

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
ESCUELA DE POSGRADO



**Incentivos monetarios para el logro de resultados en el sector
educativo: evaluación de impacto de “Bono Escuela” en el
rendimiento estudiantil**

**TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGISTER EN
ECONOMÍA**

AUTOR

Graciela Rocío Muñiz Cahuana

ASESOR:

Viviana Natali Cruzado de la Vega

Febrero, 2018

RESUMEN

La aplicación de incentivos monetarios en la educación básica es aún poco extendida a nivel regional. Por otro lado, la experiencia internacional, con varios años de estudio en la implementación de este tipo de incentivos, no es concluyente sobre su impacto en resultados educativos o el mecanismo interno más efectivo. En esta investigación se plantea una evaluación de impacto sobre la primera experiencia en el Perú de incentivo monetario a docentes asociado al desempeño estudiantil en colegios públicos: “Bono Escuela”.

Se busca evaluar si un incentivo monetario condicionado a los resultados del estudiante genera un cambio como producto de que los docentes compitan por el premio (efecto *ex ante*) o como producto de haberlo recibido (efecto *ex post*). Para ello se analiza el nivel primario, en las áreas de Comprensión Lectora y Matemática, y se aplica un modelo de emparejamiento y de triple diferencia para medir el efecto *ex ante*, y una regresión discontinua para medir el efecto *ex post*. También se analiza el efecto *ex ante* en el nivel secundaria, aunque con mayores limitaciones en la información. Además, se mide la heterogeneidad de los resultados en función a los años de implementación, área de procedencia y tipo de escuela.

Para el caso de primaria los resultados señalan que hay un efecto *ex ante* positivo que se consolida a partir del tercer año de implementación del incentivo, y que se concentra en la zona urbana y en los colegios polidocentes completos. Por otro lado, no habría efectos del incentivo de manera *ex post*, como producto de haber ganado el premio. Para el caso de secundaria, se empiezan a ver resultados positivos aunque pequeños en el segundo año de implementación, y también se concentran en la zona urbana.

Contenido

RESUMEN	2
1. Introducción	4
2. Marco Teórico	7
2.1 Teoría del principal-agente y contratos óptimos	7
2.2 Evidencia empírica sobre los mecanismos internos para el logro de resultados a través de incentivos	13
3. El Bono de Incentivo al Desempeño Escolar, “Bono Escuela”	19
4. Estrategia de identificación y evaluación	21
4.1 Metodología para la evaluación del efecto competencia en los niveles de primaria y secundaria	24
4.1.1 Identificación de grupo de tratamiento y zonas de desempeño	24
4.1.2 Identificación del grupo de control y estimación del modelo de emparejamiento	27
4.1.3 Estimación por triple diferencia para el caso del nivel primaria.....	29
4.2 Metodología para la evaluación del efecto reciprocidad en el nivel de primaria	31
4.3 Efectos diferenciados (característica de la escuela, ámbito y periodo de maduración)	33
5. Resultados	34
5.1 Resultados en el nivel de primaria	34
5.1.1 Efecto competencia	35
5.1.2 Efecto reciprocidad.....	37
5.2 Resultados en el nivel de secundaria	38
5.2.1 Efecto competencia	38
6. Conclusiones, limitaciones e implicancias de política	40
7. Bibliografía y referencias	44
8. Anexos	49

1. Introducción

El mecanismo de incentivos monetarios es un tema de estudio abordado, en primera instancia, por la economía laboral, y está estrechamente relacionado a la teoría de contratos y del principal- agente. Busca identificar los esquemas de pago óptimos que permitan fomentar el mayor nivel de esfuerzo del agente (ente contratado), y así maximizar los intereses del principal (ente contratista), bajo el supuesto de que son los objetivos de este último los de mayor relevancia. Por ello, es de particular interés en contextos donde existe un problema de asimetría de información, riesgo moral y altos costos asociados a la supervisión permanente del agente por parte del principal.

Adicionalmente, la economía del comportamiento y la psicología han brindado perspectivas complementarias al proceso bajo el cual funcionan los incentivos, tanto monetarios como no monetarios. Diversos estudios han abordado cómo el mecanismo de incentivos funciona en diferentes ámbitos, que van desde motivar a los individuos a comer más saludable (Angelucci *et al.*, 2015), bajar de peso (Cawley y Price, 2009; Babcock *et al.*, 2011), hasta cómo motivar que ciudadanos mejor capacitados y con mayor ética asuman cargos políticos en los municipios (Ferraz y Finan, 2009), cómo motivar un mayor esfuerzo laboral (DellaVigna y Pope, 2016), o cómo propiciar que estudiantes o profesores realicen un mayor esfuerzo para lograr mejores resultados académicos (Levitt *et al.*, 2012; Babcock *et al.*, 2011; Glewwe, Ilias y Kremer, 2010; Goodman y Turner, 2010; Muralidharan y Sundararaman, 2010; Rau y Contreras, 2012; Vigdor, 2008).

En este trabajo de investigación, el ámbito de interés es la aplicación de los mecanismos de incentivos monetarios la educación básica. Algunas investigaciones consideran que este sector puede ser un terreno particularmente fértil para que los mecanismos de incentivos impulsen una mayor efectividad, sobre todo teniendo en cuenta que este sector se caracteriza por tener retornos a largo plazo y, por tanto, puede presentar problemas de miopía (dificultad para anticipar los resultados) que pueden corregirse con recompensas en el corto plazo (Angelucci *et al.*, 2015; Levitt *et al.*, 2012). Otros estudios, por su parte, alegan que las características de los agentes (refiriéndose a los profesores), que serían principalmente del tipo “altruistas”, podrían generar que los mecanismos de incentivos monetarios sean contraproducentes (Dixit, 2002; Le Grand, 2003).

Lo cierto es que los diversos trabajos producidos a la fecha aún no pueden establecer conclusiones determinantes sobre los incentivos monetarios en lo que respecta al sector educativo, pues existen aquellos que han demostrado efectos positivos (Glewwe, Ilias, y Kremer, 2010; Vigdor, 2008; Lavy, 1999; Muralidharan y Sundararaman, 2010; Rau y Contreras, 2012), como aquellos que han identificado un impacto nulo o poco significativo (Atkinson *et al.*, 2004; Goodman y Turner, 2010; Martins, 2009). La literatura señala diversos factores que pueden marcar una diferencia en los efectos obtenidos: si el incentivo es individual o grupal, por ejemplo, puede implicar un efecto adicional de *spillovers*, pudiendo ello expandir o reducir el efecto neto (Babcock *et al.*; 2011). Por su parte, el esquema de incentivos puede ser planteado como una recompensa o, más bien, como una penalización (Tversky y Kahneman, 1991; en Levitt *et al.*, 2012). También puede darse el caso de plantear un esquema de recompensas monetarias o no-monetarias (Kosfeld y Neckermann, 2011; en Levitt *et al.*, 2012). Los estudios señalan que la inmediatez o, por el contrario, rezago de los beneficios obtenidos a través del

incentivo, puede afectar los resultados esperados (Levitt *et. al.*, *ibíd.*). Asimismo, es posible una heterogeneidad del efecto según el sexo del individuo incentivado, la edad, o según el área que se esté proponiendo mejorar (por ejemplo, Matemáticas o Lectura).

El principal canal a través del cual se espera que un incentivo monetario promueva un resultado es a través del llamado “efecto competencia”; es decir, el cambio que el bono genera en la conducta de los aspirantes, por el hecho de estar aptos para competir, con la finalidad de ganar el incentivo. No obstante, la definición de cuál será la acción o resultado que desea promoverse también puede resultar clave en la efectividad del incentivo. Por ejemplo, algunos autores sugieren que el incentivo es más efectivo si se plantea promover un insumo en lugar de un resultado, pues de otra manera los individuos pueden no tener claridad sobre los factores de la función de producción que deben seguir para alcanzar el objetivo final (Babcock *et al.*, 2011; Fryer, 2010). Por su parte, algunos sostienen que podría haber un efecto de reciprocidad por parte de los individuos que hayan logrado exitosamente el objetivo, a manera de retribución por el hecho de haber recibido una recompensa por su esfuerzo (DellaVigna y Pope, 2016; Rau y Contreras, 2012; Akerlof, 1984 y 1982).

En el contexto de América Latina, aunque el planteamiento de incentivos monetarios no es un tema de estudio nuevo, aún no ha sido extendido ampliamente como política pública. El caso más consolidado es el Sistema Nacional de Evaluación del Desempeño (SNED), implementado por Chile en el año 1996, y que aplica para los colegios municipales, particulares subvencionados y liceos técnico-profesionales (92% del total de colegios de Chile). Para la definición de los ganadores, los colegios son agrupados según características que reflejen un contexto o características similares, con la idea de hacer más comparable los resultados escolares (por ejemplo, variables como región, área de procedencia, nivel socioeconómico de los estudiantes, etc.). El incentivo es del tipo colectivo, en tanto el 90% del premio otorgado se reparte a los docentes y auxiliares de la escuela, para su uso individual, en función al número de horas trabajadas por semana. El 10% restante se reparte bajo criterio del director, dirigido a los docentes que considere de mejor rendimiento. Diversos estudios han demostrado la efectividad de esta política en el caso chileno (Mizala y Romaguera, 2004; Contreras *et. al.*, 2005; Gallego, 2008; Rau y Contreras, 2012; Casavilca, 2014).

En el Perú, a partir del 2014 se implementó el Bono de Incentivo el Desempeño Escolar (BDE), “Bono Escuela”, el cual es un esquema de incentivos que tiene como objetivo otorgar un reconocimiento económico anual a los docentes de instituciones educativas públicas que hayan obtenido los mejores resultados en indicadores educativos predefinidos. Tales indicadores se encuentran asociados a la superación (mejora en los resultados de la Evaluación Censal de Estudiantes – ECE- con respecto al año anterior), efectividad (promedio destacable en la última ECE), retención (estudiantes permanecen en la educación formal hasta el cierre del año escolar), registro y gestión de asistencia (estudiantes figuran en los sistemas de información oficiales del sector), entre otros.

Al igual que en el caso chileno, para el BDE los colegios públicos son agrupados en estratos más homogéneos, con el objetivo de aumentar su comparabilidad. Así, en un primer nivel, los colegios son divididos en grupos, según el tipo de ECE a la que hayan aplicado (segundo grado de primaria, cuarto grado EIB, o segundo grado de secundaria) y, en un grupo diferenciado, los colegios que por razones externas a ellos no hayan rendido la prueba. A su vez, dentro de cada grupo se plantean estratos. Tales estratos

han sido definidos bajo los criterios de característica de la escuela (polidocente completo, multigrado/unidocente) y de dependencia de una misma gestión (local o regional, en función del tamaño de la jurisdicción), entre otros. Habiendo definido los estratos, se premia a las instituciones educativas públicas que se encuentren en el percentil 35 superior. El monto asignado representa en promedio un 10% del ingreso anual de un docente. Debe precisarse que el BDE es un incentivo del tipo colectivo, pues los beneficiarios son toda la plana docente del nivel educativo premiado, si bien se hace una diferenciación en el monto asignado al docente del grado evaluado y al director de la escuela.

Siendo una intervención relativamente nueva en el Perú, no hay a la fecha publicaciones que realicen una evaluación de impacto de este incentivo. Esta investigación plantea realizar un estudio de Bono Escuela y su impacto en el rendimiento estudiantil, teniendo en cuenta que este indicador representa hasta el 65% de los criterios tomados en cuenta para el otorgamiento del bono. Para ello, se evaluará si existe un impacto del bono como producto de la competencia para la obtención del premio (efecto *ex ante* del otorgamiento). También será de interés analizar si existe un impacto del bono posterior al otorgamiento del bono; es decir, entre aquellos que efectivamente obtuvieron el premio (efecto *ex post* del otorgamiento). Cabe precisar que estos efectos también fueron analizados para el caso chileno, donde se encontró una significancia positiva del primer efecto, pero no se demostró evidencia significativa del segundo (Rau y Contreras, 2012).

Adicionalmente, en esta investigación se diferenciará el impacto en los resultados para aquellos colegios que se encuentran en una zona de estricta competencia -entendiendo por ello aquellos colegios que no son ni “seguros ganadores” ni “seguros perdedores” (demasiado distantes del punto de corte que determina el premio, para ambos sentidos)- y que, por tanto, tienen mayores incentivos para esforzarse. Para el caso chileno, esta distinción tuvo una alta significancia (Rau y Contreras, 2012; Casavilca, 2014).

De esta manera, este estudio plantea responder las siguientes preguntas:

- ¿Existe un efecto promedio *ex ante*, producto de la competencia entre escuelas para obtener el premio de Bono Escuela, sobre el rendimiento estudiantil de primaria y secundaria en el área de Comprensión Lectura? ¿Y para el caso de Matemática?
- ¿Existe un efecto promedio *ex post*, producto de la obtención del bono, sobre el rendimiento estudiantil de primaria para el área de Comprensión Lectura? ¿Y para el caso de Matemática?
- ¿Cuál es el periodo de maduración que requiere este incentivo para encontrar efectos relevantes?
- ¿Existe una diferencia en el efecto competencia cuando se analiza únicamente a los colegios en zona de competencia?
- ¿Existe una diferencia en los efectos promedios encontrados según área de procedencia (urbano o rural), y característica de la escuela (polidocente completo o multigrado/unidocente)?

Para medir el primer efecto se realizará, en primera instancia, un modelo de emparejamiento que permita identificar, entre las escuelas privadas, un grupo de control

que mejor pueda compararse al grupo de tratamiento (escuelas públicas), para luego aplicar un modelo de triple diferencia que permita aislar en cierta medida la tendencia previa al tratamiento. Por su parte, para el segundo efecto se aplicará un modelo de regresión discontinua de diseño difuso, que evalúa a los colegios públicos premiados que se encuentren más cercanos alrededor del punto de corte de la premiación, para así evaluar si el hecho de haber recibido el bono en el periodo previo tiene un impacto en el rendimiento. Debe acotarse que para el caso se secundaria, al contar con un periodo de evaluación más corto (la primera evaluación censal se realizó en el 2015), solo se mide el impacto producto de la competencia por el bono, con una variante en la metodología.

Por ello, este documento presenta la siguiente estructura: luego de la sección introductoria, se presenta el marco teórico, donde se planteará el mecanismo causal que justifica el planteamiento de incentivos monetarios en el sector educativo, a través de la teoría del principal agente. Además, en esta sección también se hará una revisión de la evidencia empírica que analiza los mecanismos internos bajo los cuales habrían operados los diferentes incentivos planteados. Posteriormente, en la tercera sección se explica con mayor detalle las características del incentivo de Bono Escuela. En la cuarta sección, se plantea la estrategia de evaluación para responder las preguntas planteadas anteriormente. Finalmente, en la quinta y sexta sección se presentarán los resultados y conclusiones, respectivamente.

2. Marco Teórico

2.1 Teoría del principal-agente y contratos óptimos

Uno de los principales marcos conceptuales para entender la herramienta de incentivos monetarios es la teoría del principal-agente. Como define Levacic (2009), se afirma que existe una relación principal-agente cuando una persona u organización (el “principal”), explícita o implícitamente, contrata a otra –el “agente”- para proporcionar un servicio o proveer un bien. Bajo la premisa de que el objetivo es maximizar el objetivo (bienestar) del principal, la teoría del principal-agente tiene como finalidad establecer el tipo de contrato óptimo en términos del retorno para este principal (Douma y Schreuder, 2002; Levacic, *op. cit.*:1). Esta teoría económica analiza tanto los servicios que se ofrecen a través del mercado, como aquellos que no se rigen estrictamente bajo las leyes de oferta y demanda, tanto en el sector privado como el público.

Principales factores por considerar en la definición de un contrato

Para ello, Levacic (*op. cit.*:2) plantea que existen algunas variables importantes de tener en cuenta para la definición de un contrato que tenga el objetivo de maximizar el nivel de esfuerzo del agente. En primer lugar, está la **naturaleza de la motivación de los agentes**, sobre la cual Le Grand (2003) distingue dos tipos de agentes: los “oportunistas” –es decir, agentes que priorizan su propio bienestar- y aquellos “altruistas”, agentes con un interés afín a los objetivos del principal. Sobre esta característica, Le Grand –que concentra su análisis en el sector público- señala que la predominancia de cierta motivación u otra no está asociada a algún sector productivo en particular, sino que puede cambiar en el tiempo. En consecuencia, él argumenta que el salario y los mecanismos de incentivos que motivan a los empleados en el sector público deben ser estructurados de tal manera que sean consistentes con ambos tipos de

comportamiento, y que no se debe asumir que un agente es del tipo altruista solo por el hecho de trabajar en entidades sin fines de lucro, o similares a estas.

Otra variable que debe tenerse en cuenta es la **posibilidad de riesgo moral**; es decir, el riesgo de que un agente no brinde su mayor esfuerzo ante la poca capacidad del principal para inferir con exactitud tal nivel de esfuerzo a partir de los resultados obtenidos (Dixit, 2002). De acuerdo a Dixit, el problema se complejiza cuando los resultados de una acción, aunque son observables, pueden depender tanto de acciones del agente como de factores aleatorios que escapan al control del agente responsable, y cuya proporcionalidad en la causalidad para el resultado final es difícil de medir. En estos casos, no sería viable asociar completamente el salario a los resultados obtenidos. En casos de riesgo moral, asimetría de información y un agente averso al riesgo (es decir, un agente al que le es más costoso asumir la incertidumbre del producto final), un contrato óptimo distribuirá el riesgo entre el principal y el agente, de tal manera que el agente reciba un monto fijo más un monto variable, sujeto al rendimiento.

Adicionalmente, otro factor relevante para la definición de un contrato es el **nivel y costos de la asimetría en la información**; en otras palabras, qué tan costoso es para el principal revelar el nivel de esfuerzo del agente. Altos costos a menudo están asociados a sectores donde no hay una alta competencia entre agentes (Levacic, *op. cit.* 2009:4). En general, se refiere a contextos donde los costos de supervisión o monitoreo del accionar del agente son difíciles de escalar, sea por la poca disponibilidad de personal con las calificaciones necesarias para cumplir esta función, o por los altos costos de traslado, visitas frecuentes y personal. Finalmente, debe considerarse la **complejidad del sector productivo** en análisis. A menudo, se evidencia que los agentes ejercen múltiples tareas, o responden a múltiples principales, que a su vez tienen objetivos diversos (Levacic, *op. cit.*; Dixit, *op. cit.*). En particular, señala Levacic, el sector público se caracteriza por tener una estructura donde hay muchos principales y muchos agentes, siendo que los principales finales son los ciudadanos que proporcionan sus impuestos a cambio de que los trabajadores de una organización pública (el agente) proporcionen los servicios que la comunidad requiere.

Factores por considerar en la definición de contratos en el sector Educación

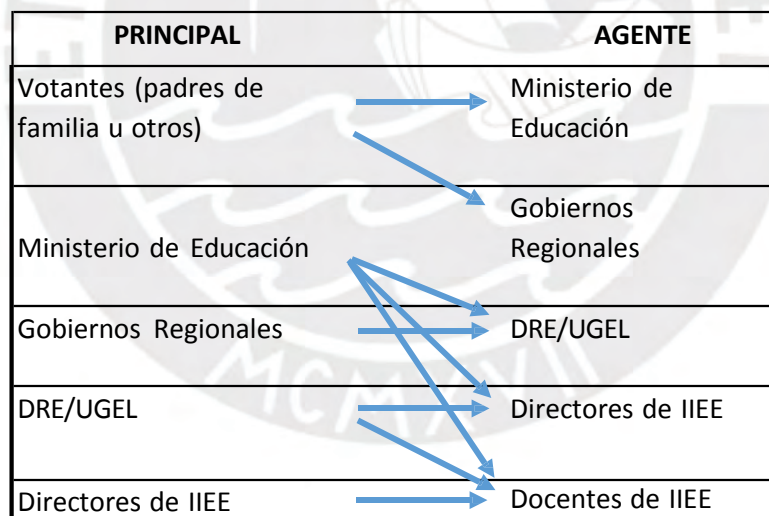
Aplicando la teoría del principal-agente al sector Educación, Dixit (2002) identifica algunas características que pueden ser de gran relevancia al momento de definir un esquema de incentivos. En primer lugar, como ya se mencionó, está el factor de tener **múltiples objetivos**. La educación pública tiene exigencias tales como: i) impartir habilidades básicas de comunicación y lectura, matemáticas, razonamiento y cálculo, etc.; ii) acompañar e influir en el desarrollo emocional y físico de los niños y adolescentes; iii) preparar a los estudiantes para una vida laboral, fortaleciendo sus habilidades vocacionales y aptitudes para una mayor empleabilidad; iv) fomentar en los estudiantes una conciencia crítica y responsable, para el ejercicio de su ciudadanía; v) dotarlos de conocimientos básicos para el cuidado de su salud y administración financiera; y vi) proveer un ambiente libre de violencia y drogas. Como señala Dixit, si bien estos objetivos no son mutuamente contradictorios en ningún caso, estos sí compiten por una priorización dadas las restricciones de personal y de tiempo disponible, y en ese sentido sí tienen un grado de sustitución.

En segundo lugar, como señaló Levacic, este sector se caracteriza por tener

múltiples partes interesadas, que representan principales externos que, a través de su influencia en el plano político, participan en la relación principal-agente del sector. Así, el primer principal serían las familias, padres (y contribuyentes a nivel nacional y local) que esperan que el servicio público les brinde una educación de calidad, con la cual sus hijos puedan alcanzar las competencias esperadas para su edad. No obstante, en esta cadena jerárquica aparecen otros protagonistas, que alternan su condición de agentes y principales dentro de la administración pública. Por ejemplo, en el Perú, el Ministerio de Educación, como representante del Poder Ejecutivo sobre el sector, responde a las expectativas de las familias y a una visión de país, de igual manera que el Gobierno Regional a un nivel más local. Por su parte, las Direcciones Regionales de Educación (DRE) y Unidades de Gestión Educativa Local (UGEL) rinden cuentas tanto a los gobiernos regionales -de los cuales dependen directamente en términos administrativos y presupuestales-, como al Ministerio de Educación, el cual ejerce una rectoría a nivel nacional en lo que compete al sector.

Complementariamente, las DRE y las UGEL de cada región administran la selección, la administración y las remuneraciones del personal directivo y docente de una institución educativa. Finalmente, el director de la IE, en última instancia, representa la máxima autoridad a nivel de la escuela, razón por la cual, aunque no intervenga en la definición de su salario, ejerce también una rectoría sobre el docente. En el Esquema 1 se presenta un esquema de lo descrito previamente:

Esquema 1: Esquema de relación principal-agente en el sector Educación: caso peruano



Elaboración propia. Adaptado de Levacic (2009)

Dixit (*op. cit.*) señala algunos *stakeholders* adicionales, como los empleadores potenciales de los graduados de la educación básica y superior, las escuelas privadas, los grupos que favorecen o se muestran en contra de algún contenido particular del currículo, los profesores y su sindicato, y la sociedad en su conjunto. Cada uno de estos agentes, muestran preferencias y ponen énfasis en diversos objetivos, que en ocasiones compiten entre sí por una priorización.

Una tercera característica del sector educativo sería la de contar con **múltiples periodos**. Con esta definición, Dixit se refiere a la construcción de la carrera profesional del docente,

que es larga, pues es en el transcurso del tiempo que sus habilidades se consolidan y son reveladas. En consecuencia, Dixit sostiene que para los docentes el desarrollo de su carrera puede ser más importante que los incentivos basados en su desempeño de corto plazo, incluso si este puede medirse con un alto grado de precisión.

En cuarto lugar, una característica importante del sector educativo es la **falta de competencia**. Aun cuando se reconoce que los colegios privados –y, en particular, los colegios parroquiales- tienen un grado de competencia con los colegios públicos, en general se evidencia que para una proporción importante de la población, la escuela pública representa un monopolio en la provisión del servicio educativo, que se sustenta principalmente en contextos donde la demanda educativa es pequeña o muy débil, sea por escasos recursos económicos o poco capital político. Finalmente, y altamente relevante, se señala como característica del sector educativo la **motivación de los agentes y su dificultad para observarla o medirla**. Al respecto, como señala Dixit, la docencia es considerada como una profesión que está motivada por la satisfacción de facilitar el aprendizaje a los jóvenes, así como por un interés genuino en el desarrollo infantil. No obstante, también es de esperar que los docentes, al igual que otras ocupaciones, tengan una motivación por el nivel de ingreso, lo cual ha sido estudiado por algunas investigaciones que evalúan la respuesta de los profesores ante variaciones en el salario (Zymelman y DeStefano, 1993; Eide *et. al.* 2004; Chevalier y Dolton, 2005; Glewwe y Kremer, 2006). En ese sentido, este concepto está alineado a lo sostenido por Le Grand (*op. cit.*) y mencionado anteriormente, acerca de no asumir un tipo de motivación en los docentes (“oportunista” o “altruista”), pues este puede variar en el tiempo o de acuerdo a las circunstancias. Asimismo, también se reconoce que la labor docente está caracterizada por el riesgo moral. Ello debido a la dificultad de observar y medir el aspecto motivacional del docente, así como en la dificultad para revelar su nivel de esfuerzo a partir de los productos medibles.

Tipo de contratos tradicionalmente considerados en el sector Educación

Todos estos factores descritos influyen en la eficacia del diseño contractual para lograr los objetivos esperados en el sector; sin embargo, la evidencia muestra que la forma en que tradicionalmente se han diseñado los contratos del sector educativo dista de tener en cuenta tales criterios. En la Tabla 1 se definen los criterios que potencialmente pueden ser considerados en una estructura salarial basada en incentivos. De acuerdo a lo señalado por Mizala y Romaguera (2004), estos indicadores pueden ser clasificados, en primera instancia, bajo un criterio de insumo, proceso o producto. El primer grupo tiene indicadores que señalan las condiciones iniciales del docente o de la escuela, que pueden ir cambiando en el tiempo, como el nivel de experiencia o las horas trabajadas. El criterio de procesos agrupa a los indicadores que se enfocan en las acciones del docente o de la escuela como colectivo, mientras que el grupo de productos se refiere a los indicadores que reflejan los resultados finales que se evalúan en el sector. A su vez, cada uno de estos indicadores puede plantearse a nivel individual o a nivel colectivo en la escuela. En la Tabla 1 se resume la posibilidad de plantear los indicadores en uno u otro nivel. Por ejemplo, criterios como el de las calificaciones y experiencia se caracterizan por evaluarse a nivel individual. Por su parte, el trabajo en alguna área geográfica socialmente desventajosa es una característica de la escuela, por lo que su inclusión como criterio dentro la estructura salarial afectaría a toda la plana docente de la escuela. Los indicadores de proceso o producto se caracterizan porque pueden ser planteados tanto a nivel individual como colectivo.

Como señalan Mizala y Romaguera (2004), el tipo de contrato que ha predominado en el sector educativo en América Latina es el salario fijo, y se ha basado principalmente en criterios de insumo, en función a variables individuales como las calificaciones personales, el nivel de experiencia o el nivel de responsabilidad. Este tipo de criterio suele estar basado en la individualidad de cada docente, a diferencia de otros esquemas donde se combina esta estructura fija con un factor variable, en función de indicadores de procesos o productos obtenidos en el aula del docente (plano individual) o como promedio en la escuela (plano colectivo).

Tabla 1: Criterios para la diferenciación de salarios de docentes

Criterios para la diferenciación de pago a docentes	Incentivos individuales	Incentivos basados en la escuela
Criterios de insumo		
Calificaciones	Sí	
Experiencia	Sí	
Horas trabajadas	Sí	
Grado que se le imparte al estudiante	Sí	
Trabajo en áreas socialmente desventajosas		Sí
Criterios de proceso		
emprender actividades de desarrollo profesional	Sí	Sí
comprenderse con actividades extracurriculares	Sí	Sí
trabajo con padres de familia	Sí	Sí
Criterios de producto		
tasa de deserción	Sí	Sí
tasa de conclusión	Sí	Sí
logro estudiantil	Sí	Sí
reportes de satisfacción del estudiante o los padres	Sí	Sí
evaluación a cargo de inspectores o directores	Sí	Sí

Mizala y Romaguera (2004).

El esquema de incentivos basados en el desenvolvimiento dentro de la escuela como una nueva propuesta de contrato en el sector Educación, y principales factores por considerar en su diseño

Bajo el tipo de contrato actualmente predominante (de pagos fijos y basado en criterios de insumo), el salario del docente no depende del nivel de esfuerzo o resultado que obtenga en el estudiante, y en consecuencia, no se alcanza un contrato óptimo, pues el riesgo de que el docente no brinde su mayor esfuerzo es asumido únicamente por el principal. La literatura empírica aún no es concluyente sobre el esquema ideal en el diseño contractual del sector, aunque sí advierte diversas consideraciones que deben tenerse en cuenta al plantear una nueva estructura salarial. Por ejemplo, debe tenerse en cuenta que, de acuerdo a lo descrito anteriormente, **en casos de riesgo moral, asimetría de información y un agente averso al riesgo** (es decir, un agente al que le es más costoso asumir la incertidumbre del producto final, p.e. el rendimiento estudiantil), un contrato óptimo debe distribuir el riesgo entre el principal (el que plantea el esquema salarial; en

el caso peruano, el Ministerio de Educación) y el agente (docente), de tal manera que el agente reciba un monto fijo más un monto variable, este último sujeto a criterios basados en procesos o productos (ver Tabla 1).

Asimismo, dada la complejidad del sector, para la efectividad del contrato es igualmente importante considerar los otros factores ya mencionados que pueden afectar los resultados esperados con este nuevo esquema salarial. Como señala Levacic (2009), en relación a la **naturaleza de la motivación de los agentes**, si los docentes son del tipo de agente puramente altruista (comparte los mismos intereses del principal), el logro estudiantil alcanzará el mayor nivel posible de acuerdo al conocimiento y habilidades de su docente, independientemente de la estructura y nivel salarial; por el contrario, si el docente es del tipo de agente oportunista, la calidad de la enseñanza impartida al estudiante será baja, de no tener los incentivos adecuados. Por ello, sostiene Levacic, una consideración adicional sobre la mejor forma de establecer el contrato del docente, sería una combinación de **incentivos monetarios (pagos asociados al rendimiento), y no monetarios**, entendido este último como un reforzamiento de la satisfacción personal del docente por el buen ejercicio de su labor. Ello con el objetivo de atraer a la carrera profesionales que serán efectivos y, sobre todo, permanecerán así a lo largo de su vida laboral. Dichos pagos asociados al rendimiento pueden plantearse bajo diferentes esquemas; por ejemplo, de manera lineal o proporcional a los logros obtenidos en el estudiante, o estableciendo un monto fijo luego de superar un umbral deseado.

Asimismo, cuando **múltiples principales** tienen **múltiples objetivos** pero plantean especificaciones en el contrato a partir de productos más fáciles de medir, el riesgo es que aquellos indicadores no evaluados sean descuidados por el agente. Por ejemplo, si se plantea un indicador de producto (como los logros de aprendizajes en el área de comprensión lectora), otros productos adicionales - como el desarrollo personal, cultural o la actividad física del estudiante-, en tanto no sean evaluados, pueden quedar relegados desde el punto de vista del docente. En otras palabras, de tener un perfil de agente prioritariamente oportunista, el docente solo se enfocará en aquellas variables directamente asociadas a las especificaciones del incentivo.

Otro aspecto que debe tenerse en cuenta en el diseño de los pagos asociados al rendimiento es si se definen de manera **individual o colectiva**; es decir, si los beneficios afectan únicamente a los docentes directamente relacionados con el producto o a un grupo de docentes de la escuela. Por un lado, los incentivos individuales pueden desincentivar el esfuerzo colectivo de la plana docente, mientras que, por el contrario, incentivos grupales pueden generar docentes *free-riders*; es decir, docentes que sin ningún esfuerzo pueden obtener el premio a costa del esfuerzo de otros docentes.

No obstante, como señala Babcock et al. (2011:19), existen varios mecanismos por los cuales un incentivo planteado de manera colectiva puede ser beneficioso. Un primer factor social sería la motivación social. Valores como el altruismo, sentimiento de culpabilidad por decepcionar las expectativas de otros, sentimiento de vergüenza por fallar, o el miedo al castigo social de sus pares, han sido considerados por una amplia literatura empírica (Charness y Kuhn, 2011; citado en Babcock et al., 2011). Un segundo factor que explicaría los beneficios de un incentivo colectivo sería las complementariedades potenciales de un trabajo en equipo. Al respecto, Lazear (2000) sostiene que este factor es imprescindible para que un incentivo colectivo sea efectivo, y que no debe utilizarse este tipo de mecanismo en caso no exista dicha

complementariedad. Acotado al sector Educación, la propuesta de un incentivo colectivo reforzaría la necesidad de un liderazgo pedagógico (asumido por el director de escuela o docente de mayores aptitudes), capaz de dar retroalimentación apropiada sobre las prácticas de enseñanza de los docentes, por un objetivo común.

Por último, otra consideración importante que Levacic resalta es la importancia de evaluar las limitaciones que puede tener el indicador escogido como criterio de incentivo, refiriéndose a si el indicador recoge exactamente lo que el principal desea obtener, y si el sistema vigente asegura la calidad en el recojo de tal información. En consecuencia, es importante establecer un sistema nacional de mediciones con un alto grado de confiabilidad y comparabilidad a lo largo del tiempo, que además recoja información del contexto del estudiante, de tal manera que periódicamente se pueda evaluar el desempeño del docente y su aporte marginal al rendimiento del estudiante.

En las siguientes secciones se hará un recuento de los mecanismos internos que se han planteado y evaluado en diversas investigaciones al momento de diseñar un esquema de incentivos, y las conclusiones obtenidas en ellas. Asimismo, se revisará la evidencia de incentivos monetarios planteados en el sector educativo, y las evaluaciones de impacto alrededor de estos, tanto a nivel internacional como regional.

2.2 Evidencia empírica sobre los mecanismos internos para el logro de resultados a través de incentivos

Varios estudios han planteado diversas hipótesis sobre cómo el mecanismo de incentivos funciona en diferentes ámbitos, público y privado. En esta subsección se mencionan los principales mecanismos que operarían detrás de un esquema de incentivos:

Efecto de la competencia por alcanzar el bono: “efecto competencia”

Se considera un efecto *ex ante*, pues comienza a operar una vez se anuncia las condiciones para el otorgamiento del incentivo. Bajo la premisa de que el incentivo opera a modo de un “torneo” por alcanzar el incentivo monetario, este factor competitivo contribuiría a tener agentes más motivados y dispuestos a ofrecer su mayor esfuerzo por alcanzar la meta. En ese sentido, todas las evaluaciones de impacto de este tipo de incentivo monetario analizarían implícita o explícitamente este efecto.

Sobre los incentivos monetarios en el sector educativo a nivel regional, existe el Sistema Nacional de Evaluación por Desempeño (SNED) que funciona en Chile desde el año 1996. Esta evaluación se aplica cada dos años a los colegios municipales y subvencionados. Se agrupan escuelas bajo un criterio de comparabilidad y se mide el logro estudiantil por cada subgrupo. Para medir el logro estudiantil se construye un índice de indicadores ponderados relacionados a los resultados en la prueba estandarizada y aplicada a nivel nacional, así como otros indicadores de menor peso, como la tasa de aprobación y retención de matrícula. Una vez asignado el bono (aprox 6% del salario anual), el 90% se reparte entre los profesores en función al número de horas trabajado, y el 10% restante se distribuye de acuerdo al criterio del director de la IE.

Los estudios de impacto para este incentivo se han centrado en evaluar el efecto de recibir el incentivo sobre los resultados en el rendimiento. Por ejemplo, Mizala y Romaguera (2004) encuentran un impacto positivo al analizar las primeras tres rondas

del SNED, identificando que el periodo de maduración sería a partir de la tercera ronda (sexto año de implementación). Por su parte, Contreras et al. (2005) y Gallego (2008) utilizan datos de la primera aplicación del SNED e identifican un impacto positivo aunque poco robusto, en el primer caso, y positivo pero menor, en el segundo caso. Asimismo, Rau y Contreras (2012), plantean explícitamente este mecanismo por el cual operaría el incentivo monetario a docentes en el caso de Chile. Al enfatizar la existencia de este efecto, Rau y Contreras proponen la necesidad de aislar a aquellos agentes que denominan “seguros ganadores y seguros perdedores”; es decir, a aquellos individuos que, en el caso de los “seguros perdedores” podrían percibirse muy lejos de la probabilidad de alcanzar la meta, o, en el caso de los “seguros ganadores”, estarían muy confiados en que con el mismo o un mínimo esfuerzo mayor (incluso menor) mantendrán su posicionamiento en el *ranking*. Al acotar el análisis bajo este criterio, encuentran un efecto positivo del bono (0.3 desviaciones estándar).

Casavilca (2014), por su parte, realiza un estudio longitudinal utilizando información desde 1996 para evaluar si el cambio en el umbral de desempeño para clasificar al bono en el 2006 –pasó de premiar al 25% de docentes, a premiar al 35%– tuvo un efecto en el rendimiento estudiantil. Encontró que esta medida favoreció el logro estudiantil, y afectó en un 75% la efectividad en las áreas evaluadas de Comprensión Lectora y Matemática. Asimismo, encuentra indicios de que la teoría alrededor de los “seguros ganadores y perdedores” de Rau y Contreras (*op. cit.*) es consistente en su análisis, pues se observó que el grupo de escuelas que estaban ligeramente por encima del umbral inicial (25%) tuvieron menor efectividad sobre el rendimiento en relación al resto de colegios, sugiriendo así que existiría un efecto de “seguros ganadores”.

De igual manera, en Estados Unidos también se han implementado diversos programas de salario basado en el desempeño docente. Por ejemplo, en Carolina del Norte, desde el año 1996 se aplica para todos los colegios públicos un esquema de pago asociado al desempeño. Se otorga un bono a los profesores si el puntaje de la escuela alcanza un umbral definido, o el porcentaje de mejora es significativo. Cada docente recibe \$ 1,500 adicional como máximo. Además, este sistema es parte del sistema de rendición de cuentas del Estado. Vigdor (2008) mostró un impacto importante en los resultados de la prueba estatal para todas las áreas, sobre todo en el área de Matemática.

Asimismo, Hudson (2010) evalúa un esquema de salario por desempeño en 17 estados. En este estudio se observó un incremento en el desempeño en matemática, en 0.15 desviaciones estándar más con respecto al grupo de control. En Comprensión Lectora se evidenció un impacto positivo aunque menor. Por su parte, Winters *et al.* (2008) analizó un programa en Arkansas implementado desde el 2004 al 2007, en el cual los profesores fueron premiados individualmente sobre la base del rendimiento de sus estudiantes. El bono se otorgó en función al nivel de mejora y el número de estudiantes a cargo del profesor. El estudio evidenció una alta significancia estadística en el impacto del programa, en Matemática (0.16 d.e.); Lectura (0.15 d.e.) y Lenguaje (0.22 d.e.). Se observó además que la mejora fue mayor en el caso de los profesores menos efectivos inicialmente. De igual forma, Springer *et al.* (2010) evalúa el programa implementado en Nashville entre el 2006 y el 2009, donde los profesores de quinto a octavo grado participaron voluntariamente en un experimento de bono por desempeño. En el estudio se observó un efecto positivo en el área de Matemática para quinto grado, en el segundo y tercer año del programa, pero esta mejora no perduró en el siguiente año. En los demás grados participantes no se observó algún efecto.

Por su parte, Angelucci *et al.* (2015), aunque enfocan su análisis de incentivos monetarios en el sector de consumo de comida saludable y fomento de la lectura, también hacen referencia a un efecto competencia dentro de un incentivo monetario, planteado como un efecto de señalización. Al respecto, los autores señalan que la señalización, en algunos casos, puede dar información sobre la dificultad de la tarea o la calidad del bien incentivado, o en otros casos, puede hacer sentir al individuo bajo una situación de control. Como ejemplo del efecto ambiguo de la señalización, plantean que, al fomentar la lectura, los efectos de señalización pueden operar a través de dos opciones: i) observando el comportamiento de los pares (“quizá mis pares leen porque es divertido”), u ii) observando los incentivos detrás de la nueva conducta de los pares (“a mis pares los motivan para leer porque leer es difícil”). A partir de los resultados de su investigación, Angelucci *et al.* (*op. cit.*) concluyen que cuando los incentivos de los pares son evidenciados (todos los participantes saben a qué se debe el mayor esfuerzo del agente incentivado), los efectos del incentivo son positivos, y que, de acuerdo a su observación, estos se deben a un sentimiento de anticonformismo y de envidia hacia los incentivos ganados por los pares que fueron expuestos al incentivo monetario; es decir, lo que podría asociarse también a un efecto competencia.

Efecto de plantear el incentivo de manera individual o colectiva

Como se señaló anteriormente, los incentivos monetarios pueden plantearse de tal manera que cada agente es premiado por su esfuerzo individual, o de manera que la obtención del bono esté sujeta a los resultados de un grupo de agentes; por ejemplo, todos los docentes de una escuela. Babcock *et al.* (2011), señalan que en círculos como la empresa privada, el militar, salud o programas de bienestar, es común ver incentivos basados en el premio colectivo. En su investigación, se aplica el incentivo colectivo en dos experimentos: pago por estudiar (asistir a la biblioteca) y pago por hacer ejercicio (asistir al gimnasio). Los hallazgos de Babcock *et al.* sostienen que los efectos sociales de un incentivo monetario son relevantes y pueden inducir a un mayor esfuerzo por costos menores que un incentivo individual directo. Por ejemplo, se evidenció que los resultados de asistencia a la biblioteca fueron mejores en los casos donde hubo un incentivo colectivo, y además, donde los individuos eran conscientes de su grupo, sugiriendo que al fomentar el sentido de equipo se obtiene un impacto mayor.

Por su parte, en relación el sector educativo, Goodman y Turner (2010) hicieron una evaluación de impacto sobre un programa de Nueva York en 181 escuelas públicas con alto nivel de pobreza, escogidas aleatoriamente, sobre un esquema de incentivo grupal de bonos asociados al rendimiento, en dos periodos académicos (2007 y 2008). Los productos medidos fueron el esfuerzo docente (ausentismo en aula), rendimiento estudiantil en Matemática y Lectura, y actividades en clase (encuestas a estudiantes y docentes sobre el clima en aula). Se trabajó una metodología para controlar el efecto de posible *free-riding* (la posibilidad de que algunos o todos los docentes confiaran en el esfuerzo del otro para obtener el bono), y luego de ello se concluyó que no hay evidencia clara de que el programa tuviera impacto en el rendimiento estudiantil. Sin embargo, sí se observó una reducción significativa en el nivel de ausentismo docente, sobre todo en aquellas escuelas menos afectadas por un factor de *free-riding*.

Por otro lado, Glewwe, Ilias, y Kremer (2010) realizaron un estudio de un programa piloto implementado en dos distritos de Kenia, en 1998 y 1999. Esta intervención consistía en otorgar premios a los docentes de una escuela que en promedio obtuviera mejores

resultados o mayores avances en ciertas áreas evaluadas. Al ganar, se dividía el bono equitativamente entre los profesores de la escuela que trabajaron con grados del 4° al 8° de primaria. El programa incrementó la participación en el examen. No incrementó puntajes en el primer año, pero en el segundo año los resultados del grupo de tratamiento crecieron en 0.14 desviaciones estándar con respecto al grupo de control. No obstante, el efecto no se mantuvo luego de finalizar el programa.

En otro estudio, Lavy (2004) evalúa un piloto implementado en el año 2000, en 49 escuelas de preparatoria. El incentivo era individual y evaluaba el porcentaje de estudiantes que aprobaban el examen de salida de preparatoria, así como el puntaje promedio obtenido por los alumnos en las áreas de Inglés, Hebreo y Matemáticas. El estudio mostró que hubo un impacto positivo significativo en las áreas evaluadas, lo cual se debió a los cambios realizados en los métodos pedagógicos, a un incremento de horas después de clase y a una mejora en la capacidad de respuesta a las necesidades de los estudiantes. Asimismo, la investigación realizó un análisis de costo eficiencia de otros tres programas piloto con el mismo objetivo de lograr un incremento en el rendimiento estudiantil: un programa de pago asociado al desempeño pero diseñado como un incentivo colectivo; un programa que dirigía el bono al estudiante; y un programa dirigido a dar refuerzo escolar a los alumnos con menor rendimiento. De acuerdo a los hallazgos de Lavy, el programa original mostró resultados más costo-efectivos que las intervenciones alternativas.

Por su parte, Muralidharan y Sundararaman (2010) evaluaron un programa piloto en la India que inició en el año 2005, donde 100 escuelas recibieron un bono basado en su rendimiento estudiantil bajo el esquema de un incentivo individual, y otras 100 escuelas lo recibieron basado en un esquema de incentivo colectivo. Bajo ambos esquemas se premió a las escuelas que excedieran el promedio de logro estudiantil en más del 5%. Después de 2 años, el estudio evidenció que, si bien en el primer año ambos esquemas mostraron un impacto positivo en el rendimiento, durante el segundo año el esquema de incentivo individual superó los resultados bajo el bono colectivo. Así, los estudiantes de las IIEE con un incentivo monetario individual mejoraron su rendimiento en 0.28 desviaciones estándar (d.e.) más en Matemática y 0.17 d.e. más en Lenguaje, con respecto al grupo de control. Los estudiantes de estas IIEE también respondían mejor las preguntas de diferentes niveles de dificultad, y mostraron mejores resultados en cursos no evaluados, sugiriendo así la presencia de *spillovers*.

Efecto de plantear incentivos monetarios (como premio o penalidad) o no monetarios

La necesidad de evaluar los efectos de una intervención que plantea incentivos monetarios o no monetarios está muy vinculado a la naturaleza de la motivación del agente, de acuerdo a lo que se planteó en la subsección anterior. Así, según sostiene Le Grand, un incentivo monetario podría ser contraproducente en los casos donde el agente es del tipo altruista, al distorsionar la intención inicial y premiar a aquellos que “solo” estarían motivados por fines económicos. Sin embargo, Le Grand también sostiene que la naturaleza motivacional de los agentes no debe asumirse en función del sector en el que se desenvuelva el agente. Por ejemplo, aunque se pensaría que los agentes de sectores con fines sociales podrían ser automáticamente considerados del tipo altruista, dicha generalización es precipitada, pues la motivación de los agentes es cambiante en el tiempo y responde a circunstancias particulares.

Un caso que podría mencionarse como un ejemplo de intervención fallida de aplicar incentivos monetarios en el sector Educación es el plan de reformas de escala de pagos en Portugal, impulsada en el 2007. Esta reforma implicó un cambio en la estructura salarial de los docentes de instituciones educativas públicas, que pasó de ser fija a tener un componente asociado al desempeño. El desempeño implicaba no solo el progreso en el rendimiento estudiantil, sino variables adicionales como la retroalimentación de los padres sobre el profesor, el nivel de asistencia a clases por parte del profesor, asistencia a capacitaciones, involucramiento en investigación y cumplimiento de obligaciones administrativas y pedagógicas. El ascenso a escalas más altas estaba sujeto no solo al cumplimiento de estos criterios sino a la disponibilidad de vacantes, definida por el Ministerio de Educación. De acuerdo a Martins (2009), se utilizó a las escuelas privadas y a las escuelas de dos provincias (que implementaron la reforma de manera más “suavizada”) como grupo de control, y no se mostró evidencias de un impacto sobre el rendimiento estudiantil; por el contrario, el efecto fue negativo y hubo un incremento en la carga administrativa de la escuela, además de una insatisfacción laboral por parte de los docentes.

De igual manera, existen estudios con resultados mixtos. En Inglaterra, a partir del año 2000 se efectivizó una reforma educativa que consistió en cambiar el esquema de salario fijo a uno que implicaba un incremento en el monto fijo más un componente variable en función al desempeño (demostrar haber logrado una mejora en el rendimiento de sus estudiantes). En Lenguaje y Ciencia se evidenció una mejora importante, pero no mostró resultados en el área de Matemática, de acuerdo a la evaluación realizada por Atkinson *et. al.* (2004).

Por otro lado, DellaVigna y Pope (2016) analizan el efecto de aplicar incentivos monetarios y no monetarios con el objetivo de obtener un mayor esfuerzo laboral. En primer lugar, los autores encuentran que los incentivos monetarios tienen un efecto fuerte y monotónico en la motivación (si es mayor el incentivo, hay un mayor esfuerzo, de manera proporcional). Sin embargo, este no sería el caso de los incentivos no monetarios, pues aunque sí se observó un efecto positivo en el esfuerzo a cambio de donaciones para más necesitados, dicho esfuerzo no se incrementó al proponer donaciones más altas dirigidas a terceros. Asimismo, evidenciaron que los incentivos no monetarios tienen un efecto moderado sobre los trabajadores, y que, al compararlo con los efectos de un incentivo monetario, este último tiene un mayor efecto, incluso con un pago mínimo.

Asimismo, Lavy analiza los esquemas de pago asociados al desempeño en Israel. En Lavy (1999) se analiza el periodo de dos años de un piloto iniciado en 1995 en 62 escuelas de educación secundaria. Las instituciones educativas se ordenaban en función de la mejora en el desempeño, ajustado por el contexto particular de los estudiantes, y se premiaba al tercio superior. El 75% del bono se destinaba como incentivo monetario directo a los profesores, y el monto restante iba a las instalaciones de la escuela. Los criterios de desempeño fueron la tasa de deserción estudiantil, el tiempo de estudio de los estudiantes, y la tasa de aprobación en el examen de apertura. El estudio mostró clara evidencia de resultados positivos en los indicadores medidos. Adicionalmente, se compararon los efectos con un programa similar que en lugar de otorgar un bono al maestro, destinaba el 100% del incentivo a mayores recursos educativos para la escuela. Lavy encontró que, si bien el último programa también resultó efectivo, fue más costo efectivo el incentivo monetario destinado al profesor, dando indicios con ello que los

docentes podrían tener una motivación económica que supera los principios altruistas (esforzarse por más recursos para la escuela).

Por su parte, otras investigaciones realizan estudios bajo un diseño experimental con el objetivo de concluir si un incentivo monetario es más efectivo cuando es planteado de manera positiva (a manera de premio tras conseguir un objetivo) o de manera negativa (como un castigo o pago de penalidad en caso de no lograr la meta). Cawley y Price (2009) analizaron los mecanismos detrás de un incentivo monetario para la reducción de peso. En este caso, se probó un mecanismo de premiación con metas trimestrales y de penalización en caso de incumplimiento del objetivo transcurrido un año. Sus resultados muestran que hubo mejor respuesta en el grupo que tuvo la combinación de los dos incentivos: la penalización de perder su bono anual y el sorteo por el bono trimestral.

Efecto de condicionar el incentivo según criterios a nivel de procesos o de productos

Por otro lado, es importante resaltar que la decisión de qué criterio utilizar (basado en características del agente, de los procesos o los productos alcanzados) es sumamente importante. Los estudios señalados previamente se caracterizan en general por condicionar el incentivo a nivel de producto (por ejemplo, el rendimiento estudiantil), y en muchos casos se ha demostrado su efectividad. No obstante, en el estudio mencionado de Babcock et al. (2011), por ejemplo, los incentivos van hacia un criterio de proceso (asistir a la biblioteca, asistir al gimnasio) en lugar de un producto, pues se postula que incentivar un proceso puede llegar a ser más efectivo que incentivar el resultado final, dado que en el último caso los individuos incentivados pueden tener poco claro cuáles son los canales de la función de producción para lograr el objetivo final. Los hallazgos del estudio de Goodman y Turner (2010) mencionado previamente también irían en línea con este planteamiento, pues aunque en el estudio no se evidenció un efecto del incentivo monetario sobre el rendimiento estudiantil, sí se observó una reducción significativa en el nivel de ausentismo docente, un indicador que podría considerarse a nivel de proceso en lugar de resultado.

De igual manera, Fryer (2010) plantea una evaluación experimental sobre la base de doscientos cincuenta escuelas urbanas, para analizar el efecto de incentivos monetarios dirigidos al estudiante sobre el rendimiento escolar. En su estudio, se encontró que los estudiantes lograron mejoras cuando el premio se definía en función a criterios de procesos, como leer una lista proporcionada de libros acordes a su edad, asistir a la escuela, cumplir normas de conducta, cumplir las tareas, aprobar exámenes intermedios, etc. Por el contrario, en los casos en los que el incentivo se condicionó a criterios de productos, este no obtuvo ningún efecto sobre el rendimiento estudiantil.

Efecto de haber recibido el bono: "efecto reciprocidad"

El segundo efecto operaría de una manera *ex post* y solo entre aquellos que obtuvieron el bono. Este efecto es explorado inicialmente por Akerlof (1982, 1984, en Rau y Contreras, *op. cit.*), y desarrollado por Rau y Contreras en su evaluación del incentivo monetario a docentes en el caso chileno. A lo que se refiere es al sentimiento de retribución que surgiría en un agente de perfil principalmente "altruista", como producto de haber recibido un premio. Rau y Contreras no encontraron significancia estadística de este mecanismo. Por su parte, en la investigación de DellaVigna y Pope (2016) los autores también analizaron la posibilidad de encontrar un efecto reciprocidad; es

decir, un esfuerzo mayor como retribución u agradecimiento por obtener el premio. Al respecto, lo que hallaron fue que hubo un efecto positivo, aunque reducido.

3. El Bono de Incentivo al Desempeño Escolar, “Bono Escuela”

Habiendo revisado el marco teórico sobre el cual se postula la efectividad de los incentivos monetarios como mecanismo para obtener mejores resultados en el sector educativo, se explica a continuación el incentivo monetario implementado en el Perú, denominado “Bono Escuela”. En este incentivo, el principal es el Ministerio de Educación, y el agente del contrato es el docente. Además, este bono se caracteriza por haber sido definido como monetario y positivo (premio económico), de tipo colectivo (se premia a todos los docentes de las escuelas ganadoras), y condicionado según un criterio a nivel de producto (sujeto principalmente al logro de aprendizajes y retención escolar).

Entrando a detallar, el Bono de Incentivo al Desempeño Escolar (BDE, “Bono Escuela”) es un incentivo monetario a cargo del Ministerio de Educación (MINEDU), anunciado y ejecutado por primera vez en el Perú desde finales del 2014, de manera retroactiva, sobre la base de los resultados en la ECE 2012 y 2013. Está dirigido a los directores, personal jerárquico¹ y docentes de las instituciones educativas públicas del nivel Primaria. Adicionalmente, a partir del 2016, se ha ejecutado este mecanismo de incentivos para el nivel Secundaria, sobre la base de los resultados en la ECE 2015.

Tiene el objetivo de incentivar un mayor esfuerzo por parte de los docentes y directivos para la mejora de resultados y procesos específicos, a nivel de la institución educativa. Así, la asignación del bono se realiza en función a un índice de desempeño escolar (IDE), construido a partir de la ponderación de cada uno de estos objetivos²: i) *superación*, asociado al logro estudiantil de aprendizajes en las áreas evaluadas, busca premiar la mejora (incremento con respecto al año anterior) en el puntaje promedio de la escuela; ii) *efectividad*, asociado al logro estudiantil de aprendizajes en las áreas evaluadas, busca premiar los mejores puntajes alcanzados como promedio de la escuela; iii) *gestión de estudiantes*, busca evitar la deserción escolar (mide la tasa de retención de estudiantes para todos los grados del colegio), así como asegurar un buen nivel de asistencia de los estudiantes; iv) *registro*, asociado al registro oportuno de los estudiantes y sus calificaciones, por grado impartido, en las fuentes oficiales de información del MINEDU; v) *gestión administrativa*, asociado al cumplimiento de ciertos procesos, como la conformación del consejo educativo institucional, rendición de cuentas del presupuesto asignado a la escuela, y aparatos sanitarios en buen estado; y vi) *acciones pedagógicas*, relacionado a la percepción promedio de los estudiantes en el aula acerca de las prácticas de enseñanza que realizan sus docentes, así como de la calidad del clima

¹ Se considera personal jerárquico a docentes que se desempeñan como asesor de área, jefe de taller, de laboratorio, de campo o de áreas funcionales equivalentes.

² Criterios definidos para el 2016 mediante DS 203-2015-EF “Decreto Supremo que modifica el Decreto Supremo N° 287-2014-EF y modificatoria que aprueba el monto del Bono de Incentivo al Desempeño Escolar, así como los criterios, requisitos y condiciones para su otorgamiento, y dicta otras disposiciones”. Se describe la información con respecto a la normativa del año 2016, aunque se identifican ligeras variaciones en la construcción del índice para años anteriores.

escolar.

Con respecto a la elegibilidad, los criterios para aplicar al bono tienen una lógica de incentivar la participación de las IIEE en la Evaluación Censal de Estudiantes, sin dejar de considerar a aquellas IIEE que por factores externos no rinden la evaluación³. En ese sentido, se han definido tres grupos para el nivel de primaria y dos grupos para el nivel secundaria. En el caso de primaria, en primer lugar se define el *Grupo A*, conformado por las escuelas que fueron programadas para aplicar la Evaluación Censal de Estudiantes (ECE) en segundo grado de primaria, bajo la modalidad de Educación Básica Regular. Todas las escuelas pertenecientes a este grupo son evaluadas en Bono Escuela a partir un IDE conformado por los indicadores de superación (45%), efectividad (20%), gestión de estudiantes (30%) y registro (5%), descritos en el párrafo precedente.

Asimismo, se define el *Grupo B*, como el conjunto de escuelas que implementan la Educación Intercultural Bilingüe (EIB) y que, por tanto, fueron programadas para rendir la ECE en cuarto grado de primaria, bajo la modalidad de castellano como segunda lengua. Los indicadores y las ponderaciones utilizadas para las escuelas de este grupo son las mismas que las del Grupo A. El tercer caso es el *Grupo C*, conformado por todas las instituciones educativas que no fueron programadas para rendir la ECE en los dos años previos al otorgamiento del bono. Por esta razón, no pueden ser evaluadas en los indicadores de superación y efectividad; sin embargo, este grupo es medido a través de sus resultados en los indicadores de gestión de estudiantes (50%), registro (10%) y gestión administrativa (40%).

Para el caso del nivel Secundaria, las escuelas han sido definidas en dos grandes grupos. El primero de ellos, denominado *Grupo D*, está conformado por las escuelas que fueron programadas para rendir la ECE de segundo grado. En este caso, las escuelas son evaluadas en los indicadores relacionados a la superación (35%), efectividad (15%), gestión de estudiantes (25%), registro (5%) y acciones pedagógicas (20%). Finalmente, se define el *Grupo E*, como el conjunto de aquellas escuelas que no fueron programadas para rendir la ECE. Por ello, no son evaluadas en los indicadores de superación, efectividad y prácticas pedagógicas⁴; sin embargo, el IDE se construye a partir de los indicadores de gestión de estudiantes (50%), registro (10%) y gestión administrativa (40%).

Complementariamente, para el otorgamiento del bono se definen estratos que tienen por finalidad definir subgrupos más homogéneos entre sí, y de esta manera poder establecer un *ranking* de escuelas en función del IDE alcanzado dentro de un subgrupo de mayor comparabilidad. Por ejemplo, para el caso del Grupo A, las escuelas son agrupadas, en primer lugar, bajo una lógica territorial, por la Unidad de

³ Al respecto es importante mencionar que, de acuerdo al "Reporte Técnico de la Evaluación Censal de Estudiantes (ECE 2015)" del Ministerio de Educación (2015:14), se excluye a los estudiantes que asisten a una IE con menos de cinco estudiantes matriculados en el grado evaluado. En primaria, esta exclusión representa alrededor del 5 % de los estudiantes a nivel nacional, y aproximadamente el 37 % de escuelas a nivel nacional. En cambio, en secundaria, la exclusión representa el 0,4 % de los estudiantes a nivel nacional, y el 7% de escuelas a nivel nacional.

⁴ Cabe precisar que el indicador de acciones pedagógicas se construye a partir de un cuestionario incluido en la ECE de segundo grado de secundaria, en el cual los estudiantes describen la frecuencia con la que observan determinadas prácticas en sus docentes, así como una percepción del clima escolar. Dado que los estudiantes del Grupo E no rinden la ECE, tampoco completan el cuestionario.

Gestión Educativa Local (UGEL) a la que pertenecen⁵ y, en segundo lugar, en función de su característica (polidocente completo o multigrado/unidocente). De igual manera, en el Grupo B también se clasifican a las escuelas bajo una lógica territorial y de característica de la IE, con la diferencia de que dicha lógica territorial se define solo a nivel de Dirección Regional de Educación⁶ (DRE) y ya no a nivel de UGEL. Para el caso de las escuelas del Grupo C, las escuelas son sub-agrupadas únicamente bajo una lógica territorial, a nivel de UGEL.

Por su parte, para el caso del Grupo D, del nivel Secundaria, el primer criterio que se ha considerado para agrupar a las escuelas es en función de si estas fueron parte del nuevo modelo de secundaria que está siendo impulsado por el gobierno central, o son parte del modelo regular⁷. En el caso en que la escuela sea parte del nuevo modelo, las escuelas son sub-agrupadas adicionalmente en función de la DRE a la que pertenecen; mientras que, si son parte del modelo regular, las escuelas son sub-agrupadas en función de la UGEL a la que pertenecen, y también en función de si pertenecen al área urbana o rural. Por último, las escuelas del Grupo E son definidas únicamente en función de la DRE a la que están adscritas.

Habiendo definido los estratos de escuelas comparables entre sí y la forma en que se construirá el indicador de desempeño escolar, se establece un ranking a partir del cual se premia a las escuelas que se encuentran en el percentil 35 superior de su estrato, en el caso de primaria, o que se encuentran en el percentil 20 superior, en el caso de secundaria. Es importante precisar que tanto para primaria como para secundaria, se hace una distinción en términos del monto asignado para el caso de las escuelas que se encuentren en el percentil 25 o percentil 10, respectivamente. Adicionalmente, es importante considerar que Bono Escuela es un incentivo del tipo colectivo; es decir, que premia el esfuerzo grupal de los docentes y director de una escuela, y no solo al docente del grado evaluado. El monto asignado oscila entre los S/ 1,000 y los S/ 3,000, dependiendo del nivel de responsabilidad directa sobre los resultados y el grupo elegible⁸.

4. Estrategia de identificación y evaluación

En esta sección se explica la metodología que se usó para la evaluación de impacto de Bono Escuela. En primer lugar es importante apreciar que, siendo una intervención universal (aplicada de la misma manera a todas las escuelas públicas a nivel nacional), no es posible evaluar su efectividad en términos de los siguientes mecanismos: ser un

⁵ Las UGEL son unidades descentralizadas de administración, incluida la administración financiera en la mayoría de los casos. En general se puede afirmar que las UGEL se definen a nivel provincial, salvo para el caso de Lima Metropolitana donde, al concentrar una cantidad importante de escuelas, las UGEL se definen también bajo una lógica territorial, pero agrupando cierto número de distritos aledaños.

⁶ Las DRE también representan una unidad descentralizada de administración, salvo que su jurisdicción no se define a nivel provincial como las UGEL, sino a un nivel superior (nivel regional). Las UGEL tienen una dependencia con respecto a la DRE en el plano normativo, aunque no en el plano administrativo.

⁷ El nuevo modelo de educación secundaria, denominado Jornada Escolar Completa, es una nueva definición de educación secundaria que inició en el 2015 con 1,000 escuelas a nivel nacional y que, a diferencia del modelo regular, incluye mayores horas pedagógicas, más horas destinadas al idioma inglés, mayor número de profesionales de enseñanza, etc.

⁸ Aunque el incentivo es del tipo colectivo, para el caso de primaria se establece una diferencia en la asignación correspondiente al docente directamente responsable del grado evaluado, así como para el director de la IE. Por su parte, los docentes responsables del grado previo al evaluado también reciben un monto diferenciado, aunque ligeramente menor. Finalmente, los docentes de los demás grados, así como los docente del nivel sin grado asignado, también reciben un incentivo, aunque ligeramente menor. Se muestra el detalle de la asignación por grado y según grupo elegible, en la Tabla 4, en la sección Anexos.

incentivo colectivo en lugar de individual; y monetario positivo (premio) en lugar de negativo (penalidad) o no monetario. Por otro lado, aunque Bono Escuela plantea diversos indicadores sobre los cuales se condiciona el incentivo (revisar capítulo 3 del presente documento), puede considerarse que es un incentivo basado en indicadores a nivel de producto, en tanto el mayor peso para el otorgamiento del bono responde a los resultados en el rendimiento estudiantil y la retención escolar.

De los indicadores mencionados, la evaluación de impacto que se planteará en esta investigación se acota únicamente a los efectos sobre el rendimiento estudiantil. Ello debido a que este indicador (dividido en los objetivos de *superación* y *efectividad*) concentra hasta el 65% del peso en el índice elaborado para el otorgamiento del bono en el caso de primaria, y hasta el 50% para el caso del nivel secundario. Además, en términos metodológicos, es el indicador con mejor información longitudinal (la Evaluación Censal de Estudiantes se aplica desde el 2007 en el caso de primaria, y a partir del 2015 para el caso de secundaria), lo cual será importante para el aislamiento de una probable tendencia previa al tratamiento. Por otro lado, se precisa que en esta evaluación se analizarán solo los grupos A (segundo grado de primaria) y D (segundo grado de secundaria). Se excluye a los grupos C y E, al no contar con resultados de la ECE por razones externas a ellos, y tampoco se incluye al grupo B, por la dificultad de definir un grupo de control⁹.

En consecuencia, sobre la base del marco teórico señalado en el segundo capítulo, se analizarán dos de los mecanismos bajo los cuales puede operar el incentivo Bono Escuela: el primero es aquel definido como “efecto competencia”; es decir, el efecto producido por el anuncio del bono, que propiciaría un mayor esfuerzo entre los aspirantes al incentivo. Por otro lado, el segundo mecanismo analizado es el “efecto reciprocidad”, referido al efecto *ex post* al bono, que se produciría entre los docentes que obtuvieron efectivamente el premio, por el hecho de haberlo recibido. Para ambos casos, se medirá el efecto promedio del tratamiento sobre los tratados (ATT, por sus siglas en inglés).

En particular, sobre la metodología utilizada para la evaluación del efecto competencia, se plantea como grupo de tratamiento a aquellas instituciones educativas públicas que se encuentren en zona de estricta competencia; es decir, aquellas entre las que posiblemente haya un mayor efecto del incentivo, como consecuencia de estar más cerca al punto de corte en el índice que define el otorgamiento del bono (y excluyendo así a los “seguros ganadores” y “seguros perdedores” descritos en el marco teórico). Para determinar ello se compararán los resultados excluyendo e incluyendo al grupo de escuelas que encuentran en el extremo del puntaje en el índice.

Acerca del grupo de control establecido para esta evaluación, las escuelas consideradas son las de gestión privada. Sobre ello, es necesario reconocer las diferencias entre el grupo de tratamiento y de control como una primera fuente de variación para la medición del impacto del incentivo (ver Tabla 7 en el Anexo) y un factor limitante importante para obtener un efecto neto de la intervención. Sin embargo, siendo que es un incentivo cuya

⁹ A nivel nacional, todas las escuelas de Educación Intercultural Bilingüe (EIB), que constituyen el grupo B, son parte del grupo de tratamiento de la intervención, no habiendo escuelas privadas para este tipo de oferta educativa. La comparación con otro tipo de escuela pública no sería adecuado, por tener una propuesta pedagógica diferente (bajo la EIB, los estudiantes tienen clases en su propia lengua y en castellano). Al 2016, a nivel nacional solo hubo dos escuelas privadas que ofrecieron el mismo tipo de servicio.

participación se propone de manera universal (para todas las escuelas públicas), se plantea esta opción como la mejor alternativa para evaluar su impacto, siguiendo el planteamiento de estudios previos, como Martins (2009), Rau y Contreras (2012), y Casavilca (2014). Para reducir este sesgo, se planteará un modelo de emparejamiento que tiene el objetivo de mejorar el balance entre ambos grupos. Posteriormente, se aplicará una estimación bajo el modelo de diferencias en diferencias, con la finalidad de reducir el sesgo por la tendencia previa al tratamiento.

Complementariamente, es importante mencionar que las diferencias en la información disponible para el caso del nivel primario y secundaria han sido aprovechadas de tal manera que se puedan reducir los sesgos de la metodología aplicada. Así, por ejemplo, de la información disponible para el nivel primario no ha sido posible contar con variables complementarias sobre el estudiante que permitan trabajar la información a este nivel. Por tanto, el nivel de análisis utilizado para este caso ha sido la escuela. Sin embargo, siendo que la Evaluación Censal de Estudiantes (ECE) se aplica para este nivel desde el 2007, se ha utilizado información desde este periodo para aplicar una estimación paramétrica por triple diferencia, que permita reducir el sesgo encontrado por el no cumplimiento de tendencias paralelas en el periodo previo a la intervención. Siendo que para este nivel Bono Escuela se implementó desde finales del año 2014, se considera que el efecto de la intervención sobre dicho año es nulo, razón por la cual será considerado el año de línea de base. Por tanto, se cuenta con dos periodos consecutivos del incentivo (2015-2016), que permitirá evaluar si existe un efecto diferenciado debido a un periodo de maduración de la intervención.

Por su parte, para el caso del nivel secundaria, debe tenerse en cuenta que la primera evaluación censal se aplicó en el año 2015, razón por la cual Bono Escuela se anunció dicho año y empezó a ejecutarse a partir del 2016 para este nivel. Considerando el antecedente de Bono Escuela para el caso de primaria (desde el 2014) sería difícil afirmar que existe una línea de base para este nivel, considerando las expectativas de los docentes de ser evaluados para la obtención de un bono. Asimismo, dado que no se cuenta con información sobre el rendimiento estudiantil en periodos previos al anuncio del tratamiento (año 2015), no es posible reducir la tendencia previa. En cambio, a diferencia del caso de primaria, existe información valiosa recogida a nivel de estudiante, que permite tener mejores covariables para un modelo de emparejamiento que logre un mejor balance en factores exógenos a la evaluación. Por ello, para este nivel, la estimación se trabajará a nivel de estudiante, y se utilizará el modelo de emparejamiento como la mejor alternativa disponible para medir el impacto de la intervención, reconociendo las limitaciones de esta metodología para eliminar el sesgo de factores no observables. Complementariamente, para el segundo año de aplicación de la ECE en este nivel (periodo 2016), se estima además un modelo de diferencias en diferencias, con el objetivo de reducir las tendencias del año previo y obtener un resultado más limpio del sesgo mencionado inicialmente.

Finalmente, para la medición del efecto reciprocidad, conceptualmente se esperaría que el otorgamiento del bono tenga un efecto en los resultados del año posterior al otorgamiento del incentivo. Por ello -habiendo precisado que el otorgamiento del primer bono se ejecutó en el año 2014 para el caso de primaria, y en el año 2016 para el caso de secundaria-, se analizará la evidencia del efecto de los bonos 2014 y 2015 sobre el rendimiento estudiantil en los años 2015 y 2016, respectivamente, para el caso de primaria, mientras que para el caso de secundaria no será posible la evaluación del efecto

reciprocidad. Para ello se empleará la metodología de regresión discontinua bajo el diseño difuso, bajo una estimación no paramétrica, donde se comparará a las escuelas premiadas con respecto a aquellas que estuvieron muy cerca de alcanzar el bono, en función de su proximidad al punto de corte del índice de desempeño escolar (IDE). Sobre este punto cabe precisar que, a fin de reducir algún problema de identificación en las escuelas ganadoras y no ganadoras del bono, se ha acotado el análisis a aquellas escuelas públicas que mantuvieron al menos el 80% de su plana docente. Ello con el objetivo de no adjudicar como efecto del bono el cambio en el rendimiento en escuelas que, habiendo ganado en un periodo –y, por tanto, habiéndose premiado a sus docentes– no puedan considerarse “las mismas” el año siguiente, debido a la rotación de los docentes que participaron directamente en el éxito de la escuela. Se identificó que, entre las escuelas públicas, aproximadamente el 40% de escuelas tienen una rotación mayor al 20% de sus docentes.

Las bases de datos empleadas para la presente estimación fueron principalmente: la Evaluación Censal de Estudiantes (que recoge información sobre el rendimiento estudiantil y variables del contexto del estudiante para el caso de secundaria), el Censo Escolar (que recoge información de las características del docente del nivel educativo y la infraestructura escolar), Bono Escuela (que recoge el puntaje obtenido por cada escuela en el IDE sobre el cual se otorga el incentivo, y grupo al que pertenece), así como información complementaria que permite identificar las escuelas públicas que se encuentran sujetas a intervenciones pedagógicas, a manera de control.

4.1 Metodología para la evaluación del efecto competencia en los niveles de primaria y secundaria

4.1.1 Identificación de grupo de tratamiento y zonas de desempeño

Precisiones sobre la información trabajada

Para efectos del presente estudio, en lo que corresponde al nivel de primaria, la unidad de análisis será la escuela. Ello principalmente debido a que el bono se asigna a este nivel, y los indicadores son definidos como un promedio para cada institución educativa. Adicionalmente, no se cuenta con variables de control a nivel de estudiante, de manera censal. No obstante, para la definición del impacto del bono se considerará el uso de factores de expansión, donde el peso de cada colegio será el número de estudiantes evaluados. Por otro lado, para el caso de secundaria se presentarán los resultados a nivel de alumno, debido a que para este nivel se cuenta con un cuestionario a nivel de estudiante que recoge información valiosa sobre el contexto particular de cada estudiante, así como algunas variables sobre la participación de sus padres en su desarrollo académico, y algunas prácticas pedagógicas de sus profesores. Asimismo, para el caso de segundo grado de primaria, se cuenta con un panel data para el periodo 2007-2016, que recoge los resultados promedio de las instituciones educativas públicas y privadas¹⁰, para el área de Comprensión Lectora y Matemáticas. Por su parte, para el caso de

¹⁰ Como se observa en la Tabla 5, los colegios públicos representan aproximadamente el 73% del total de colegios de segundo grado de primaria que rinden la prueba (es decir, sin considerar los colegios con menos de cinco estudiantes, que no entran en el marco de la ECE). Para el caso de EIB, los colegios públicos representan el 99.9%; y para el caso de secundaria, el 66%.

secundaria, el panel data se construye con los periodos 2015 y 2016.

Es importante mencionar que, aunque el panel no se encuentra balanceado, se ha optado por mantener todas las observaciones disponibles. Se ha analizado los posibles sesgos que se podrían presentar y, si bien en las primeras evaluaciones no se alcanzó la cobertura suficiente para ser declarada información censal a nivel de estudiante (en el 2008 se evaluó al 90% de instituciones educativas y al 71% de estudiantes programados¹¹), a partir del 2009 se alcanzó cobertura censal tanto a nivel de institución educativa como de estudiante, al 90% y 81%, respectivamente. Además, en los años siguientes la cobertura fue en aumento, llegando a evaluar el 95.6% de estudiantes y 99.7% de colegios programados a nivel nacional en el año 2016.

Al analizar la cobertura por estratos, se puede observar cierto sesgo hacia la zona rural y los colegios unidocentes/multigrado. Aunque este sesgo no es significativo a nivel nacional, sí se observa en algunas regiones (UMC 2015, 2016). Por ello, se concluye i) que la evolución de la mejora en el rendimiento estudiantil puede estar subestimada en alguna medida para los primeros años, siendo que en los años iniciales de la prueba hubo un sesgo hacia regiones donde predominaba un mayor nivel de pobreza¹², y ii) que la subestimación de los resultados y las mejoras interanuales se ha reducido progresivamente al ampliarse anualmente la cobertura de la prueba, para todos los estratos. Por su parte, para el caso de secundaria, habiendo alcanzado una cobertura estudiantil superior al 90% en los dos periodos evaluados, solo cabría considerar ligeros sesgos en la zona rural de algunas regiones (aunque no representativos a nivel nacional).

Sobre la información de Bono Escuela, para las escuelas de segundo de primaria se acumulan tres rondas ininterrumpidas desde su creación en el 2014. Puede observarse que, en promedio, el 49% de escuelas nunca ha recibido el bono, un 40% lo recibió una vez, un 11% lo recibió dos veces y solo un 1% recibió el bono en las tres rondas. Finalmente, para el caso de segundo grado de secundaria, solo existe una ronda de Bono Escuela publicada, habiendo ganado un 17% del total de colegios públicos del nivel secundario (8,321 IIEE como total en el 2015).

Tabla 2: Nro. de bonos recibidos por escuelas públicas de primaria durante el periodo 2014-2016

Nro de bonos recibidos	Frecuencia de IIEE	% del total	Promedio de veces que participó en el bono
0	9,349	49%	2.1
1	7,666	40%	2.7
2	2,106	11%	2.9
3	141	1%	3.0
Total	19,262	100%	2.4

Fuente: Bono Escuela 2014-2015. Elaboración propia.

¹¹ De acuerdo a los estándares de la Unidad de la Medición de la Calidad (UMC), se puede afirmar cobertura censal cuando se ha cubierto a, por lo menos, el 90% de instituciones educativas o más del 80% de estudiantes, con respecto a lo programado.

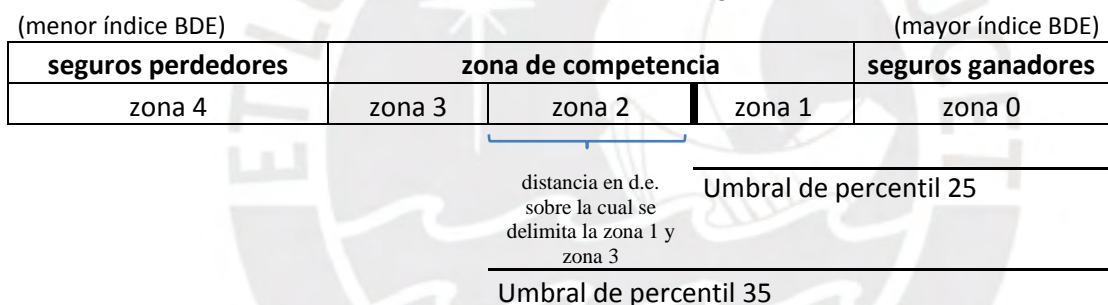
¹² Siendo Cajamarca, Apurímac, Puno y Ayacucho las regiones que se retrasaron en alcanzar cobertura regional considerada censal.

Identificación de las escuelas en estricta zona de competencia

El alto porcentaje de instituciones educativas que se concentra en la categoría de nunca haber recibido un bono nos da indicios de la existencia de un grupo de colegios que posiblemente se encuentre lejos de mejorar sus resultados en el corto plazo, y en ese sentido tendría muy poca motivación para cambiar su nivel de esfuerzo. Por otro lado, habría un grupo de instituciones educativas que, teniendo resultados muy altos, tiene seguridad de su posición ventajosa sobre el promedio y por ello el umbral para acceder al bono (percentil 35) le es indiferente. Este es el concepto detrás de lo que Rau y Contreras (2012) definen como “seguros perdedores” y “seguros ganadores”.

Aun cuando el número de rondas acumuladas es menor al caso chileno, se puede observar una tendencia similar a los hallazgos de Rau y Contreras (*ibid.*:14) y Casavilca (2014), que también evidencian una concentración mayor de “seguros perdedores”, pero un porcentaje muy pequeño de “seguros ganadores”. Dado lo anterior, en este estudio se define como grupo de interés a aquellas escuelas que se encuentran en una posición expectante del bono. Para su identificación, se utilizará la metodología empleada por Casavilca (*ibid.*:17), en la cual se clasifican las escuelas en 5 zonas diferentes, en función del puntaje alcanzado en el índice que determina la asignación del bono, de acuerdo al siguiente esquema:

Esquema 2: Grupos de desempeño y zona de competencia



Adaptado de Casavilca (2014).

Así, en primer lugar, para cada estrato de los grupos definidos por Bono Escuela¹³, se estandariza el puntaje del índice BDE con respecto al valor mínimo del percentil 25. Habiendo clasificado de esta manera a todos los colegios -en función de su distancia con dicho valor en términos de desviaciones estándar- se procede a calcular la distancia promedio del valor mínimo del percentil 25 con respecto al valor mínimo del percentil 35, que viene a ser la delimitación de la zona 2. Cabe precisar que este promedio se calculó para cada año, y se observó que en promedio fue de 0.3 desviaciones estándar para el caso de primaria y de 0.5 desviaciones estándar para el caso de secundaria¹⁴. Una vez calculada esta diferencia promedio, se aplica dicho valor a la delimitación de la zona 1, zona 3 y, por defecto, las zonas 0 y 4.

Así, formalizando la metodología señalada, en cada año y para cada grado evaluado se define el índice BDE estandarizado (x^*) como:

¹³ Cabe precisar que en el 2016, el número de estratos para el caso del grupo A fue de 368; y para el grupo D, de 274. El detalle por año se puede encontrar en la Tabla 6.

¹⁴ En el caso de la investigación de Casavilca (2014), este promedio también fue de 0.3 desviaciones estándar.

$$x_{ij}^* = \frac{x_{ij} - x_j^{\min 25}}{\sigma_j} \dots(1)$$

Donde x_{ij}^* es el valor estandarizado del colegio i en el estrato j, x es el valor inicial del índice BDE, $x_j^{\min 25}$ es el valor mínimo del percentil 25 del estrato j, y σ_j es la desviación estándar del índice BDE en el estrato j. Por su parte, se define la variable y como el promedio del valor estandarizado mínimo en el percentil 35 (donde n es el número de estratos)¹⁵, en valor absoluto:

$$y = \left| \frac{\sum_{j=1}^n x_j^{*,\min 35}}{n} \right| \dots(2)$$

Con estas definiciones, se clasifican las escuelas en las siguientes zonas, teniendo en cuenta que la población considerada “en zona de competencia” será aquella que se encuentra en la zona expectante al bono; es decir, entre las zonas 1 y 3:

Zona 0	Si $y \leq x_{ij}^*$
Zona 1	Si $0 < x_{ij}^* < y$
Zona 2	Si $-y < x_{ij}^* < 0$
Zona 3	Si $-2y < x_{ij}^* < -y$
Zona 4	Si $x_{ij}^* < -2y$

4.1.2 Identificación del grupo de control y estimación del modelo de emparejamiento

Para evaluar el impacto de este programa sobre el rendimiento de los estudiantes, una de las primeras dificultades es la definición de un grupo de control adecuado. Al ser una intervención cuyo objetivo es fomentar la competencia en todas las escuelas públicas, el grupo de control necesariamente deberá identificarse entre los colegios privados. Sin embargo, antes de poder considerar este tipo de escuelas como grupo de control, es importante reconocer intuitivamente que existen factores observables y no observables que pueden afectar la comparabilidad de los grupos. La información de la Tabla 7 (ver Anexo) nos presenta indicios importantes de una situación de partida diferente entre el grupo de control y de tratamiento. En consecuencia, es de esperar que existan además factores no observables que diferencien ambos tipos de gestión, en cuanto a contexto particular de los estudiantes y variables asociadas al funcionamiento de la IE. En ese sentido, se reconoce la limitación que puede tener la presente investigación originada por este sesgo. Sin embargo, siendo que es un incentivo aplicado de manera universal (para todas las escuelas públicas), se plantea esta opción como la

¹⁵ Para el periodo 2007-2013, considerando que no existía el BDE, la definición de las zonas de competencia se construyó a partir de un índice “ficticio”, basado únicamente en la ECE (se ponderaron con nuevos pesos los indicadores de superación y efectividad). Además, se tomó un promedio de y para todo el periodo, de 0.33 desviaciones estándar. El detalle por año se puede encontrar en la Tabla 6.

mejor alternativa para evaluar su impacto, siguiendo el planteamiento de estudios previos, como Martins (2009), Rau y Contreras (2012), y Casavilca (2014).

Para reducir esta probabilidad de sesgo, se aprovechará la información disponible para cada uno de los niveles evaluados, primaria y secundaria, que difiere en cada caso. Por ejemplo, para el caso de primaria, aunque no se cuenta con información complementaria a nivel de estudiante que permita plantear un nivel de emparejamiento a este nivel, sí existe información de periodos previos al tratamiento, desde el año 2007. Por ello, en este caso, el modelo de emparejamiento se definirá con variables promedio a nivel de escuela, como una manera de reducir el desbalance en las condiciones iniciales de las variables observables. Sin embargo, como se detallará más adelante, la disponibilidad de información previa al tratamiento permitirá utilizar complementariamente un modelo de triple diferencia, con el objetivo de reducir el sesgo de factores no observables.

Para ello, se implementa un primer filtro de variables para una coincidencia exacta, utilizando las variables de año, área de procedencia (urbano/rural), característica de la IE (polidocente completo/multigrado o unidocente) y zona de competencia en la que se encuentra la IE, esta última variable de acuerdo a lo descrito en la subsección 4.1.1. Con ello se busca asegurar un primer nivel de comparabilidad y soporte común, según la pertenencia de la IE. A partir de ahí, se aplica la metodología de emparejamiento por vecindario más cercano, de uno a uno con posibilidad de reemplazo, por regresión logística. Según lo propuesto por Abadie e Imbens (2006; en Rau y Contreras, 2012), el emparejamiento con reemplazo produciría un *match* de mejor calidad, al incrementar el set de opciones para encontrar la mejor pareja para los individuos tratados. Las covariables que se utilizarán para la estimación del *propensity score* son: el tamaño de la escuela (aproximado con el número de alumnos), proporción de hombres en la IE, zona de competencia a la que pertenece la IE y tamaño de la jurisdicción a la que pertenece la IE (número de IE por provincia). El detalle de la distribución del puntaje de propensión para los tratados y no tratados se detalla en el Gráfico 1 de la sección de Anexos.

Para el caso del nivel de secundaria las limitaciones de análisis son mayores, por la dificultad de no contar con información previa al tratamiento. Por ello, la evaluación del efecto competencia no podrá plantearse de la misma manera que para el caso de primaria. No obstante, se aprovechará la información más extensa con la que se cuenta para el nivel de secundaria a nivel de estudiante, con el objetivo de aumentar la precisión de la comparabilidad de los grupos de tratamiento y de control y así usar el modelo de emparejamiento como método de evaluación del efecto competencia.

Para ello, se utilizará la información de la encuesta realizada a los estudiantes de segundo grado de secundaria. A diferencia del nivel primaria, para este caso se cuenta con información adicional sobre el contexto socioeconómico del estudiante, interés de sus padres por su vida académica, aspiraciones y percepciones personales sobre la escuela, y calificación sobre la frecuencia de algunas acciones pedagógicas de sus profesores. Se espera que con esta información se pueda reducir el efecto producido por factores no observables que pudieran sesgar los resultados. Siendo que hay una encuesta para cada periodo disponible de la ECE, se analizará tanto el año 2015 como 2016.

Para ello, al igual que para el nivel primario, se aplicará un primer filtro de variables de coincidencia exacta, basado en las siguientes variables: área de procedencia, zona de

competencia, variable dicotómica de la lengua materna del estudiante, nivel educativo de la madre del estudiante, acceso al alumbrado eléctrico en el hogar, acceso a Internet dentro del hogar, acceso a la red pública de desagüe dentro del hogar, condición de repitente del estudiante y percepción sobre las acciones pedagógicas de sus docentes. Finalmente, para la identificación del efecto competencia en el nivel secundario se estimará un *propensity score*, bajo el modelo de vecindario más cercano de uno a uno, con posibilidad de reemplazo, mediante una regresión logística. Las variables utilizadas en este modelamiento son: condición de repitente del estudiante, percepción sobre la atención de los padres por su vida personal y académica, percepción sobre el clima escolar, percepción sobre las prácticas pedagógicas del docente, número de instituciones educativas en la UGEL a la que está adscrita la escuela, y zona de competencia de la IE.

4.1.3 Estimación por triple diferencia para el caso del nivel primaria

Como se mencionó en la sección anterior, en primer lugar se reconoce la condición de tratado como una primera fuente de variación en los resultados del estudiante¹⁶. En segundo lugar, en la medida que las diferencias entre colegios también evolucionan en el tiempo, este factor temporal también constituye una fuente adicional de variación (Casavilca, 2014:36). Por ello, siendo que para el caso del nivel primaria se cuenta con información para periodos previos al tratamiento, se considera que la estimación por diferencias en diferencias (DD) es la mejor alternativa para reducir la probabilidad de sesgo. Este modelo permitiría aislar la tendencia temporal del grupo de tratamiento que puede haberse presentado por efectos ajenos al incentivo. Asimismo, permitiría aislar el sesgo producto de diferencias permanentes y no observables (efectos fijos) entre el grupo de tratamiento y de control (Imbens y Wooldridge, 2008:64, en Mesinas, 2010)¹⁷.

No obstante, un supuesto clave para validar el uso de esta metodología es el de tendencias paralelas entre el grupo de control y tratamiento, en ausencia del programa. Sin embargo, como sostiene Ravallion (2005:47, en Mesinas, 2010), este supuesto es poco realista para programas dirigidos a áreas donde la infraestructura y las condiciones previas (incluyendo el nivel de ingresos) poseen una desventaja frente al grupo no focalizado (en este caso, el grupo de control). Aunque el autor se refiere en particular a programas de lucha contra la pobreza, se considera una justificación aplicable a este caso.

En el Gráfico 2 de la sección Anexos se puede visualizar un primer análisis de tendencia bajo esta metodología. Gráficamente, habría indicios de que el supuesto de tendencias paralelas en la etapa previa al tratamiento no se sostiene en este caso. Por ejemplo, al analizar la evolución promedio de los resultados de Comprensión Lectora en la etapa previa a la intervención, se observa, en primer lugar, que el promedio de las IIEE públicas siempre se ha mantenido por debajo del promedio de escuelas privadas, para todos los años evaluados. Por otro lado, las tendencias de ambos tipo de IE no han ido a la par; por el contrario, se puede apreciar que la tendencia ha sido más fluctuante en el caso de las escuelas públicas. Una conclusión similar se observa al analizar la evolución de los resultados en el área de Matemáticas, aunque en este caso, tanto las IIEE

¹⁶ En la Tabla 7 se podrá revisar la caracterización promedio de las escuelas, diferenciando por tipo de gestión (público o privada).

¹⁷ Para una breve explicación de la intuición detrás del modelo de DD, y ver el análisis incondicional previo a la estimación del efecto, revisar la Tabla 24 en la sección Anexos.

públicas como privadas han fluctuado considerablemente, aunque no en el mismo sentido para cada año. Ello aplica tanto a la figura que considera todos los colegios (posterior al emparejamiento), como a la figura que acota la muestra a la zona estricta de competencia (zona 1, 2 y 3, según lo descrito en la sección 4.1). Adicionalmente, para verificar el análisis gráfico, se estimó un modelo de diferencias en diferencias en un escenario placebo; es decir, sin considerar los años de intervención y considerando el año 2014 como el año de línea de base. Los resultados de esta estimación confirman una tendencia fluctuante en el tiempo en la etapa previa al programa. Estos resultados se detallan en la Tabla 8, donde el estimador de la variable “txl” confirma la tendencia no-paralela.

Por tanto, dados los resultados obtenidos en la estimación placebo, es un reto adicional para la evaluación de este programa tener en cuenta esta fuente adicional de variación. En tal sentido, el modelo de triple diferencia explota la disponibilidad de información pre línea de base (en este caso, antes del 2014) para aislar el efecto previo al programa. Bajo esta estimación se busca reducir el sesgo al introducir controles que identifiquen la diferencia entre la tendencia previa a la línea de base y la tendencia de la línea de base. Por ello, en esta investigación se usará una metodología de triple diferencia similar a la utilizada por Mesinas (2010), que implica la diferencia entre el grupo de tratamiento y un grupo de control en tres periodos diferentes (pre línea de base, línea de base y después de la implementación).

De acuerdo a lo señalado por Chaudhury y Parajuli (2006, citado en Mesinas, 2010), el supuesto de identificación no testeable detrás de esta metodología es que la tendencia observada previamente a la implementación del programa continúa en la etapa ex post. En ese sentido, aun cuando es difícil de asumir que la triple diferencia removerá todo el sesgo, esta metodología sí provee una manera más limpia de separar parte del sesgo, que surge por una brecha en las condiciones iniciales del rendimiento estudiantil. Adicionalmente, otra ventaja de esta metodología sería controlar la presencia de programas que operaron en la etapa previa al programa y que podrían haber afectado la tendencia previa de los resultados de aprendizaje.

Bajo esta metodología, el impacto neto del programa es dado por:

$$(\Delta y_{tr,1} - \Delta y_{c,1}) - (\Delta y_{tr,0} - \Delta y_{c,0}) \dots (3)$$

donde el subíndice *tr* y *c* se refieren al grupo de tratamiento y de control, respectivamente. Por su parte, Δy_1 se refiere a la diferencia en el resultado entre el periodo ex post y el periodo de línea de base del programa. Por su parte, Δy_0 se refiere a la diferencia en el resultado entre el periodo de línea de base del programa y el periodo previo pre línea de base. En términos paramétricos, el modelo se planteará de la siguiente forma:

$$Y_{i,t} = X_{i,t}\beta + \alpha_1 Intervencion_{i,t} + \alpha_2 Post2014_{i,t} + \alpha_3 Post2015_{i,t} + \alpha_4 Post2014 * Intervencion + \alpha_5 Post2015_{i,t} * Intervencion_{i,t} + \varepsilon_{i,t} \dots (4)$$

Donde $Y_{i,t}$ es el resultado promedio de la ECE para el colegio *i* en el periodo *t*, tanto en

Comprensión Lectora como en Matemática. Por su parte, la variable $Intervencion_{i,t}$ es una dicotómica que toma el valor de 1 en caso el colegio haya recibido el tratamiento (escuela estatal) o el valor de 0 en caso el caso pertenezca al grupo de control (escuela privada). Asimismo, $Post2014_{i,t}$ es una variable dicotómica que toma el valor de 1 si el periodo es igual o posterior al 2014 (año línea de base) o valor de 0 en caso el periodo observado sea previo al 2014 (pre línea de base). Adicionalmente, $Post2015_{i,t}$ es una variable dicotómica que toma el valor de 1 si el periodo observado es igual o posterior al año 2015 (ex post al programa) o valor de 0 para periodos previos al año 2015 (línea de base). Por otro lado, las variables $Post2014 * Intervencion$ y $Post2015_{i,t} * Intervencion_{i,t}$ son variables dicotómicas interactivas de la condición de tratamiento y el periodo observado. La matriz $X_{i,t}$ se refiere a las covariables que recogen las características promedio de la escuela, estudiantes y docentes, que servirán de control. Finalmente, el término $\varepsilon_{i,t}$ es el término de error que adopta la siguiente estructura $\varepsilon_{i,t} = \eta_i + \mu_{i,t}$ para reflejar los efectos fijos por escuela en la ecuación (4). El efecto neto del programa producto de la triple diferencia se reflejaría en el coeficiente α_5 , explicado intuitivamente mediante la ecuación (3) y en la Tabla 3.

Acerca de las variables de control utilizadas ($X_{i,t}$), se utilizaron variables promedio de la institución educativa, como: área de procedencia, característica de la IE (polidocente completo o multigrado/unidocente), número de alumnos de segundo de primaria, proporción de estudiantes de sexo masculino en el grado, proporción de estudiantes en edad normativa para el grado, proporción de estudiantes del grado (entre 6 y 8 años de edad) que han cursado Educación Inicial, total de docentes para el nivel primaria, proporción de docentes hombres para el nivel, proporción de docentes con estudios superiores concluidos, y número de computadoras con acceso a Internet que se usan para fines pedagógicos. Se incluye además un rezago del rendimiento alcanzando en Comprensión Lectora y Matemática. Asimismo, se considera información a nivel agregado, como el porcentaje de pobreza distrital, la pertenencia a un distrito JUNTOS, la pertenencia a un distrito del VRAEM, y la pertenencia a un centro poblado donde predomina el castellano. Finalmente, se incluye como control la intervención de algunos programas a cargo del Ministerio de Educación, que tienen el objetivo de mejorar los resultados estudiantiles, como Acompañamiento Pedagógico y Fortalecimiento de la Educación Física, y se incluyen variables dicotómicas regionales y temporales.

4.2 Metodología para la evaluación del efecto reciprocidad en el nivel de primaria

Como se mencionó anteriormente, en la presente investigación también se busca identificar si existe algún efecto del bono entre los que ganaron el incentivo. Esto es lo que está detrás del denominado efecto reciprocidad, que postula que los profesores podrían tener un estímulo adicional para un esfuerzo mayor una vez recibido el bono. Al respecto, es importante precisar que, si bien este efecto no representa el mecanismo principal a través del cual operaría el diseño de la intervención, se considera un efecto indirecto importante de analizar, de acuerdo a los hallazgos de la literatura empírica.

Así pues, existe una diferencia en el grupo de tratamiento para la medición de este efecto, pues solo se acota al grupo de escuelas ganadoras del incentivo. Por su parte, el grupo de control serían las escuelas públicas que se encontraron muy cerca de obtener el bono. Tal como se explicó, debido a la limitación en el número de periodos disponibles de la intervención para el caso de secundaria, este efecto solo se medirá en el caso del nivel primario.

Por otro lado, es importante tener en cuenta que la rotación de docentes entre escuelas, de ser alta, puede representar un problema de identificación. Por ejemplo, en caso de que una escuela resultara ganadora y, por tanto, se premiara a su plana docente, sería difícil evaluar un efecto recíproco del bono en el año siguiente sobre el rendimiento promedio de dicha escuela si para entonces los docentes ganadores migraron a otra IE. Por tal motivo, el análisis se centrará en aquellas escuelas públicas (ganadoras y no ganadoras) que hayan mantenido igual por lo menos el 80% de su plana docente durante el año siguiente al otorgamiento del bono. Se identificó que en promedio este grupo de escuelas representa el 60% del total de escuelas públicas.

Se plantea que este efecto puede ser medido a través de una regresión discontinua. La idea principal detrás de un diseño de regresión discontinua es que la asignación del tratamiento está determinada, ya sea completa o parcialmente, por el valor de una variable particular, en función del lado donde se encuentre el individuo tratado en la función de distribución, estableciendo como punto de referencia el valor determinado de dicha variable. De esta manera, este valor límite generaría una discontinuidad en la probabilidad de recibir el tratamiento, en función de esta variable particular. En consecuencia, como señala Van Der Klaauw (2008, citado en Mesinas, 2010), la discontinuidad de la distribución condicional del resultado evaluado, como una función de esta variable de delimitación, es interpretada como un efecto causal del tratamiento.

Para el caso de la intervención evaluada, la condición de tratamiento se define en este caso como el grupo de escuelas que ganó el bono, y está en función del índice de cumplimiento de los indicadores evaluados y descritos en la sección 3, referidos a la efectividad y superación en el rendimiento estudiantil, retención de estudiantes y registro de información, entre otros. Sobre la base de estos indicadores estandarizados y ponderados en un solo índice, se establece un ordenamiento de los puntajes por estrato y se define el top de escuelas que se declararán ganadoras. En consecuencia, el puntaje mínimo definido en cada estrato para hacerse acreedor al bono es el límite que determina el grupo de tratamiento y de control.

Para efectos de esta evaluación, dicho puntaje ha sido estandarizado con respecto al valor mínimo del percentil 35 de cada estrato, de tal manera que el punto de corte se centraliza en cero para todos los estratos, facilitando así su comparabilidad. Teniendo como referencia este valor, se procedió a analizar los datos de Bono Escuela para el caso de primaria, y se encontró que, aunque este es el indicador determinístico, existen algunos casos de escuelas que no cumplen con este criterio. El detalle por año se puede observar en la Tabla 9 de la sección de Anexos. Por ejemplo, en lo que respecta al año 2016, se observa 40 escuelas que, habiendo obtenido un puntaje positivo (mayor a cero), no fueron consideradas en la lista de escuelas ganadoras. Dichos colegios representaron el 1% del total de escuelas con un puntaje superior al punto de corte. Por su parte, también en el 2016 hubo 44 escuelas que, habiendo obtenido un puntaje negativo (menor al valor límite del percentil 35, centralizado en cero), sí fueron incluidas

dentro del bono. Estas escuelas representaron el 0.58% del total de escuelas por debajo del umbral.

En consecuencia, se considerará un caso de estimación no paramétrica de regresión discontinua bajo un diseño difuso, donde la asignación de tratamiento no está 100% determinada por el punto de corte del índice, y en consecuencia, la probabilidad de ser tratado no varía de cero a uno en el punto de corte. En su defecto, bajo este diseño solo se requiere una discontinuidad de la probabilidad de asignación en el puntaje límite de la variable establecida como delimitante (Imbens y Wooldrige, 2008:60). En el caso de Bono Escuela, la probabilidad de ser tratado entre el grupo con un puntaje menor a cero, es de 0.58%, mientras que en el grupo con puntaje mayor a cero, esta probabilidad pasa a ser 99%, sobre la base de los resultados 2016.

4.3 Efectos diferenciados (característica de la escuela, ámbito y periodo de maduración)

Sobre la implementación de Bono Escuela, es importante tener en consideración el periodo de maduración de la intervención como posible primer factor de heterogeneidad en los resultados. Como muestra la evidencia internacional (sección 2.2), los resultados pueden variar según el periodo de la intervención que se evalúe (Glewwe, Ilias y Kremer, 2010; Muralidharan y Sundararaman, 2010; Springer *et. al*, 2010). Para el caso del bono en el Perú, habiendo iniciado a finales del año 2014 para el caso de primaria, la construcción del IDE para el otorgamiento del primer bono se construyó sobre la base de los indicadores obtenidos en el 2013, e incluso sobre la base del año 2012 para la medición del indicador de superación en la ECE. Siendo de carácter retroactivo, la intervención no puede haber tenido impacto en las evaluaciones previas al año 2014. Incluso durante dicho año, al haber formalizado la intervención en el segundo semestre del año, el tiempo disponible para que una escuela haya podido plantear una estrategia de competencia por el bono es muy limitado, razón por la cual se considera el año de línea de base para la intervención y, en consecuencia, se cuenta con dos periodos consecutivos del incentivo.

Debido a ello, al analizar el efecto competencia de la intervención, también se realizará un análisis sobre el periodo de maduración, identificando si un año adicional de intervención tiene un efecto diferenciado en el efecto del incentivo, y teniendo en cuenta que en el caso de primaria se tienen disponibles dos años consecutivos de intervención. Un caso diferente puede ser el del nivel secundaria, pues, al haber antecedentes de Bono Escuela antes del primer año de implementación de la ECE para este nivel, es probable que los resultados del primer año de evaluación (2015) se vean afectados por la expectativa de una competencia, aun cuando la formalización de ella se haya dado recién en el segundo semestre del año 2015. Por ello, en el caso de secundaria se habla de dos periodos consecutivos para la intervención. En consecuencia, para el nivel primario se analizará el efecto de la intervención en cada periodo, bajo el modelo de triple diferencia descrito en la sección 4.1.3. Así, para la identificación de heterogeneidad la especificación del modelo será replanteada de la siguiente manera:

$$\begin{aligned}
Y_{i,t} = & X_{i,t}\beta + \alpha_1 Intervencion_{i,t} + \alpha_2 Post2014_{i,t} + \alpha_3 Post2015_{i,t} \\
& + \alpha_4 Post2016_{i,t} + \alpha_5 Post2014_{i,t} * Intervencion_{i,t} + \alpha_6 Post2015_{i,t} * Intervencion_{i,t} \dots (5) \\
& + \alpha_7 Post2016_{i,t} * Intervencion_{i,t} + \varepsilon_{i,t}
\end{aligned}$$

De esta manera, a la ecuación (5) se añaden las variables tendenciales y las variables interactivas de la condición de tratamiento y el periodo observado, para los periodos 2015 y 2016. Los estimadores α_6 y α_7 representan el efecto marginal de cada año de la intervención, neto de la tendencia previa a la línea de base y la tendencia anual de cada año de la intervención.

Un modelo similar se planteará para el caso de secundaria, a nivel de estudiante, con los dos únicos años con información de rendimiento disponible para el nivel, que son a su vez los años de la intervención. Como se mencionó anteriormente, a diferencia del caso de primaria, para Secundaria no se cuenta con información de la línea y pre-línea de base, por lo cual debe considerarse una limitación importante en el impacto de la intervención sobre el primer año de evaluación. Sin embargo, haciendo una distinción en el efecto marginal de la intervención por año evaluado, se plantea evaluar si hay un efecto de maduración de la intervención en el segundo año del incentivo.

Asimismo, una de las preguntas importantes que se plantea en esta investigación es la diferenciación de los resultados para aquel grupo de escuelas que se encuentra en zona de estricta competencia, excluyendo a aquellos colegios denominados “seguros ganadores” o “seguros perdedores”, de acuerdo a lo planteado en la sección 4.1. Finalmente, también se hará una distinción según área de procedencia (urbano/rural) y característica de la IE¹⁸ (polidocente completo/multigrado o unidocente), con la finalidad de analizar si los resultados del impacto de la intervención varían para el caso de las escuelas urbanas (o polidocentes completas) en estricta competencia, en relación a sus pares en zona rural (o multigrado/unidocente).

5. Resultados

En esta sección se presentan los resultados producto de las preguntas planteadas sobre los canales a través de los cuales un incentivo monetario como Bono Escuela tendría efecto sobre el rendimiento estudiantil, a partir de la metodología detallada en la sección anterior. Así, para el caso de primaria, se presentan los resultados de la triple diferencia para la evaluación del efecto competencia, y los resultados de la estimación por regresión discontinua para la evaluación del efecto reciprocidad. Por su parte, para el caso de secundaria se presentan los resultados obtenidos a partir del modelo de emparejamiento, para la medición del efecto competencia. Asimismo, para ambos niveles se analizan los resultados de la consideración de una posible heterogeneidad, relacionada al periodo de maduración de la intervención, el área de procedencia y característica de la IE.

5.1 Resultados en el nivel de primaria

¹⁸ Cabe mencionar que la definición de IE polidocente completo, multigrado o unidocente solo se utiliza para el nivel primaria, pues para el nivel Secundario todas las escuelas son polidocentes completas. En ese sentido, esta posible heterogeneidad solo se analizará para el primer caso

5.1.1 Efecto competencia

Como se mencionó en la sección 4, para la medición del efecto competencia fue importante, en primer lugar, la clasificación de las escuelas en función de los resultados obtenidos en el índice de desempeño (construido por el Ministerio de Educación a partir de los parámetros y ponderaciones definidos por Bono Escuela, descritos en la sección 3), y su distanciamiento con respecto al umbral de calificación al otorgamiento del bono. La idea principal detrás de esta categorización es considerar que las escuelas pueden tener distintos grados de motivación en la competencia por el bono, en función de su posición con respecto a sus pares y sus posibilidades de ganar. De esta manera, la variable zona de competencia está compuesta de cinco categorías, donde 0 es la mejor posición y 4 es la peor, y las categorías del medio son las consideradas las zonas en estricta competencia. Para aquellos periodos o colegios (privados) donde no se aplicó el bono, se hizo una simulación de la construcción del índice, basado únicamente en los resultados de la ECE.

Sobre el grupo de tratamiento, para el caso de primaria se consideró únicamente al grupo A como el objeto de estudio; es decir, las escuelas públicas de la Educación Básica Regular, evaluadas en segundo grado de primaria. Por su parte, se construyó un grupo de control a partir de las instituciones educativas no participantes del Bono Escuela; es decir, las escuelas privadas. Para ello, se planteó en primer lugar un filtro de variables con coincidencia exacta y un modelo de emparejamiento que permitiera delimitar una región de soporte común –es decir, escuelas del grupo de tratamiento y de control con un mejor balance entre sí- a partir de los factores observables. Las variables utilizadas para este modelo de emparejamiento fueron el tamaño de la escuela (aproximado con el número de alumnos), proporción de hombres en la IE, zona de competencia a la que pertenece la IE y tamaño de la jurisdicción a la que pertenece la IE (número de IE por provincia), mediante la estimación de un *propensity score* y un modelo de emparejamiento de vecindario más cercano, uno a uno, con posibilidad de reemplazo. En el Gráfico 1 de la sección Anexos, puede apreciarse la distribución del grupo de tratamiento y grupo de control.

A partir de ello, se propuso el modelo de diferencias en diferencias para reducir el potencial sesgo por las diferencias de los factores no observables entre el grupo de tratamiento y de control. No obstante, al observar que el supuesto de tendencias paralelas en la etapa previa al tratamiento no era aplicable en este caso (el Gráfico 2 permite realizar un análisis visual de la tendencia, y la Tabla 8 muestra los resultados de la prueba placebo), se consideró que una estimación por triple diferencia podía contribuir a reducir este sesgo, al diferenciar además por la tendencia previa a la línea de base. Las variables de control utilizadas fueron: área de procedencia, característica de la IE (polidocente completo o multigrado/unidocente), número de alumnos de segundo de primaria, proporción de estudiantes de sexo masculino en el grado, proporción de estudiantes en edad normativa para el grado, proporción de estudiantes del grado (entre 6 y 8 años de edad) que han cursado Educación Inicial, total de docentes para el nivel primaria, proporción de docentes hombres para el nivel, proporción de docentes con estudios superiores concluidos, y número de computadoras con acceso a Internet que se usan para fines pedagógicos. Se incluye además un rezago del rendimiento

alcanzando en Comprensión Lectora y Matemática. Asimismo, se considera información a nivel agregado, como el porcentaje de pobreza distrital, la pertenencia a un distrito JUNTOS, la pertenencia a un distrito del VRAEM, y la pertenencia a un centro poblado donde predomina el castellano. Finalmente, se incluye como control la intervención de algunos programas a cargo del Ministerio de Educación, que tienen el objetivo de mejorar los resultados estudiantiles, como Acompañamiento Pedagógico y Fortalecimiento de la Educación Física, y se incluyen variables dicotómicas regionales y temporales.

Como resultado, las Tablas 10 y 11 muestran los resultados de la estimación por triple diferencia y el efecto sobre los tratados (ATT), tanto para el rendimiento en Comprensión Lectora, como para el rendimiento en Matemática. Nótese que este primer análisis incluye a todas las escuelas, sin excluir a aquellos colegios considerados “seguros ganadores” y “seguros perdedores” (ver sección 4.1.1. para más detalle). En el primer caso (Tabla 10), se analiza el efecto competencia de manera agregada, considerando todo los periodos intervenidos como un *pool* de información. Por su parte, en el segundo caso (Tabla 11), el efecto competencia es desagregado por periodo intervenido, con la finalidad de analizar si existe alguna heterogeneidad en los resultados explicada por un periodo de maduración de la intervención, que pueda complementar el análisis del primer caso. Para ambas variaciones en el modelo, se consideran tres opciones: i) un modelo de efectos fijos a nivel de IE básico sin covariables, ii) un modelo de efectos fijos a nivel de IE que incluye variables de control, y iii) un modelo de efecto fijos a nivel de IE que además de incluir covariables incorpora una ponderación por IE, utilizando para ello el número de alumnos por escuela.

La evidencia de la Tabla 10 sustenta que hay un efecto competencia importante y significativo sobre los resultados en el área de Matemática. En el escenario más conservador (el caso del segundo modelo), el efecto competencia tendría un impacto positivo de 0.11 desviaciones estándar sobre los resultados de Matemática. Este efecto podría alcanzar 0.22 desviaciones estándar en el escenario más optimista. Por otro lado, los resultados parecen ser de menor impacto aunque significativos al analizar el efecto sobre el área de Comprensión Lectora, pues indican un efecto de 0.08 desviaciones estándar más sobre el puntaje promedio como producto del efecto competencia.

Los resultados de la Tabla 11, por otro lado, incluyen el análisis sobre el periodo de maduración de la intervención como factor que debe considerarse al momento de establecer su impacto sobre el rendimiento. Es importante precisar que el indicador del R-cuadrado muestra que este caso representa una mejor especificación. Los resultados consolidan la evidencia de un impacto mayor y con alta significancia del efecto competencia sobre el rendimiento estudiantil, para el caso del área de Matemáticas. Se observa además una diferencia en el efecto marginal sobre los resultados para cada uno de los años de la intervención, siendo que el año 2015 presenta un impacto menor en magnitud (0.11 desviaciones estándar), aunque en todos los periodos el área de Matemática se ve influenciada positivamente y de manera significativa por el efecto del bono. En el 2016, el efecto del bono fue de 0.21 desviaciones estándar, de acuerdo al modelo de efectos fijos con ponderación.

Asimismo, bajo este escenario resulta particularmente interesante el impacto de la competencia por el bono sobre el rendimiento en el área de Comprensión Lectora. Los resultados muestran una ambigüedad durante el primer año de la intervención, con un

efecto positivo aunque no significativo. Sin embargo, a partir del año 2016, el efecto del bono sobre el rendimiento se consolida en términos de significancia estadística, con un impacto de 0.09 desviaciones estándar bajo el modelo de efectos fijos con ponderación.

Por otro lado, uno de los objetivos del presente estudio es analizar si el efecto del bono es afectado al acotar el análisis a un grupo particular del tratamiento, aquel que se encuentra estrictamente en la zona de competencia. Por esta razón, en la Tabla 12 se muestran los resultados de la estimación por triple diferencia, excluyendo a los seguros ganadores (grupo 0) y seguros perdedores (grupo 4). Por implicar una mejor especificación, la comparación se realiza en relación al modelo que analiza los periodos de maduración.

Como puede observarse, las estimaciones realizadas evidencian que el impacto del efecto competencia se agudiza al acotar el análisis a la zona estrictamente en competencia. Sobre el efecto en el área de Matemática, en el 2016, el impacto fue de 0.27 desviaciones estándar, y de 0.16 desviaciones estándar para el caso de Comprensión Lectora. No obstante, en este grupo de análisis también es importante observar que en el año 2015 el impacto del bono muestra un efecto negativo para ambas áreas de evaluación, siendo que a partir del segundo año de intervención se revierte esta situación.

De lo mencionado anteriormente, se concluye que hay evidencia importante para considerar que el periodo de maduración del bono se consolidó a partir de su segundo año de implementación. El año 2015 presenta resultados ambiguos e incluso en varios escenarios significativamente negativos. Sin embargo, los resultados durante el año 2016 son sistemáticamente positivos y significativos en todos los casos analizados, para ambas áreas evaluadas, con una acentuación del efecto positivo al acotar el análisis a la zona de estricta competencia.

Por último, también se analizó la heterogeneidad de los resultados al desagregar la información por área de procedencia y característica de la IE. Nuevamente, estos análisis se realizaron sobre la base del modelo de triple diferencia con periodos de maduración, para la zona en estricta competencia. Con respecto al área de procedencia, los resultados indican que el efecto competencia del bono no es estadísticamente significativo en el área rural, en ninguna de las áreas evaluadas (Tabla 13). En ese sentido, la relevancia del bono se concentra en el área urbana, donde también se corrobora la consolidación del periodo de maduración del bono en el año 2016. En este periodo, el impacto del bono sobre el rendimiento estudiantil en el área de Matemática fue de 0.29 desviaciones estándar, mientras que en el área de Comprensión Lectora, fue de 0.17 desviaciones estándar (Tabla 14). Por su lado, en lo que se refiere a la característica de la IE, los resultados evidencian que el impacto del incentivo se concentra en las escuelas polidocentes completas, con una maduración a partir del segundo periodo. Así, en el año 2016 se puede observar un efecto sobre el rendimiento estudiantil de 0.28 desviaciones estándar para el caso de Matemática, mientras que para el caso de Comprensión Lectora, el impacto es de 0.17 desviaciones estándar (Tabla 16). Sin embargo, para el caso de las escuelas unidocentes o multigrado, los efectos no resultan significativos en ninguna de las áreas evaluadas o periodo de intervención (Tabla 15).

5.1.2 Efecto reciprocidad

Acerca de este segundo efecto, el objetivo es analizar si los profesores ganadores del bono en el periodo t-1 obtienen mejores resultados en el periodo t. Para ello, el grupo de control considerado son aquellos colegios públicos que se encontraron muy cerca del punto de corte en el índice de desempeño escolar estimado para alcanzar el bono; sin embargo, al no superar dicho límite, no ganaron. En ese sentido, se considera que estos colegios son muy similares en términos de desempeño con el grupo de tratamiento (escuelas ganadoras), y por tanto, la diferencia en los resultados obtenidos en el periodo siguiente al otorgamiento del bono puede entenderse como el efecto *ex post* al bono.

Para ello se utilizó un modelo de regresión discontinua, sobre la base de una regresión no paramétrica. El diseño utilizado fue el de regresión discontinua de diseño difuso, al haber evidencia de que, aunque sí existe un cambio en la probabilidad de tratamiento alrededor del punto de corte, esta probabilidad no pasa de cero a uno. En otras palabras, existen algunos casos de escuelas que, habiendo tenido un puntaje mejor al umbral, se hicieron acreedores del Bono, y de igual manera a la inversa. De la información disponible se observa que esta probabilidad de tratamiento pasa de 0.58% a 99% en el umbral. En el Gráfico 3 se puede visualizar este cambio en la probabilidad alrededor del punto de corte, centralizado en cero, para cada año.

Teniendo en cuenta que Bono Escuela tiene tres periodos de otorgamiento del bono, es posible analizar el efecto reciprocidad en dos periodos, 2015 y 2016, tanto para el área de Comprensión Lectora como de Matemática. Asimismo, es sabido que bajo el método de regresión discontinua, los resultados son altamente sensibles al ancho de banda que se defina alrededor del punto de corte. Por tal razón, se presentan los resultados sobre el ancho de banda óptimo estimado, así como de los resultados al considerar el 50% y el 200% de esta banda. De igual manera, con el fin de una mejor identificación, los resultados se acotaron al total de colegios públicos que mantuvieron igual al menos el 80% de su plana docente, que representan aproximadamente el 60% de escuelas públicas.

Así, tanto para el año 2015 como para el año 2016, y para ambas áreas de evaluación, la evidencia sugiere que el efecto reciprocidad sería positivo, relativamente pequeño y con muy poca significancia estadística (ver Tabla 17). Ello se observa tanto para el cálculo bajo el ancho de banda óptimo, como para un grupo más reducido alrededor del punto de corte (50% del ancho de banda original). Por su parte, al analizar los resultados bajo un ancho de banda más extenso (200% del ancho de banda original), los resultados aumentan en magnitud y significancia estadística. Sin embargo, al representar un grupo de escuelas más extenso, dicho resultado debe observarse con mayor cuidado, pues puede incluir un grupo de escuelas de menor comparabilidad. En consecuencia, en términos generales se puede afirmar que la evidencia no respalda la hipótesis de un efecto reciprocidad.

5.2 Resultados en el nivel de secundaria

5.2.1 Efecto competencia

De manera similar al nivel primaria, “Bono Escuela” también aplica a todas las escuelas

públicas para el caso de secundaria. Por consiguiente, el grupo de control también debe ser hallado entre las escuelas privadas. Sin embargo, a diferencia de primaria, en el caso de secundaria los logros de aprendizaje empezaron a evaluarse al cierre del año 2015. En ese sentido, surge una dificultad particular de no contar con una línea de base que permita estimar los efectos fijos de los factores no observables, recoger la tendencia previa y de esa manera reducir el riesgo de sesgo de los estimadores. No obstante, una característica adicional del proceso de evaluación en el nivel secundaria fue la posibilidad de incluir un cuestionario dirigido al estudiante. En este cuestionario se incluyeron variables relevantes relacionadas al entorno particular del estudiante, así como del entorno escolar. Se considera que esta información puede reducir parcialmente el sesgo producido por los efectos fijos de los factores no observables que diferencian al grupo de tratamiento del grupo de control. Además, contar con información relevante a nivel de estudiante permitirá plantear el emparejamiento a este nivel, lo cual podría contribuir a una mejor comparabilidad.

Dadas las limitaciones señaladas, se estima un *propensity score*, a partir de las siguientes variables: condición de repitente del estudiante, percepción sobre la atención de los padres por su vida personal y académica, percepción sobre el clima escolar, percepción sobre las prácticas pedagógicas del docente, número de instituciones educativas en la UGEL a la que está adscrita la escuela, y zona de competencia de la IE. El método utilizado fue de vecindario más cercano, de uno a uno, con posibilidad de reemplazo. El Gráfico 4 permite visualizar la distribución de la medida de propensión entre el grupo de tratamiento y el tratado. Asimismo, en las Tablas 18 y 19 de la sección de Anexos se brindan mayores detalles de la estimación del ATT bajo la estimación por *propensity score*.

Los resultados para el año 2015 (Tabla 18) no parecen evidenciar un efecto competencia favorable en este nivel. Por el contrario, el grupo de tratamiento sufre un efecto adverso, tanto en el área de Comprensión Lectora como en Matemática. Esta situación no es diferente al acotar la población a la zona de colegios en estricta competencia, ni tampoco al desagregarla según área de procedencia. Para el caso del año 2016, los resultados no son diferentes, evidenciando un efecto negativo de la intervención cuando se analiza el total de escuelas, así como cuando se acota el análisis a la zona en estricta competencia y según área de procedencia. Desde el punto de vista metodológico, los resultados adversos pueden deberse a las limitaciones de la información y un posible sesgo de los factores no observables. Otra explicación podría estar asociada al periodo de maduración de la intervención, la cual, si se comporta de manera similar al caso de primaria, podría no evidenciar resultados positivos sino hasta un tercer año.

Complementariamente, con la finalidad de analizar la heterogeneidad según el periodo de maduración y reducir las limitaciones que se enfrentan para este nivel al no contar con una línea de base que permita corregir el riesgo de factores no observables, se plantea un modelo de diferencia en diferencia, que permitirá retirar en alguna medida el factor tendencial de los resultados en el segundo año, y así evaluar el efecto marginal de la intervención para este año, con respecto al año 2015. Al aplicar esta metodología los resultados muestran evidencia de que en el año 2016 se empieza a ver un impacto positivo y significativo de la intervención, aunque pequeño. Tanto en Comprensión Lectora como en Matemática, la intervención tendría un impacto positivo de 0.03 desviaciones estándar en el segundo año de la intervención (Tabla 20). Al acotar el análisis a la zona en estricta competencia, los resultados son similares para el caso de Comprensión Lectora (0.011 desviaciones estándar, no significativo), dando indicios de

que no se da la presencia de heterogeneidad para esta área. Sin embargo, para el caso de Matemática, el impacto es ligeramente mayor, de 0.044 desviaciones estándar (Tabla 21). Finalmente, al desagregar el análisis según área de procedencia, nuevamente los resultados se inclinan a no mostrar ningún impacto de la intervención en la zona rural – igual que en el caso de primaria- y, sin embargo, concentrar los efectos de la intervención en la zona urbana. Estos últimos resultados pueden revisarse con mayor detalle en la Tabla 22 y 23 de la sección Anexos.

6. Conclusiones, limitaciones e implicancias de política

El sector educativo se ha caracterizado por plantear esquemas salariales que no se ajustan a una lógica de remuneración condicionada a resultados. Por ejemplo, en la región de Latinoamérica, el esquema de contratos en el sector está planteado tradicionalmente como un monto fijo, diferenciado en función a criterios de insumo; es decir, criterios basados principalmente en las características del docente (nivel de experiencia, títulos pedagógicos), sin ninguna relación a su desempeño en el aula, sea por el cumplimiento de procesos intermedios, o por los resultados obtenidos en sus estudiantes. Como consecuencia, el riesgo de que un agente (en este sector, el docente) no brinde su mayor esfuerzo ha sido asumido enteramente por el principal (en este caso, la sociedad y padres de familia en su conjunto, representados por las autoridades nacionales que contratan directamente al docente).

Debido a los problemas que implican estos esquemas de pago fijo, desde hace algunas décadas que varias experiencias internacionales plantean la necesidad de migrar hacia nuevos esquemas de pago en el sector educativo, que distribuyan la incertidumbre entre el principal y el agente, incluyendo un factor variable en el salario sujeto a resultados objetivos. Al respecto, hay un amplio espectro de variaciones en estos mecanismos de incentivos, pues son diferentes componentes que deben considerarse, como la proporción del incentivo en relación al salario fijo, su planteamiento de manera individual o colectiva, la definición de qué metas evaluar –bajo criterios de proceso o resultado- y cómo medirlas, con qué temporalidad definir la meta evaluada y el incentivo, así como su planteamiento como recompensa o penalidad (monetarias o no monetaria). El conjunto de decisiones que se tome en relación a estos factores afectará los resultados obtenidos en diversos sentidos. Por esta razón, como ha mostrado la evidencia internacional, las evidencias alrededor de la efectividad de un esquema de incentivos son aún poco concluyentes sobre la combinación ideal para una mayor efectividad.

A ello debe añadirse que, como señala Levacic (2009), la naturaleza de la motivación del docente promedio influirá en el éxito de un esquema de incentivos monetarios, pudiendo ser contraproducente en el caso de ser agentes del tipo “altruista” (con objetivos personales alineados a los del principal), y favorecedor en el caso de ser agentes del tipo “oportunista” (con objetivos personales particulares). No obstante, como plantea Le Grand (2003), es importante no categorizar anticipadamente a un determinado tipo de profesional como “altruista” u “oportunista”, pues la naturaleza de su motivación puede variar en el tiempo y según el contexto. Diversos estudios han demostrado la efectividad de incentivos monetarios en el sector educativo, mientras que otros han demostrado un efecto nulo o intrascendente en el tiempo, razón por la cual no es posible anticipar una

caracterización de la motivación intrínseca del docente. Sin perjuicio de ello, la estrategia de plantear un nuevo esquema salarial que incluya un factor variable sujeto a resultados implica un reto importante en el proceso de transformación hacia una cultura de rendición de cuentas, de definición de indicadores objetivos y medibles, de competitividad y de meritocracia.

En el Perú, la introducción de un esquema de incentivos es aún muy reciente dentro del sector. Implementado a finales del 2014, el Bono de Incentivos al Desempeño Escolar, “Bono Escuela” es la primera propuesta que tiene el objetivo de otorgar una compensación adicional a aquellas escuelas (director y docentes) que logran destacar en indicadores asociados a procesos y resultados propios del servicio educativo, en relación a un grupo de escuelas que pueden considerarse comparables entre sí. La apuesta económica por esta inversión ha sido considerable, al premiar a 10,529 escuelas en el año 2016 (64,394 docentes), entre escuelas del nivel Primaria y Secundaria. Por ello, es altamente relevante evaluar su impacto en los diferentes resultados que propone priorizar, como el logro estudiantil, retención de estudiantes y asistencia a clases, así como algunos indicadores de procesos, como ciertas prácticas docentes, clima escolar y gestión administrativa. En esta investigación se centra el análisis en el impacto de la intervención sobre el logro estudiantil, por ser el indicador que mejor recoge la evolución en la calidad escolar, además de ser el componente que representa hasta el 65% en la ponderación del indicador para la asignación del bono en el caso de primaria, y hasta el 50% para el caso de secundaria.

En “Bono Escuela”, el principal es el Ministerio de Educación, y el agente del contrato es el docente. Además, este bono se caracteriza por haber sido definido como monetario y positivo (premio económico), de tipo colectivo (se premia a todos los docentes de las escuelas ganadoras), y condicionado bajo un criterio a nivel de producto (sujeto principalmente al logro de aprendizajes y retención escolar). Considerando que este incentivo ha sido planteado como política universal (aplica a todas las escuelas públicas de primaria y secundaria a nivel nacional), no es posible identificar el efecto marginal de cada uno de estos factores considerados. No obstante, el propósito de esta investigación fue evaluar dos potenciales mecanismos a través de los cuales operaría el bono. Uno de ellos ha sido denominado “efecto competencia”, y alude al efecto del incentivo como producto de aspirar a obtener la bonificación y, para ello, replantear la estrategia y el nivel de esfuerzo que pone el docente para superar sus propios resultados y el de otras escuelas. Por otro lado, se evalúa otro efecto, denominado “efecto reciprocidad”, que postula la posibilidad de que el docente ganador del bono alcance un resultado superior en el periodo siguiente a la compensación, como una suerte de retribución social por el premio recibido. Al respecto, es importante precisar que aunque el diseño del incentivo ha sido originalmente pensado para actuar bajo el mecanismo de competencia, la literatura empírica sustenta como relevante evaluar el efecto reciprocidad como un efecto indirecto sobre el rendimiento escolar que puede producirse por la obtención del bono.

En el caso de Primaria, con respecto al efecto competencia, los resultados revelan que el periodo de maduración es un factor importante por tener en cuenta al momento de determinar un impacto de la intervención, pues se evidencia que la consolidación de este incentivo monetario se concreta en el segundo año de implementación (en el 2016), tanto para el área de Comprensión Lectora como de Matemática, siendo mayor el efecto en esta última área. Los resultados también muestran que el efecto es mayor en aquellas

escuelas que se encuentran en una zona expectante del bono; es decir, que se perciben como capaces de poder acceder al bono (no son “seguros perdedores”), pero sin llegar a estar del todo confiadas en que lo recibirán (no son “seguros ganadores”). La evidencia sugiere además que el impacto positivo y significativo del incentivo se ha concentrado en el área urbana y entre los colegios polidocentes completos, y que no ha tenido un efecto significativo en el área rural ni en las escuelas multigrado o unidocentes. Por su lado, acerca del efecto reciprocidad, los resultados obtenidos sugieren que no hay un impacto de la intervención a través de este canal.

En el caso de Secundaria, aunque se reconocen las mayores limitaciones para la estimación del impacto de la intervención como producto de no poder controlar los factores no observables con el modelo de emparejamiento, los resultados al aplicar una doble diferenciación parecen ser consistentes con la importancia de considerar un periodo de maduración, pues ha sido posible ver resultados positivos aunque pequeños en el segundo año de implementación, y es incierta su evolución durante los siguientes periodos. La evidencia en el caso de Secundaria también es consistente con la idea de que es en la zona urbana donde se concentra el efecto positivo del bono, aunque no parece respaldar la hipótesis de que haya una diferencia en los resultados para el caso de las escuelas en la zona de estricta competencia. Acerca del efecto reciprocidad, este canal no ha podido medirse, dado el corto tiempo de implementación que tiene Bono Escuela en este nivel.

En relación a los resultados en zona rural o en escuelas multigrado o unidocentes, es posible que ello nos de indicios de un docente que tiene una caracterización de agente diferente. En línea con lo señalado con Le Grand, podría ser un agente con una naturaleza de motivación tal, que un incentivo monetario puede ser contraproducente en sus resultados. De ser así, es posible considerar que en estos contextos sea necesario replantear el esquema de incentivos de manera más consistente con la teoría de la naturaleza motivacional, reemplazándolo por otro tipo de incentivos, como aquellos no monetarios, donde el reconocimiento se vuelve más importante. Por otro lado, a la luz de estos resultados, también es posible considerar que en las escuelas de este tipo sea particularmente importante rediseñar los incentivos en términos de qué indicadores medir. Como han señalado algunas experiencias externas, es posible que cuando se plantean indicadores de resultados, el impacto no sea el esperado, por el hecho de que el agente no tenga una idea clara de qué factores se asocian más a una mejora en el indicador (en este caso, el rendimiento estudiantil). Al no tener claridad de la función de producción y de cuáles de sus acciones están estrechamente asociadas a ella, puede fallar en su intento por obtener el premio. En ese sentido, algunas evidencias internacionales señalan la posibilidad de que un incentivo basado en procesos (acciones que dependen directamente de la decisión y capacidad del docente) puede terminar siendo más beneficioso para el producto final que se desea impactar.

Acerca de las limitaciones del presente estudio, es importante resaltar que el uso de las escuelas privadas como grupo de control puede representar un riesgo de sesgo importante para la medición del efecto competencia. Como señalan algunos estudios (Ryan, 2014; Epple, Figlio y Romano, 2000), existe evidencia de que el nivel de ingresos del hogar y las habilidades académicas del estudiante, así como la valoración de la educación formal por los padres, pueden ser predictores del tipo de gestión en el cual los padres matriculan a sus hijos, sugiriendo así que las características no observables asociadas al contexto particular del estudiante pueden contribuir sustancialmente en las

diferencias de los resultados académicos observables de los alumnos. A ello se suma las diferencias en las características propias de la escuela pues, como sugieren Alcázar y Valdivia (2009), que analizaron el caso de las escuelas de Fe y Alegría, la gestión privada puede tener ventajas no observables, como un mayor grado autonomía con respecto a su plan de estudios y sus decisiones administrativas, que conllevaría a una mayor eficacia en el uso de los recursos disponibles.

En este estudio se han planteado medidas complementarias para la mitigación de dicho riesgo, como la aplicación de un modelo de emparejamiento que mejore la comparabilidad entre escuelas, sobre la base de variables consideradas en un estudio de Rau y Contreras (2012) para el caso chileno, que incluye variables sobre las características de la escuela, como el área de procedencia, tamaño de la escuela, proporción de hombres en la IE, zona de competencia a la que pertenece la IE y número de IIEE por provincia. Adicionalmente, para el caso del nivel secundario, la información disponible permitió incluir variables de emparejamiento asociadas al contexto particular del estudiante y su escuela, como la condición de repitente del estudiante, la percepción del estudiante sobre la atención de sus padres por su vida personal y académica, su percepción sobre el clima escolar, y su percepción sobre las prácticas pedagógicas del docente. Complementariamente, se propuso una estimación bajo un modelo de triple diferenciación y efectos fijos a nivel de IE que reduzca el sesgo por factores no observables de la escuela y por las diferencias en la trayectoria previa a la intervención entre escuelas públicas tratadas y escuelas privadas de control (incumplimiento del supuesto de tendencias comunes del modelo de doble diferencia). Pese a ello, no es posible asegurar que todo el potencial sesgo por no observables se haya eliminado. En el momento previo al inicio del bono escuela, las escuelas públicas mostraron mejoras en sus desempeño producto quizá de intervenciones previas que implementó el sector en escuelas públicas para mejorar los logros de aprendizaje (acompañamiento pedagógico), que no permitirían deducir la dirección del sesgo. A ello debe agregarse la dificultad de controlar por las intervenciones complementarias que implementa el sector con la finalidad de mejorar los productos y procesos en la gestión pública de las escuelas, como las políticas universales de revalorización docente o de Compromisos por Desempeño (que premia a todas las UGEL que cumplan determinadas metas del sector). Los niveles de R cuadrado podrían dar indicios de que futuras investigaciones pueden explorar la posibilidad de una mejor especificación para reducir el sesgo por variables omitidas.

Con respecto a los hallazgos sobre el efecto reciprocidad, aunque estos no resultaron significativos, es posible que los resultados reflejen la necesidad de un mayor tiempo de maduración de la intervención. En todo caso, se considera importante seguir explorando este canal en futuras investigaciones. De igual manera, el incentivo propuesto también tiene el objetivo de promover otros indicadores de resultado y de proceso, si bien tienen un peso menor en el índice generado para el otorgamiento del bono. En ese sentido, sería importante que futuras evaluaciones analicen el efecto de “Bono Escuela” en estas otras variables, como retención escolar, o ciertas acciones pedagógicas.

Finalmente, aunque no es posible aún saber si los efectos del bono perdurarán en el tiempo, se considera que este nuevo planteamiento de esquema salarial debe ser considerado como una herramienta de necesidad permanente, en la medida que permite una mejor distribución de la incertidumbre de los resultados del sector, y un cambio en la lógica de asignación. En ese sentido, se considera altamente relevante

seguir invirtiendo en este tipo de incentivos, y concentrar los esfuerzos a indicadores de resultados como el rendimiento estudiantil, además de otros procesos claves para el aprendizaje, como lo pueden ser las prácticas pedagógicas dentro del aula, u otras acciones concretas y medibles que contribuyan a la mejora en el clima escolar.



7. Bibliografía y referencias

- Abadie, A.; Imbens, G. (2006): "Large Sample Properties of Matching Estimators for Average Treatment Effects". *Econometrica*, 74(1), 235–267.
- Akerlof, G. (1982). "Labor Contracts as Partial Gift Exchange". *The Quarterly Journal of Economics*.
- Akerlof, G. (1984). "Gift Exchange and Efficiency-Wage Theory: Four Views". *American Economic Review* N° 74(2), pp. 79–83.
- Alcázar, L.; Valdivia, N. (2009). "Escuelas de Fe y Alegría en el Perú: Análisis del Modelo de Gestión Institucional y Pedagógica y Lecciones para la Educación Pública". Lima: Banco Mundial.
- Angelucci, M; Prina, S.; Heather, R.; y Samek, A. (2015). "When incentives backfire: spillover effects in food choice". NBER, documento de investigación N° 21481.
- Atkinson, A.; Burgess, S.; Croxson, B.; Gregg, P.; Propper, C.; Slater, H.; Wilson, D. (2004). "Evaluating the Impact of Performance-related Pay for Teachers in England". *Labour Econ* N° 16(3), pp. 251–261.
- Babcock, P.; Bedard, K.; Charness, G.; Hartman, J.; y Royer, H. (2011) "Letting down the team? Evidence of social effects of team incentives". NBER, documento de investigación N° 16687.
- Benhabib, J.; Spiegel, M. (1994). "The role of human capital in economic development evidence from aggregate cross-country data". *Journal of Monetary Economics*, 34, N° 2, pp. 143-173.
- Blackwell, M.; Iacus, S.; King, G.; y Porro, G. (2010). "Coarsened exact matching in Stata".
- Casavilca, P. (2014). "Sobre sistemas de incentivos al desempeño docente: ¿más ganadores induce mayor efectividad? El caso del SNED en Chile". Tesis para el grado de Magíster en Economía. Instituto de Economía. Pontificia Universidad Católica de Chile.
- Cawley, J.; Price, J. (2009). "Outcomes in a program that offers financial rewards for weight loss". NBER, documento de investigación N° 14987.
- Charness, G.; Kuhn, P. (2011). "Lab Labor: What Can Labor Economists Learn from the Lab?". *Handbook of Labor Economics*, Volume 4. Ed. Orley Ashenfelter y David Card. North Holland, primera edición.
- Chaudhury, N.; Parajuli, D. (2006) "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan." Documento de investigación N° 4102, Banco Mundial. Washington, DC.
- Chevalier, A.; Dolton, P. (2005). "The labour market for teachers". Disponible en: S. Machin y A. Vignoles (Editores). "What's the Good of Education?". *Princeton University Press*.
- Contreras, D.; Larrañaga, O.; Flores, L.; Lobato, F; y Macías, V. (2005): Políticas educacionales en Chile: vouchers, concentración, incentivos y rendimiento. En: *Uso e*

impacto de la información educativa en América Latina. Chile: PREAL, pp. 61-110.

DellaVigna, S.; Pope, D. (2016). "What motivates effort? Evidence and expert forecasts". NBER, documento de investigación N° 22193. Dixit, A. (2002). "Incentives and organizations in the public sector: an interpretative review". *Journal of Human Resources*, 37, pp. 696–728.

Douma, S.; Schreuder, H. (2002). "Economic Approaches to Organizations". Pearson Education. Londres.

Epple, D.; Figlio, D.; y Romano, R. (2000). "Competition between private and public schools: testing stratification and pricing predictions". NBER, documento de investigación N° 7956.

Eide, E.; Goldhaber, D.; Brewer, D. (2004). "The teacher labour market and teacher quality". *Oxford Review of Economic Policy*, 20, pp. 230–243.

Ferraz, C.; Finan, F. (2009). "Motivating Politicians: The Impacts Of Monetary Incentives On Quality And Performance". NBER, documento de investigación N° 14906.

Fryer, R. (2010) "Financial Incentives and Student Achievement: Evidence From Randomized Trials". NBER. Documento de investigación N° 15898.

Gallego, F. (2008). "Efectos del SNED en resultados del proceso educativo." Manuscrito. Pontificia Universidad Católica de Chile.

Glewwe, P.; Kremer, M. (2006). "Schools, teachers and educational outcomes in developing countries". Disponible en: E. Hanushek y F. Welch (Editores). "Handbook on the Economics of Education". Amsterdam: North Holland.

Glewwe, P.; Ilias, N.; Kremer, M. (2010). "Teacher Incentives". *American Economic Journal: Applied Economics* 2(3), pp. 205–227.

Goodman, S.; Turner, L. (2010). "Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program". Universidad de Columbia.

Hanushek, E.; Kain, J.; Rivkin, S. (1999). "Do higher salaries buy better teachers?". *National Bureau of Economic Research. Working paper* N° 7082.

Hudson, S. (2010). "The Effects of Performance-Based Teacher Pay on Student Achievement". SIEPR. Paper de discusión. N° 09-023.

Imbens, G.; Wooldridge, J. (2008). "Recent Developments in the Econometrics of Program Evaluation." NBER, documento de investigación N° 14251. Cambridge, MA.

Imbens, G.; Kalyanaraman, K. (2009). "Optimal Bandwidth Choice For The Regression Discontinuity Estimator". NBER, documento de investigación N° 14726.

Khander, S.; Koolwal, G.; y Samad, H. (2010). "Handbook on Impact Evaluation. Quantitative Methods and Practices". Banco Mundial.

Kosfeld, M.; Neckermann, S. (2011). "Getting More Work for Nothing? Symbolic Awards and Worker Performance". *American Economic Journal: Microeconomics*, 3(3): 86-99.

Lavy, V. (1999). "Evaluating the Effect of Teachers' Performance Incentives on Pupils'

- Achievements". Universidad Hebrea de Jerusalem. Lavy, V. (2004). "Performance Pay And Teachers' Effort, Productivity And Grading Ethics". *National Bureau Of Economic Research. Working Paper* N°10622.
- Lazear, E. (2000) "Personnel Economics and Economic Approaches to Incentives," *HKCER Letters*, 61.
- Le Grand, J. (2003). "Motivation, Agency and Public Policy". *Oxford University Press*. Universidad de Oxford.
- León, J. (2016). "Evaluación de Impacto del Bono de Incentivo al Desempeño Escolar o Bono Escuela". Informe Preliminar. GRADE.
- Levacic, R. (2009). "Teacher Incentives and Performance: An Application of Principal-Agent Theory". *Estudios de Desarrollo de Oxford*, Vol. 37, No. 1, Marzo 2009.
- Levitt, S.; List, J.; Neckermann, S.; Sadoff, S. (2012). "The behavioralist goes to school: leveraging behavioral economics to improve educational performance". NBE. Documento de investigación N° 18165.
- Lucas, R. (1988). "On the Mechanics of Economic Development", *Journal of Monetary Economics*, No. 22, pp. 3-42.
- Martins, S.P. (2009). "Individual Teacher Incentives, Student Achievement and Grade Inflation". IZA Documento de discusión N° 4051.
- Mesinas, J. (2010). "Impact Evaluation of the Peruvian Conditional Cash Transfer Programme "JUNTOS" on Educational Performance ". *Katholieke Universiteit Leuven*.
- Mizala, A.; Romaguera, P. (2004). "School and teacher performance incentives: the Latin American experience". *International Journal of Educational Development*, 24, pp. 739-754.
- Muralidharan, K.; Sundararaman, V. (2010). "Teacher Performance Pay: Experimental Evidence from India". Documento de trabajo N° 15323. *National Bureau of Economic Research*.
- Organización para la Cooperación y el Desarrollo Económico – OCDE (2013). "PISA 2012 results. Vol I. Student Performance in Mathematics, Reading and Science". Paris: OCDE.
- Porter, J. (2003): "Estimation in the Regression Discontinuity Model". Departamento de Economía. Universidad de Harvard.
- Rau, T.; Contreras, D. (2012). "Tournaments, gift exchanges, and the effect of monetary incentives for teachers: the case of Chile". Universidad de Chile.
- Ravallion, M. (2005) "Evaluating Anti-Poverty Programs." *Handbook of Agricultural Economics* vol. 4. Robert E. Evenson y T. Paul Schultz. North-Holland.
- Ryan, C. (2014). "Private school 'effects' on student achievement in Australian schools". Instituto de Economía Aplicada e Investigación Social. Universidad de Melbourne.
- Springer, M.G.; Ballou, D.; Hamilton, L.; Le, V.; Lockwood, J.R.; McCafrey, D.; Pepper, M.; Stecher, B. (2010). "Teacher Pay for Performance: Experimental Evidence from the

Project on Incentives in Teaching”. *National Center on Performance Incentives*.
Universidad de Vanderbilt, Nashville, TN.



Tversky, A.; Kahneman, D. (1991). "Loss Aversion in Riskless Choice: A Reference-Dependent Model". *The Quarterly Journal of Economics*, 106(4): 1039-1061.

Unidad de Medición de la Calidad Educativa – UMC (2013). "PISA 2012: Primeros resultados. Informe Nacional del Perú". Primera edición. Lima.

Unidad de Medición de la Calidad Educativa – UMC (2015). "Reporte Técnico de la Evaluación Censal de Estudiantes (ECE 2015)". Ministerio de Educación. Lima.

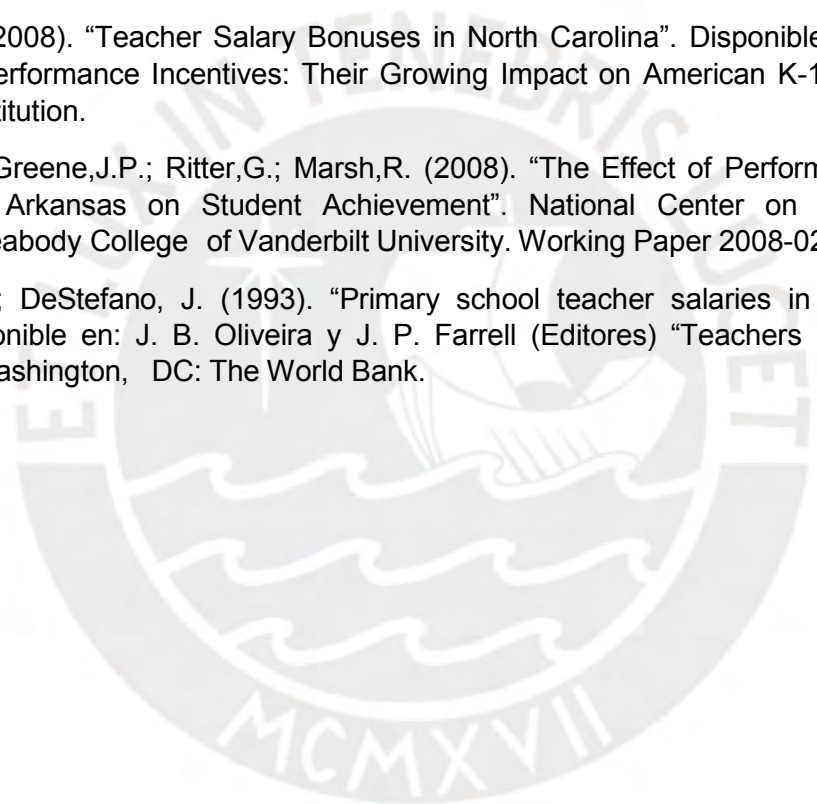
Van der Klaauw, W. (2008). "Regression Discontinuity Analysis: A Survey of Recent Developments in Economics," *Labour*, 22(2): 219-245.

Vegas, E. (2006). "Incentivos docentes y sus efectos en el aprendizaje del alumnado en Latinoamérica". Basado en el libro "Incentives to Improve Teaching: Lessons from Latin America". Banco Mundial, 2005.

Vigdor, J.L. (2008). "Teacher Salary Bonuses in North Carolina". Disponible en Springer, M. (Ed.), "Performance Incentives: Their Growing Impact on American K-12 Education". Brookings Institution.

Winters, M.; Greene, J.P.; Ritter, G.; Marsh, R. (2008). "The Effect of Performance-Pay in Little Rock, Arkansas on Student Achievement". National Center on Performance Incentives. Peabody College of Vanderbilt University. Working Paper 2008-02.

Zymelman, M.; DeStefano, J. (1993). "Primary school teacher salaries in Sub-Saharan Africa" Disponible en: J. B. Oliveira y J. P. Farrell (Editores) "Teachers in Developing Countries". Washington, DC: The World Bank.



8. Anexos

Tabla 4: Monto asignado a los ganadores de Bono Escuela, por grupo elegible y percentil

	Grupo A		Grupo B		Grupo C		Grupo D		Grupo E	
	Top 25	Top 25-35	Top 25	Top 25-35	Top 25	Top 25-35	Top 10	Top 10-20	Top 10	Top 10-20
Director	3,000	2,000	3,000	2,000	1,500	1,000	2,500	2,000	2,000	1,500
1°	2,500	1,667	2,500	1,667	1,500	1,000	2,000	1,500	1,500	1,000
2°	3,000	2,000	2,500	1,667	1,500	1,000	2,000	1,500	1,500	1,000
3°	1,500	1,000	2,500	1,667	1,500	1,000	2,000	1,500	1,500	1,000
4°	1,500	1,000	3,000	2,000	1,500	1,000	2,000	1,500	1,500	1,000
5°	1,500	1,000	1,500	1,000	1,500	1,000	2,000	1,500	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	2,000	1,500	1,500	1,000



Tabla 5: Distribución de escuelas que rindieron la ECE, por año evaluado

Segundo grado de primaria, Educación Básica Regular

Año	Nro. IIEE Privadas	Nro. IIEE		% IIEE públicas
		Públicas	Total	
2007	5,715	18,354	24,069	76%
2008	5,803	17,734	23,537	75%
2009	5,761	16,971	22,732	75%
2010	5,939	16,736	22,675	74%
2011	6,149	16,309	22,458	73%
2012	6,528	18,892	25,420	74%
2013	6,486	17,245	23,731	73%
2014	6,593	16,114	22,707	71%
2015	6,739	14,656	21,395	69%
2016	6,765	14,218	20,983	68%

Cuarto grado de primaria, Educación Intercultural Bilingüe

Año	Nro. IIEE Privadas	Nro. IIEE Públicas	Total	% IIEE públicas
2012	2	1658	1660	99.9%
2013	0	1632	1632	100.0%
2014	5	1861	1866	99.7%
2015	3	1835	1838	99.8%
2016	2	2014	2016	99.9%

Segundo grado de secundaria, Educación Básica Regular

Año	Nro. IIEE Privadas	Nro. IIEE Públicas	Total	% IIEE públicas
2015	4,365	8,321	12,686	66%
2016	4,300	8,575	12,875	67%

Tabla 6: Estratos considerados para los grupo A y D

(para los años previos a Bono Escuela , se incluye información de la simulación de la metodología planteada en esta intervención)

Segundo grado de primaria

Año del Bono	Nro. Estratos	Distancia entre el umbral del percentil 25 y 35, en desviaciones estándar
2008	372	-0.361
2009	380	-0.331
2010	379	-0.345
2011	378	-0.339
2012	387	-0.328
2013	387	-0.312
2014	369	-0.348
2015	366	-0.313
2016	368	-0.263

Segundo grado de secundaria

Año del Bono	Nro. Estratos	Distancia entre el umbral del percentil 25 y 35, en desviaciones estándar
2015	282	-0.751
2016	274	-0.505

Gráfico 1: Distribución del *propensity score* antes y después del modelo de emparejamiento, para el caso de primaria

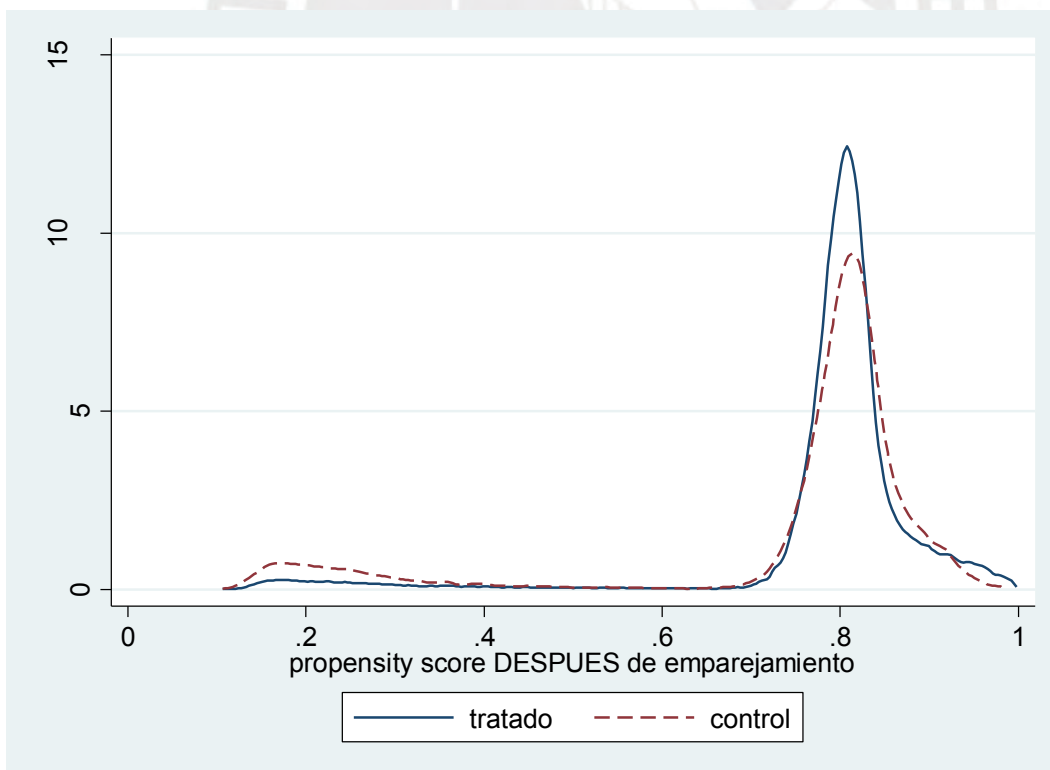
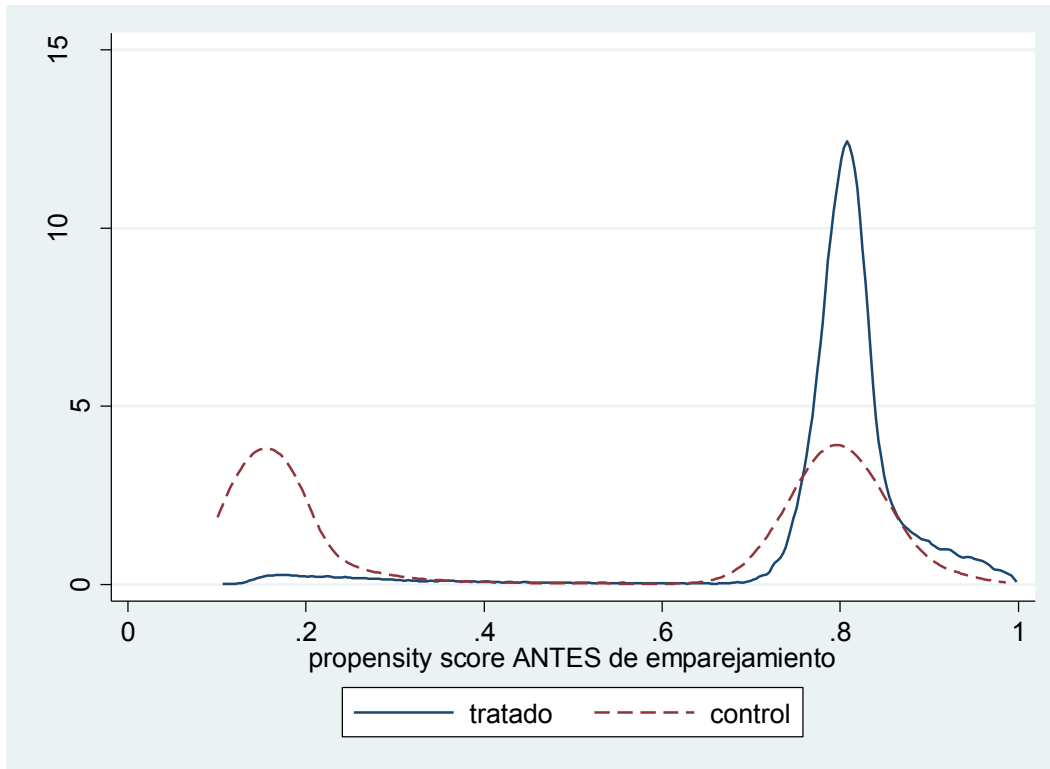


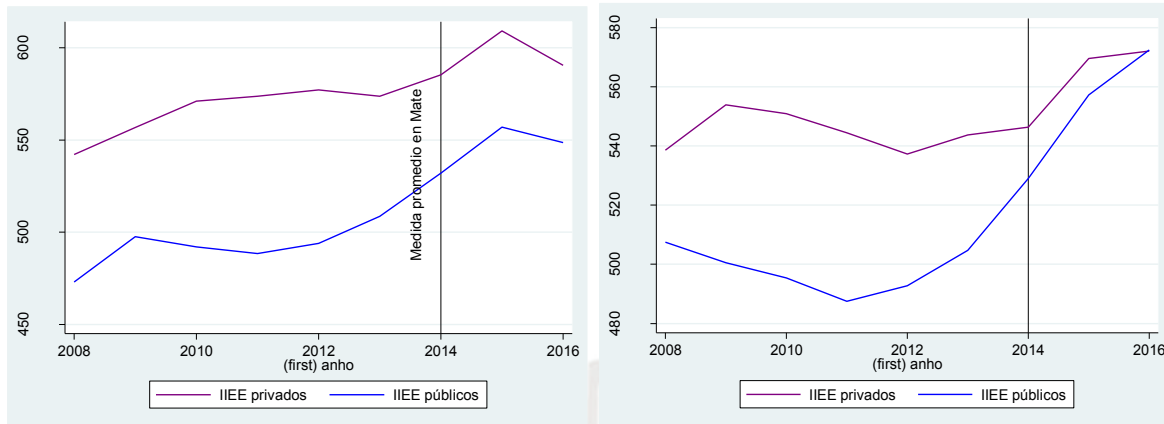
Tabla 7: Características promedio de los colegios públicos y privados de segundo grado de primaria, en la etapa previa y posterior a la intervención
(análisis sobre la zona de competencia)

	Periodo 2008-2013		Periodo 2014-2016	
	Privado	Público	Privado	Público
Medida ECE, en desviaciones estándar				
Comprensión Lectora	8.27 (0.76)	7.23 (1.02)	9.39 (0.96)	8.72 (1.25)
Matemática	6.95 (0.85)	6.48 (1.13)	6.24 (0.90)	6.32 (1.19)
Características del entorno				
Área (1=urbano)	0.96 (0.20)	0.34 (0.47)	0.98 (0.13)	0.44 (0.50)
Característica de la IE (1: polidocente completo)	0.81 (0.40)	0.38 (0.49)	0.82 (0.39)	0.46 (0.50)
Nro secciones de segundo de primaria	1.22 (0.59)	1.39 (0.96)	1.26 (0.63)	1.48 (1.05)
Nro de alumnos de segundo de primaria	20.31 (20.61)	22.16 (30.74)	21.56 (20.78)	25.04 (33.61)
Proporción de hombres en segundo de primaria	0.53 (0.18)	0.51 (0.19)	0.53 (0.17)	0.51 (0.18)
Variación anual de puntaje en CL	0.01 (0.09)	0.03 (0.17)	0.01 (0.08)	0.04 (0.13)
Variación anual de puntaje en Matemática	0.01 (0.11)	0.03 (0.21)	0.02 (0.12)	0.06 (0.20)
% población pobre en el distrito	23.37 (15.83)	53.94 (23.58)	22.26 (16.03)	49.97 (24.54)
Total de docentes en el nivel educativo	6.40 (4.55)	5.65 (7.30)	6.74 (5.12)	6.49 (7.70)
Proporción de docentes hombres en el nivel educativo	0.24 (0.22)	0.44 (0.33)	0.24 (0.21)	0.42 (0.31)
Total de docentes con estudios superiores concluidos	5.35 (4.51)	5.40 (7.10)	5.76 (5.08)	6.34 (7.62)
Proporción de docentes con estudios superiores concluidos	0.81 (0.29)	0.94 (0.19)	0.83 (0.29)	0.96 (0.15)
Nro de computadoras con acceso a Internet para uso pedagógico	12.07 (22.41)	5.36 (15.41)	13.38 (19.27)	7.93 (17.81)
Proporción de estudiantes de segundo de primaria que han cursado Educación Inicial	0.92 (0.21)	0.72 (0.38)	0.95 (0.21)	0.91 (0.25)
Total de estudiantes de segundo de primaria	21.83 (21.54)	25.67 (34.13)	22.23 (21.20)	26.75 (35.57)
Proporción de estudiantes de segundo de primaria en edad normativa	0.98 (0.08)	0.87 (0.16)	0.99 (0.06)	0.93 (0.12)
Distrito JUNTOS (1=sí pertenece)	0.18 (0.38)	0.50 (0.50)	0.07 (0.26)	0.60 (0.49)
Distrito VRAEM (1=sí pertenece)	0.01 (0.07)	0.03 (0.18)	0.00 (0.07)	0.04 (0.19)
Centro poblado donde predomina una lengua distinta al castellano (1= sí pertenece)	0.00 (0.07)	0.15 (0.36)	0.00 (0.04)	0.12 (0.32)
N° observaciones	20,097	40,488	36,666	93,321
N° IIEE	7,867	16,364	8,085	19,704

Gráfico 2: Análisis de tendencia en el rendimiento promedio de la IE, para el caso de segundo grado de primaria, en las áreas de Comprensión Lectora y Matemática

(Incluyendo todas las escuelas)

Medida promedio en CL



(Incluyendo solo las escuelas en zona de competencia)

Medida promedio en CL

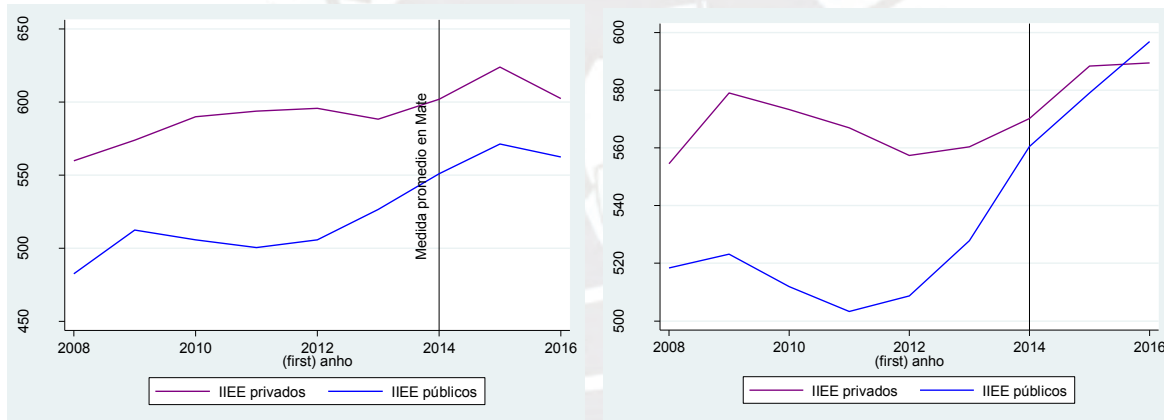


Tabla 8: Estimación placebo en doble diferencia para el análisis de tendencia previa, para las áreas de Comprensión Lectora y Matemáticas en el grado de segundo de primaria

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Post 2014 (t)	0.455*** -(41.35) [0.01]	-1.184*** (-82.64) [0.01]	0.855*** -(66.67) [0.01]	-1.968*** (-117.18) [0.02]	0.940*** -(79.38) [0.01]	-1.789*** (-102.70) [0.02]
Tratamiento (1=público=tratado)	-0.004 (-0.05) [0.08]	-0.294*** (-3.08) [0.10]	-0.018 (-0.19) [0.09]	-0.068 (-0.57) [0.12]	0.029 (-0.34) [0.09]	0.002 (-0.01) [0.12]
Post 2014 x tratado	0.273*** -(20.79) [0.01]	0.371*** -(22.42) [0.02]	0.304*** -(22.54) [0.01]	0.393*** -(22.94) [0.02]	0.266*** -(20.83) [0.01]	0.391*** -(22.02) [0.02]
Area (1=urbano)			0.009 (-0.63) [0.01]	-0.003 (-0.16) [0.02]	0 (-0.03) [0.01]	0.005 (-0.31) [0.02]
Característica de la IE (1: polidocente completo)			0.019 (-1.39) [0.01]	0.02 (-1.18) [0.02]	-0.006 (-0.43) [0.01]	-0.024 (-1.34) [0.02]
Nro de alumnos de 2do de primaria % de hombres en 2do de primaria % población pobre en el distrito			0 (-0.26) [0.00]	-0.000* (-1.71) [0.00]	0 (-1.15) [0.00]	0 (-1.59) [0.00]
Total de docentes en el nivel educativo % de docentes hombres en el nivel educativo % de docentes con estudios superiores concluidos			0.006*** (-5.39) [0.00]	0.005*** (-3.6) [0.00]	0.003*** (-3.11) [0.00]	0.001 (-0.84) [0.00]
Nro de computadoras con acceso a Internet para uso pedagógico % de estudiantes de 2do de primaria con Educ. Inicial			-0.001*** (-5.51) [0.00]	-0.001*** (-4.69) [0.00]	-0.001*** (-6.27) [0.00]	-0.001*** (-5.03) [0.00]
Proporción de estudiantes de 2do de primaria en edad normativa			0 (-1.49) [0.00]	0 (-1.37) [0.00]	0 (-0.77) [0.00]	0 (-0.55) [0.00]
Distrito JUNTOS (1=sí pertenece)			0 (-1.22) [0.00]	0.001*** (-2.87) [0.00]	0 (-0.49) [0.00]	0.000** (-2.24) [0.00]
Distrito VRAEM (1=sí pertenece)			0 (-0.06) [0.00]	0 (-0.76) [0.00]	0 (-1.06) [0.00]	-0.001*** (-2.28) [0.00]
CCPP donde predomina una lengua distinta al castellano (1= sí pertenece)			0 (-0.09) [0.00]	0 (-1.36) [0.00]	0 (-0.1) [0.00]	-0.001*** (-2.36) [0.00]
Rezag de un periodo de puntaje ECE – CL			-0.074*** (-6.88) [0.01]	0.001 (-0.11) [0.01]	-0.061*** (-5.59) [0.01]	0.003 (-0.22) [0.01]
Rezag de un periodo de puntaje ECE – Matemática			0.023 (-0.45) [0.05]	-0.095 (-1.58) [0.06]	-0.037 (-0.77) [0.05]	-0.123** (-1.97) [0.06]
Constante			-0.146 (-1.11) [0.13]	-0.051 (-0.30) [0.17]	-0.264** (-2.07) [0.13]	-0.068 (-0.37) [0.18]
			-0.001*** (-15.29) [0.00]	0.000*** (-2.8) [0.00]	-0.001*** (-15.12) [0.00]	0 (-0.92) [0.00]
			0.000*** (-6.88) [0.00]	-0.001*** (-12.12) [0.00]	0.001*** (-9.51) [0.00]	-0.000*** (-5.07) [0.00]
	7.241***	6.709***	7.457***	7.639***	7.736***	7.815***

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Tendencia	-(105.95) [0.07]	-(80.37) [0.08]	-(78.15) [0.10]	-(63.4) [0.12]	-(89.2) [0.09]	-(66.39) [0.12]
Efecto fijo IE	No	No	Sí	Sí	Sí	Sí
R-cuadrado ajustado	Sí	Sí	Sí	Sí	Sí	Sí
Nro. observaciones	0.093	0.096	0.307	0.314	0.446	0.411
	128,423	128,423	107,381	107,381	2,744,921	2,744,921

t-score en parentesis; error estandar robusto en llaves

* p<0.10, ** p<0.05, *** p<0.010



Tabla 9: Probabilidad de otorgamiento del bono para las escuelas públicas por debajo y encima del umbral.

(Escuelas que NO superaron el puntaje mínimo de su estrato para el otorgamiento del bono)

	2014		2015		2016	
	N° escuelas	%	N° escuelas	%	N° escuelas	%
No recibieron el bono	8,424	100	7,674	99.75	7,601	99.42
Recibieron el bono	0	0	19	0.25	44	0.58
Total	8,424	100	7,693	100	7,645	100

(Escuelas que superaron el puntaje mínimo de su estrato para el otorgamiento del bono)

	2014		2015		2016	
	N° escuelas	%	N° escuelas	%	N° escuelas	%
No recibieron el bono	69	1.53	40	1.02	40	1.04
Recibieron el bono	4451	98.47	3864	98.98	3,814	98.96
Total	4,520	100	3,904	100	3,854	100



Tabla 10: Estimación por triple diferencia, para todas las escuelas evaluadas en segundo grado de primaria

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2015 en adelante)	0.912*** -(71.34) [0.01]	0.833*** -(58.21) [0.01]	0.885*** -(67.42) [0.01]	0.857*** -(57.3) [0.01]	0.850*** -(67.37) [0.01]	0.801*** -(53.38) [0.02]
Periodo tratado (1=2014 en adelante)	0.449*** -(41.35) [0.01]	-1.193*** (-86.49) [0.01]	0.447*** -(39.57) [0.01]	-1.204*** (-83.50) [0.01]	0.480*** -(44.09) [0.01]	-1.168*** (-77.45) [0.02]
Tratamiento (1=público=tratado)	0.086 -(1.35) [0.06]	-0.254*** (-3.36) [0.08]	0.238*** -(3.17) [0.08]	-0.111 (-1.24) [0.09]	0.295*** -(4.04) [0.07]	-0.073 (-0.91) [0.08]
post2015 x tratado	0.089*** (-5.89) [0.02]	0.130*** (-7.7) [0.02]	0.043*** (-2.72) [0.02]	0.114*** (-6.42) [0.02]	0.079*** (-5.27) [0.02]	0.215*** (-12.26) [0.02]
post2014 x tratado	0.287*** -(22.15) [0.01]	0.376*** -(23.43) [0.02]	0.115*** -(8.29) [0.01]	0.375*** -(21.89) [0.02]	0.093*** -(6.83) [0.01]	0.368*** -(20.78) [0.02]
Area (1=urbano)			0.023* (-1.86) [0.01]	-0.053*** (-3.53) [0.01]	-0.015 (-1.30) [0.01]	-0.074*** (-4.77) [0.02]
Característica de la IE (1: polidocente completo)			0.025** (-2) [0.01]	-0.066*** (-4.25) [0.02]	0.060*** (-4.59) [0.01]	-0.104*** (-5.86) [0.02]
Nro de alumnos de 2do de primaria			0.002*** (-7.39) [0.00]	-0.003*** (-10.97) [0.00]	0.002*** (-8.4) [0.00]	-0.002*** (-8.18) [0.00]
% de hombres en 2do de primaria			-0.001*** (-6.55) [0.00]	0 (0) [0.00]	-0.001*** (-8.06) [0.00]	0 (-0.46) [0.00]
Soporte Pedagógico (1=sí recibe)			0.285*** -(24.52) [0.01]	0.240*** (-18.09) [0.01]	0.207*** (-18.65) [0.01]	0.138*** (-10.56) [0.01]
Plan de Fortalecimiento de Educación Física (1=sí recibe)			0.143*** -(10.75) [0.01]	0.101*** (-6.6) [0.02]	0.104*** (-7.56) [0.01]	0.042*** (-2.78) [0.02]
Acompañamiento Rural o EIB (1=sí recibe)			0.389*** -(24.07) [0.02]	0.140*** (-8.09) [0.02]	0.301*** (-18.54) [0.02]	0.093*** (-5.43) [0.02]
% población pobre en el distrito			-0.003*** (-9.82) [0.00]	0 (-0.68) [0.00]	-0.004*** (-12.39) [0.00]	-0.002*** (-4.52) [0.00]
Total de docentes en el nivel educativo			0.012*** (-10.18) [0.00]	0.005*** (-4.19) [0.00]	0.007*** (-6.29) [0.00]	0.002 (-1.49) [0.00]
% de docentes hombres en el nivel educativo			-0.001*** (-5.65) [0.00]	-0.001*** (-4.74) [0.00]	-0.001*** (-6.72) [0.00]	-0.001*** (-4.57) [0.00]
% de docentes con estudios superiores concluidos			0 (-0.93) [0.00]	0 (-0.23) [0.00]	-0.000** (-2.33) [0.00]	0 (-0.16) [0.00]
Nro de computadoras con acceso a Internet para uso pedagógico			-0.000** (-2.20) [0.00]	-0.001*** (-6.14) [0.00]	-0.000** (-2.47) [0.00]	-0.000** (-2.05) [0.00]
% de estudiantes de 2do de primaria con Educ Inicial			0.000*** (-3.82) [0.00]	0.001*** (-12.36) [0.00]	0.000*** (-4.57) [0.00]	0.001*** (-5.49) [0.00]
Proporción de estudiantes de 2do de primaria en edad normativa			0.002*** (-7.92) [0.00]	-0.002*** (-7.17) [0.00]	0.003*** (-11.04) [0.00]	-0.002*** (-5.47) [0.00]

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Distrito JUNTOS (1=sí pertenece)			0.408*** -(64.76) [0.01]	-0.214*** (-29.08) [0.01]	0.421*** -(70.77) [0.01]	-0.218*** (-30.61) [0.01]
Distrito VRAEM (1=sí pertenece)			0.171*** -(3.67) [0.05]	0.130** (-2.17) [0.06]	0.144*** -(3.28) [0.04]	0.183*** -(3.12) [0.06]
CCPP donde predomina una lengua distinta al castellano (1= sí pertenece)			-0.15 (-1.59) [0.09]	-0.166* (-1.68) [0.10]	-0.079 (-0.89) [0.09]	-0.114 (-1.04) [0.11]
Rezago de un periodo de puntaje ECE – CL			0 (-1.54) [0.00]	-0.001*** (-7.05) [0.00]	0.002*** (-20.95) [0.00]	0.000*** (-3.94) [0.00]
Rezago de un periodo de puntaje ECE – Matematica			0.001*** (-10.34) [0.00]	0 (-0.32) [0.00]	0.000*** (-4.56) [0.00]	-0.001*** (-7.99) [0.00]
Constante	7.182*** -(130.35) [0.06]	6.682*** -(102.01) [0.07]	6.458*** -(80.33) [0.08]	7.288*** -(76.27) [0.10]	5.736*** -(68.23) [0.08]	7.348*** -(77.46) [0.09]
Efecto Fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.441	0.093	0.490	0.118	0.630	0.170
Nro. observaciones	158,898	158,898	134,669	134,669	3,527,303	3,527,303

t-score en parentesis; error estandar robusto en llaves

* p<0.10, ** p<0.05, *** p<0.010



Tabla 11: Estimación por triple diferencia, para todas las escuelas evaluadas en segundo grado de primaria, desagregado por año de intervención

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2014 en adelante)	0.451*** (-42.01) [0.01]	-1.193*** (-86.47) [0.01]	0.454*** (-40.62) [0.01]	-1.203*** (-83.37) [0.01]	0.497*** (-46.71) [0.01]	-1.165*** (-77.06) [0.02]
Periodo tratado (1=2015 en adelante)	0.363*** (-27.47) [0.01]	0.830*** (-51.49) [0.02]	0.362*** (-26.81) [0.01]	0.851*** (-50.79) [0.02]	0.335*** (-26.7) [0.01]	0.799*** (-47.91) [0.02]
Periodo tratado (1=2016 en adelante)	1.121*** (-82.91) [0.01]	0.006 (-0.34) [0.02]	1.118*** (-79.61) [0.01]	0.018 (-1.03) [0.02]	1.115*** (-82.18) [0.01]	0.016 (-0.99) [0.02]
Tratamiento (1=público=tratado)	-0.016 (-0.26) [0.06]	-0.265*** (-3.51) [0.08]	0.081 (-1.08) [0.07]	-0.131 (-1.45) [0.09]	0.113* (-1.7) [0.07]	-0.106 (-1.31) [0.08]
post2014 x tratado	0.275*** (-21.34) [0.01]	0.374*** (-23.33) [0.02]	0.163*** (-11.93) [0.01]	0.381*** (-22.23) [0.02]	0.137*** (-10.43) [0.01]	0.376*** (-21.18) [0.02]
post2015 x tratado	0.013 (-0.84) [0.02]	0.061*** (-3.2) [0.02]	-0.015 (-0.89) [0.02]	0.045** (-2.28) [0.02]	0.012 (-0.8) [0.02]	0.106*** (-5.44) [0.02]
post2016 x tratado	0.042*** (-2.62) [0.02]	0.128*** (-6.78) [0.02]	0.029* (-1.77) [0.02]	0.131*** (-6.57) [0.02]	0.094*** (-5.93) [0.02]	0.213*** (-11.19) [0.02]
Constante	7.270*** (-133.27) [0.05]	6.692*** (-102.03) [0.07]	6.933*** (-87.19) [0.08]	7.298*** (-75.92) [0.10]	6.754*** (-87.92) [0.08]	7.550*** (-78.53) [0.10]
Efecto fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.509	0.094	0.557	0.119	0.713	0.175
Nro. observaciones	158,898	158,898	134,669	134,669	3,527,303	3,527,303

t-score en parentesis; error estandar robusto en llaves

se omiten variables de control

* p<0.10, ** p<0.05, *** p<0.010



Tabla 12: Estimación por triple diferencia, para las escuelas en zona de competencia evaluadas en segundo grado de primaria, desagregado por año de intervención

	Modelo simple - CL b/t/se	Modelo simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2014 en adelante)	0.443*** -(32.72) [0.01]	-1.178*** (-65.67) [0.02]	0.437*** -(34.64) [0.01]	-1.218*** (-75.93) [0.02]	0.442*** -(37.12) [0.01]	-1.234*** (-70.29) [0.02]
Periodo tratado (1=2015 en adelante)	0.374*** -(20.91) [0.02]	0.826*** -(38.07) [0.02]	0.310*** -(19.04) [0.02]	0.784*** -(39.36) [0.02]	0.314*** -(20.3) [0.02]	0.771*** -(36.72) [0.02]
Periodo tratado (1=2016 en adelante)	1.113*** -(54.12) [0.02]	0.008 (-0.33) [0.02]	0.978*** -(55.46) [0.02]	-0.128*** (-6.22) [0.02]	0.974*** -(57.75) [0.02]	-0.120*** (-5.51) [0.02]
Tratamiento (1=público=tratado)	-0.12 (-1.31) [0.09]	-0.239* (-1.82) [0.13]	0.043 (-0.48) [0.09]	-0.062 (-0.63) [0.10]	0.159*** (-2.81) [0.06]	0.013 (-0.13) [0.09]
post2014 x tratado	0.341*** -(19.5) [0.02]	0.433*** -(19.58) [0.02]	0.101*** -(6.19) [0.02]	0.312*** -(15.66) [0.02]	0.104*** -(6.33) [0.02]	0.347*** -(15.78) [0.02]
post2015 x tratado	-0.039* (-1.71) [0.02]	0.002 (-0.09) [0.03]	-0.134*** (-6.41) [0.02]	-0.103*** (-4.17) [0.02]	-0.112*** (-5.41) [0.02]	-0.069** (-2.56) [0.03]
post2016 x tratado	0.150*** -(6.1) [0.02]	0.248*** -(8.62) [0.03]	0.122*** -(5.81) [0.02]	0.230*** -(9.37) [0.02]	0.158*** -(7.64) [0.02]	0.268*** -(10.38) [0.03]
Constante	7.595*** -(98.27) [0.08]	6.905*** -(62.51) [0.11]	3.891*** -(40.01) [0.10]	3.763*** -(32.69) [0.12]	3.739*** -(42.84) [0.09]	4.213*** -(29.77) [0.14]
Efecto fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.669	0.156	0.795	0.341	0.869	0.350
Nro. observaciones	53,167	53,167	47,762	47,762	1,426,581	1,426,581

t-score en parentesis; error estandar robusto en llaves

se omiten variables de control

* p<0.10, ** p<0.05, *** p<0.010



Tabla 13: Estimación por triple diferencia, para las escuelas en zona de competencia y en área rural evaluadas en segundo grado de primaria, desagregado por año de intervención

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2014 en adelante)	0.515*** (-2.59) [0.20]	-0.918*** (-3.72) [0.25]	0.424** (-2.07) [0.20]	-0.919*** (-4.64) [0.20]	0.428** (-2.2) [0.19]	-0.943*** (-5.65) [0.17]
Periodo tratado (1=2015 en adelante)	-0.517 (-1.39) [0.37]	0.12 (-0.38) [0.32]	-0.268 (-0.81) [0.33]	0.167 (-0.64) [0.26]	-0.308 (-0.84) [0.37]	0.153 (-0.56) [0.27]
Periodo tratado (1=2016 en adelante)	1.682*** (-5.52) [0.30]	0.661* (-1.9) [0.35]	1.354*** (-4.64) [0.29]	0.419 (-1.45) [0.29]	1.475*** (-4.98) [0.30]	0.523** (-2) [0.26]
Tratamiento (1=público=tratado)	-0.499*** (-2.65) [0.19]	-0.796*** (-3.04) [0.26]	-1.042** (-2.53) [0.41]	-0.937** (-2.42) [0.39]	-0.833** (-2.01) [0.41]	-0.784 (-1.55) [0.50]
post2014 x tratado	0.255 (-1.28) [0.20]	0.145 (-0.59) [0.25]	0.075 (-0.37) [0.21]	-0.004 (-0.02) [0.20]	0.072 (-0.37) [0.20]	0.047 (-0.28) [0.17]
post2015 x tratado	0.790** (-2.12) [0.37]	0.634** (-1.99) [0.32]	0.448 (-1.36) [0.33]	0.498* (-1.91) [0.26]	0.498 (-1.36) [0.37]	0.524* (-1.92) [0.27]
post2016 x tratado	-0.477 (-1.56) [0.31]	-0.449 (-1.29) [0.35]	-0.275 (-0.94) [0.29]	-0.341 (-1.18) [0.29]	-0.417 (-1.40) [0.30]	-0.446* (-1.69) [0.26]
Constante	7.550*** (-40.54) [0.19]	7.227*** (-27.92) [0.26]	4.606*** (-11.02) [0.42]	4.368*** (-10.94) [0.40]	4.275*** (-10.24) [0.42]	4.379*** (-8.58) [0.51]
Efecto fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.542	0.094	0.714	0.322	0.729	0.323
Nro. observaciones	27,609	27,609	22,822	22,822	251,348	251,348

t-score en parentesis; error estandar robusto en llaves

se omiten variables de control

* p<0.10, ** p<0.05, *** p<0.010



Tabla 14: Estimación por triple diferencia, para las escuelas en zona de competencia y en área urbana evaluadas en segundo grado de primaria, desagregado por año de intervención

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2014 en adelante)	0.443*** -(32.82) [0.01]	-1.184*** (-66.58) [0.02]	0.428*** -(34.01) [0.01]	-1.221*** (-76.41) [0.02]	0.436*** -(36.31) [0.01]	-1.233*** (-67.97) [0.02]
Periodo tratado (1=2015 en adelante)	0.382*** -(21.65) [0.02]	0.838*** -(38.89) [0.02]	0.310*** -(19.06) [0.02]	0.777*** -(39.56) [0.02]	0.313*** -(20.18) [0.02]	0.768*** -(36.05) [0.02]
Periodo tratado (1=2016 en adelante)	1.117*** -(55.08) [0.02]	0.002 -(0.07) [0.02]	0.978*** -(55.84) [0.02]	-0.139*** (-6.72) [0.02]	0.967*** -(56.79) [0.02]	-0.134*** (-5.91) [0.02]
Tratamiento (1=público=tratado)	0.03 -(0.34) [0.09]	-0.022 (-0.18) [0.13]	0.143** -(2.29) [0.06]	0.012 -(0.12) [0.09]	0.153*** -(2.66) [0.06]	0.009 -(0.1) [0.09]
post2014 x tratado	0.364*** -(19.34) [0.02]	0.486*** -(20.37) [0.02]	0.145*** -(7.9) [0.02]	0.352*** -(15.93) [0.02]	0.128*** -(7.02) [0.02]	0.364*** -(14.81) [0.02]
post2015 x tratado	0.009 -(0.36) [0.02]	0.029 -(1) [0.03]	-0.096*** (-4.17) [0.02]	-0.074*** (-2.74) [0.03]	-0.084*** (-3.73) [0.02]	-0.046 (-1.55) [0.03]
post2016 x tratado	0.209*** -(8.31) [0.03]	0.317*** -(10.55) [0.03]	0.140*** -(6.34) [0.02]	0.275*** -(10.53) [0.03]	0.170*** -(7.8) [0.02]	0.292*** -(10.61) [0.03]
Constante	7.944*** -(132.63) [0.06]	7.004*** -(81.08) [0.09]	3.964*** -(36.53) [0.11]	3.826*** -(25.28) [0.15]	3.604*** -(30.18) [0.12]	4.238*** -(23.27) [0.18]
Efecto fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.779	0.257	0.854	0.369	0.892	0.368
Nro. observaciones	25,558	25,558	24,940	24,940	1,175,319	1,175,319

t-score en parentesis; error estandar robusto en llaves

se omiten variables de control

* p<0.10, ** p<0.05, *** p<0.010



Tabla 15: Estimación por triple diferencia, para las escuelas en zona de competencia y de tipo multigrado o unidocente, evaluadas en segundo grado de primaria, desagregado por año de intervención

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2014 en adelante)	0.389*** (-6.15) [0.06]	-1.306*** (-15.06) [0.09]	0.314*** (-4.59) [0.07]	-1.249*** (-16.06) [0.08]	0.337*** (-4.5) [0.07]	-1.241*** (-14.58) [0.09]
Periodo tratado (1=2015 en adelante)	0.201 (-1.58) [0.13]	0.641*** (-4.38) [0.15]	0.367*** (-3.76) [0.10]	0.758*** (-5.43) [0.14]	0.332*** (-3.02) [0.11]	0.722*** (-5.48) [0.13]
Periodo tratado (1=2016 en adelante)	1.295*** (-8.16) [0.16]	0.028 (-0.16) [0.18]	1.021*** (-8.39) [0.12]	-0.245 (-1.48) [0.17]	1.030*** (-7.99) [0.13]	-0.229 (-1.53) [0.15]
Tratamiento (1=público=tratado)	-0.406** (-2.40) [0.17]	-0.692*** (-2.96) [0.23]	-0.731** (-2.19) [0.33]	-0.676** (-2.09) [0.32]	-0.457 (-1.39) [0.33]	-0.461 (-1.29) [0.36]
post2014 x tratado	0.372*** (-5.64) [0.07]	0.531*** (-5.96) [0.09]	0.177** (-2.52) [0.07]	0.325*** (-4.05) [0.08]	0.154** (-2.01) [0.08]	0.339*** (-3.88) [0.09]
post2015 x tratado	0.058 (-0.44) [0.13]	0.107 (-0.72) [0.15]	-0.193* (-1.93) [0.10]	-0.091 (-0.64) [0.14]	-0.147 (-1.31) [0.11]	-0.055 (-0.41) [0.13]
post2016 x tratado	-0.094 (-0.59) [0.16]	0.162 (-0.9) [0.18]	0.06 (-0.49) [0.12]	0.306* (-1.84) [0.17]	0.03 (-0.23) [0.13]	0.295* (-1.95) [0.15]
Constante	7.468*** (-46.17) [0.16]	7.102*** (-31.77) [0.22]	4.424*** (-13.29) [0.33]	4.116*** (-12.47) [0.33]	4.138*** (-12.67) [0.33]	4.089*** (-11.41) [0.36]
Efecto fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.543	0.093	0.714	0.318	0.727	0.324
Nro. observaciones	27,817	27,817	22,817	22,817	213,882	213,882

t-score en parentesis; error estandar robusto en llaves

se omiten variables de control

* p<0.10, ** p<0.05, *** p<0.010

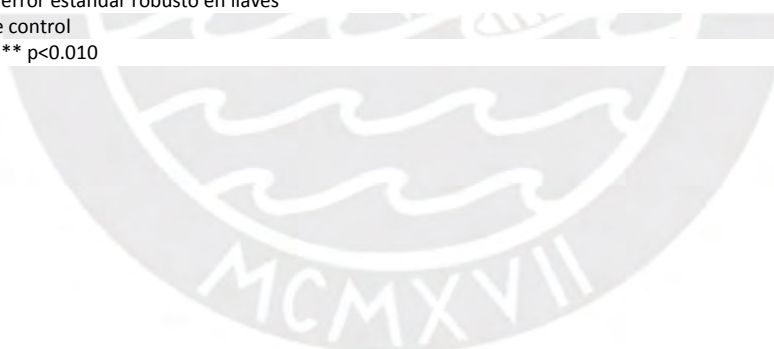


Tabla 16: Estimación por triple diferencia, para las escuelas en zona de competencia y de tipo polidocente completo, evaluadas en segundo grado de primaria, desagregado por año de intervención

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se	Modelo EF con covariables y peso - CL b/t/se	Modelo EF con covariables y peso - Mate b/t/se
Periodo tratado (1=2014 en adelante)	0.442*** -(33.4) [0.01]	-1.160*** (-64.45) [0.02]	0.430*** -(34.33) [0.01]	-1.207*** (-75.16) [0.02]	0.437*** -(35.42) [0.01]	-1.226*** (-67.45) [0.02]
Periodo tratado (1=2015 en adelante)	0.393*** -(22.69) [0.02]	0.830*** -(38.37) [0.02]	0.316*** -(19.22) [0.02]	0.762*** -(38.32) [0.02]	0.316*** -(19.66) [0.02]	0.763*** -(35.46) [0.02]
Periodo tratado (1=2016 en adelante)	1.106*** -(54.7) [0.02]	0.024 -(1) [0.02]	0.971*** -(54.6) [0.02]	-0.112*** (-5.37) [0.02]	0.967*** -(55.2) [0.02]	-0.123*** (-5.31) [0.02]
Tratamiento (1=público=tratado)	0.043 -(0.52) [0.08]	0.032 -(0.28) [0.12]	0.172*** -(2.89) [0.06]	0.022 -(0.25) [0.09]	0.159*** -(2.77) [0.06]	-0.001 (-0.01) [0.10]
post2014 x tratado	0.358*** -(19.46) [0.02]	0.457*** -(19.4) [0.02]	0.137*** -(7.61) [0.02]	0.329*** -(15.17) [0.02]	0.120*** -(6.64) [0.02]	0.352*** -(14.66) [0.02]
post2015 x tratado	0.001 -(0.04) [0.02]	0.062** -(2.17) [0.03]	-0.114*** (-5.00) [0.02]	-0.032 (-1.21) [0.03]	-0.096*** (-4.22) [0.02]	-0.037 (-1.29) [0.03]
post2016 x tratado	0.225*** -(8.95) [0.03]	0.296*** -(9.86) [0.03]	0.148*** -(6.7) [0.02]	0.248*** -(9.63) [0.03]	0.168*** -(7.63) [0.02]	0.277*** -(10.02) [0.03]
Constante	7.924*** -(136.42) [0.06]	6.974*** -(83.11) [0.08]	3.990*** -(37.99) [0.11]	3.904*** -(27.11) [0.14]	3.783*** -(32.96) [0.11]	4.252*** -(24) [0.18]
Efecto fijo IE	Sí	Sí	Sí	Sí	Sí	Sí
R-cuadrado ajustado	0.786	0.264	0.860	0.379	0.890	0.369
Nro. observaciones	25,350	25,350	24,945	24,945	1,212,571	1,212,571

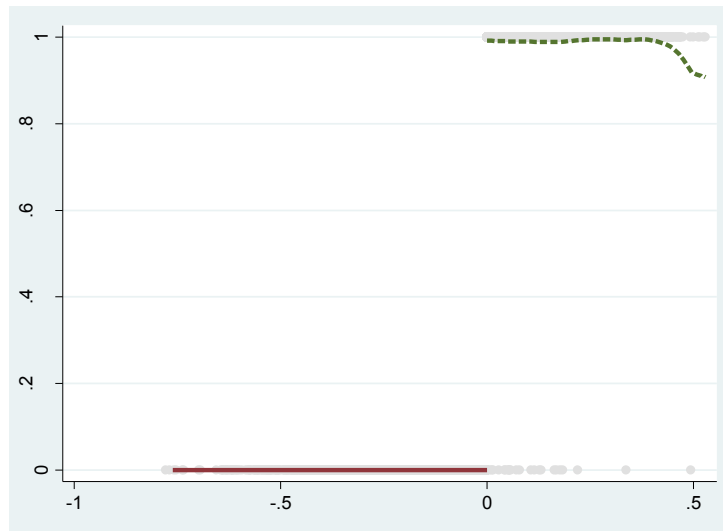
t-score en parentesis; error estandar robusto en llaves

se omiten variables de control

* p<0.10, ** p<0.05, *** p<0.010

Gráfico 3: Distribución de la probabilidad de tratamiento alrededor del punto de corte para el otorgamiento del bono

Bono otorgado en el 2014



Bono otorgado en el 2015

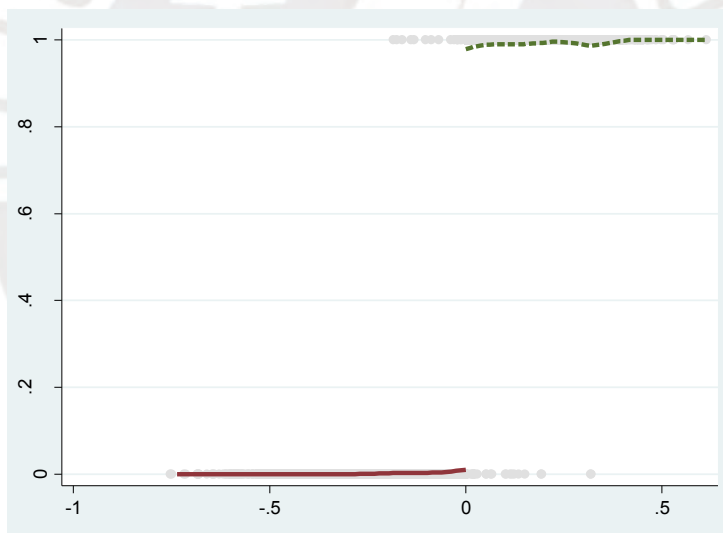


Tabla 17: Resultados de la estimación por regresión discontinua, diseño difuso, con respecto a las áreas evaluadas de Comprensión Lectora y Matemáticas

Efecto del otorgamiento de Bono Escuela 2014 sobre los resultados en la ECE 2015

Para Comprensión Lectora

Estimador	Coefficiente	Error estándar	Z-test	P> z
lwald	0.068	0.075	0.900	0.366
lwald50	0.019	0.100	0.190	0.848
lwald200	0.106	0.058	1.830	0.067

Ancho de banda 0.111

Para Matemática

Estimador	Coefficiente	Error estándar	Z-test	P> z
lwald	0.063	0.082	0.770	0.440
lwald50	0.012	0.110	0.110	0.916
lwald200	0.107	0.063	1.700	0.089

Ancho de banda 0.100

Efecto del otorgamiento de Bono Escuela 2015 sobre los resultados en la ECE 2016

Para Comprensión Lectora

Estimador	Coefficiente	Error estándar	Z-test	P> z
lwald	0.012	0.070	-0.01	0.99
lwald50	-0.056	0.092	-0.28	0.783
lwald200	0.094	0.053	1.48	0.138

Ancho de banda 0.161

Para Matemática

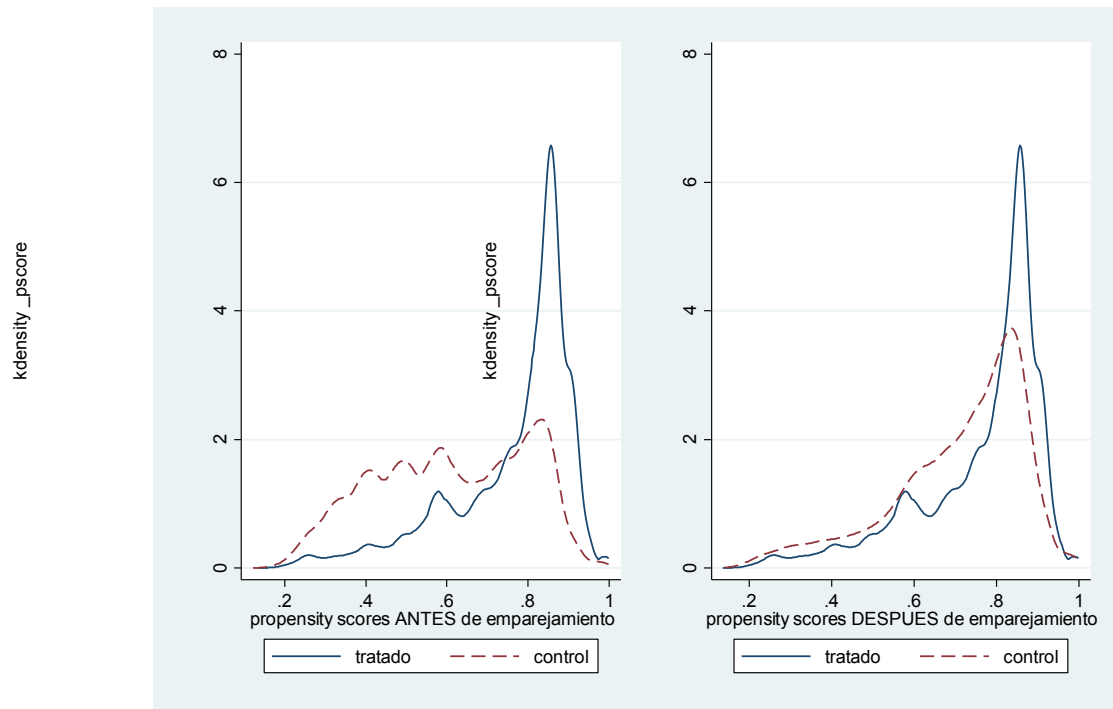
Estimador	Coefficiente	Error estándar	Z-test	P> z
lwald	0.046	0.072	0.63	0.528
lwald50	0.035	0.094	0.38	0.706
lwald200	0.103	0.057	1.8	0.071

Ancho de banda 0.156

Gráfico 4: Distribución del *propensity score* antes y después del modelo de

emparejamiento, para el caso de secundaria, años 2015 y 2016

Año 2015



Año 2016

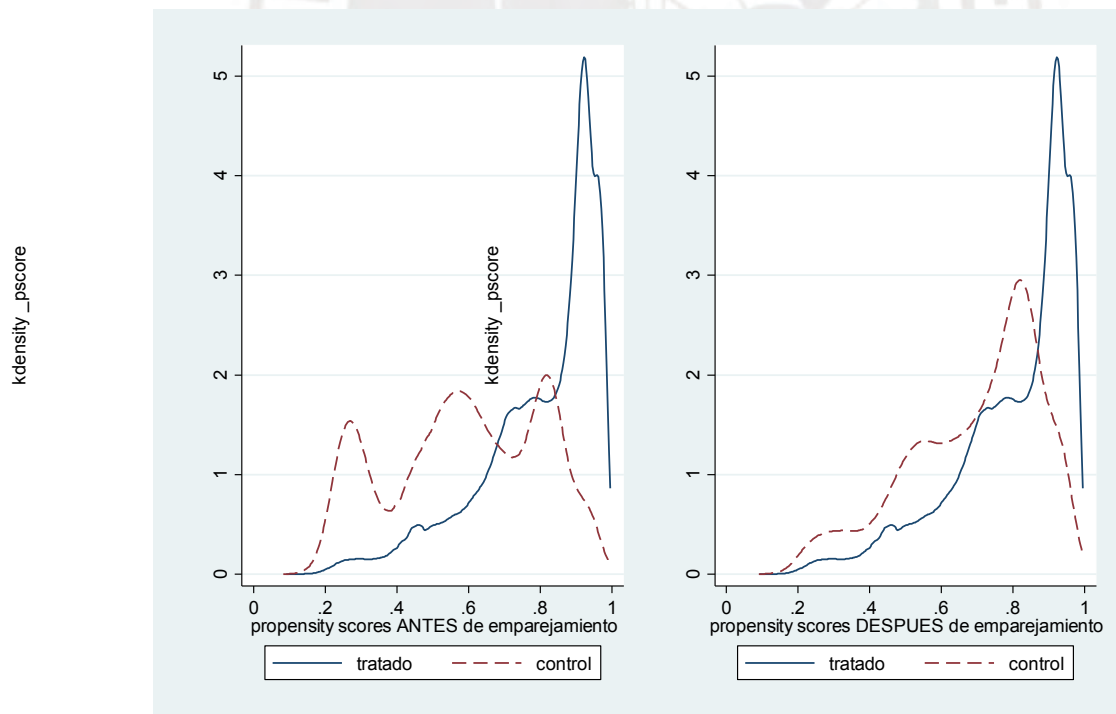


Tabla 18: Resultado ATT de la estimación por *propensity score* para la evaluación del

efecto competencia en las áreas de Comprensión Lectora y Matemática, para el caso de secundaria (2015)

Incluyendo a todas las escuelas

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	7.325	7.603	-0.278	0.009	-31.1	360,944
Comprensión Lectora	7.494	7.912	-0.418	0.008	-50.8	360,944

Incluyendo solo a las escuelas en zona de competencia

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	7.509	7.984	-0.475	0.011	-44.0	154,748
Comprensión Lectora	7.690	8.278	-0.588	0.009	-62.2	154,748

Incluyendo solo a las escuelas en zona de competencia y en área urbana

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	7.545	8.014	-0.469	0.011	-41.8	142,728
Comprensión Lectora	7.744	8.314	-0.570	0.010	-58.6	142,728

Incluyendo solo a las escuelas en zona de competencia y en área rural

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	7.124	7.309	-0.186	0.060	-3.1	11,628
Comprensión Lectora	7.111	7.322	-0.210	0.058	-3.6	11,628

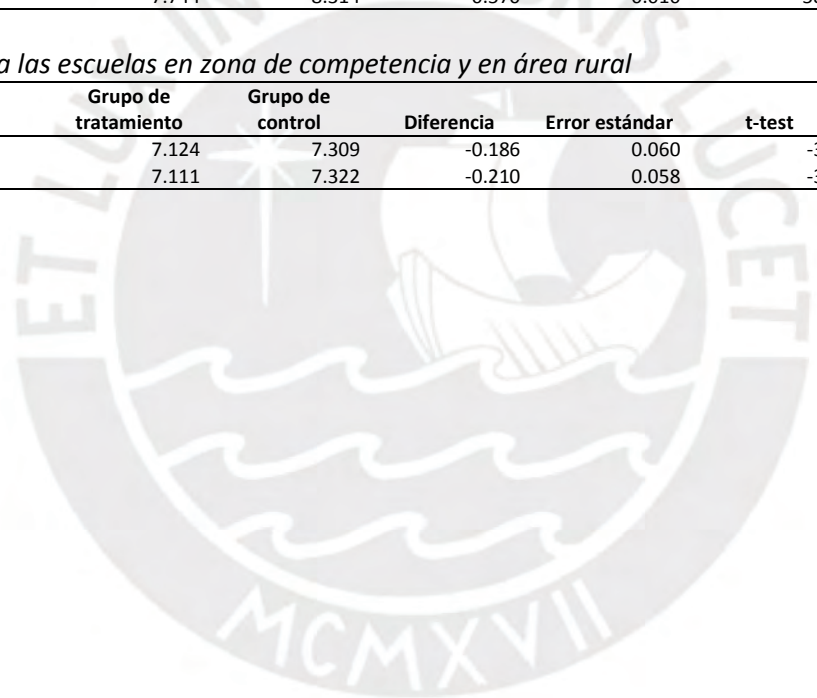


Tabla 19: Resultado ATT de la estimación por *propensity score* para la evaluación del efecto competencia en las áreas de Comprensión Lectora y Matemática, para el caso de secundaria (2016)

Incluyendo a todas las escuelas

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	6.911	6.999	-0.088	0.009	-10.1	418,041
Comprensión Lectora	7.978	8.141	-0.163	0.008	-19.2	418,041

Incluyendo solo a las escuelas en zona de competencia

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	7.107	7.328	-0.220	0.012	-17.8	214,858
Comprensión Lectora	8.187	8.481	-0.293	0.012	-25.5	214,858

Incluyendo solo a las escuelas en zona de competencia y en área rural

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	6.635	6.826	-0.192	0.113	-1.7	11,925
Comprensión Lectora	7.538	7.539	-0.001	0.123	0.0	11,925

Incluyendo solo a las escuelas en zona de competencia y en área urbana

	Grupo de tratamiento	Grupo de control	Diferencia	Error estándar	t-test	Nro. Observaciones
Matemática	7.140	7.365	-0.225	0.012	-18.6	203,411
Comprensión Lectora	8.232	8.536	-0.305	0.011	-27.1	203,411

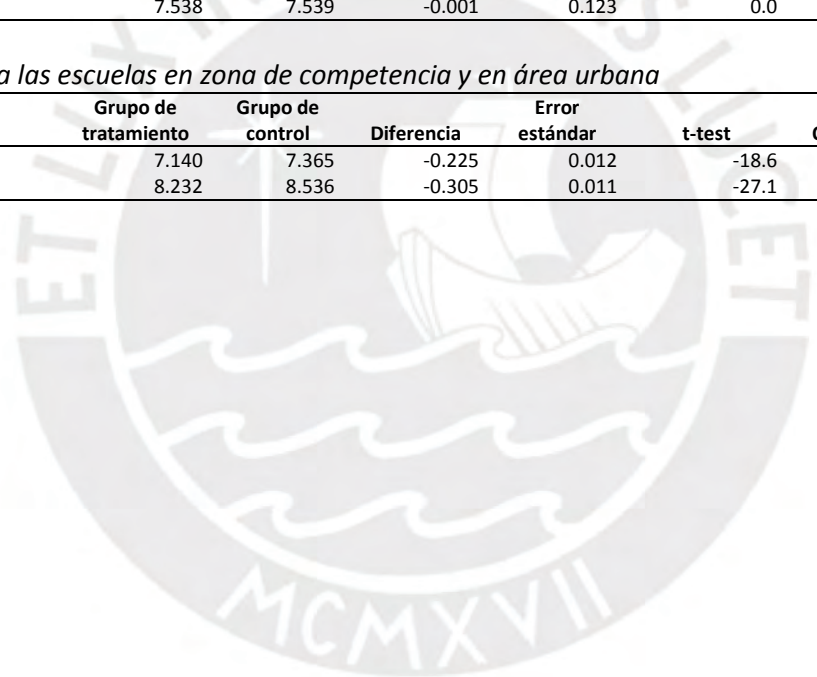


Tabla 20: Estimación por diferencia en diferencia, para todas las escuelas evaluadas en segundo grado de secundaria

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se
Periodo tratado (1=2016 en adelante)	0.436*** (-66.95) [0.01]	-0.455*** (-56.81) [0.01]	-0.594*** (-13.33) [0.04]	-1.594*** (-30.12) [0.05]
Tratamiento (1=público=tratado)	-0.089* (-1.93) [0.05]	-0.130** (-2.16) [0.06]	-0.062 (-1.43) [0.04]	-0.095* (-1.68) [0.06]
post2016 x tratado	0.068*** (-9.41) [0.01]	0.070*** (-7.68) [0.01]	0.019** (-2.42) [0.01]	0.034*** (-3.55) [0.01]
Área (1=urbano)			0.006 (-0.16) [0.04]	0.012 (-0.24) [0.05]
Sexo del estudiante (1=hombre)			-0.009*** (-3.78) [0.00]	0.179*** (-74.6) [0.00]
Internado (1=sí recibe)			0.106** (-2.08) [0.05]	0.05 (-0.8) [0.06]
Tutorial (1=sí recibe)			-0.05 (-0.63) [0.08]	-0.286*** (-3.18) [0.09]
Nivel educativo de la madre (1= Sin Educación; 10= Estudios de posgrado)			0.031*** (-55.64) [0.00]	0.026*** (-43.66) [0.00]
Acceso a red pública de desagüe en hogar (1=sí accede)			0.164*** (-45.6) [0.00]	0.126*** (-35.45) [0.00]
Tenencia de libros en el hogar (1: No hay libros; 7: 100 o más libros)			0.088*** (-106.1) [0.00]	0.092*** (-101.08) [0.00]
Tenencia de horno microondas en hogar (1=sí tiene)			-0.085*** (-33.77) [0.00]	-0.100*** (-38.11) [0.00]
Tenencia de refrigeradora en hogar (1=sí tiene)			0.016*** (-5.59) [0.00]	0.013*** (-4.31) [0.00]
Tenencia de lavadora en hogar (1=sí tiene)			-0.033*** (-11.06) [0.00]	-0.043*** (-14.77) [0.00]
Conexión a Internet en hogar (1=sí tiene)			0.021*** (-8.19) [0.00]	0.004 (-1.36) [0.00]
Percepción del aula sobre prácticas pedagógicas			0.001 (-0.86) [0.00]	0.001 (-1.22) [0.00]
Percepción del aula sobre clima escolar			0.013*** (-25.94) [0.00]	0.014*** (-24.25) [0.00]
Total de docentes en el nivel educativo			0 (-1.19) [0.00]	0.001*** (-2.74) [0.00]
% de docentes hombres en el nivel educativo			0 (-0.05) [0.00]	0 (-0.67) [0.00]
Constante	7.640***	7.485***	6.843***	6.556***

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se
	-(192.3) [0.04]	-(145.05) [0.05]	-(129.49) [0.05]	-(94.11) [0.07]
R-cuadrado ajustado	0.083	0.051	0.125	0.092
Efecto Fijo IE	Sí	Sí	Sí	Sí
Nro. observaciones	778,985	778,985	723,687	723,687

t-score en paréntesis; error estándar robusto en llaves

* p<0.10, ** p<0.05, *** p<0.010



Tabla 21: Estimación por diferencia en diferencia, para todas las escuelas en zona de competencia, evaluadas en segundo grado de secundaria

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se
Periodo tratado (1=2016 en adelante)	0.435*** -(46.84) [0.01]	-0.465*** (-38.64) [0.01]	-0.611*** (-8.28) [0.07]	-1.647*** (-18.47) [0.09]
Tratamiento (1=público=tratado)	0.079*** -(2.92) [0.03]	0.064 -(0.63) [0.10]	0.065 -(1.55) [0.04]	0.073 -(0.6) [0.12]
post2016 x tratado	0.073*** -(6.47) [0.01]	0.090*** -(6.03) [0.01]	0.011 -(0.93) [0.01]	0.044*** -(2.75) [0.02]
Área (1=urbano)			0.054 -(0.66) [0.08]	0.021 -(0.25) [0.09]
Sexo del estudiante (1=hombre)			-0.020*** (-5.64) [0.00]	0.187*** -(49.98) [0.00]
Internado (1=sí recibe)			0.08 -(1.48) [0.05]	-0.104 (-0.42) [0.25]
Tutorial (1=sí recibe)			-1.121*** (-109.71) [0.01]	-1.219*** (-95.46) [0.01]
Constante	7.721*** -(344.22) [0.02]	7.529*** -(88.45) [0.09]	6.776*** -(77) [0.09]	6.465*** -(47.83) [0.14]
R-cuadrado ajustado	0.053	0.031	0.097	0.078
Efecto Fijo IE	Sí	Sí	Sí	Sí
Nro. observaciones	370,576	370,576	346,609	346,609

Se omiten variables adicionales de control.

t-score en paréntesis; error estándar robusto en llaves

* p<0.10, ** p<0.05, *** p<0.010



Tabla 22: Estimación por diferencia en diferencia, para todas las escuelas en zona de competencia de área rural, evaluadas en segundo grado de secundaria

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se
Periodo tratado (1=2016 en adelante)	0.483*** -(8.73) [0.06]	-0.400*** (-6.08) [0.07]	0.088 -(0.44) [0.20]	-0.768*** (-3.01) [0.26]
Tratamiento (1=público=tratado)	0.103 -(0.78) [0.13]	0.486*** (-3.46) [0.14]	0.088 -(0.45) [0.19]	0.533*** (-3.02) [0.18]
post2016 x tratado	0.004 -(0.08) [0.06]	-0.022 (-0.31) [0.07]	-0.035 (-0.47) [0.07]	-0.138 (-1.42) [0.10]
Sexo del estudiante (1=hombre)			0.01 -(0.97) [0.01]	0.170*** -(16.37) [0.01]
Internado (1=sí recibe)			-0.037* (-1.75) [0.02]	-0.336*** (-12.30) [0.03]
Tutorial (1=sí recibe)			-1.005*** (-28.64) [0.04]	-1.032*** (-23.81) [0.04]
Constante	6.992*** -(55.07) [0.13]	6.630*** -(49.54) [0.13]	6.448*** -(30.19) [0.21]	5.935*** -(27.73) [0.21]
R-cuadrado ajustado	0.059	0.044	0.099	0.08
Efecto Fijo IE	Sí	Sí	Sí	Sí
Nro. observaciones	23,425	23,425	22,031	22,031

Se omiten variables adicionales de control.

t-score en paréntesis; error estándar robusto en llaves

* p<0.10, ** p<0.05, *** p<0.010



Tabla 23: Estimación por diferencia en diferencia, para todas las escuelas en zona de competencia de área urbana, evaluadas en segundo grado de secundaria

	Modelo EF simple - CL b/t/se	Modelo EF simple - Mate b/t/se	Modelo EF con covariables - CL b/t/se	Modelo EF con covariables - Mate b/t/se
Periodo tratado (1=2016 en adelante)	0.433*** (-46.64) [0.01]	-0.465*** (-38.43) [0.01]	-0.637*** (-8.31) [0.08]	-1.679*** (-18.08) [0.09]
interv	0.078*** (-3.97) [0.02]	-0.012 (-0.14) [0.08]	0.058* (-1.89) [0.03]	-0.018 (-0.18) [0.10]
post2016 x tratado	0.075*** (-6.56) [0.01]	0.093*** (-6.1) [0.02]	0.012 (-0.97) [0.01]	0.046*** (-2.83) [0.02]
Sexo del estudiante (1=hombre)			-0.022*** (-5.99) [0.00]	0.188*** (-47.72) [0.00]
Internado (1=sí recibe)			0.095 (-1.57) [0.06]	-0.068 (-0.24) [0.29]
Constante	7.771*** (-491.96) [0.02]	7.625*** (-112.62) [0.07]	6.861*** (-167.45) [0.04]	6.579*** (-73.92) [0.09]
R-cuadrado ajustado	0.053	0.031	0.098	0.078
Efecto Fijo IE	Sí	Sí	Sí	Sí
Nro. observaciones	347,151	347,151	324,578	324,578

Se omiten variables adicionales de control.

t-score en paréntesis; error estándar robusto en llaves

* p<0.10, ** p<0.05, *** p<0.010



Tabla 24: Análisis incondicional de diferencias en diferencias

El concepto principal detrás del modelo de diferencias en diferencias se detalla en la tabla presentada a continuación. Acotando el tratamiento a la zona de competencia, aquí se muestra que, al analizar los resultados de Comprensión Lectora, los resultados señalan que aquellos que participaron en Bono Escuela mejoraron en 0.37 desviaciones estándar por encima de aquellos que no participaron (escuelas privadas). Aun cuando ambos grupos mejoraron en el tiempo (periodo 2014-2016) con respecto a ellos mismos, las escuelas participantes habrían mejorado a una tasa mayor (cambio en la tendencia), lo que originó que en neto mejoren en mayor medida que los no participantes. De igual manera, en el caso de Matemática, las escuelas tratadas mejoraron su situación con respecto a las escuelas no tratadas, en 0.56 desviaciones estándar. Incluso viendo una tendencia a la baja en ambos grupos, las escuelas públicas redujeron en menor medida su promedio, lo que en términos marginales mejoró su situación con respecto al periodo inicial, llegando a superar a los colegios privados en promedio, durante el periodo 2014-2016.

**Análisis incondicional de diferencias en diferencias
(puntaje de la ECE en desviaciones estándar)**

	2008-2013 (1)	2014-2016 (2)	Diferencias (3)=(2)-(1)
Resultados en Comprensión Lectora			
IIEE Públicos	7.23 (0.00)	8.72 (0.01)	1.49 (0.01)
IIEE Privados	8.27 (0.00)	9.39 (0.01)	1.12 (0.01)
Diferencia	-1.04 (0.01)	-0.67 (0.01)	0.37 (0.01)
Resultados en Matemáticas			
IIEE públicos	6.48 (0.00)	6.32 (0.01)	-0.16 (0.01)
IIEE Privados	6.95 (0.00)	6.24 (0.01)	-0.71 (0.01)
Diferencia	-0.48 (0.01)	0.08 (0.01)	0.56 (0.01)