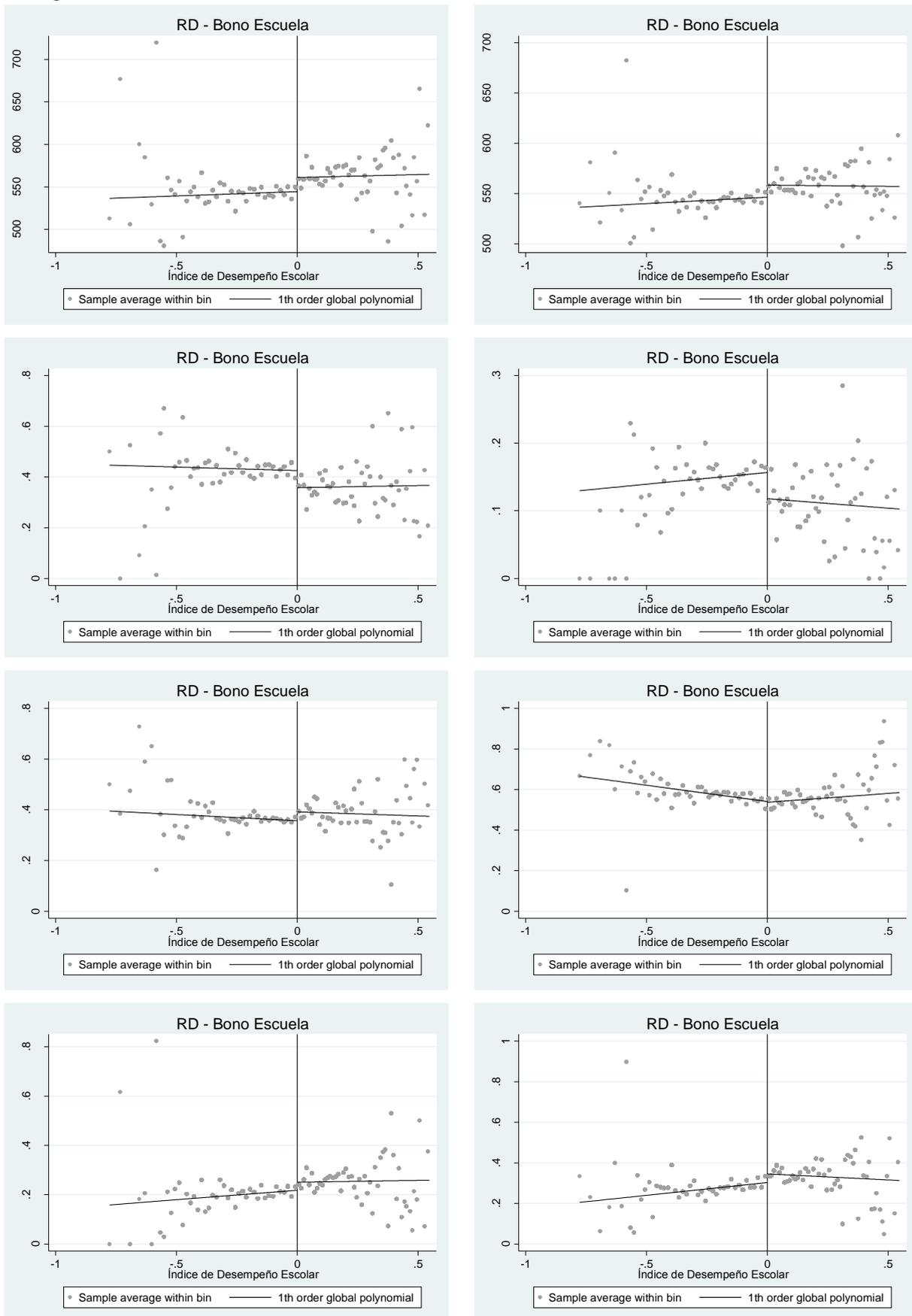
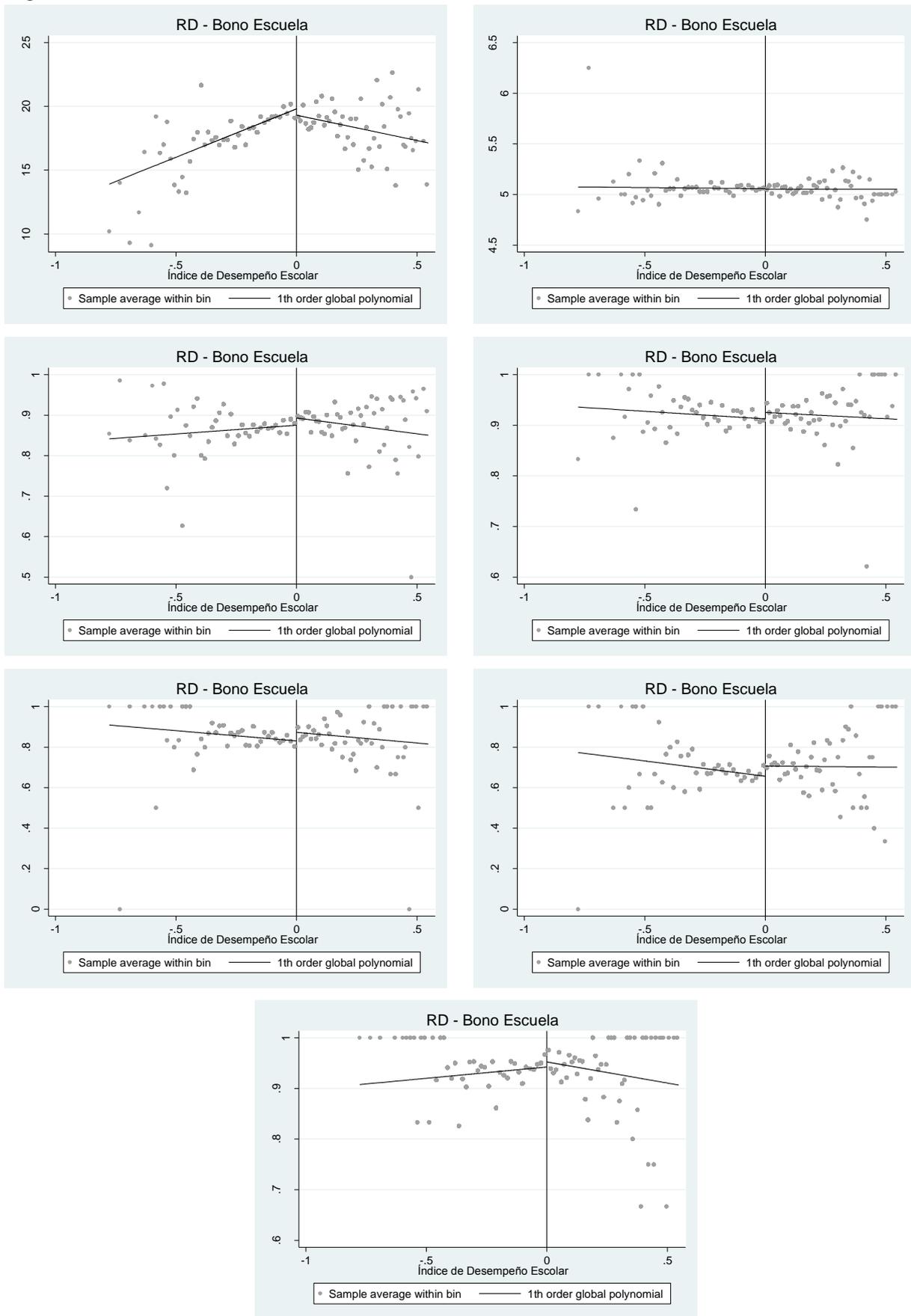


Figura 2. Discontinuidad de las variables intermedias (2015) alrededor del punto de corte de la variable de asignación.



Una vez verificado que se cumplen los diferentes supuestos del diseño de evaluación seleccionado (RD) para ver los efectos del *Bono escuela* a un año de su entrega, se procedió a estimar los efectos del *bono escuela* usando las dos estrategias planteadas anteriormente: i) estimación paramétrica, y ii) estimación no paramétrica. Una de las ventajas de usar ambas estrategias para ver el impacto de programa está en que la regresión discontinua paramétrica cuenta con mayor validez externa dado que usa todos los datos de la muestra de análisis, mientras la regresión discontinua no paramétrica tiene mayor validez interna porque usa los datos dentro de un vecindario local; así, ambas estrategias se complementan y le dan una mayor validez al efecto del programa.

Resultados de la Estimación Paramétrica

Para la estimación paramétrica del efecto del *bono escuela*, se procedió primero a determinar cuál es la forma funcional que mejor ajusta la relación entre la variable de asignación (índice de desempeño) y las variables de resultados, dado que esta es una de las piezas claves para este tipo de estimación. Así, se estimaron relaciones lineales, cuadráticas y cúbicas entre las variables de resultado educativo y la variable de asignación para luego proceder a comparar los índices de información de Akaike y Schwarz. Así, modelos que contaban con un menor valor en ambos índices indican que presenta un mejor ajuste (Benardi & Macías, 2012). En el caso de ambos índices, se pudo apreciar que la forma funcional que mejor ajustaba la relación entre la variable de resultado y la variable de asignación era la relación lineal³⁷, la cual fue usada para todos los análisis paramétricos que se presentan en la presente sección.

Como se mencionó anteriormente, se va emplear un RD Difuso porque existen instituciones educativas que debiendo haber recibido el incentivo no lo recibieron debido a que sus docentes no cumplían con los criterios de elegibilidad. La RD Difusa es similar a un Mínimos Cuadrados en 2 Etapas (MC2E) donde primero se estima la probabilidad de recibir el incentivo, usando como instrumento la variable de asignación para luego en una segunda etapa estimar el efecto del programa usando la variable predicha en la primera etapa. En el caso de las variables cualitativas (o dicotómicas) como la asistencia de los directores, contar con un comité de tutoría o contar con los documentos de gestión, se utiliza un Mínimo Cuadrado en 2 Etapas No Lineal, es decir, se emplea un modelo *probit* en ambas etapas para la estimación del efecto del programa dado que las variables de resultado son cualitativas o dicotómicas.

En primer lugar, se tienen los efectos en las variables relacionadas a resultados finales como son el rendimiento o niveles de desempeño en Matemática y Comprensión de lectura. En relación a Matemática, se aprecia en el siguiente cuadro que existe un efecto positivo y significativo a favor del grupo de instituciones educativas que reciben el incentivo monetario tanto en el puntaje promedio como en el nivel más bajo de desempeño de los estudiantes (en inicio). Por otro lado, los análisis de sensibilidad realizados para ver la robustez de los resultados, mediante la reducción de los extremos de la muestra de acuerdo a la variable de asignación, muestran que el efecto de programa se mantiene incluso eliminando el 10% o 20% de la muestra; sin embargo, cuando se recorta el 30% de la muestra, el efecto del programa se desvanece³⁸. Este último aspecto se puede deber al número de observaciones disponibles cuando se realiza el recorte del 15% (inferior y superior) de la muestra debido a que hace sea difícil poder detectar tamaños de efecto pequeños³⁹ como los que intervenciones similares al *Bono Escuela* obtienen en países en vías de desarrollo (McEwan, 2015).

³⁷ Ver Anexo F para detalles de los análisis adicionales de la regresión discontinua paramétrica realizada.

³⁸ Similar resultado se aprecia cuando se usan cortes simétricos a la izquierda y a la derecha con relación a la variable de asignación. Ver el Anexo F para los resultados de este ejercicio.

³⁹ Ver Anexo F para detalles del tamaño de la muestra a ambos lados del punto de corte.

Cuadro 12. Resultados de la Regresión Discontinua Paramétrica en Matemática

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	11.24 ** (4.324)	-0.05 ** (0.016)	0.03 + (0.015)	0.02 (0.014)
Eliminando 10%	10.44 * (5.248)	-0.05 * (0.020)	0.02 (0.017)	0.02 (0.017)
Eliminando 20%	11.48 + (6.287)	-0.06 * (0.024)	0.03 (0.019)	0.02 (0.020)
Eliminando 30%	8.00 (7.376)	-0.03 (0.028)	0.01 (0.024)	0.02 (0.024)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

En el caso de Comprensión de lectura, a diferencia de Matemáticas, no se encuentran efectos significativos del otorgamiento de incentivos en los niveles de desempeño de los estudiantes. Sin embargo, sí se observa que existe un efecto positivo y significativo en el puntaje promedio, a favor de las instituciones educativas que reciben el incentivo. Sin embargo, este efecto no es consistente conforme se va reduciendo la muestra para el análisis de regresión⁴⁰.

Cuadro 13. Resultados de la Regresión Discontinua Paramétrica en Comprensión de Lectura

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	6.60 * (3.029)	-0.02 + (0.011)	0.00 (0.014)	0.02 (0.014)
Eliminando 10%	5.41 (3.603)	-0.02 (0.012)	0.00 (0.018)	0.01 (0.018)
Eliminando 20%	7.47 + (4.245)	-0.01 (0.016)	-0.01 (0.020)	0.03 (0.020)
Eliminando 30%	5.72 (5.023)	-0.01 (0.019)	-0.01 (0.025)	0.02 (0.024)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Finalmente, se procedió a ver el efecto del programa en variables intermedias, es decir, variables relacionadas con la gestión y/o procesos escolares al interior de las instituciones educativas. En el siguiente cuadro se aprecia que existe un efecto positivo y significativo del otorgamiento de los incentivos en la asistencia de los docentes

⁴⁰ Ver Anexo F para detalles del tamaño de muestra a ambos lados del punto de corte.

y directores. Es decir, aquellas instituciones educativas que han recibido el incentivo en el 2014, cuentan con un mayor porcentaje de docentes presentes y una mayor probabilidad que el director se encuentre en la institución educativa al momento de la visita realizada por los encuestadores del Semáforo Escuela 2015. Sin embargo, no se encontró efecto en variables como asistencia de los estudiantes, carga docente, horas de clase al día, contar con comité de tutoría o contar con los documentos de gestión escolar (Plan Anual de Trabajo y Matriz de Objetivos). Por otro lado, al ver el análisis de sensibilidad de los resultados, reduciendo la muestra por los extremos, se aprecia que estos efectos se mantienen incluso cuando se reduce la muestra en un 30%, mostrando la robustez de los resultados encontrados en cuanto a la asistencia de los docentes y el director.

Cuadro 14. Resultados de la Regresión Discontinua Paramétrica para las variables Intermedias

	Asistencia del Director	Asistencia Docente	Asistencia Estudiantes	Carga docente	Horas de clase	Comité de tutoría	Documentos de gestión
Toda la muestra	0.23 * (0.094)	0.04 (0.025)	0.03 (0.025)	-0.01 (0.026)	0.00 (0.025)	0.07 (0.080)	0.08 (0.114)
Eliminando 10%	0.28 * (0.115)	0.07 * (0.031)	0.06 + (0.031)	-0.02 (0.029)	-0.01 (0.032)	0.10 (0.097)	0.12 (0.167)
Eliminando 20%	0.27 * (0.134)	0.06 + (0.034)	0.03 (0.037)	-0.04 (0.031)	-0.02 (0.036)	0.07 (0.120)	0.04 (0.212)
Eliminando 30%	0.36 * (0.170)	0.09 * (0.041)	0.04 (0.046)	-0.05 (0.043)	-0.02 (0.051)	0.10 (0.157)	0.17 (0.278)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos *probit* para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Resultados de la Estimación no paramétrica

Una segunda estrategia que se utilizó para ver la validez de los resultados, fue estimar los efectos del programa usando estimaciones no paramétricas o regresiones lineales locales en un vecindario determinado dado que cuentan con una mayor validez interna. Uno de los principales aspectos es determinar el ancho de banda (distancia con respecto al punto de corte) adecuado para realizar la estimación local. Cabe señalar que existen diferentes métodos (Lee & Lemiux, 2010) para el cálculo del ancho de banda óptimo, pero para el presente estudio se utilizó el procedimiento planteado por Imbens y Kalyanaraman (2012) que busca minimizar el error cuadrático medio de la diferencia entre el valor predicho y el observado. Usando este método, los anchos de banda óptimos para cada una de las variables de resultados (finales e intermedios) que se van a analizar se presentan en el siguiente cuadro.

Cuadro 15. Anchos de banda óptimos para las variables de resultado de acuerdo al método de Imbens y Kalyanaraman (2012)

Dependiente	Ancho de Banda (h)
Comprensión de Lectura	0.14
% Satisfactorio en Comprensión de Lectura	0.10
% En proceso en Comprensión de Lectura	0.08
% En inicio en Comprensión de Lectura	0.10
Matemática	0.11
% Satisfactorio en Matemática	0.18
% En proceso en Matemática	0.08
% En inicio en Matemática	0.10
Asistencia Estudiantes	0.10
Asistencia Docente	0.13
Carga docente	0.13
Horas de clase al día	0.12

Nota: Para el cálculo del ancho de banda óptimo se utilizó el procedimiento planteado por Imbens y Kalyanaraman (2012). Este algoritmo minimiza el error cuadrático medio de la diferencia entre el valor predicho y el observado.

Se aprecia que existe variación en los anchos de banda para las diferentes variables de resultado consideradas, oscilando entre 0.18 (% suficiente en matemática) y 0.08 (% en proceso en Comprensión de lectura o matemática). De esta manera, con la finalidad de poder ver la sensibilidad de los efectos de acuerdo al ancho de banda seleccionado, se van a estimar los efectos usando los anchos de banda siguientes: 0.25 (n=5,089), 0.20 (n=4,483) y 0.15 (n=3,777). Así mismo, para la estimación del modelo se van a usar el método de Kernel Triangular que permite ponderar las observaciones, dándoles un mayor peso a aquellas observaciones que se ubican cerca del punto de corte y un menor peso a aquellas que se encuentran más alejadas.

En relación a los resultados en Matemática, el siguiente cuadro muestra que existe un efecto positivo y significativo del incentivo monetario en el rendimiento en matemática. En el caso del puntaje promedio, se aprecia que este efecto es robusto y se mantiene significativo incluso al reducir la muestra de análisis de acuerdo a los diferentes anchos de banda considerados. En relación a los niveles de desempeño, se aprecia que este efecto se da en los niveles “en inicio” y “en proceso”. Este resultado estaría indicando que aquellas instituciones educativas que recibieron el incentivo monetario en el 2014, estarían logrando que sus estudiantes consoliden los contenidos iniciales y básicos para su grado de estudios. Sin embargo, al realizar el análisis en el ancho de banda óptimo, se aprecia que los niveles de desempeño, el efecto se desvanece, mientras para el puntaje promedio, el efecto se mantiene al 10% de significancia. Este aspecto se puede deber a que el número de observaciones no es el adecuado para poder detectar tamaños de efecto pequeños como los que tienen este tipo de intervenciones.

Cuadro 16. Resultados de la Regresión Discontinua No Paramétrica en Matemática

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
h=0.25	18.20 ** (6.65)	-0.08 ** (0.02)	0.04 * (0.02)	0.03 + (0.02)
h=0.20	18.76 * (7.54)	-0.08 ** (0.03)	0.04 * (0.02)	0.03 (0.02)
h=0.15	18.88 * (9.05)	-0.07 * (0.03)	0.04 (0.02)	0.03 (0.03)
Óptimo	18.99 + (11.34)	-0.06 (0.04)	0.02 (0.04)	0.03 (0.02)

***p<0.001,**p<0.01,*p<0.05,+p<0.10

Nota: Para la estimación no paramétrica se usa la estimación de Kernel Triangular. Está estimación pondera las observaciones y le da un mayor a aquellas cercanas al punto de corte. No se considera anchos de banda menores a 0.15 debido a que reduce significativamente el número de observaciones del grupo de tratamiento.

En Comprensión de lectura, se aprecia la misma tendencia que en Matemática, las instituciones educativas que recibieron el incentivo monetario en el 2014 tienen un rendimiento promedio mayor que sus pares que no lo recibieron. Este efecto se mantiene significativo incluso en los diferentes anchos de banda considerados, apreciándose de esta manera la robustez de los resultados. En relación a los niveles de desempeño, al igual que en matemática, se aprecia que los efectos más robustos se dan en el nivel “en inicio”. Se tiene que son los estudiantes de escuelas estatales que recibieron el *bono escuela* quienes menores deficiencias presentan de comprender los contenidos mínimos de comprensión de lectura y de ahí el bajo porcentaje en este nivel. Por otro lado, se aprecian efectos al 10% en el nivel “satisfactorio” a favor de las escuelas que recibieron el *bono escuela*, sin embargo, este efecto solo se mantiene para dos de los tres anchos de banda seleccionados y en el ancho de banda óptimo tampoco se mantiene.

Cuadro 17. Resultados de la Regresión Discontinua No Paramétrica en Comprensión de lectura

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
h=0.25	13.35 ** (4.884)	-0.05 ** (0.017)	0.01 (0.020)	0.04 + (0.022)
h=0.20	14.48 ** (5.529)	-0.06 ** (0.019)	0.01 (0.022)	0.05 + (0.025)
h=0.15	14.41 * (6.608)	-0.06 * (0.023)	0.01 (0.026)	0.04 (0.030)
Óptimo ^{1/}	14.29 * (6.777)	-0.06 + (0.032)	0.05 (0.040)	0.03 (0.039)

***p<0.001,**p<0.01,*p<0.05,+p<0.10

1/ El ancho de banda fijado para este análisis ha sido fijado con el método de Imbens y Kalyanaraman (2012)

Nota: Para la estimación no paramétrica se usa la estimación de Kernel Triangular. Está estimación pondera las observaciones y le da un mayor a aquellas cercanas al punto de corte. No se considera anchos de banda menores a 0.15 debido a que reduce significativamente el número de observaciones del grupo de tratamiento.

Finalmente, se exploró los efectos en las variables intermedias al igual que en el caso de la estimación paramétrica. Se puede apreciar que solo hay efecto del *bono escuela* en el caso de la asistencia de los docentes, mientras no efecto se aprecia en el caso de la carga docente, horas de clase al día o asistencia de los estudiantes. En el caso del efecto en la asistencia de los docentes, se aprecia que el efecto es robusto dado que se mantiene significativo para los diferentes anchos de banda utilizados e incluso en el ancho de banda óptimo fijado mediante el método de Imbens y Kalyanaraman (2012).

Cuadro 18. Resultados de la Regresión Discontinua No Paramétrica en Resultados intermedios

	Asistencia Estudiantes	Asistencia Docente	Carga docente	Horas de clase
h=0.25	0.03 + (0.016)	0.03 * (0.014)	-0.94 + (0.557)	-0.01 (0.038)
h=0.20	0.03 + (0.018)	0.03 * (0.016)	-0.98 (0.628)	-0.01 (0.043)
h=0.15	0.03 + (0.021)	0.04 * (0.019)	-0.94 (0.745)	0.00 (0.050)
Óptimo ^{1/}	0.03 (0.028)	0.04 * (0.021)	-0.87 (0.806)	0.00 (0.059)

***p<0.001,**p<0.01,*p<0.05,+p<0.10

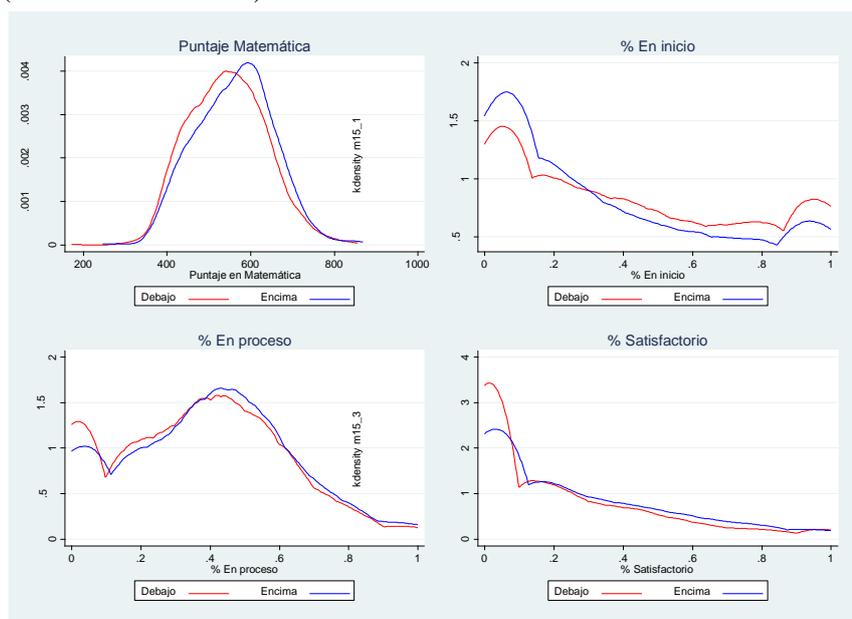
1/ El ancho de banda fijado para este análisis ha sido fijado con el método de Imbens y Kalyanaraman (2012)

Nota: Para la estimación no paramétrica se usa la estimación de Kernel Triangular. Está estimación pondera las observaciones y le da un mayor a aquellas cercanas al punto de corte. No se considera anchos de banda menores a 0.15 debido a que reduce significativamente el número de observaciones del grupo de tratamiento.

De esta manera, tanto los análisis paramétricos y no paramétricos, nos han mostrado que existen efectos del *bono escuela* en los puntajes promedios de los estudiantes tanto en comprensión de lectura como en matemática, mientras en el caso de los niveles de desempeño se aprecia que las diferencias son estadísticamente significativas en el nivel más bajo, siendo robustos estos resultados para los diferentes cortes o anchos de banda seleccionados, mientras para los niveles intermedios o más altos, los efectos no son consistentes o robustos. Estos resultados se pueden apreciar de manera gráfica en las siguientes figuras.

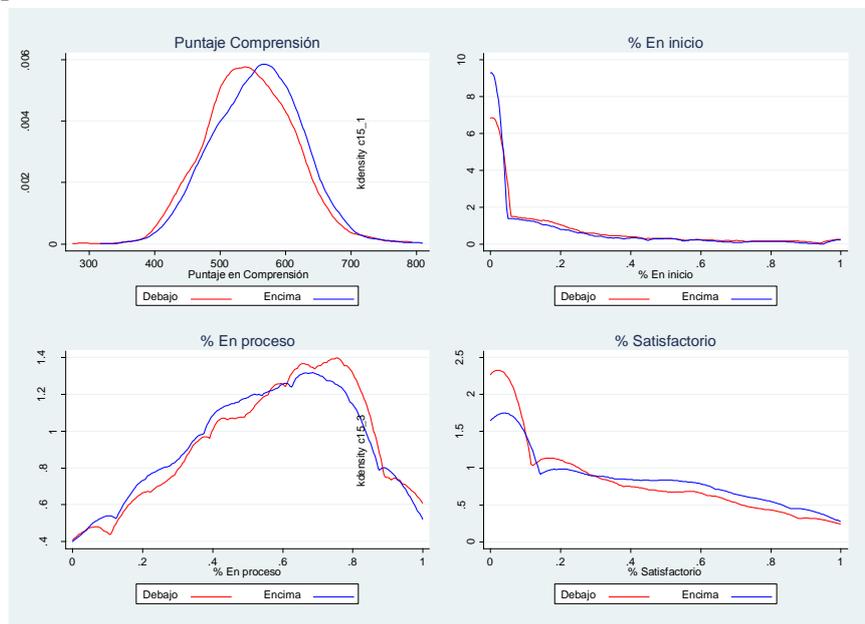
En la figura 8 se aprecia que existe una clara diferencia en la distribución del puntaje promedio y el porcentaje de estudiantes en el nivel de desempeño “En inicio”. En ambos casos, son las escuelas que recibieron el *bono escuela* quienes tienen un mayor promedio en el puntaje de Matemáticas o un menor porcentaje de escuelas con un promedio alto de estudiantes que se ubican en el nivel de desempeño “En inicio”. En el caso de los otros dos niveles, no se observan mayores diferencias en la distribución entre ambos grupos de estudio, tal como se pudo observar en los resultados del análisis de regresión.

Figura 9. Distribución de los puntajes en Matemáticas para las escuelas por encima y debajo del punto de corte (ancho de banda 0.25)



En el caso de comprensión de lectura, no se observan mayores diferencias en la distribución de las variables. Sólo se puede apreciar ligeras diferencias en el puntaje promedio, a favor de las escuelas que recibieron el *bono escuela*.

Figura 10. Distribución de los puntajes en Comprensión de lectura para las escuelas por encima y debajo del punto de corte (ancho de banda 0.25)



Finalmente, no se presentan resultados para el incremento en la diferencia en los puntajes debido a que la muestra para la regresión discontinua ($n=5,858$) es menor al tamaño de muestra para el análisis de emparejamiento ($n=8,855$), motivo por el cual el número de observaciones para los análisis de regresión discontinua no permitirían detectar los tamaños de efectos que la literatura indica se tienen para este tipo de intervenciones.

7. Conclusiones y Recomendaciones de Política

El presente estudio presenta los resultados de la evaluación de impacto del *Bono Escuela* en diferentes resultados educativos de las instituciones educativas estatales. Uno de los principales retos que se afrontó para el desarrollo del presente estudio, fue usar la información administrativa disponible en el Ministerio de Educación, motivo por el cual para la evaluación se hace uso de la información generada por diferentes áreas del Ministerio de Educación: i) Unidad de Medición de Calidad Educativa, y ii) Unidad de Estadística e Informática y Secretaría de Planificación Estratégica. Además, se usó las bases de datos del Sistema Focalizado de Hogares (SISFOH) para el periodo 2012-2013 del Ministerio de Desarrollo e Inclusión Social, con la finalidad de tener información contextual (distrito) que nos permita controlar mejor los efectos del programa. Así, mediante el uso de las bases de datos administrativas del MINEDU y del MIDIS, se pudo evaluar el impacto del *Bono Escuela*, mostrándose de esta manera los beneficios del contar con información administrativa por parte del mismo gobierno.

Los análisis realizados muestran que el *Bono Escuela* tiene un impacto positivo y significativo en los resultados educativos de los estudiantes peruanos. En primer lugar, se pudo apreciar que existe un efecto solo por el hecho de generar la competencia entre las instituciones educativas estatales. Los resultados del análisis de emparejamiento nos muestran que las instituciones educativas estatales urbanas tienen en promedio un mayor rendimiento que sus pares comparables de instituciones educativas no estatales en el rendimiento promedio y niveles de desempeño en matemática, mientras en el caso de comprensión lectora no se encuentran diferencias entre ambos grupos de instituciones educativas. Este resultado estaría mostrando que los docentes y directores de instituciones educativas estatales estarían siendo motivados por el incentivo monetario que representa el bono dado que se estaría premiando el esfuerzo que vienen realizando con sus estudiantes. La motivación de los docentes por el *bono escuela* fue uno de los temas que emergió en la Evaluación de procesos realizada a esta

herramienta (Alcazar, 2016). En dicho estudio se encontró que los docentes percibían al *bono escuela* como un reconocimiento al trabajo que vienen realizando con sus estudiantes, apreciándose así la motivación que genera esta herramienta en los docentes estatales.

En relación a la continuidad de los efectos, se aprecia que existe un impacto positivo y significativo del *bono escuela* en los resultados educativos (estimación paramétrica y no paramétrica). Los análisis anteriores muestran que aquellas instituciones educativas que recibieron el incentivo monetario en el 2014, obtienen un mejor rendimiento promedio en Matemática y Comprensión lectora en el 2015, es decir, sus estudiantes mantienen el buen performance a pesar de que son cohortes diferentes. En relación a cambios en los niveles de desempeño, sólo en el caso de Matemática, se aprecia que los estudiantes de 2do de primaria de las instituciones educativas intervenidas muestran mejoras en su desempeño pero en el nivel en inicio.

En cuanto a las variables intermedias (o de procesos), se apreció en el caso de los efectos a un año de entregado el incentivo monetario que existe un efecto positivo y significativo en la asistencia de los docentes y directores durante las visitas del equipo del Sistema de Monitoreo del 2015. En otras palabras, aquellas instituciones educativas que han recibido el *bono escuela*, son aquellas donde se aprecia una menor inasistencia de los docentes y directores, a diferencia de aquellas instituciones educativas donde no recibieron el bono. Este aspecto resulta importante dado que la mayor asistencia de los docentes permite que se pierdan menos días de clase y los docentes puedan desarrollar más contenidos del currículo, aspecto que favorece el aprendizaje de los estudiantes (Das et al, 2005; Suryadarma et ál., 2006; Cueto et al, 2008).

Por otro lado, la mayor asistencia de los directores tiene un efecto positivo e indirecto en el rendimiento de los estudiantes, el cual se da a través de la norma social que se genera al interior de las instituciones educativas. Estudios como los de Bradley et al. (2007) en Australia y Gaziell (2004) en Israel, muestran que aquellas instituciones educativas donde el faltar no es aceptado o no es parte de la norma social (mayor asistencia de los directores) son aquellos donde se observa una menor inasistencia de los docentes y por ende mejora los aprendizajes de los estudiantes. Así mismo, la mayor asistencia de los directores permite que el Director pueda desarrollar en mayor medida su rol de líder pedagógico y administrativo al interior de las instituciones educativas, ya sea promoviendo espacios para la interacción de los diferentes actores educativos o mediante la gestión de los recursos escolares, aspectos que promueven los aprendizajes de los estudiantes (Wallace Foundation, 2013).

Un aspecto que llama la atención de los resultados obtenidos, es el efecto robusto en matemática y no en comprensión lectora. Estos resultados pueden estar relacionados al nivel de dominio de los contenidos a enseñar por parte de los docentes. En el caso de Matemática, de acuerdo a la UMC (2015), el 31% de los estudiantes de 2do grado se ubica en el nivel en inicio, mientras en comprensión de lectura se ubica en este mismo nivel el 13% de los estudiantes. Este hecho podría estar haciendo que los docentes estén trabajando más contenidos matemáticos dado que lograr mejoras en esta área a los estudiantes es mucho más factible dado que se requiere simplemente que dominen de forma parcial los contenidos para el grado que están siendo evaluados. Mientras en comprensión de lectora, si requeriría que los docentes dominen mejor las estrategias de enseñanza para ubicar a sus estudiantes en un nivel satisfactorio dado el bajo porcentaje de estudiantes en el nivel en inicio.

En cuanto al tamaño del efecto encontrado, si usamos los criterios o *benchmarks* planteados por Cohen (1988), se considerarían los efectos encontrados como pequeños debido a que están por debajo de 0.50 desviaciones estándar⁴¹. Sin embargo, el problema que presenta considerar este criterio para evaluar el effect size radica en que Cohen no toma en consideración el contexto en el cual se desarrollan las intervenciones. Motivo por el cual, Hill et al (2007) plantean que para considerar un tamaño de efecto como pequeño, mediano o grande, se pueden usar tres diferentes criterios: i) cambios esperados de acuerdo al grupo de población estudiado, ii) metas políticamente relevantes (reducción de brechas), iii) tamaños de efecto en estudios similares. Así, para el presente estudio aplica considerar el último criterio y ver qué tan diferentes son los efectos en relación a estudios similares. De esta manera, al comparar los efectos del *bono escuela* sobre el rendimiento con los resultados en estudios similares (McEwan, 2015; Imberman, 2015), se puede apreciar que los tamaños de efecto son similares

⁴¹ De acuerdo a Cohen (1988), un tamaño de efecto pequeño es de 0.20 desviaciones estándar, un efecto mediano es de 0.50 desviaciones estándar y un efecto grande es de 0.80 desviaciones estándar.

tanto para Matemática como para Comprensión de lectura. Por otro lado, en el caso de la asistencia docente, los tamaños de efecto encontrados (0.23 desviaciones estándar) son similares a aquellos encontrados en el meta-análisis desarrollado por Guerrero et al (2013) para países en desarrollo, donde los efectos oscilaban entre 0.15 a 0.30 desviaciones estándar.

Finalmente, en el caso de los análisis de regresiones discontinuas, ambos métodos de estimación nos permiten validar el efecto del *Bono Escuela* en las diferentes variables de resultados educativos; en otras palabras, ambas metodologías son complementarias y nos permiten asegurar la robustez de los efectos encontrados. Sin embargo, los resultados de la estimación paramétrica son más robustos dado que se cuenta con un mayor número de observaciones, se encuentra la mejor forma funcional que ajuste la variable de asignación y resultado educativo, a la vez que se realiza los análisis de sensibilidad cortando la muestra de análisis en los extremos de la distribución, de forma tal de ver la sensibilidad de los resultados.

En relación a las implicancias de política, en primer lugar, resulta primordial desarrollar programas que ayuden a los docentes a mejorar sus prácticas pedagógicas para la enseñanza en matemática y comprensión de lectura, pues –como señalan los resultados de este estudio– al parecer los docentes están teniendo éxito en mejorar el puntaje promedio de los estudiantes y mejorar los niveles de desempeño de los estudiantes que se ubican en los niveles más bajos, pero es necesario que también puedan lograr incrementar los aprendizajes de los estudiantes que se ubican en el nivel suficiente. En ese sentido, resulta promisorio que, actualmente, el Ministerio de Educación esté implementando programas como Acompañamiento Pedagógico en zonas rurales y Soporte pedagógico en zonas urbanas, dado que esto puede facilitar que se pueda capacitar continuamente a los docentes en el manejo y uso de prácticas pedagógicas para la enseñanza matemática y de comprensión lectora.

En segundo lugar, el *bono escuela* está teniendo un efecto positivo y significativo en la asistencia de los docentes y directores, actores claves para que se puedan dar los procesos pedagógicos dentro del aula y la escuela. Motivo por el cual, es necesario el desarrollar estrategias para incrementar la asistencia de los docentes en las instituciones educativas estatales, especialmente en instituciones educativas unidocentes y multigrados donde los docentes son directores y cuentan con una carga administrativa. Así, es necesario desarrollar estrategias para facilitar que los docentes no pierdan días de clase por aspectos administrativos (p.ej.: recojo de materiales educativos de la UGEL) y de esta forma se incremente el número de días de clase efectivo en las instituciones educativas.

En términos de estudios a futuro, como se mencionó anteriormente, el *bono escuela* tiene un efecto positivo y significativo en la asistencia de los docentes y directores, a pesar que un buen porcentaje de docentes enseñan en instituciones educativas rurales en nuestro país. En este sentido, sería bueno poder realizar estudios cualitativos⁴² en diferentes instituciones educativas rurales que recibieron el *bono escuela* con la finalidad de identificar que estrategias están usando los docentes para evitar ausentarse de las instituciones educativas (p.ej.: trámites administrativos)⁴³, así como indagar en los padres de familia de las comunidades si han podido apreciar cambios en la conducta de los docentes en los últimos años para poder identificar no solo a través de la visión de los mismos docentes sino de otros actores educativos, cuales son los aspectos que estarían motivando esta mejora en el desempeño de los docentes.

En cuanto a las limitaciones para la evaluación de impacto, una de las principales ha sido el no contar con información a nivel de los estudiantes dado que realizar el análisis a nivel de instituciones educativas limita de manera significativa la muestra para el estudio. Sin embargo, para futuras evaluaciones de impacto del programa, una vez limpia y consistenciada, se puede usar la información disponible en el Sistema de Información de Apoyo a la Gestión de la Institución Educativa (SIAGIE) dado que se cuenta con información de los estudiantes del 2013 al 2015. De esta manera, se podría ver los efectos a nivel individual de la implementación del *bono escuela*.

⁴² Donde se usen diferentes herramientas cualitativas como las observaciones participantes, entrevistas a profundidad y grupos focales, para de esa manera poder triangular mejor la información que dan los actores que son objeto de estudio.

⁴³ Si bien se cuenta con una evaluación de procesos del *Bono Escuela*, el foco de esta no estaba concentrada en ver los cambios de la práctica diaria de los docentes donde la asistencia a la institución educativa se puede apreciar mejor.

Finalmente, otra limitación es la presencia de diferentes intervenciones en simultáneo en las instituciones educativas dado que el *Bono Escuela* tiene un carácter universal. Este aspecto se pudo observar en los filtros empleados para seleccionar aquellas instituciones educativas que solo contaban con la influencia del *bono escuela*, pasándose de aproximadamente 13 mil instituciones educativas a cerca de 6 mil para propósitos de la evaluación. Este aspecto origina que no se pueda evaluar el impacto de la herramienta para diferentes grupos al interior de la muestra de análisis dado el número de observaciones, llegándose simplemente a poder hacer análisis de factores asociados o estudios correlacionales.

8. Referencias

- Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1), 1-19.
- Ashenfelter, O. & D. Card (1985), "Using the Longitudinal Structure of Earnings to Estimate the Effects of Training Programs", *Review of Economics and Statistics*, 67, 648-660.
- Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics*, 10(2), 150-161.
- Baker, George (1992): Incentive Contracts and Performance Measurement, *Journal of Political Economy* 100, 598-614
- Baker, George (2000): Distortion and Risk in Optimal Incentive Contracts, *American Economic Review*, or working paper version?
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., & Shepard, L. A. (2010). Problems with the Use of Student Test Scores to Evaluate Teachers. EPI Briefing Paper# 278. Economic Policy Institute.
- Balarin, M. (2016) El contexto importa: reflexiones acerca de cómo los contextos y la composición escolar afectan el rendimiento y la experiencia educativa de los estudiantes. Mimeo. GRADE.
- Bernardi, F., & Macías, E. F. (2012). Análisis de datos con Stata (Vol. 45). CIS.
- Benavides, M., León, J., & Etesse, M. (2014). Desigualdades educativas y segregación en el sistema educativo peruano: una mirada comparativa de las pruebas PISA 2000 y 2009 (No. 0015). Grupo de Análisis para el Desarrollo, Lima, Peru.
- Blundell, R. y M. Costa Dias. (2000). "Evaluation Methods for Non-experimental Data". *Fiscal Studies*, Volumen 21(4), pp. 427-468.
- Bloom, B. (1980). *All Our Children Are Learning*. New York: McGrawHill.
- Bradley, S., Green, C. y Leeves, G. (2007). Workers Absence and Shirking: Evidence from Matched Teacher-School Data. *Labour Economics*, 14, 319-334.
- Bryk, A.S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.
- Carroll, J. (1963). A Model of School Learning. *Teachers College Record*, 64:723-733.
- Cerqua, A., & Pellegrini, G. (2014). Do subsidies to private capital boost firms' growth? A multiple regression discontinuity design approach. *Journal of Public Economics*, 109, 114-126.
- Cattaneo, M. D., Jansson, M., & Ma, X. (2015). Simple Local Regression Distribution Estimators with an Application to Manipulation Testing. Working paper, University of Michigan.
- Cohen. J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum
- Cueto, S. (2007). Las evaluaciones nacionales e internacionales de rendimiento escolar en el Perú: Balance y perspectivas. En: *Investigación, políticas y desarrollo en el Perú-GRADE*, 405-455.
- Cueto, S., Torero, M., León, J., & Deustua, J. (2008). Asistencia docente y rendimiento escolar: el caso del Programa META. Documento de Trabajo N° 53, Lima.
- Cueto, S., Guerrero, G., León, J., Zapata, M., & Freire, S. (2013). ¿La cuna marca las oportunidades y el rendimiento educativo? Una mirada al caso peruano. Grupo de Análisis para el Desarrollo (GRADE).
- Cueto, S., León, J. y A. Miranda (2015). Peru: Impact of Socioeconomic Gaps in Educational Outcomes. En: Schwartzman, S. *Education in South America*, London, Bloomsbury.
- Deke, J., & Dragoset, L. (2012). Statistical Power for Regression Discontinuity Designs in Education: Empirical Estimates of Design Effects Relative to Randomized Controlled Trials. Working Paper. Mathematica Policy Research, Inc.
- Das, J.; Dercon, S., Habyarimana, J. y Krishnan, P. (2007). Teacher Shocks and Student Learning: Evidence from Zambia. *The Journal of Human Resources*, 52(4), 820-862.
- de la Cuesta, B., & Imai, K. (2016). Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections. *Annual Review of Political Science*, 19, 375-396.
- Díaz, H. y J. Saavedra (2000). *La carrera del maestro en el Perú. Factores institucionales, incentivos económicos y desempeño*. Lima: GRADE.
- Duflo E, y Hanna R (2005) *Monitoring works: getting teachers to come to school*. Working Paper No.11880. Cambridge: NBER.
- Ehrenberg, R. G., y D. J. Brewer (1995). Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman Revisited. *Economics of Education Review* 14 (1): 121.

- Fama, F., 1980, Agency problems and the theory of the firm, *Journal of Political Economy* 88, 288-307.
- Ferguson, R. F. y H. F. Ladd. (1996). *How and Why Money Matters: An Analysis of Alabama Schools*. In *Holding Schools Accountable: Performance-based Reform in Education*, ed. Helen F. Ladd, 265–98. Washington, DC: Brookings Institution.
- Figlio, D. y L. Kenny (2006). Individual teacher incentives and student performance National Bureau of Economic Research.
- Fraser Institute (2014) *Teacher Incentive Pay that Works: A Global Survey of Programs that Improve Student Achievement*, Barbara Mitchell Centre for Improvement in Education, Vancouver, Canada.
- Gaziel, H. H. (2004). Predictors of Absenteeism among Primary School Teachers. *Social Psychology of Education*, 7, 421–434.
- Glazerman, S., & Seifullah, A. (2012). *An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP) after Four Years. Final Report*. Mathematica Policy Research, Inc.
- Guadalupe, C., León, J., & Cueto, S. (2013). *Charting progress in learning outcomes in Peru using national assessments*. UNESCO. Paper commissioned for the EFA Global Monitoring Report 2013/4, *Teaching and learning: Achieving quality for all*.
- Guerrero, G., & León, J. (2015). Ausentismo docente en Perú: Factores asociados y su efecto en el rendimiento, *Revista Peruana de Investigación Educativa*, v7,31-68pp.
- Guerrero, G., Leon, J., Zapata, M., & Cueto, S. (2013). Getting teachers back to the classroom. A systematic review on what works to improve teacher attendance in developing countries. *Journal of Development Effectiveness*, 5(4), 466-488.
- Glewwe, P., N. Ilias y M. Kremer (2003). *Teacher Incentives*. Working Paper 9671. NBER Working Paper Series.
- Heckman, James; Ichimira Hidehiko; Todd, Petra; (1997). “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program”. *Review of Economic Studies*, 64, pp. 605-654.
- Hendricks, M. D. (2014). Does it pay to pay teachers more? Evidence from Texas. *Journal of Public Economics*, 109, 50-63.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172-177.
- Imbens, Guido, and Karthik Kalyanaraman (2009). "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." IZA Discussion paper No. 3995 (February).
- Imbens, G., and Lemieux, T., (2008), “Regression discontinuity designs,” *Journal of Econometrics*, 142, 615-635.
- Imberman, S. A., & Lovenheim, M. F. (2013). Incentive strength and teacher productivity: evidence from a group-based teacher incentive pay system. *Review of Economics and Statistics*, 97(2), 364-386.
- Imberman, S. A. (2015). How effective are financial incentives for teachers?. *IZA World of Labor*.
- Jacob, Robin, Pei Zhu, Marie-Andree Somers and Howard Bloom (2012). “A Practical Guide to Regression Discontinuity.” MDRC.
- Kelley, C., Heneman, H. & Milanowski, A. (2000). School-based performance award programs, teacher motivation, and school performance: Finding from a study of three programs. CPRE (Consortium for Policy Research in Education) Research Report Series 44.
- Lavy, V. (2004). *Performance Pay and Teachers’ Effort, Productivity, and Grading Ethics*. NBER Working Paper w10622, National Bureau of Economic Research, Cambridge, MA.
- Lee, H., & Munk, T. (2008). Using regression discontinuity design for program evaluation. In *Proceedings of the 2008 Joint Statistical Meeting* (pp. 3-7).
- Levacic, R. (2009). Teacher incentives and performance: An application of principal–agent theory. *Oxford Development Studies*, 37(1), 33-46.
- Lincove, J. A. (2012). *Can Teacher Incentive Pay Improve Student Performance on Standardized Tests?.* Artículo presentado en la Conferencia AEFP.
- Lohr, S.L. (1999). *Sampling: Design and analysis*. Pacific Grove, CA: Brooks/Cole Publishing Company.
- Loyalka, P., Sylvia, S., Liu, C., Chu, J., & Shi, Y. (2015). *Pay By Design: Teacher Performance Pay Design and the Distribution of Student Achievement*. In 2015 Fall Conference: The Golden Age of Evidence-Based Policy. Appam.
- Lunt, M. (2014): *Propensity Analysis in Stata Revision: 1.1*. Documento disponible en: http://personalpages.manchester.ac.uk/staff/mark.lunt/propensity_guide.pdf

- McEwan, P. J. (2014). Improving Learning in Primary Schools of Developing Countries A Meta-Analysis of Randomized Experiments. *Review of Educational Research*, 0034654314553127.
- McCrary, J., (2008), "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698-714.
- Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics*, 13(2), 151-161.
- Miranda, L. (2008). Factores asociados al rendimiento escolar y sus implicancias para la política educativa del Perú. Análisis de programas, procesos y resultados educativos en el Perú. *Contribuciones empíricas para el debate*, 11-39.
- Mizala, A. y Romaguera, P. (2005). Rendimiento escolar y premios por desempeño. La experiencia latinoamericana y el SNED en Chile. En Cueto, S. (ed.). *Uso e Impacto de la Información Educativa en América Latina*. Santiago de Chile: PREAL (Pp. 23-60).
- Muralidharan K y Sundararaman V (2009) Teacher performance pay: experimental evidence from India. Working Paper No. 15323. Cambridge: NBER.
- Navarro, J. C., ed. (2002). *¿Quiénes son los maestros? Carreras e incentivos en América Latina*. Washington, DC: The Inter-American Development.
- OECD (2011), *Establishing a Framework for Evaluation and Teacher Incentives: Considerations for Mexico*, OECD Publishing. <http://dx.doi.org/10.1787/9789264094406-en>
- OECD (2014a), *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014)*, PISA, OECD Publishing.
- Prendergast, C. (2000). The tenuous tradeoff between risk and incentives (No. w7815). National bureau of economic research.
- Pomeranz, D. (2011). *Métodos de evaluación*. Harvard Business School.
- Rivkin, S., E.Hanushek y J. Kain (1998) Teachers, schools and Academic Achievement. NBER Working Paper N°. w6691, National Bureau of Economic Research, Cambridge, MA.
- Rockoff, J. 2004. The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review* 94 (2): 247–57
- Rosenbaum, P. y D. Rubin; (1983). "The Central Role of the Propensity Score in the Observational Studies for Causal Effects" en *Biometrika*, 70: 41-55.
- Shadish, W., Cook, T. & Campbell, D. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. New York, NY: Houghton Mifflin Company.
- Springer, Matthew G., Dale Ballou, Laura Hamilton, Vi-Nhuan Le, J. R. Lockwood, Daniel F. McCaffrey, Matthew Pepper, & Brian M. Stecher (2010). "Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching (POINT)." Society for Research on Educational Effectiveness.
- Stuart, E. A., & Rubin, D. B. (2008). Best practices in quasi-experimental designs. En: *Best practices in quantitative methods*, Capítulo 11: 155-176.
- Solmon, L. C., & Podgursky, M. (2000). *The Pros and Cons of Performance-Based Compensation*. Milken Family Foundation, Santa Monica, CA.
- Suryadarma, D., Suryahadi, A., Sumarto, S. y Rogers, F. H. (2006). Improving Student Performance in Public Primary Schools in Developing Countries: Evidence from Indonesia. *Education Economics*, 14(4), 401-429.
- UMC (2004). Factores asociados al rendimiento estudiantil. Resultados de Evaluación Nacional 2001. En: Documento de Trabajo UMC No. 9. Lima: Ministerio de Educación del Perú.
- UMC (2006). Evaluación nacional del rendimiento estudiantil 2004: ¿cómo disminuir la inequidad del sistema educativo peruano y mejorar el rendimiento de sus estudiantes? Lima: Ministerio de Educación del Perú.
- UMC (2013). Estudio de educación inicial: un acercamiento a los aprendizajes de las niñas y los niños de cinco años de edad. En: Informe de resultados. Lima: Ministerio de Educación del Perú.
- UMC (2015). Resultados de la Evaluación Censal de Estudiantes 2014. Presentación General de los Resultados de la ECE 2014. Disponible en: <http://umc.minedu.gob.pe/evaluacion-censal-de-estudiantes-2014-ece-2014/>.
- UMC (2016). Resultados de la Evaluación Censal de Estudiantes 2014. Presentación General de los Resultados de la ECE 2015. Disponible en: <http://umc.minedu.gob.pe/evaluacion-censal-de-estudiantes-ece-2015/>
- Vegas E. y I. Umansky (2005) Mejorar la enseñanza y el aprendizaje por medio de incentivos ¿Qué lecciones nos entregan las reformas educativas de América latina? Washington, DC: Banco Mundial.

- Vegas, Emiliana (2006). Incentivos docentes y sus efectos en el aprendizaje del alumnado en Latinoamérica. *Revista de Educación*, 340, 213-241.
- Vivanco, E. (2013). Evaluating the effects of school and teacher incentives: Quasi-Experimental Evidence from Mexico. Manuscrito inédito.
- Wallace Foundation (2013). *The School Principal as Leader: Guiding Schools to Better Teaching and Learning*. Working Paper, New York, NY.
- Yang G, Stemkowsky S, Saunders W. (2007) A review of propensity score application in healthcare outcome and epidemiology. Premier Inc. Working Paper PR02.

Anexo A. Programas de Incentivos en países en vías de desarrollo

N°	Autor	Título	Año de publicación	País	Tipo de incentivo	Mecanismo	Objetivo de la Evaluación	Muestra	Metodología	Tipo de diseño	Resultados
1	Lavy	Performance pay and teacher's effort, productivity and grading ethics	2004	Israel	Incentivo financiero escuela	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Medir el efecto del programa de incentivos en el desempeño de los docentes y rendimiento de los estudiantes.	La muestra está compuesta por alumnos de secundaria entre los grados 10 y 12, en un total de 97 escuelas. De las 97 escuelas se eligieron 49 para el grupo tratamiento. La muestra de estudiantes está compuesta por 2,405 en el grupo de tratamiento y 1,773 para el grupo de control en el año 2000. Para el 2001, la muestra estuvo compuesta por 2,350 estudiantes para el grupo de tratamiento y 1,678 para el grupo de control en el año 2001.	* Estimaciones de diferencias-en-diferencias. * Modelo de Emparejamiento. * Regresión lineal multivariada (análisis de robustez). * Regresión discontinua (análisis de robustez).	No Experimental	* Se muestra que los incentivos docentes incrementan los logros de los estudiantes. * Los resultados indican que no existen tergiversación de notas por parte de los profesores, dado el efecto del programa sobre las tasas de matrícula.
2	Duflo y Hanna	Monitoring works: getting teachers to come to school.	2005	India	Incentivo financiero individual	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Medir el efecto del programa de incentivos en el ausentismo docente y el	Se considera como muestra a 60 escuelas en la India, elegidas aleatoriamente de 120 escuelas del programa. Se tiene	* Mínimos cuadrados en dos etapas * Regresiones lineales multivariadas	Experimental	* Se encuentran efectos del sistema de incentivos en la tasa de ausentismo docente, siendo la tasa promedio en las escuelas control de

N°	Autor	Título	Año de publicación	País	Tipo de incentivo	Mecanismo	Objetivo de la Evaluación	Muestra	Metodología	Tipo de diseño	Resultados
							rendimiento de los estudiantes.	una muestra de estudiantes de primer grado, considerando 29,489 niños en tratamiento y 26,695 en control. Y en relación al análisis de asistencia de los profesores se utiliza una muestra de 1575 profesores en tratamiento y 1496 de control.			42% y 21% en escuelas de tratamiento. * Se tiene un impacto positivo y significativo sobre el rendimiento de los estudiantes (0.17 desviaciones estándar).
3	McEwan y Santibañez	Teacher incentives and students achievement: Evidende from a Mexican reform	2005	México	Incentivo financiero individual	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Analizar si los docentes incluidos en el programa mejoran los puntajes de los estudiantes.	Profesores de los grados de 3 - 6 de educación primaria. Con 76,567 profesores en 27,213 escuelas (los profesores pueden pertenecer a más de una de escuela) del Programa de la Carrera Magisterial de México.	* Regresiones lineales multivariadas * Regresión discontinua	No experimental	* Se halla evidencia de un efecto positivo y significativo del incentivo a los docentes (0.05 - 0.08 desviaciones estándar).
4	Mendoza del Solar	Efecto del Incentivo Docente sobre el rendimiento de Estudiantes de Escuelas Rurales	2008	Perú	Incentivo individual y colectivo	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Analizar el efecto de los incentivos monetarios para docentes de zonas rurales sobre el	La muestra consiste en 214 alumnos, de los cuales 169 pertenecen al grupo de tratamiento A (Cuzco) y 45 al grupo de tratamiento B (Ayacucho). Asimismo, se cuenta	* Análisis de Varianzas de dos vías.	No experimental	* No se encontraron efectos del programa de incentivos a los docentes en el rendimiento de los estudiantes.

N°	Autor	Título	Año de publicación	País	Tipo de incentivo	Mecanismo	Objetivo de la Evaluación	Muestra	Metodología	Tipo de diseño	Resultados
							rendimiento en comprensión de textos de los estudiantes de tercer grado de primaria.	con 18 docentes pertenecientes al grupo de tratamiento A y 9 al grupo de tratamiento B.			
5	Cueto, Torero, León y Deutsua	Asistencia docente y rendimiento escolar: el caso del Programa META	2008	Perú	Incentivo individual y colectivo	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Medir el efecto del programa de incentivos en el ausentismo docente y el rendimiento de los estudiantes.	La muestra considera 178 escuelas y 354 docentes de los grados 4to, 5to y 6to de primaria.	* Modelo de Emparejamiento a dos niveles (escuela y estudiantes)	Cuasi experimental	* Los resultados muestran que el programa tuvo un impacto positivo en la asistencia de los docentes (17 días más de asistencia). * Efectos mixtos del programa de incentivos en el rendimiento de los estudiantes. En matemática existe un efecto a favor del grupo de tratamiento para 5to grado, mientras en Comprensión Lectora el efecto se aprecia en 4to y 6to grado.
6	Rau y Contreras	Tournaments, gift exchanges, and the effect	2009	Chile	Incentivo individual y colectivo	Pago de bono para profesores que tienen	Medir el efecto del programa de incentivos a	Para los análisis se hacen uso de las bases de datos de las pruebas SIMCE para	* Modelo de Emparejamiento * Regresión discontinua	Cuasi experimental	* Existe un efecto positivo y significativo de participar en el

N°	Autor	Título	Año de publicación	País	Tipo de incentivo	Mecanismo	Objetivo de la Evaluación	Muestra	Metodología	Tipo de diseño	Resultados
		of monetary incentives for teachers: the case of Chile				estudiantes con buenas calificaciones	docentes en el rendimiento de los estudiantes.	el periodo 1989-2006. La información del rendimiento se utiliza a nivel de escuela dado que a este nivel se recibe el bono.			esquema de incentivos mediante el análisis de diferencias con emparejamiento (0.19 - 0.36 desviaciones estándar). * No existe un efecto de haber ganado el incentivo a un año después de haberse implementado.
7	Muralidharan y Sundararaman	Teacher performance pay: experimental evidence from India	2009	India	Incentivo financiero individual y escuela	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Estudiar el efecto individual y grupal de los incentivos sobre el rendimiento de los estudiantes.	Se tiene una muestra de 300 escuelas in Andhra Pradesh (India). De estas 100 forman parte del control y 200 son parte de tratamiento (100 escuelas para el bonus grupal y las otras 100 para el bonus individual). Además, se tiene una muestra de 42,145 niños en el primer año. Y 37,617 en el segundo año del programa.	* Regresión lineal multivariada.	Experimental	* Al final del segundo año del programa, los estudiantes que participaron en los incentivos escuela obtuvieron mejores rendimientos que las escuelas control; obteniendo 0.27 y 0.17 desviaciones estándar en las pruebas de matemática y lenguaje, respectivamente. * Tanto las escuelas con incentivos individuales como grupales tuvieron

N°	Autor	Título	Año de publicación	País	Tipo de incentivo	Mecanismo	Objetivo de la Evaluación	Muestra	Metodología	Tipo de diseño	Resultados
											un buen desarrollo en el primer año del programa, pero al final del segundo año, las escuelas con incentivos individuales superaron a las escuelas con incentivos grupales.
8	Glewwe, Ilias y Kremer	Teacher incentives.	2010	Kenia	Incentivo financiero escuela	Pago de bono para profesores que tienen estudiantes con buenas calificaciones	Medir el efecto del programa de incentivos a docentes en el rendimiento de los estudiantes.	Se usa los datos de un programa dirigido por el International Child Support (ICS) en Busia y Teso (distritos del oeste de Kenya). Se seleccionaron 50 escuelas para el programa, de un grupo de 100 escuelas entre los años 1996 al 2000. Se considera una muestra de 24,716 (línea de base) y 15,893 (post test) estudiantes; y 466 profesores (línea de base) y 320 (en programa) entre los grados de 4to a 8vo.	* Mínimos Cuadrados Generalizados (GLS). * Diferencias-en-diferencias	Experimental	* Los alumnos de las escuelas que siguen el programa de incentivos eran más propensos a que les tomen los exámenes y tuvieron resultados más altos en el corto plazo. * La asistencia de los profesores no mejoró, las tareas no aumentaron, y las prácticas pedagógicas no cambiaron.

N°	Autor	Título	Año de publicación	País	Tipo de incentivo	Mecanismo	Objetivo de la Evaluación	Muestra	Metodología	Tipo de diseño	Resultados
9	Vivanco	Evaluating the effects of school and teacher incentives: quasi experimental evidence from Mexico	2013	México	Incentivo financiero escuela	i) Pago de bono para profesores que tienen estudiantes con buenas calificaciones y ii) Pago de bono para los profesores que muestran una mejora en las calificaciones.	Evaluar el efecto del programa mexicano (PECD) que brinda bonos salariales a profesores de escuelas públicas de primaria y secundaria.	Datos de la primera ronda del Programa del Estímulo para la Calidad Docente (PECD) 2010 y la Evaluación Nacional de Logro académico en centros escolares (ENLACE) 2011. Se considera una muestra de 123,067 escuelas.	*Regresión Discontinua	Cuasi-experimental	*No existe evidencia de algún efecto claro del programa de incentivos sobre el rendimiento de los estudiantes; sin embargo, se aprecia un efecto negativo y significativo (0.25 desviaciones estándar) para las escuelas primarias indígenas.

Anexo B. Tamaño de Muestra para los análisis a nivel de estudiantes

En caso se requiera ver los efectos del programa a nivel de estudiantes, se tiene que tomar en consideración el efecto de agrupación de las observaciones (estudiantes agrupados en instituciones educativas); motivo por el cual, la fórmula para determinar el tamaño de muestra necesario debe ser modificado y debe incluir el tamaño aproximado de cada agrupación y el coeficiente de correlación entre grupos o ICC por sus siglas en inglés (Lohr 1999), siendo la formula la siguiente:

$$n = \frac{2\sigma^2(Z_\beta + Z_{\alpha/2})^2(1 + (m - 1)\rho)}{ES^2}$$

Donde:

n	:	Tamaño de muestra de cada grupo de estudio
σ^2	:	Desviación estándar de la variable de resultado
Z_β	:	Poder estadístico
$Z_{\alpha/2}$:	Nivel de significancia
ES	:	Es el tamaño del efecto de la intervención
m	:	Es el número de estudiantes por institución educativa
ρ	:	Es el coeficiente de correlación entre grupos (ICC)

A continuación, se presentan los tamaños de muestra necesarios para un diseño RCT y de RD, para los cálculos se asume un nivel de significancia del 95%, un número promedio de 21⁴⁴ estudiantes por institución educativa y un coeficiente de correlación entre grupos de 0.44⁴⁵

Cuadro 1. Tamaño de muestra por grupo para un diseño RCT y RD

Poder	RCT					RD con efecto de diseño de 3				
	ES: 0.10	ES: 0.15	ES: 0.20	ES: 0.25	ES: 0.30	ES: 0.10	ES: 0.15	ES: 0.20	ES: 0.25	ES: 0.30
70%	12,094	5,375	3,023	1,935	1,344	36,281	16,125	9,070	5,805	4,031
80%	15,366	6,830	3,842	2,459	1,707	46,099	20,489	11,525	7,376	5,122
90%	20,575	9,145	5,144	3,292	2,286	61,726	27,434	15,431	9,876	6,858
95%	25,402	11,290	6,350	4,064	2,822	76,205	33,869	19,051	12,193	8,467

Finalmente, dado que el tamaño de muestra es igual al número de grupos o cluster por el número de estudiantes promedio ($n=J*m$), el siguiente cuadro muestra el número de instituciones educativas para cada tamaño de muestra.

Cuadro 2. Número de Clusters o instituciones educativas para el diseño RCT y RD

Poder	RCT					RD con efecto de diseño de 3				
	ES: 0.10	ES: 0.15	ES: 0.20	ES: 0.25	ES: 0.30	ES: 0.10	ES: 0.15	ES: 0.20	ES: 0.25	ES: 0.30
70%	576	256	144	92	64	1,728	768	432	276	192
80%	732	325	183	117	81	2,195	976	549	351	244
90%	980	435	245	157	109	2,939	1,306	735	470	327
95%	1,210	538	302	194	134	3,629	1,613	907	581	403

⁴⁴ Para el número de estudiantes promedio por institución educativa, se usó el promedio de estudiantes evaluados en la Evaluación Censal de Estudiantes del 2014.

⁴⁵ Para determinar el ICC se usó el promedio del ICC en Comprensión Lectora (0.44) y Matemática (0.43) estimados a partir de la Evaluación Censal de Estudiantes del 2014.

Anexo C. Análisis de verificación del modelo de Emparejamiento

Cuadro 1. Balance de las variables antes y después del emparejamiento

		Promedio		% sesgo	% reducción del sesgo	T-test	
		Estatal	No Estatal			t	p>t
Proporción de estudiantes en extra- edad (pext_e)	Sin Emparejar	0.08	0.01	129.4		63.23	0.00
	Emparejado	0.04	0.05	-9.8	92.4	-1.32	0.19
Proporción de estudiantes mujeres (pmuj_e)	Sin Emparejar	0.49	0.47	16.4		6.51	0.00
	Emparejado	0.49	0.50	-4.9	69.9	-0.59	0.56
Proporción de estudiantes indígena (pind_e)	Sin Emparejar	0.10	0.00	47.5		26.54	0.00
	Emparejado	0.02	0.03	-1.4	97	-0.34	0.73
Alumnos por sección (alum_seccion_prim)	Sin Emparejar	17.63	15.87	23.7		9.47	0.00
	Emparejado	17.89	18.10	-2.8	88.2	-0.42	0.68
Proporción de docentes que son mujeres (pmuj_d)	Sin Emparejar	0.61	0.77	-76		-32.61	0.00
	Emparejado	0.71	0.69	9.3	87.8	1.48	0.14
Proporción de docentes con título pedagógico (propdoc_pedit)	Sin Emparejar	0.96	0.83	62.8		22.79	0.00
	Emparejado	0.94	0.95	-4	93.7	-0.85	0.40
Proporción de docentes que se formaron en educación primaria (p_dprim)	Sin Emparejar	0.93	0.81	60.6		22.24	0.00
	Emparejado	0.87	0.89	-8.2	86.5	-1.50	0.13
La I.E. es multigrado (mult)	Sin Emparejar	0.22	0.12	26.4		11.62	0.00
	Emparejado	0.16	0.19	-7.8	70.4	-1.25	0.21
La I.E. cuenta con los tres servicios básicos (servicios)	Sin Emparejar	0.63	0.94	-80.7		-38.92	0.00
	Emparejado	0.81	0.79	5.5	93.2	0.86	0.39
La I.E. cuenta con computadoras con internet (internet)	Sin Emparejar	0.42	0.81	-88.7		-38.54	0.00
	Emparejado	0.67	0.65	3.5	96.1	0.53	0.60
La I.E. cuenta con sala de computo (e13_2)	Sin Emparejar	0.48	0.62	-29.4		-12.27	0.00
	Emparejado	0.55	0.57	-3.8	87.2	-0.60	0.55
La I.E. cuenta con sala de profesores (e13_6)	Sin Emparejar	0.18	0.27	-21.9		-8.79	0.00
	Emparejado	0.24	0.22	4.1	81.4	0.65	0.52
La I.E. cuenta con losa deportiva (e13_7)	Sin Emparejar	0.27	0.12	36.4		16.22	0.00
	Emparejado	0.23	0.27	-10.1	72.2	-1.48	0.14
La I.E. cuenta con auditorio (e13_10)	Sin Emparejar	0.14	0.09	15.5		6.74	0.00
	Emparejado	0.12	0.15	-9.5	38.5	-1.46	0.15
La I.E. cuenta con oficinas administrativas (e13_11)	Sin Emparejar	0.49	0.69	-42.4		-17.91	0.00
	Emparejado	0.55	0.52	5.6	86.7	0.88	0.38
Proporción de hogares que cuentan con agua (agua_s)	Sin Emparejar	0.63	0.83	-93.2		-44.31	0.00
	Emparejado	0.74	0.72	8.9	90.4	1.38	0.17
Proporción de hogares que estan hacinados (hacina_s)	Sin Emparejar	0.19	0.13	81.3		38.92	0.00
	Emparejado	0.16	0.17	-8.9	89	-1.49	0.14
Proporción de la población de 18 a más que tiene educación secundaria completa o más (p_edu_sys)	Sin Emparejar	0.61	1.09	-150.2		-66.70	0.00
	Emparejado	0.90	0.86	10.5	93	1.58	0.12
Proporción de la población que es rural en el distrito (ruralidad_s)	Sin Emparejar	0.43	0.05	153.1		78.52	0.00
	Emparejado	0.20	0.22	-10.8	92.9	-1.69	0.09

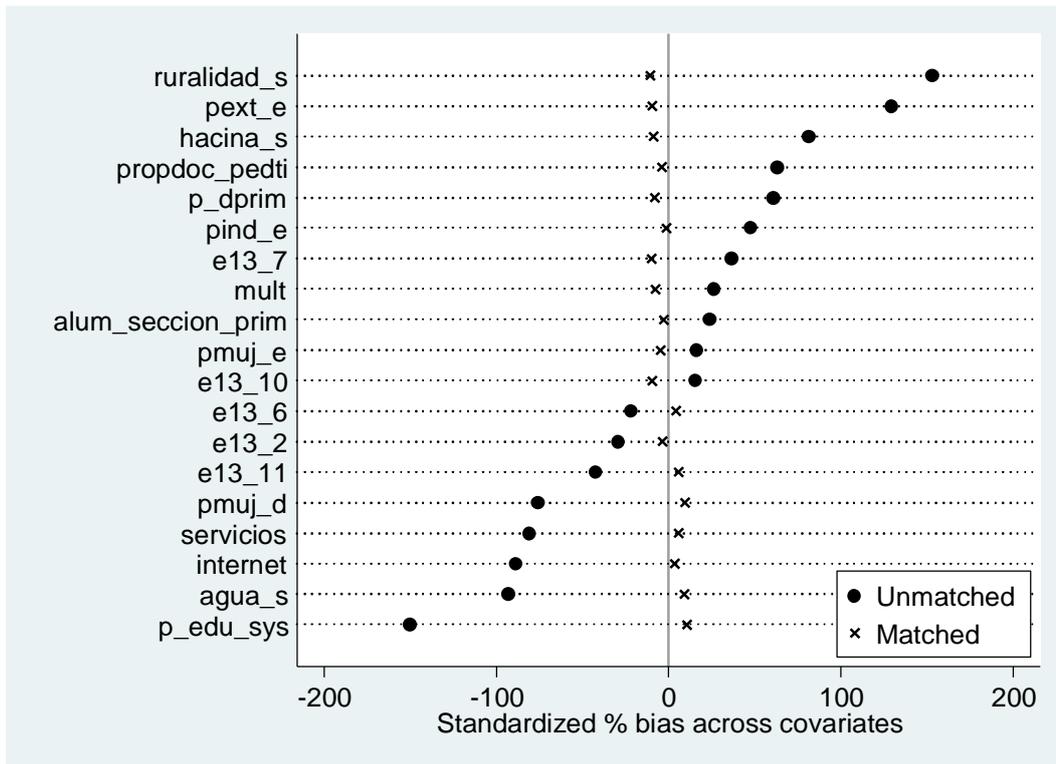
Nota: Se utilizó el emparejamiento *one-to-one* sin reemplazo y con una distancia (caliper) de 0.05.

Cuadro 2. Diferencias estandarizadas^{1/} entre el grupo de tratamiento y control antes y después del emparejamiento

	Sin emparejar (Muestral total)			Emparejada (Soporte Comun)		
	Estatal	No Estatal	Media estandarizada	Estatal	No Estatal	Media estandarizada
Proporción de estudiantes en extra- edad (pext_e)	0.08	0.01	1.31	0.04	0.04	0.08
Proporción de estudiantes mujeres (pmuj_e)	0.49	0.47	0.15	0.49	0.49	0.02
Proporción de estudiantes indígena (pind_e)	0.10	0.00	0.48	0.02	0.02	0.04
Alumnos por sección (alum_seccion_prim)	17.62	15.96	0.22	17.89	18.15	-0.03
Proporción de docentes que son mujeres (pmuj_d)	0.61	0.77	-0.77	0.71	0.70	0.01
Proporción de docentes con título pedagógico (propdoc_pedti)	0.96	0.83	0.63	0.94	0.95	-0.04
Proporción de docentes que se formaron en educación primaria (p_dprim)	0.93	0.81	0.61	0.87	0.86	0.03
La I.E. es multigrado (mult)	0.21	0.11	0.28	0.16	0.16	-0.01
La I.E. cuenta con los tres servicios básicos (servicios)	0.63	0.94	-0.81	0.81	0.83	-0.05
La I.E. cuenta con computadoras con internet (internet)	0.42	0.81	-0.89	0.67	0.68	-0.02
La I.E. cuenta con sala de computo (e13_2)	0.48	0.63	-0.30	0.55	0.59	-0.07
La I.E. cuenta con sala de profesores (e13_6)	0.18	0.28	-0.23	0.24	0.24	-0.01
La I.E. cuenta con losa deportiva (e13_7)	0.27	0.13	0.36	0.23	0.27	-0.09
La I.E. cuenta con auditorio (e13_10)	0.14	0.10	0.15	0.12	0.14	-0.06
La I.E. cuenta con oficinas administrativas (e13_11)	0.49	0.69	-0.42	0.55	0.57	-0.05
Proporción de hogares que cuentan con agua (agua_s)	0.63	0.84	-0.93	0.74	0.74	-0.02
Proporción de hogares que están hacinados (hacina_s)	0.19	0.13	0.82	0.16	0.16	-0.01
Proporción de la población de 18 a más que tiene educación secundaria completa o más (p_edu_sys)	0.61	1.09	-1.50	0.90	0.90	-0.01
Proporción de la población que es rural en el distrito (ruralidad_s)	0.42	0.05	1.53	0.20	0.19	0.03

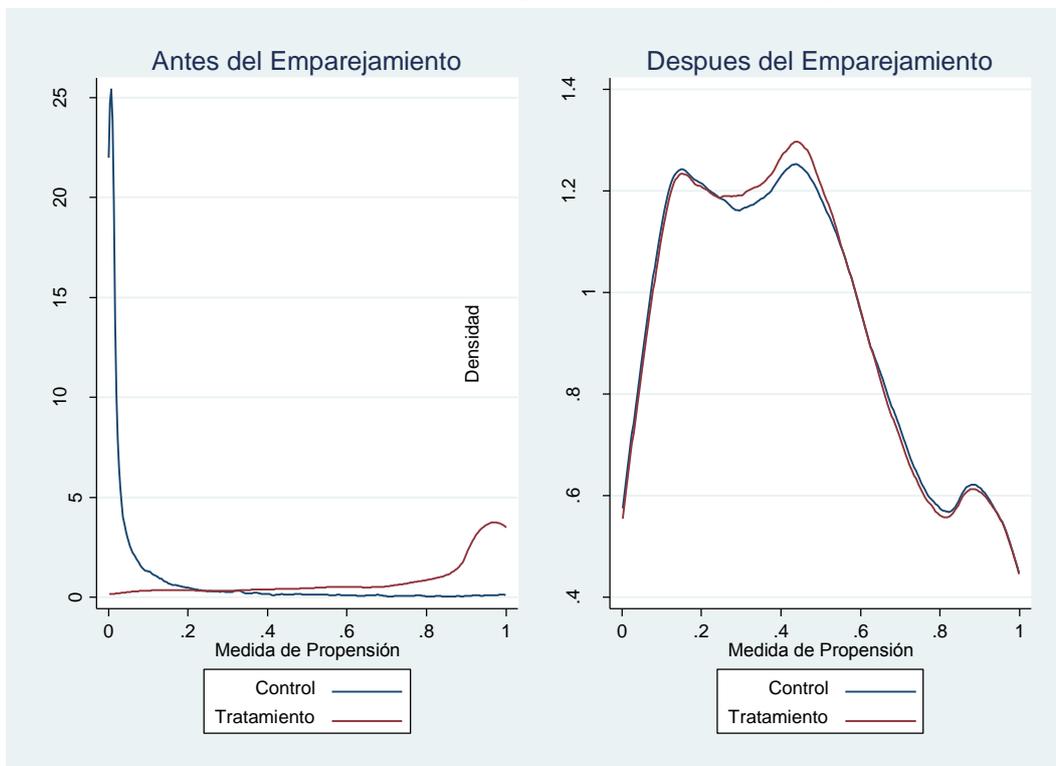
^{1/}Dado que el t-test puede verse influenciado por el tamaño de la muestra e indicar que cualquier diferencia es estadísticamente significativa. Se usa la diferencia estandarizada entre las medias o proporciones. La estandarización de los promedios se hace en función a la muestra de análisis y se considera que una variable esta balanceada entre ambos grupos si la diferencia estandarizada es menor a 0.10 desviaciones estándar en valor absoluto (Lunt, 2014).

Figura 1. Balance de las variables antes y después del emparejamiento



Nota: Se utilizó el emparejamiento *one-to-one* sin reemplazo y con una distancia (caliper) de 0.05.

Figura 2. Puntajes de propensión antes y después del emparejamiento por grupo de estudio



Nota: Se utilizó el emparejamiento *Uno-a-Uno* sin reemplazo y con una distancia (caliper) de 0.05.

Anexo D. Resultados adicionales para los modelos de Emparejamiento

Cuadro 1. Diferencias en el rendimiento y niveles de desempeño (2015) ajustados en Matemática y Comprensión de lectura por grupo de estudio (coeficientes estandarizados)

	Vecino más cercano ^{1/}			Uno a Uno		
	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia
	(E – NE) ^{2/} (n=4)	(E – NE) (n=3)	(E – NE) (n=2)	(E – NE) (caliper: 0.10)	(E – NE) (caliper:0.05)	(E – NE) (caliper:0.01)
Comprensión de lectura						
Puntaje	-0.10 (0.063)	-0.10 (0.064)	-0.10 (0.065)	-0.11 * (0.057)	-0.15 ** (0.053)	-0.08 (0.065)
Satisfactorio	-0.09 (0.062)	-0.08 (0.063)	-0.08 (0.063)	-0.06 (0.055)	-0.10 * (0.051)	-0.04 (0.062)
En proceso	0.09 (0.060)	0.08 (0.061)	0.08 (0.062)	0.04 (0.055)	0.07 (0.053)	0.02 (0.061)
En inicio	0.03 (0.065)	0.03 (0.066)	0.04 (0.066)	0.06 (0.058)	0.09 (0.056)	0.06 (0.062)
Matemática						
Puntaje	0.45 *** (0.064)	0.45 *** (0.066)	0.44 *** (0.068)	0.38 *** (0.055)	0.35 *** (0.058)	0.38 *** (0.066)
Satisfactorio	0.44 *** (0.069)	0.45 *** (0.071)	0.44 *** (0.073)	0.37 *** (0.056)	0.34 *** (0.062)	0.35 *** (0.069)
En proceso	0.12 + (0.071)	0.12 + (0.071)	0.12 + (0.072)	0.13 * (0.049)	0.13 ** (0.045)	0.18 *** (0.052)
En inicio	-0.44 *** (0.062)	-0.45 *** (0.063)	-0.44 *** (0.064)	-0.38 *** (0.052)	-0.36 *** (0.049)	-0.41 *** (0.055)
Observaciones						
Total	3,000	3,000	3,000	1,300	1,277	1,200
Estatad	1,129	1,129	1,129	637	633	601
No Estadad	1,871	1,871	1,871	663	644	599

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se usó como distancia (caliper) entre vecinos 0.01.

2/ E – NE = Estadad – No Estadad.

Errores estándar robustos entre paréntesis, los cuales fueron calculados mediante la metodología de *bootstrapping* con 100 réplicas.

Nota: Las variables usadas en el modelo de regresión *logit* para el emparejamiento son: porcentaje de estudiantes mujeres, porcentaje de estudiantes en extra-edad, porcentaje de estudiantes indígenas, alumnos por sección, porcentaje de docentes mujeres, porcentaje de docentes con título pedagógico, porcentaje de docentes que se han formado en educación primaria, I.E. multigrado, la I.E. cuenta con los tres servicios básicos, la I.E. cuenta con computadoras con conexión a internet, la I.E. tiene sala de cómputo, la I.E. tiene sala de profesores, la I.E. tiene losa deportiva, la I.E. tiene patio, la I.E. tiene oficinas administrativas, proporción de hogares con agua potable en el distrito, proporción de hogares que se encuentran hacinados en el distrito (3 o más personas por habitación), proporción de la población de 18 años o más con educación secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito y efectos fijos por región.

Cuadro 2. Diferencias en Diferencias en el rendimiento y niveles de desempeño (2015-2014) ajustados en Matemática y Comprensión de lectura por grupo de estudio (coeficientes estandarizados)

	Vecino más cercano			One to One		
	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia	Diferencia
	(E – NE)* (n=4)	(E – NE) (n=3)	(E – NE) (n=2)	(E – NE) (caliper: 0.10)	(E – NE) (caliper:0.05)	(E – NE) (caliper:0.01)
Comprensión de lectura						
Puntaje	0.14 * (0.066)	0.13 + (0.068)	0.12 + (0.069)	0.13 * (0.060)	0.06 (0.055)	0.11 (0.070)
Satisfactorio	0.12 + (0.065)	0.12 + (0.067)	0.11 (0.069)	0.15 ** (0.057)	0.08 (0.054)	0.08 (0.066)
En proceso	-0.02 (0.061)	-0.02 (0.063)	-0.02 (0.065)	-0.08 (0.050)	-0.02 (0.053)	0.01 (0.062)
En inicio	-0.20 ** (0.069)	-0.19 ** (0.069)	-0.18 ** (0.069)	-0.15 * (0.057)	-0.11 * (0.051)	-0.17 ** (0.056)
Matemática						
Puntaje	0.15 * (0.060)	0.14 * (0.061)	0.13 * (0.063)	0.11 * (0.052)	0.06 (0.051)	0.08 (0.065)
Satisfactorio	0.17 ** (0.065)	0.17 * (0.067)	0.16 * (0.068)	0.14 * (0.060)	0.12 * (0.056)	0.03 (0.016)
En proceso	-0.07 (0.060)	-0.08 (0.061)	-0.08 (0.062)	-0.10 (0.054)	-0.14 * (0.057)	-0.12 * (0.062)
En inicio	-0.08 (0.060)	-0.07 (0.061)	-0.06 (0.062)	-0.03 (0.050)	0.01 (0.050)	0.01 (0.063)
Observaciones						
Total	2,635	2,635	2,635	1,215	1,202	1,111
Estatad	984	984	984	596	595	549
No Estadad	1,651	1,651	1,651	619	607	562

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se usó como distancia (caliper) entre vecinos 0.01.

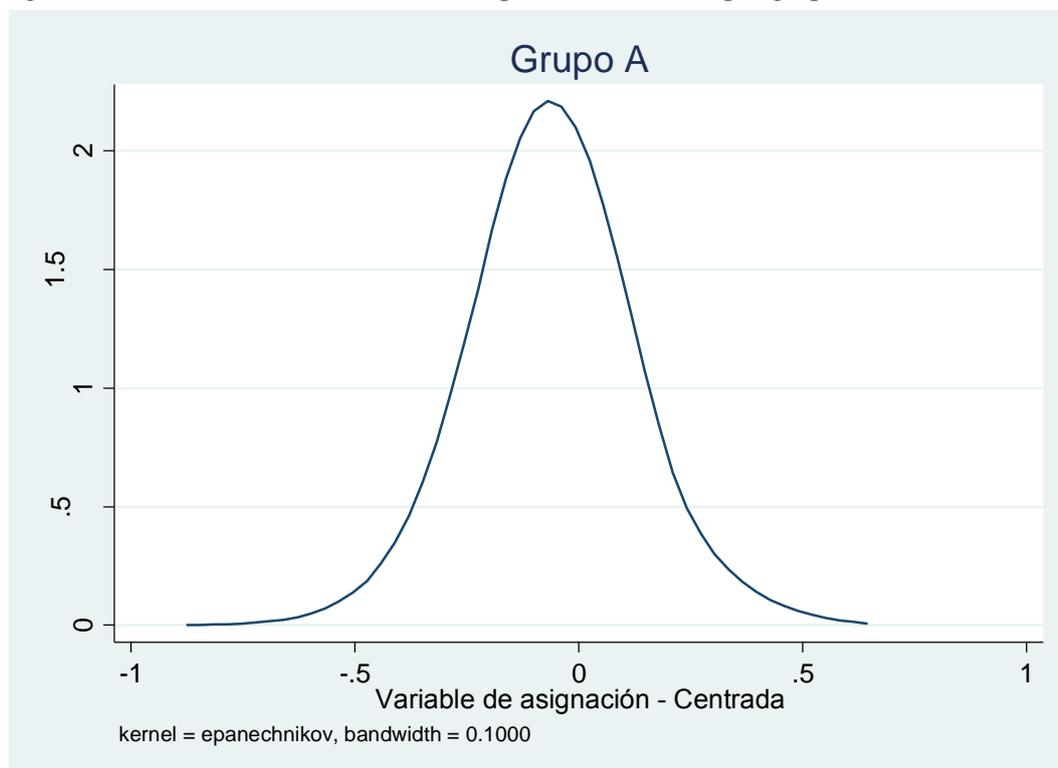
2/ E – NE = Estadad – No Estadad.

Errores estándar robustos entre paréntesis, los cuales fueron calculados mediante la metodología de *bootstrapping* con 100 réplicas.

Nota: Las variables usadas en el modelo de regresión *logit* para el emparejamiento son: porcentaje de estudiantes mujeres, porcentaje de estudiantes en extra-edad, porcentaje de estudiantes indígenas, alumnos por sección, porcentaje de docentes mujeres, porcentaje de docentes con título pedagógico, porcentaje de docentes que se han formado en educación primaria, I.E. multigrado, la I.E. cuenta con los tres servicios básicos, la I.E. cuenta con computadoras con conexión a internet, la I.E. tiene sala de computo, la I.E. tiene sala de profesores, la I.E. tiene losa deportiva, la I.E. tiene patio, la I.E. tiene oficinas administrativas, proporción de hogares con agua potable en el distrito, proporción de hogares que se encuentran hacinados en el distrito (3 o más personas por habitación), proporción de la población de 18 años o más con educación secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito y efectos fijos por región.

Anexo E. Gráficos de verificación del modelo de Regresión Discontinua

Figura 1. Distribución del índice de desempeño escolar 2014 por grupo^{1/}.



1/ El índice de desempeño escolar fue centrado alrededor del punto de corte para recibir el incentivo (p35)

Cuadro 1. Test de Manipulación de la variable de asignación

	Polinomio de orden 1	Polinomio de orden 2
Ancho de banda	0.04	0.04
Estimador de la matriz de Varianzas y Covarianzas	Asintótica	Asintótica
Número de observaciones a la izquierda	3,864	3,864
Número de observaciones a la izquierda	1,994	1,994
Observaciones a la izquierda	610	610
Observaciones a la derecha	685	685
T-estadístico	-0.16	0.29
p-value	0.87	0.77

Figura 2. Continuidad de variables observables alrededor del punto de corte

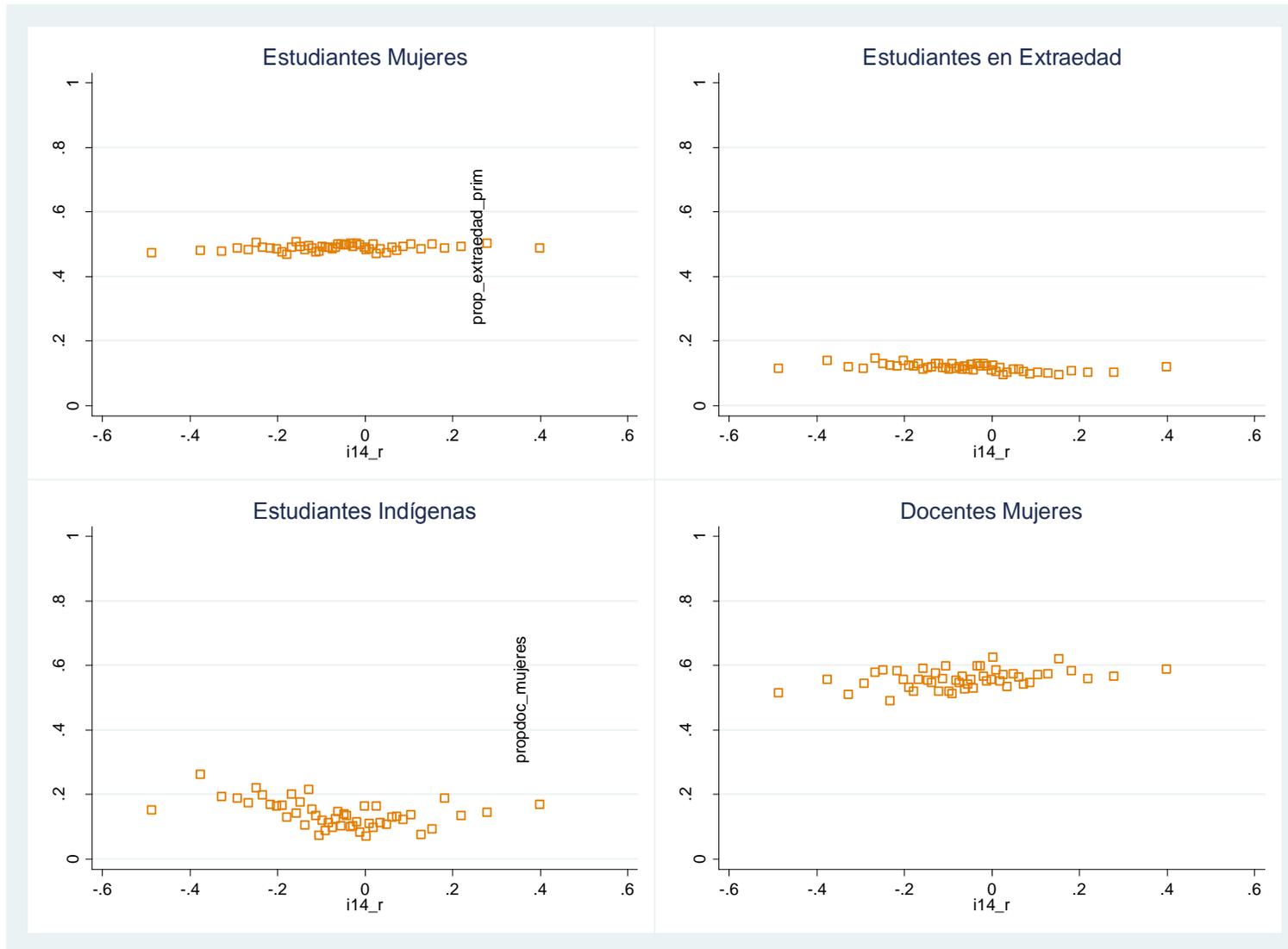


Figura 3. Continuidad de variables observables alrededor del punto de corte

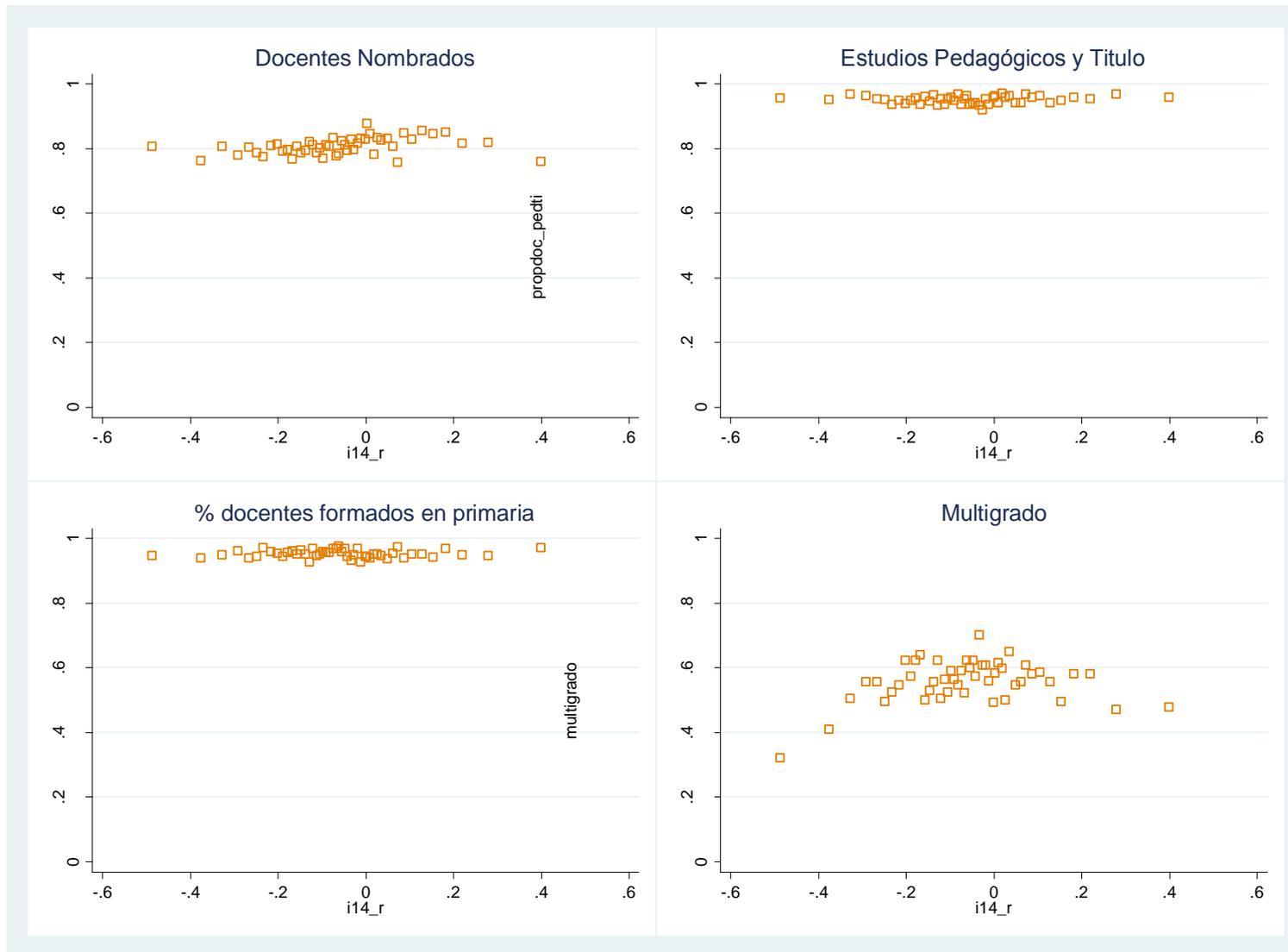


Figura 4. Continuidad de variables observables alrededor del punto de corte

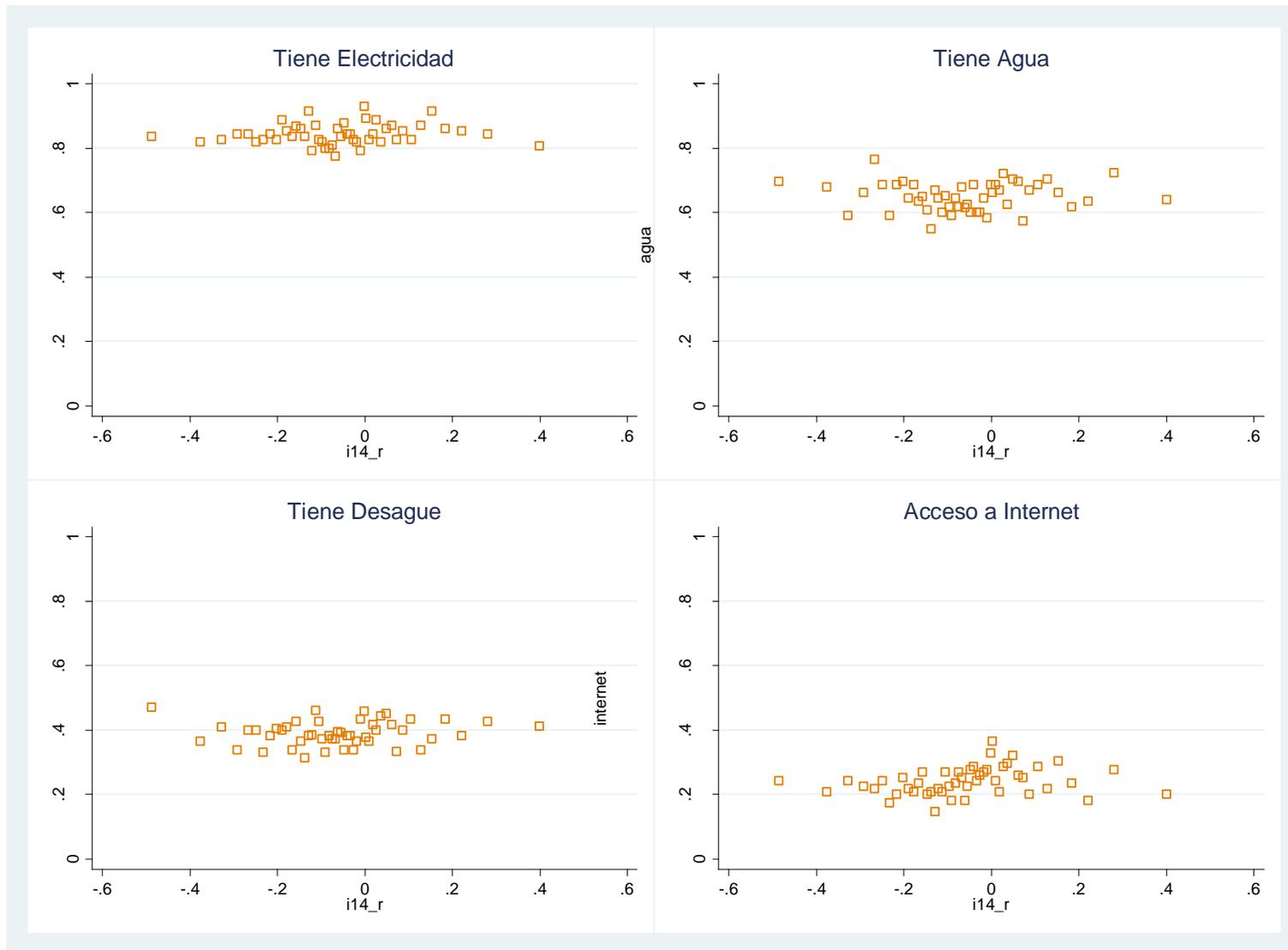


Figura 5. Continuidad de variables observables alrededor del punto de corte

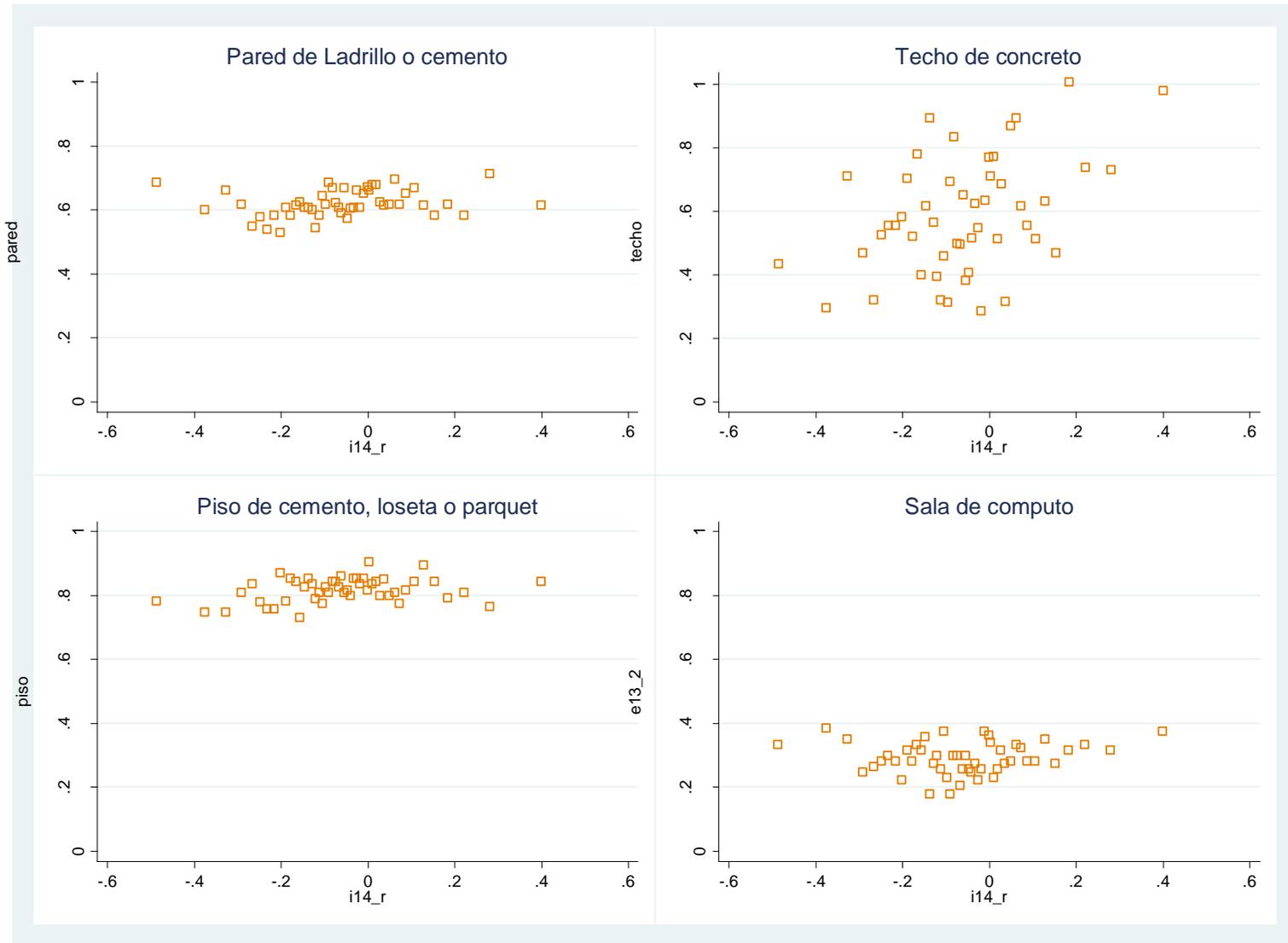


Figura 6. Continuidad de variables observables alrededor del punto de corte

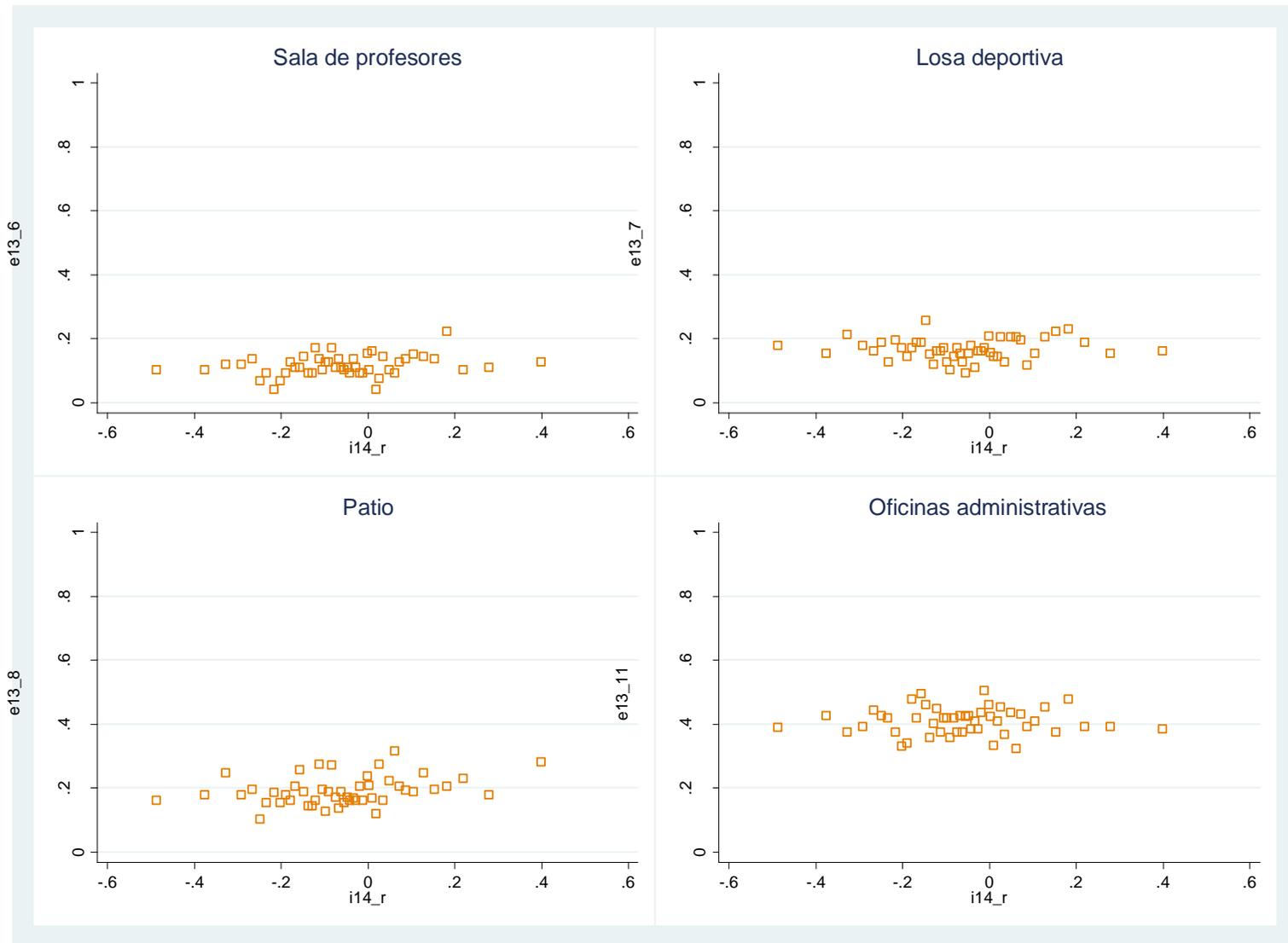


Figura 7. Continuidad de variables observables alrededor del punto de corte

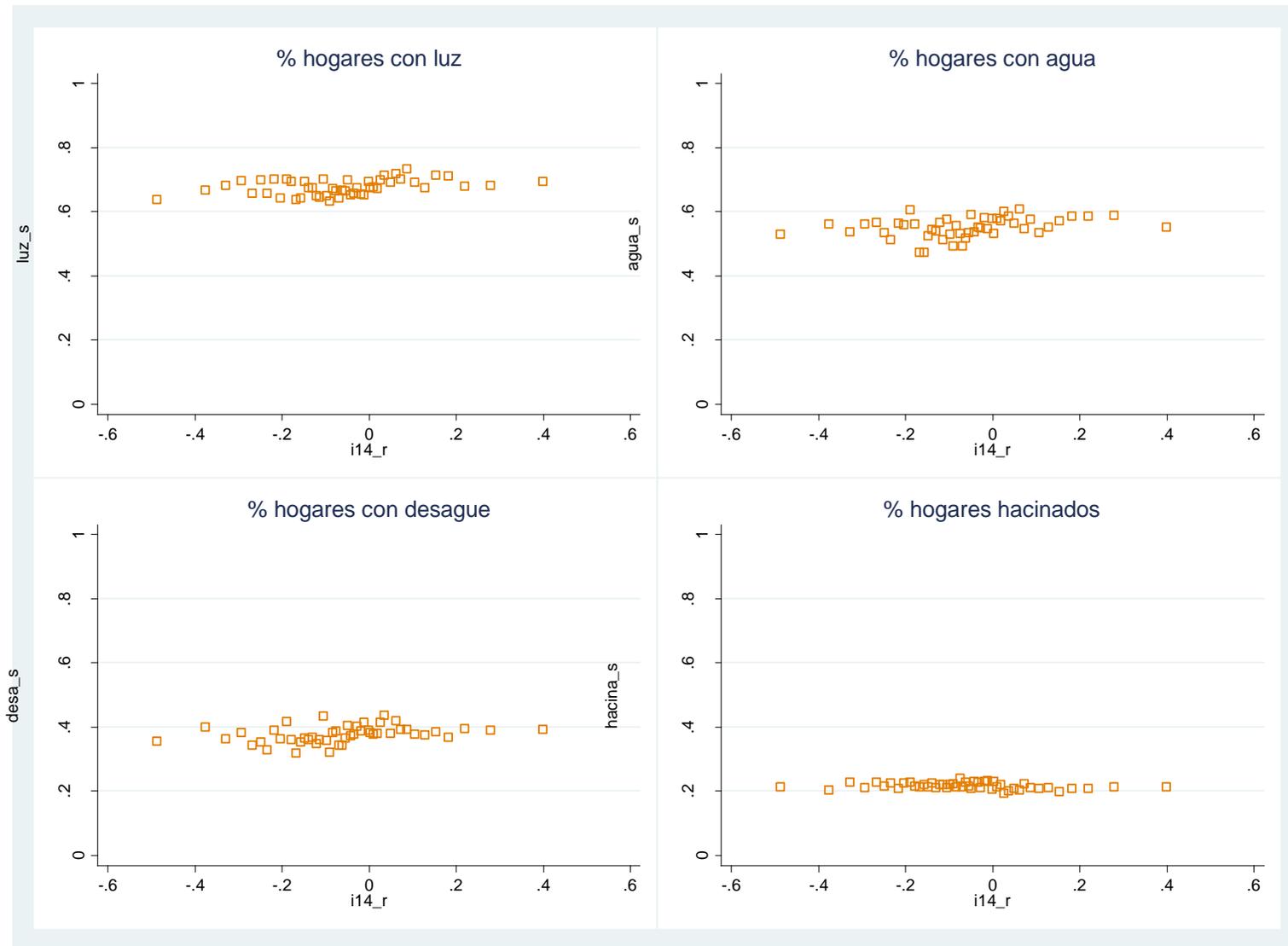
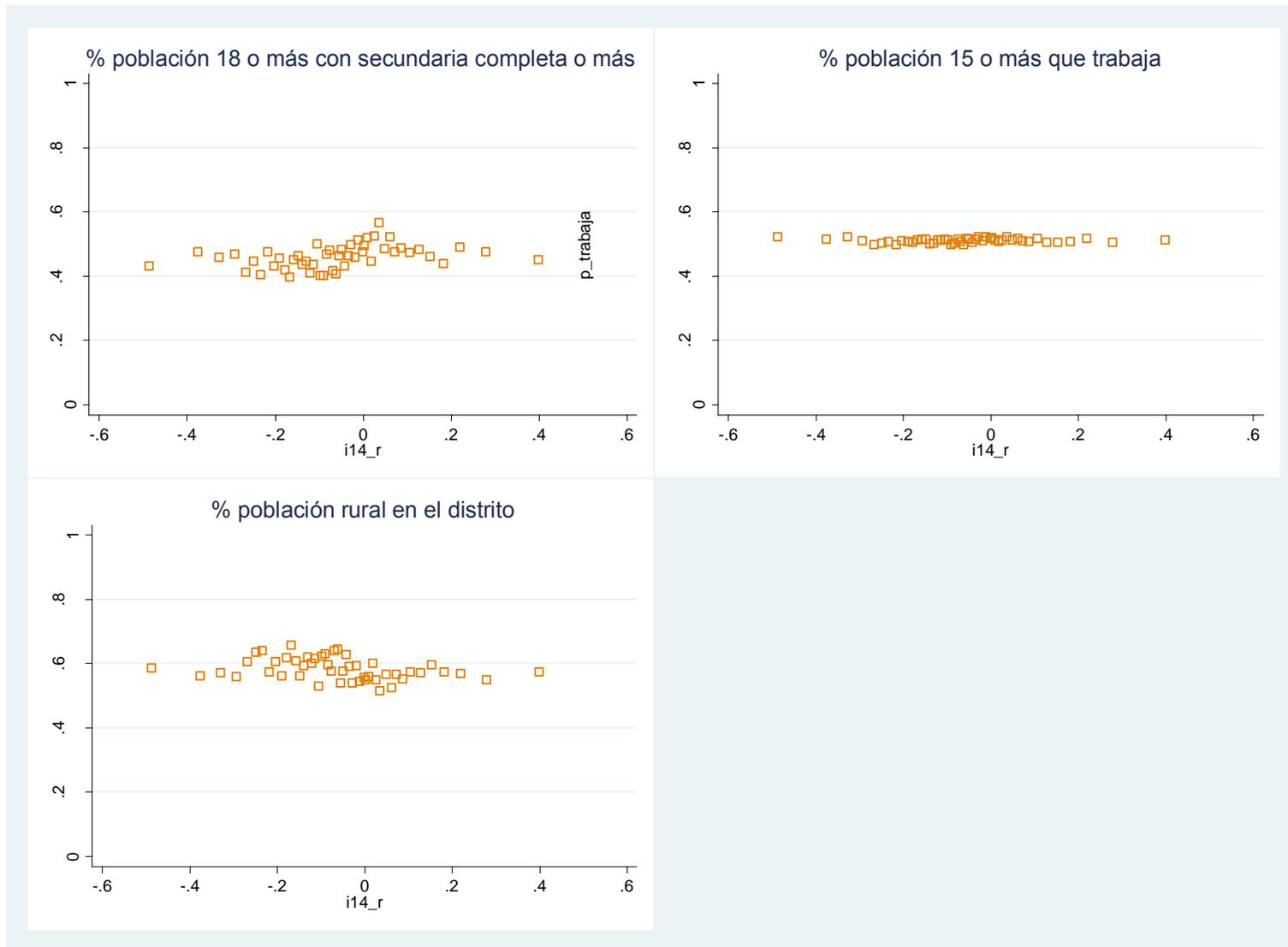


Figura 8. Continuidad de variables observables alrededor del punto de corte



Cuadro 2. Test de discontinuidad de las variables observadas alrededor del punto de corte

		Proporción de estudiantes Mujeres		Proporción de estudiantes en Extra edad		Proporción de estudiantes indígenas	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	-0.02 *	0.00	-0.03 *	-0.04 *	0.05	0.08
	ee	(0.007)	(0.015)	(0.012)	(0.020)	(0.039)	(0.069)
	t	-2.38	-0.12	-2.23	-2.17	1.30	1.13
N		3777	3777	3772	3772	3777	3777
		Proporción de docentes mujeres		Proporción de docentes nombrados		Proporción de docentes con títulos pedagógicos	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.01	0.03	0.01	0.01	0.02	0.01
	ee	(0.028)	(0.046)	(0.025)	(0.043)	(0.014)	(0.025)
	t	0.38	0.57	0.39	0.35	1.42	0.58
N		3777	3777	3777	3777	3777	3777
		Proporción de docentes formados en E. Primaria		I.E. Multigrado		Cuenta con acceso a internet	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	-0.01	-0.01	-0.14	-0.19	0.03	0.04
	ee	(0.013)	(0.021)	(0.081)	(0.121)	(0.047)	(0.083)
	t	-0.81	-0.25	-1.68	-1.60	0.59	0.52
N		3777	3777	3777	3777	3707	3707
		Cuenta con electricidad		Cuenta con agua		Cuenta con desagüe	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.06	0.09	0.11 *	0.20 *	0.09	0.10
	ee	(0.046)	(0.070)	(0.057)	(0.092)	(0.060)	(0.098)
	t	1.35	1.25	2.00	2.17	1.48	1.03
N		3706	3706	3706	3706	3706	3706
		Pared es de cemento o ladrillo		Techo es de concreto		Piso es de cemento, loseta, vinílico o parquet	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.06	0.10	0.38	0.57	0.00	0.03
	ee	(0.050)	(0.090)	(0.232)	(0.422)	(0.036)	(0.061)
	t	1.14	1.12	1.63	1.34	-0.07	0.06
N		3701	3701	3701	3701	3701	3701
		Sala de Computo		Sala de Profesores		Losa deportiva	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.04	0.04	0.04	0.09	0.03	0.02
	ee	(0.053)	(0.088)	(0.039)	(0.074)	(0.042)	(0.068)
	t	0.76	0.40	0.93	1.19	0.78	0.25
N		3772	3772	3772	3772	3772	3772
		Patio		Oficinas administrativas		Proporción de hogares con luz	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.04	0.03	0.00	0.00	0.03	0.04
	ee	(0.044)	(0.073)	(0.047)	(0.085)	(0.025)	(0.045)
	t	0.97	0.47	-0.10	0.03	1.37	0.81
N		3772	3772	3772	3772	3746	3746
		Proporción de hogares con agua		Proporción de hogares con desagüe		Proporción de hogares hacinados	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.02	-0.02	0.00	-0.03	-0.02	-0.03
	ee	(0.027)	(0.048)	(0.026)	(0.047)	(0.013)	(0.020)

t	0.68	-0.41	-0.06	-0.73	-1.82	+	-1.46
N	3746	3746	3746	3746	3746		3746
		Proporción de la población con secundaria completa o más		Proporción de la población de 15 a más que trabaja		Proporción de la población rural en el distrito	
		Lineal	Cuadrático	Lineal	Cuadrático	Lineal	Cuadrático
Corte	β	0.03	0.00	0.00	-0.01	-0.03	-0.01
	ee	(0.032)	(0.059)	(0.007)	(0.012)	(0.032)	(0.056)
	t	0.89	0.02	-0.72	-1.12	-0.89	-0.23
	N	3746	3746	3746	3746	3746	3746

***p<0.001, **p<0.01, *p<0.05, +p<0.10

Nota: Se utilizó para los análisis de continuidad un análisis no-paramétrico (regresión lineal local) de Kernel triangular. El ancho de banda seleccionado fue de 0.15. Así mismo, se probó tendencias lineales y cuadráticas alrededor del punto de corte. Errores estándar entre paréntesis.

Cuadro 3. Discontinuidad en las variables de resultado en puntos de corte donde no debería haber discontinuidad.

	p50		p75		p80		
	β	ee	β	ee	β	ee	
Comprensión de Lectura	1.9	(5.066)	6.4	(4.309)	-1.0	(5.523)	
% Satisfactorio en Comprensión de Lectura	0.02	(0.019)	0.02	(0.024)	-0.02	(0.025)	
% En proceso en Comprensión de Lectura	-0.01	(0.023)	0.00	(0.023)	0.04	(0.022)	+
% En inicio en Comprensión de Lectura	0.00	(0.018)	-0.02	(0.017)	-0.02	(0.017)	
Matemática	0.9	(6.868)	11.3	(6.483)	-3.5	(8.346)	+
% Satisfactorio en Matemática	0.02	(0.019)	0.02	(0.017)	0.00	(0.023)	
% En proceso en Matemática	0.00	(0.017)	0.01	(0.017)	-0.02	(0.020)	
% En inicio en Matemática	-0.01	(0.026)	-0.03	(0.024)	0.03	(0.030)	

Nota: El ancho de banda fijado para este análisis ha sido fijado con el método de Imbens y Kalyanaraman (2012)

Anexo F. Análisis adicionales de la estimación Paramétrica del Modelo de Regresión Discontinua

Cuadro 1. Indicadores de Ajuste de Akaike y Schwarz para las diferentes formas funcionales entre las variables de resultado y asignación.

	Matemática		Comprensión		Asistencia Director	
	BIC	AIC	BIC	AIC	BIC	AIC
Lineal	69,629.44	69,602.73	65,922.97	65,896.26	3,463.01	3,437.40
Cuadrático	69,643.84	69,603.78	65,934.49	65,894.44	3,479.60	3,441.19
Cúbico	69,659.00	69,605.60	65,950.70	65,897.29	3,495.48	3,444.27
Asistencia Docente		Asistencia Estudiantes		Mejor modelo: menor valor del AIC o BIC (Benardi & Macías, 2012)		
BIC	AIC	BIC	AIC			
Lineal	-2,843.50	-2,869.00	-1,642.60			-1,667.61
Cuadrático	-2,828.50	-2,866.74	-1,627.34			-1,664.87
Cúbico	-2,812.23	-2,863.23	-1,611.38	-1,661.42		

Nota: Para cada una de las formas funcionales se incluyó interacciones entre la variable de tratamiento y el componente lineal, cuadrático o cúbico. Modelos con el menor AIC (Akaike Information Criterion) y/o BIC (Bayesian Information Criterion) son aquellos que presentan un mejor ajuste.

Cuadro 2. Resultados de la Regresión Discontinua Paramétrica en Matemática (coeficientes estandarizados)

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	0.06 ** (0.022)	-0.06 ** (0.022)	0.05 + (0.027)	0.04 (0.025)
Eliminando 10%	0.05 * (0.026)	-0.06 * (0.026)	0.04 (0.031)	0.04 (0.029)
Eliminando 20%	0.05 + (0.030)	-0.07 * (0.031)	0.06 (0.035)	0.04 (0.033)
Eliminando 30%	0.04 (0.034)	-0.03 (0.034)	0.01 (0.041)	0.03 (0.038)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extraedad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por región entre paréntesis.

Cuadro 3. Resultados de la Regresión Discontinua Paramétrica en Comprensión de Lectura (coeficientes estandarizados)

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	0.05 * (0.021)	-0.04 + (0.022)	0.00 (0.024)	0.03 (0.022)
Eliminando 10%	0.04 (0.024)	-0.04 (0.024)	0.01 (0.030)	0.02 (0.026)
Eliminando 20%	0.05 + (0.028)	-0.03 (0.030)	-0.02 (0.033)	0.04 (0.029)
Eliminando 30%	0.04 (0.031)	-0.02 (0.035)	-0.01 (0.038)	0.03 (0.034)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extraedad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito. Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por región entre paréntesis.

Cuadro 4. Resultados de la Regresión Discontinua Paramétrica en Variables Intermedias (coeficientes estandarizados)

	Asistencia del Director	Asistencia Docente	Asistencia Estudiantes	Carga docente	Horas de clase	Comité de tutoría	Documentos de gestión
Toda la muestra	0.23 * (0.094)	0.04 (0.025)	0.03 (0.025)	-0.01 (0.026)	0.00 (0.025)	0.07 (0.080)	0.08 (0.114)
Eliminando 10%	0.28 * (0.115)	0.07 * (0.031)	0.06 + (0.031)	-0.02 (0.029)	-0.01 (0.032)	0.10 (0.097)	0.12 (0.167)
Eliminando 20%	0.27 * (0.134)	0.06 + (0.034)	0.03 (0.037)	-0.04 (0.031)	-0.02 (0.036)	0.07 (0.120)	0.04 (0.212)
Eliminando 30%	0.36 * (0.170)	0.09 * (0.041)	0.04 (0.046)	-0.05 (0.043)	-0.02 (0.051)	0.10 (0.157)	0.17 (0.278)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Cuadro 5. Número de Observaciones para las Estimaciones de la Regresión Discontinua Paramétrica

	Asistencia del Director	Asistencia Docente	Asistencia Estudiantes	Carga docente	Horas de clase	Comité de tutoría	Documentos de gestión	Matemática	Comprensión
Total									
Derecha	1,481	1,440	1,275	1,929	1,472	1,479	1,480	1,940	1,940
Izquierda	2,860	2,788	2,475	3,724	2,853	2,857	2,857	3,745	3,745
90% muestra									
Derecha	1,262	1,228	1,080	1,646	1,255	1,260	1,261	1,655	1,655
Izquierda	2,627	2,563	2,271	3,440	2,620	2,624	2,624	3,460	3,460
80% muestra									
Derecha	1,035	1,008	882	1,361	1,031	1,033	1,034	1,369	1,369
Izquierda	2,413	2,352	2,082	3,159	2,407	2,410	2,411	3,177	3,177
70% muestra									
Derecha	807	785	683	1,074	803	805	806	1,081	1,081
Izquierda	2,202	2,148	1,903	2,875	2,196	2,199	2,200	2,892	2,892

Cuadro 6. Resultados de la Regresión Discontinua Paramétrica en Matemática (coeficientes no estandarizados) con intervalos simétricos

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	11.24 ** (4.324)	-0.05 ** (0.016)	0.03 + (0.015)	0.02 (0.014)
h=0.35	12.18 * (4.720)	-0.05 ** (0.018)	0.02 (0.015)	0.03 (0.015)
h=0.25	11.89 * (5.283)	-0.05 * (0.021)	0.03 (0.017)	0.02 (0.017)
h=0.15	10.27 (6.661)	-0.05 + (0.026)	0.03 (0.020)	0.02 (0.019)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Cuadro 7. Resultados de la Regresión Discontinua Paramétrica en Comprensión de lectura (coeficientes no estandarizados) con intervalos simétricos

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	6.60 * (3.029)	-0.02 + (0.011)	0.00 (0.014)	0.02 (0.014)
h=0.35	6.26 * (3.161)	-0.02 + (0.012)	0.00 (0.016)	0.02 (0.015)
h=0.25	6.58 + (3.475)	-0.02 + (0.013)	0.00 (0.017)	0.02 (0.017)
h=0.15	7.71 + (4.495)	-0.03 (0.017)	0.00 (0.021)	0.02 (0.020)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Cuadro 8. Resultados de la Regresión Discontinua Paramétrica en variables de resultados intermedios (coeficientes estandarizados) con intervalos simétricos

	Asistencia del Director	Asistencia Docente	Asistencia Estudiantes	Carga docente	Horas de clase	Comité de tutoría	Documentos de gestión
Toda la muestra	0.05 * (0.021)	0.01 (0.009)	0.01 (0.011)	-0.15 (0.428)	0.00 (0.025)	0.02 (0.025)	0.01 (0.013)
h=0.35	0.06 ** (0.023)	0.02 * (0.010)	0.02 (0.012)	-0.13 (0.455)	-0.01 (0.029)	0.02 (0.026)	0.01 (0.016)
h=0.25	0.07 ** (0.027)	0.02 * (0.011)	0.01 (0.013)	-0.33 (0.465)	-0.01 (0.031)	0.03 (0.031)	0.00 (0.019)
h=0.15	0.08 * (0.032)	0.03 + (0.015)	0.02 (0.017)	-0.48 (0.579)	-0.03 (0.048)	0.01 (0.039)	0.00 (0.022)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Cuadro 9. Resultados de la Regresión Discontinua Paramétrica en Matemática (coeficientes estandarizados) con intervalos simétricos

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	0.06 ** (0.022)	-0.06 ** (0.022)	0.05 + (0.027)	0.04 (0.025)
h=0.35	0.06 * (0.023)	-0.06 ** (0.024)	0.04 (0.028)	0.05 (0.026)
h=0.25	0.06 * (0.025)	-0.06 * (0.027)	0.04 (0.030)	0.04 (0.028)
h=0.15	0.05 (0.031)	-0.06 + (0.032)	0.05 (0.034)	0.03 (0.031)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Cuadro 10. Resultados de la Regresión Discontinua Paramétrica en Comprensión de lectura (coeficientes estandarizados) con intervalos simétricos

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
Toda la muestra	0.05 * (0.021)	-0.04 + (0.022)	0.00 (0.024)	0.03 (0.022)
h=0.35	0.04 * (0.021)	-0.04 + (0.023)	0.00 (0.026)	0.03 (0.023)
h=0.25	0.04 + (0.023)	-0.04 + (0.024)	0.01 (0.028)	0.03 (0.025)
h=0.15	0.05 + (0.028)	-0.05 (0.031)	0.01 (0.032)	0.03 (0.029)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Cuadro 11. Resultados de la Regresión Discontinua Paramétrica en variables de resultados intermedios (coeficientes estandarizados) con intervalos simétricos

	Asistencia del Director	Asistencia Docente	Asistencia Estudiantes	Carga docente	Horas de clase	Comité de tutoría	Documentos de gestión
Toda la muestra	0.23 * (0.094)	0.04 (0.025)	0.03 (0.025)	-0.01 (0.026)	0.00 (0.025)	0.07 (0.080)	0.08 (0.114)
h=0.35	0.28 ** (0.102)	0.06 * (0.026)	0.04 (0.028)	-0.01 (0.027)	-0.01 (0.028)	0.08 (0.083)	0.06 (0.141)
h=0.25	0.31 ** (0.118)	0.06 * (0.029)	0.03 (0.030)	-0.02 (0.026)	-0.01 (0.029)	0.09 (0.100)	0.03 (0.168)
h=0.15	0.34 * (0.137)	0.06 + (0.037)	0.05 (0.038)	-0.03 (0.032)	-0.02 (0.043)	0.04 (0.127)	0.04 (0.212)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ Se estimaron modelos probit para poder obtener en cuanto incrementa la probabilidad dado que la variable dependiente era binaria o cualitativa.

Nota: Los controles fueron la I.E. es multigrado, proporción de estudiantes mujeres, proporción de estudiantes en extra edad, proporción de estudiantes indígenas, proporción de docentes mujeres, proporción de docentes nombrados, proporción de docentes con estudios pedagógicos y título, proporción de docentes que se formaron en educación primaria, la I.E. cuenta con computadoras conectadas al internet, la I.E. cuenta con piso de cemento, la I.E. cuenta con techos de concreto, la I.E. cuenta con paredes de ladrillo y concreto, la I.E. cuenta con luz, la I.E. cuenta con agua, la I.E. cuenta con desagüe, la I.E. cuenta con sala de computo, la I.E. cuenta con sala de profesores, la I.E. cuenta con losa deportiva, la I.E. cuenta con patio, la I.E. cuenta con oficinas administrativas, proporción de hogares que cuentan con agua en el distrito, proporción de hogares que cuentan con luz en el distrito, proporción de hogares que cuentan con desagüe en el distrito, proporción de hogares que se encuentran hacinados (3 o más personas por habitación) en el distrito, proporción de la población de 18 a más que tiene secundaria completa o más en el distrito, proporción de la población que vive en zonas rurales en el distrito.

Todos los modelos estimados incluyen la interacción entre la variable de tratamiento y la de asignación.

Errores ajustados por estrato entre paréntesis.

Anexo G. Análisis adicionales de la estimación no paramétrica del Modelo de Regresión Discontinua

Cuadro 1. Resultados de la Regresión Discontinua No Paramétrica en Matemática (coeficientes estandarizados)

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
h=0.25	0.20 ** (0.07)	-0.22 ** (0.07)	0.18 * (0.07)	0.12 + (0.07)
h=0.20	0.20 * (0.08)	-0.22 ** (0.08)	0.18 * (0.08)	0.12 (0.08)
h=0.15	0.20 * (0.10)	-0.21 * (0.10)	0.15 (0.10)	0.12 (0.10)
Óptimo ^{1/}	0.21 + (0.12)	-0.17 (0.12)	0.07 (0.15)	0.12 (0.09)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ El ancho de banda fijado para este análisis ha sido fijado con el método de Imbens y Kalyanaraman (2012)

Nota: Para la estimación no paramétrica se usa la estimación de Kernel Triangular. Está estimación pondera las observaciones y le da un mayor a aquellas cercanas al punto de corte. No se considera anchos de banda menores a 0.15 debido a que reduce significativamente el número de observaciones del grupo de tratamiento. Los resultados usando controles son similares.

Cuadro 2. Resultados de la Regresión Discontinua No Paramétrica en Comprensión de Lectura (coeficientes estandarizados)

	Promedio	Niveles de Desempeño		
		En inicio	En proceso	Satisfactorio
h=0.25	0.20 ** (0.073)	-0.22 ** (0.073)	0.03 (0.071)	0.14 + (0.073)
h=0.20	0.22 ** (0.082)	-0.24 ** (0.084)	0.03 (0.080)	0.15 + (0.083)
h=0.15	0.21 * (0.098)	-0.24 * (0.101)	0.05 (0.096)	0.14 (0.099)
Óptimo ^{1/}	0.21 * (0.101)	-0.27 + (0.140)	0.18 (0.147)	0.10 (0.131)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ El ancho de banda fijado para este análisis ha sido fijado con el método de Imbens y Kalyanaraman (2012)

Nota: Para la estimación no paramétrica se usa la estimación de Kernel Triangular. Está estimación pondera las observaciones y le da un mayor a aquellas cercanas al punto de corte. No se considera anchos de banda menores a 0.15 debido a que reduce significativamente el número de observaciones del grupo de tratamiento. Los resultados usando controles son similares.

Cuadro 3. Resultados de la Regresión Discontinua No Paramétrica en Variables Intermedias (coeficientes estandarizados)

	Asistencia Estudiantes	Asistencia Docente	Carga docente	Horas de clase
h=0.25	0.15 + (0.081)	0.16 * (0.081)	-0.12 + (0.072)	-0.03 (0.083)
h=0.20	0.16 + (0.090)	0.18 * (0.091)	-0.13 (0.081)	-0.03 (0.093)
h=0.15	0.18 + (0.106)	0.24 * (0.109)	-0.12 (0.096)	-0.01 (0.110)
Óptimo ^{1/}	0.17 (0.142)	0.26 * (0.119)	-0.11 (0.104)	0.01 (0.130)

***p<0.001, **p<0.01, *p<0.05, +p<0.10

1/ El ancho de banda fijado para este análisis ha sido fijado con el método de Imbens y Kalyanaraman (2012)

Nota: Para la estimación no paramétrica se usa la estimación de Kernel Triangular. Esta estimación pondera las observaciones y le da un mayor a aquellas cercanas al punto de corte. No se considera anchos de banda menores a 0.15 debido a que reduce significativamente el número de observaciones del grupo de tratamiento. Los resultados usando controles son similares.

Cuadro 4. Número de observaciones para la estimación de los resultados de la Regresión Discontinua No Paramétrica

	h=0.25		h=0.20		h=0.15	
	Izquierda	Derecha	Izquierda	Derecha	Izquierda	Derecha
Comprensión de Lectura	3,240	1,779	2,841	1,642	2,317	1,460
% Satisfactorio en Comprensión de Lectura	3,240	1,779	2,841	1,642	2,317	1,460
% En proceso en Comprensión de Lectura	3,240	1,779	2,841	1,642	2,317	1,460
% En inicio en Comprensión de Lectura	3,240	1,779	2,841	1,642	2,317	1,460
Matemática	3,240	1,779	2,841	1,642	2,317	1,460
% Satisfactorio en Matemática	3,240	1,779	2,841	1,642	2,317	1,460
% En proceso en Matemática	3,240	1,779	2,841	1,642	2,317	1,460
% En inicio en Matemática	3,240	1,779	2,841	1,642	2,317	1,460
Asistencia Estudiantes	2,119	1,159	1,860	1,066	1,527	942
Asistencia Docente	2,391	1,310	2,102	1,213	1,723	1,073
Carga docente	3,222	1,769	2,824	1,633	2,302	1,452
Horas de clase al día	2,447	1,341	2,148	1,236	1,763	1,097