



Universidad de San Andrés

Departamento de Economía

Maestría en Economía

¿Funciona la asesoría pedagógica en escuelas multigrado?

Evidencia experimental utilizando data administrativa

Autor: Antonio Gerardo Campos Flores

DNI: 44275338

Mentor de Tesis: Walter Esteban Sosa Escudero

Victoria, Buenos Aires, 25 de Julio de 2018



Universidad de
SanAndrés

Universidad de San Andrés

Departamento de Economía

Maestría en Economía

¿Funciona la asesoría pedagógica en escuelas multigrado?

Evidencia experimental utilizando data administrativa

Autor: Antonio Gerardo Campos Flores

DNI: 44275338

Universidad de
SanAndrés

Mentor: Walter Esteban Sosa Escudero

Victoria,

25 de julio de 2018

Tesis de Maestría en Economía de
Antonio Gerardo Campos Flores

“¿Funciona la asesoría pedagógica en escuelas multigrado? Evidencia experimental utilizando data administrativa”

Resumen

El trabajo estudia una intervención de asesoría pedagógica escalada a nivel nacional en escuelas primaria multigrado implementada en Perú. La intervención impacta los rendimientos académicos en Español y Matemáticas, en 0.098 y 0.125 desviaciones estándar, respectivamente. Los efectos encontrados son similares a los que muestra la literatura para intervenciones de asesoría pedagógica en contextos rurales o desfavorables. Además, se encuentra que en Matemáticas tiene efectos distributivos en los percentiles superiores. Asimismo, se exploran los mecanismos y se haya evidencia a favor de que los impactos en los aprendizajes provienen por las mejoras en las prácticas pedagógicas y retroalimentación entre docentes.

Palabras clave: retroalimentación a docentes, entrenamiento, aprendizaje de estudiantes, ensayo de control aleatorio

“Does coaching work in multigrade schools? Experimental evidence using administrative data”

Abstract

The paper studies a coaching program scaled-up at the national level in multigrade primary schools implemented in Peru. The intervention impacts the academic performance in Spanish and Mathematics, in 0.098 and 0.125 standard deviations, respectively. The effects found are similar to those shown in the literature for coaching interventions in rural or unfavorable contexts. In addition, the program has distributive effects in the upper percentiles in Mathematics. Likewise, the mechanisms are explored and there is evidence that the impacts on learning are due to improvements in pedagogical practices and feedback among teachers.

Keywords: teacher feedback, coaching, student learning, and randomized control trial

Códigos JEL: I21, I28, C93

1. Introducción

La calidad docente es uno de los principales determinantes de la calidad educativa y, por tanto, de la mejora de los resultados educativos (Chetty et al., 2014; Hanushek 2002, 2005). Incluso, se reconoce a la calidad docente como un factor clave que contrarresta la situación desfavorable que enfrentan los estudiantes que provienen de familias más pobres o de escuelas con pocos recursos (Hanushek, 2002). Sin embargo, a pesar de su importancia, la calidad docente es un concepto complejo y difícil de medir (Hanushek 2002, 2005) pero que en la práctica, se asocia a los conocimientos y prácticas pedagógicas de los docentes (si conocen lo que tienen que enseñar y la manera de enseñanza), pero también a su comportamiento en el aula (el esfuerzo que realizan para poner en práctica sus conocimientos y prácticas pedagógicas).

Si bien existen diversas estrategias para mejorar la calidad docente (incentivos, monitoreo), una de las alternativas más efectivas es la asesoría pedagógica (Kraft et al. 2018; Bruns y Luque, 2015). La asesoría pedagógica es un entrenamiento personalizado que tiene por objetivo mejorar las habilidades docentes a través de la transferencia de conocimientos y recursos que ayuden a mejorar sus prácticas pedagógicas (Joyce y Shower, 2002). Así, en un estudio que sistematiza 60 evaluaciones de impacto (experimentales y no experimentales), Kraft et al. (2018) encuentra que las asesorías pedagógicas incrementan, en promedio, 0.49 desviaciones estándar (DE) las prácticas pedagógicas de los docentes intervenidos y 0.18 DE los aprendizajes de sus estudiantes.

En el Perú, desde el año 2010 se implementa una intervención de asesoría pedagógica en aula denominada Acompañamiento Pedagógico Multigrado (en adelante, APM). La intervención se desarrolla en escuelas multigrado; es decir, donde un mismo docente enseña a estudiantes de distintos grados al mismo tiempo. Si bien la intervención tuvo un incremento progresivo durante los años previos, el año 2016 se incrementó significativamente a nivel nacional, siendo una parte de esa expansión realizada de manera aleatoria respetando las cuotas regionales establecidas por el Ministerio de Educación.

El APM es un programa escalado a nivel nacional (897 distritos en 25 regiones), y atiende aproximadamente 14,391 docentes en 6,404 escuelas (alrededor de 37% de las escuelas multigrados del país) donde estudian 174 mil estudiantes. En términos de presupuesto, el APM tiene un costo de S/.119 millones de soles, siendo su presupuesto por estudiante de S/.683 soles (aproximadamente US\$195 dólares).

El APM consiste en tres componentes: (i) asesoría personalizada a través de visita y observación en aula y retroalimentación; (ii) grupos de inter-aprendizajes entre docentes y (iii) talleres de actualización docente. Un asesor pedagógico atiende en promedio (de acuerdo a norma oficial) 8 escuelas y el número de horas varía según componente. En total, en el APM los docentes reciben casi 141 horas de asesoría pedagógica durante el año escolar, de las cuales 45 horas son personalizadas con observación de aula y retroalimentación y el resto de manera grupal.

Para analizar el impacto del APM, se utiliza diversas fuentes administrativas disponibles en el Ministerio de Educación del Perú. La principal fuente de información es la Evaluación Censal de Estudiantes (en adelante, ECE), una prueba estandarizada nacional y anual que brinda información de los logros de aprendizajes de los estudiantes del nivel primaria. Un aspecto que limita el estudio es que la ECE no cubre a todos los estudiantes, pero el attrition

no genera mayores problemas en los resultados (tal como se verá en la sección 4.3).

La intervención impacta los rendimientos académicos en Español y Matemáticas, en 0.098 y 0.125 DE, respectivamente. Además, se encuentran efectos distributivos en los percentiles superiores en Matemáticas. Los efectos encontrados son similares a los que muestra la literatura para intervenciones de asesoría pedagógica en contextos rurales o desfavorables. Los resultados también indican que no identifican efectos diferenciados por sexo de los estudiantes, pobreza donde está ubicada la escuela o por características de los directores.

Además, a modo exploratorio, se estudian algunos mecanismos que podrían estar explicando los resultados encontrados. A nivel de docentes, se observan algunos efectos sobre las prácticas pedagógicas de los docentes en algunas estrategias de enseñanza pero sorpresivamente no se encuentran efectos sobre indicadores de organización y planificación de las sesiones, que representan dimensiones no complicadas de mejorar y que frecuentemente son el punto de partida de las asesorías pedagógicas. Un punto interesante es la mejora en indicadores de interacción e involucramiento entre docentes que podrían estar asociados a los grupos de inter-aprendizajes promovidos por la intervención. Así, considerando las mejoras de las prácticas pedagógicas de los docentes, así como la interacción e involucramiento entre ellos, es posible identificar los mecanismos que podrían estar asociados a las mejoras de los aprendizajes de los estudiantes.

Un último punto que explora el estudio está referido a la costo-efectividad de la intervención. A pesar de los efectos encontrados, la costo-efectividad es menor a otras intervenciones de asesoramiento pedagógico. Esto puede explicarse por los mayores costos relacionados a las zonas rurales donde se desarrolla la intervención pero también a los menores efectos por fallas en la implementación propios de una intervención escalada a esta magnitud y por el periodo de maduración de la intervención (Glazerman et al, 2010).

Así, el presente documento brinda diversos aportes a la literatura de asesoría pedagógica. Primero, a conocimiento del autor, este es el primer estudio que muestra evidencia experimental de una intervención de asesoría pedagógica escalada a nivel nacional (ver mapa de distribución de escuelas a nivel nacional en el Anexo 1). Por lo general, las intervenciones de asesoría pedagógica son implementadas por organizaciones no gubernamentales que tienen incentivos para demostrar la eficiencia de la intervención (Cilliers, et al., 2018), o por gobiernos regionales en contextos específicos o espacios geográficos limitados (Bruns et al. 2017; Albornoz, et al. 2017; Zhang, et al. 2013; Angrist y Lavy, 2001) o en pequeña escala (Garet, et al, 2008; Yoshikawa et al. 2015). Tal como indica Kraft et al (2018), en estos casos, el reclutamiento de los asesores pedagógicos es mediante docentes voluntarios y condiciones favorables para su desarrollo; en contraste, las intervenciones de asesoría pedagógica escaladas requieren reclutar asesores, entrenarlos a gran escala y trabajar en condiciones no necesariamente deseables. En Perú el APM tiene presencia en todas las regiones del país y representa casi el 30% de escuelas multigrado públicas lo que implica diversos retos y complejidades de implementación, así como enfrentar la natural burocracia del sector público en las regiones del país.

Segundo, a conocimiento del autor, el presente documento es el primer estudio que muestra evidencia de los efectos de la asesoría pedagógica en un contexto de escuelas multigrado (también conocido como multiedad o clases compuestas). Una escuela multigrado se define como un espacio pedagógico donde un docente es responsable de la enseñanza de estudiantes de más de un grado a la vez en una misma aula de clases. Esto sucede,

principalmente, en áreas rurales, remotas o de difícil acceso y en educación primaria.

En este tipo de escuelas, la enseñanza requiere diversas habilidades, destrezas y estrategias que le permitan al docente enseñar todas las competencias a todos los grados por igual. En este contexto, el rol pedagógico del docente toma vital importancia y la asesoría pedagógica ayuda a enfrentar este problema. Lamentablemente, la literatura brinda pocas soluciones para mejorar la calidad docente en el contexto de escuelas multigrado¹.

Tercero, a conocimiento del autor, es uno de los pocos trabajos que brindan evidencia de asesoría pedagógica en América Latina y en países en desarrollo. La mayoría de estudios han sido desarrollados en Estados Unidos (55 de los 60 estudios analizados por Kraft et al.). La escasa literatura en países de la región se ha desarrollado en Argentina (Albornoz, et al, 2017), Brasil (Bruns, et al, 2017) y Chile (Yoshikawa, et al, 2015).

El documento se encuentra organizado de la siguiente manera: la sección 2 brinda el contexto en el cual se desarrolla la intervención, la sección 3 describe la intervención y brinda algunas ideas de los mecanismos que podrían estar explicando los resultados, la sección 4 muestra el diseño de la evaluación, incluyendo la estrategia empírica y la data disponible, la sección 5 muestra los resultados y análisis costo-efectividad de la intervención y, finalmente, la sección 6 brinda las conclusiones y consideraciones finales.

2. Contexto de la intervención

El Perú es un país ubicado en América del Sur, tiene aproximadamente 32.5 millones de habitantes y un área de 1'285,215 km² (vigésimo país más grande en tamaño del planeta). La Cordillera de los Andes atraviesa el territorio de norte a sur, lo que genera su riqueza natural y climática pero al mismo tiempo una geografía accidentada. Es debido a esta geografía que el Perú tiene un servicio educativo atomizado, con alrededor de 75% de escuelas primarias públicas tipo multigrado (de un total 29,700 escuelas públicas primaria); es decir, escuelas cuyos estudiantes de distintos grados comparten clases con un mismo docente.

En términos de organización del sistema educativo, en el Perú el Ministerio de Educación es el ente encargado de formular, implementar y supervisar la política nacional de educación. Sin embargo, la gestión educativa está descentralizada a los gobiernos regionales y locales (25 gobiernos regiones); es decir, son los encargados de la prestación o ejecución de los bienes y servicios públicos educativos. Es así que las intervenciones diseñadas y promovidas por el Ministerio de Educación son implementadas por las regiones a través de recursos propios.

Al igual que la mayoría de sistemas educativos, en Perú la matrícula en el nivel primaria es obligatoria, lo que asegura una alta tasa de matrícula (99.3% de niños y niñas de 6 a 11 año); sin embargo, los aprendizajes de los estudiantes es bastante bajo. Para ilustrar ello, si tomamos como referencia los resultados de la prueba PISA², el Perú ocupa el puesto 64 de un total de 70 países. Más aun, el Perú ha participado en cuatro pruebas PISA, en los años 2000 y 2012, ocupando el último lugar y en el 2009 el penúltimo. En la última edición el

¹ La evidencia generada en torno a escuelas multigrado está enfocada en la re-asignación de estudiantes por niveles de aprendizaje (Banerji and Chavan, 2016) como alternativa más que como una intervención que busque mejora de la calidad docentes en contextos geográficos dispersos (donde naturalmente se originan las escuelas multigrado).

² Program for International Student Assessment.

Perú mejoró relativamente, pero se debe tener en cuenta que se sumaron 5 nuevos países, todos ubicados por debajo de Perú.

Si se utilizan los resultados de la Evaluación Censal de Estudiantes (ECE), prueba interna estandarizada que se aplica a estudiantes de segundo grado de primaria, los resultados educativos son igual de pobres. Del total de estudiantes evaluados, solo el 46.4% de estudiantes alcanzan un nivel de aprendizaje esperado para su grado en el área de comunicación, mientras que solo 34.1% lo hace para el área de matemáticas. Si se considera los aprendizajes dentro del territorio nacional, el rezago y la desigualdad educativa se ahonda. Del total de estudiantes evaluados en escuelas multigrado, solo el 22.1% alcanza el nivel esperado para su grado en comunicación, mientras que solo el 18.9% lo alcanza para matemáticas.

Una posible explicación de los bajos resultados educativos es la baja inversión pública en este sector. En el Perú, el gasto en educación representa solo 3% del PBI, siendo el más bajo en toda América del Sur. Por ejemplo, las economías más grandes de la región como son Brasil y Argentina destinan alrededor del 8% del PBI, mientras que el resto de economías más pequeñas invierten cifras ligeramente menores: Venezuela (6.4%), Bolivia (5.6%), Ecuador (5.2%) y Chile (4.2%).

Como consecuencia de ello, el sistema educativo peruano está atrasado tanto en infraestructura como en calidad docente. Si analizamos lo primero, se calcula una brecha de infraestructura educativa de S/.100 mil millones de soles (S/.30 mil millones de dólares) lo que representa el aproximadamente el 15% de su PBI actual. En algunas cifras, 7 de cada 10 escuelas requieren ser reforzados o reconstruidas y 2 de cada 3 no cuentan con alguno de los siguientes servicios: agua, desagüe o electricidad (Censo de Infraestructura Educativa).

Si analizamos la calidad docente medida a través de las prácticas pedagógicas, los resultados no mejoran. De acuerdo al Monitoreo de Prácticas Escolares (MPE), herramienta de observación en aula recogida por el Ministerio de Educación, solo el 25% de docentes promueven efectivamente el pensamiento crítico de los estudiantes, mientras que solo el 24% de docentes promueven efectivamente el involucramiento de estudiantes. Similar porcentaje (24%) de docentes realizan de manera efectiva la retroalimentación en la clase pero más aún solo un 3% lo hace de manera adecuada en el trabajo escrito.

En resumen, en el contexto educativo en el Perú es complejo y adverso, mientras que las condiciones y resultados educativos, tanto a nivel estudiantes como a nivel docente, son aún precarios.

3. Descripción de la intervención

El Acompañamiento Pedagógico Multigrado (APM) es una intervención del Ministerio de Educación del Perú a través del cual se brinda asesoría personalizada a los docentes de escuelas multigrado. El APM está escalado a nivel nacional (897 distritos en 25 regiones), y atiende aproximadamente 14,391 docentes en 6,404 escuelas (aproximadamente 37% de las escuelas multigrados del país) donde estudian alrededor de 174 mil estudiantes. En términos de presupuesto, el APM tiene un costo de S/.119 millones de soles, siendo su presupuesto por estudiante de S/.683 soles (aproximadamente US\$195 dólares).

El asesor pedagógico realiza tres tipos de actividades en durante el año: (i) asesoría personalizada a través de visita y observación en aula y retroalimentación; (ii) grupos de inter-aprendizajes entre docentes y (iii) talleres de actualización docente. Un asesor pedagógico atiende en promedio (de acuerdo a norma oficial) 8 escuelas y el número de horas varía según componente.

La principal actividad que desarrolla el asesor pedagógico es la asesoría personalizada a través de visitas y observación en aula. En esta actividad, el asesor pedagógico realiza visitas de observación en aula de una sesión pedagógica de los docentes seguidas por una sesión de reflexión sobre las fortalezas y debilidades en el desarrollo de sesión y soporte teórico sobre las mejoras que deben implementarse hasta la próxima visita. De acuerdo al procedimiento establecido por las normativas del Ministerio de Educación, el asesor realiza una primera visita de diagnóstico, que le sirve para evaluar la condición inicial del docente a través de una rúbrica de observación que contiene las competencias priorizadas a desarrollar en los docentes. Después de la visita diagnóstica, el acompañante prepara el plan individualizado del docente y las aplica durante las siguientes siete visitas. Al concluir el año escolar, el asesor pedagógico realiza una visita de evaluación, cuya finalidad es determinar los avances logrados por el docente. Así, en total son nueve visitas durante el año (una al mes), con una duración de 5 horas cronológicas por cada visita.

Una segunda actividad que desarrolla el asesor es liderar los grupos de inter-aprendizaje entre docentes, que son espacios de intercambio entre pares docentes, y cuyo fin es intercambiar reflexiones sobre prácticas pedagógicas. Consiste en la realización de reuniones programadas entre el asesor pedagógico y los docentes para abordar temáticas variadas que van desde la reflexión de la práctica pedagógica hasta la profundización de aspectos de interés extraído de las visitas (RSG N° 008-2016-MINEDU). Estos espacios se realizan ocho veces al año por una duración de 2 horas cada una.

La tercera actividad son los talleres de actualización docente que buscan capacitar a los docentes en procesos pedagógicos, didácticos, planificación curricular y otros que se consideren relevantes para la mejora de la práctica pedagógica en las diferentes áreas priorizadas. Las asesorías personalizadas sirven para hacer seguimiento a la aplicación de lo aprendido en los talleres. Esta actividad diseñada y ejecutada por el asesor pedagógico con una frecuencia de 2 veces al año, el primero de 16 horas pedagógicas y el segundo de 64 horas.

A modo de resumen, el APM consiste en 3 actividades en las cuales los docentes reciben casi 141 horas de asesoría pedagógica durante el año escolar, de las cuales 45 horas son personalizadas (incluye observación de aula y retroalimentación) y el resto de manera grupal. El anexo 2 muestra una línea de tiempo de los principales hitos de la intervención.

3.1. Mecanismos

Esta sección busca presentar los potenciales canales de transmisión de la asesoría pedagógica; es decir, los canales por los que opera la asesoría pedagógica para mejorar las variables de resultado a nivel estudiante.

3.1.1. Prácticas pedagógicas docentes

Un primer mecanismo es la mejora de las prácticas pedagógicas a través de la observación de sesiones pedagógicas en aula, la retroalimentación personalizada y seguimiento continuo durante al menos 45 horas durante todo el año escolar, lo que finalmente mejora los aprendizajes de los estudiantes (Garet et al., 2008; Kretlow, et al, 2011; Landry, et al, 2011; Allen, et al 2011; Yoshikawa, et al. 2015; Blazar y Kraft, 2015; Cilliers, et al 2018).

La asesoría pedagógica parte de la observación de prácticas pedagógicas en aula, y se centra en reforzar las habilidades docentes con debilidades. De este modo, el APM brinda insumos necesarios para el mejoramiento de las prácticas pedagógicas.

De acuerdo a la literatura, si bien la relación entre prácticas pedagógicas y aprendizajes es positiva, se debe tener en cuenta que un incremento en las prácticas pedagógicas no se transmite completamente hacia los aprendizajes. Kraft et al. (2018) muestran que los programas de asesoría pedagógica incrementan considerablemente las prácticas docentes (en promedio, 0.49 DE) pero que se requieren importantes mejoras para poder lograr incrementos en los aprendizajes. Así, los autores estiman que un incremento de 1 desviación estándar se transmite solo en 0.21 DE de los aprendizajes.

En esa misma línea, un aspecto relevante para situar los mecanismos de las prácticas pedagógicas es el tamaño de la intervención. En Kraft et al. (2018) se muestra una relación inversa entre el tamaño de la intervención y los efectos en prácticas pedagógica de los estudios analizados. Para intervenciones pequeñas y más controladas, se observan efectos similares a 0.80 DE, mientras que para intervenciones más grandes y menos controladas los efectos en la literatura caen a menos 0.20 DE.

Otro aspecto importante para comprender los mecanismos de las asesorías pedagógicas es periodo de maduración de la intervención. En Glazerman et al (2010), se muestra que las mejoras en las prácticas pedagógicas requieren un periodo de maduración que hace que los efectos recién se visibilizan, en su estudio, en el tercer año de implementación.

En síntesis, se espera que el APM mejore las prácticas pedagógicas de los docentes intervenidos. De acuerdo a la literatura, estos efectos deberían ser mayores a los encontrados sobre los aprendizajes pero menores a la literatura por el periodo de maduración (solo un año) y la escala de la intervención (nivel nacional).

Para analizar este canal, se utilizan el cuestionario a directivos que pregunta sobre las prácticas pedagógicas que desarrollan los docentes a su cargo. Esta información no es recogida a través de rúbricas o instrumentos especializados pero representa la mejor aproximación disponible utilizando data administrativa.

3.1.2. Trabajo en conjunto e involucramiento entre docentes

Un segundo mecanismo a analizar es el trabajo en conjunto entre docentes, promovido por los grupos de inter-aprendizajes, cuyo objetivo es que los docentes intercambien reflexiones sobre sus prácticas pedagógicas. De acuerdo a Kraft et al (2018), la mayoría de intervenciones de asesoría pedagógica es complementado con un tratamiento adicional (talleres, videos, entre otro) y estas potencian los efectos de las asesorías: por ejemplo, los talleres grupales incrementan, en promedio, en 0.12 DE los efectos sobre los aprendizajes de los estudiantes. Un complemento novedoso del APM es son los grupos de inter-aprendizaje que motivan la interacción e involucramiento entre docente y que podrían potenciar los efectos del APM. Lamentablemente, la literatura no dispone de evidencia de este tipo de complementos pero sí es posible analizar los canales utilizando el cuestionario complementario a directivos.

El cuestionario a directivos realiza preguntas de frecuencia con las cuales los docentes realizan actividades de retroalimentación entre docentes sobre formas de enseñanza, trabajo conjunto para proyectos específicos, observación de clases entre docentes, intercambio de materiales de enseñanza, debates sobre las prácticas de evaluación e intercambios de ideas sobre el manejo de conductas inadecuadas.

3.1.3. Motivación de los docentes

Un tercer mecanismo, poco explorado en la literatura, es la motivación de los docentes por la asesoría y la mejora de sus prácticas. A través de este canal se espera que los docentes incrementen su productividad y, de este modo, los aprendizajes de los estudiantes. La escasa evidencia (Albornoz et al., 2017 y Glazerman et al. 2010) es mixta pero tampoco miden las mismas dimensiones de satisfacción, por ello, se debe tomar con cuidado. Albornoz et al. (2017) muestra mejoras en la percepción docente sobre su trabajo y en la satisfacción con sus propias prácticas pedagógicas, así como mejoras en la percepción de aprendizajes de sus estudiantes. Por otro lado, Glazerman et al. (2010) no encuentra cambios en las actitudes de los docente, incluyendo su satisfacción con su decisión de ser docente y su percepción sobre su preparación para afrontar diferentes aspectos de su trabajo.

Para analizar este canal, se utilizan dos aproximaciones de satisfacción a partir de la data administrativa disponible (cuestionario a directivos): frecuencia de inasistencia de los docentes y esfuerzo docente (medido como preparación adecuada de sus clases).

4. Diseño de la evaluación

4.1. Diseño de la aleatorización

El universo de escuelas elegibles está definido por los criterios de focalización (principalmente, ser escuela multigrado). La selección de escuelas fue estratificada por región y consistió en una parte determinística y una parte aleatoria. En la parte determinística, en cada región se continuó interviniendo a las escuelas multigrado que ya recibían la intervención al año anterior y se focalizó unas nuevas ubicadas en distritos donde ninguna escuela ha rendido la Evaluación Censal de Estudiantes. En este grupo determinístico se asignó al tratamiento en total 2,607 escuelas (29.5% del total de escuelas elegibles).

En la parte aleatoria, sobre el resto de escuelas que pertenecen al universo de elegibles, se

seleccionó aleatoriamente un grupo de escuelas, de modo que en cada región se respete una cuota regional establecida por el Ministerio de Educación. Cabe indicar que la asignación al tratamiento fue realizado desde el Ministerio de Educación sin interferencias políticas (nacionales o regionales) y la implementación de la intervención se dio a nivel regional. A pesar de tener instancias distintas de asignación e implementación, la asignación inicial fue respetada; es decir, se implementó sin modificaciones. En este grupo aleatorio se asignó en total 3,797 escuelas tratadas y 2,421 escuelas controles (43.0% y 27.4% del total de escuelas elegibles, respectivamente).

En el presente trabajo solo se analiza los impactos de la intervención del APM en el grupo aleatorio estratificado por regiones que representa más del 70% de la población elegible por la intervención a nivel nacional.

4.2. Data

En los últimos años, se ha incrementado notablemente el uso de fuentes administrativas para investigación debido a que sus bajos costos, su alta cobertura o cobertura universal y su menor susceptibilidad a sesgos en comparación a las fuentes primarias. Para el presente estudio, la data proviene de diversas fuentes de información secundarias a partir de los registros administrativos del Ministerio de Educación.

La principal fuente de datos proviene de la Evaluación Censal de Estudiantes (ECE) que recoge cada año la Oficina de Medición de la Calidad de los Aprendizajes (OMCA) del Ministerio de Educación. La ECE es una prueba estandarizada que mide los aprendizajes de las áreas de lectura y matemáticas en estudiantes de segundo grado de primaria³. Los resultados de la ECE se presentan de dos maneras: i) puntaje de cada materia que obtiene cada estudiante y ii) niveles de aprendizaje (satisfactorio, en proceso o inicio). Para el año 2016, la ECE adicionó un cuestionario dirigido a directores de las escuelas que recogió información de características generales del director, percepción del dominio y prácticas pedagógicas de los docentes a cargo⁴ y frecuencia con las que los docentes desarrollan un listado de actividades. Este cuestionario también será utilizado para analizar los posibles mecanismos de transmisión de los efectos del APM sobre los aprendizajes.

El campo de ECE se llevó a cabo entre los días 29 y 30 de noviembre de 2016; es decir, casi al finalizar el año educativo. Un detalle importante es que, si bien la ECE es censal, por motivos operativos solo visita a escuelas que tengan más de 5 estudiantes en el grado evaluado lo que implica que de las 29,742 escuelas públicas primaria en el Perú, solo se recoge información para 14,227 escuelas en el año 2016.

Además, se utiliza información administrativa del Censo Educativo, un cuestionario que el Ministerio de Educación solicita llenar de manera virtual a los directores de todas las escuelas. La información recopilada incluye materiales educativos, infraestructura de los locales escolares, número de docentes, secciones, entre otros. La data del Censo Educativo sirve para controlar y comprobar balance de características entre los grupos.

³ En lectura, la ECE evalúa si los estudiantes recuperan información del texto, infiere el significado del texto y reflexiona sobre la forma y contenido, mientras que en matemáticas evalúa si los estudiantes resuelven problemas asociados a operaciones de adición y sustracción, así como el uso del número y del sistema de numeración decimal.

⁴ El director tiene entre sus funciones supervisar y monitorear a los docentes a cargo. Para ello realiza observaciones de aula, no estandarizadas a criterio del director.

Finalmente, la última fuente de información proviene del Barrido Censal que contiene data socioeconómica recogida durante los años 2012-2013 con la finalidad de mejorar los algoritmos de pobreza y focalización de programas sociales. La data del Barrido Censal ha sido agregada a nivel de distrito para poder controlar y comprobar balance de características entre los grupos.

4.3 Muestra y pruebas de balance

La muestra final de la aleatorización, tal como se detalló en la sección 4.1, es de 6,218 escuelas (3,797 tratadas y 2,421 controles), pero la muestra final del estudio está acotada a las escuelas que disponen de data administrativa. Es así que la muestra se reduce a aquellas escuelas rindieron la Evaluación Censal de Estudiantes el año 2016 (solo 2,559 escuelas) y, adicionalmente, se excluyó aquellas regiones que tuvieran excesiva poca muestra⁵. De este modo, luego de realizar los dos filtros mencionados, la muestra final del estudio es de 2,363 escuelas (1,354 tratadas y 1,009 controles); es decir, alrededor de 38.0% de la muestra inicial.

Lo que sigue en la presente sub-sección es analizar el balance en características observables de la muestra inicial de la aleatorización, de la muestra que se pierde por uso de data administrativa (attrition) y la muestra final del estudio.

4.3.1. Balance de la muestra inicial

La tabla A del Anexo 3 muestra el análisis de balance para la muestra inicial de la aleatorización, siendo la unidad de análisis la escuela. La columna (1) y (2) muestran los promedios de cada variable de las escuelas tratadas y controles, respectivamente; la columna (3) la diferencia de ambos grupos y, finalmente, la columna (4) el p-value asociado a la diferencia de las medias de ambos grupos.

Como se observa en la tabla A las escuelas tratadas y controles son bastante similares. De un total de 19 variables analizadas a nivel escuela, solo 5 muestran ligeros desbalances pero en ninguna variable de resultado importante como los aprendizajes de Español y Matemáticas, ratios de aprobación o deserción del año anterior. Las diferencias solo se encuentran en variables como tamaño de la escuela (número de alumnos y docentes), material de techos del local escolar y ruralidad, siendo en todos los casos a favor de las escuelas del grupo control⁶.

Adicionalmente, la tabla A permite analizar las características de la muestra de estudio. En general, son escuelas pequeñas (en promedio menos de 30 estudiantes en toda la escuela), ubicadas en zonas rurales (más del 90% de escuelas) y en distritos pobres (más del 60% de hogares pobres).

⁵ Por ejemplo, luego de excluir a las escuelas que no rindieron la ECE hubo 1 región que no tenía escuelas controles, 2 regiones que solo tenían una escuela control, 3 regiones que solo tenían dos escuelas controles, y así sucesivamente. Para definir la muestra final se excluyó a las regiones que no tuvieran al menos 10 escuelas en cada grupo. Si bien este último paso no es obligatorio, los resultados tampoco varían significativamente a diferentes cortes.

⁶ Adicionalmente, se incluyó la variable número de estudiantes en segundo grado de primaria, por ser el grado evaluado. Los resultados son los mismos que al medir el número de estudiantes o número de docentes de la escuela, pues todas son variables que aproximan el tamaño de la escuela. Por tanto, al ya ajustar por número de estudiantes y docentes de la escuela, ya no se incluye al número de estudiantes del grado evaluado como variable control cuando se ajusta por el desbalance.

4.3.2. Balance de la muestra de attrition

En presencia de attrition es importante analizar el balance entre los grupos de la muestra excluida; es decir, de la muestra de escuelas que no disponen de data administrativa, con la finalidad de descartar que los grupos pierdan muestra de manera diferenciada.

La tabla B del Anexo 3 muestra los resultados de las prueba de balance para las mismas 19 variables analizadas sobre la muestra inicial. Como puede observarse, los resultados son los mismos a la tabla A; es decir, las variables que muestran y no muestran balance son las mismas que la tabla B. De este modo, se puede comprobar que la pérdida de muestra no representa un mayor problema para el estudio.

Un aspecto importante para analizar la validez externa de los resultados es mostrar que las escuelas que salieron de la muestra son similares a las que salieron de la muestra experimental. Para tales fines, se compara las características observables de las muestra saliente (muestra attrition) y la muestra restante (muestra final). La tabla D del anexo 3 muestra dicha comparación, encontrando que las muestras difieren, siendo favorables a las escuelas que se quedan en la muestra final del estudio: son más grandes (mayor número de estudiantes y docentes), tienen mayores ratios de alumnos por docente, tienen mejor infraestructura (piso, paredes y techos) y están áreas más urbanas pero las diferencias en resultados en aprendizajes (matemáticas y comprensión lectora) y ratios de aprobación no son significativas.

En este sentido, se debe considerar que las diferencias encontradas entre estas dos muestras son esperables debido a que la principal fuente de exclusión de la ECE (que determina su presencia en la muestra final) son los criterios logísticos para aplicar dicha prueba: tener al menos 5 estudiantes en el grado evaluado. Por tanto, de manera natural, las escuelas que se pierden (attrition) son escuelas más pequeñas y, probablemente, más rurales y con peor infraestructura.

De este modo, por lo mostrado en la tabla B, el attrition no representa un problema de validez interna de los resultados, pero sí podría representar un problema para la validez externa de los mismos (tabla D), incluso aun luego de probar que las variables de resultados no están desbalanceadas.

4.3.3. Balance de la muestra final del estudio

Una vez que se acota la muestra a la data administrativa disponible, la muestra se reduce a 2,363 escuelas (1,354 tratadas y 1,009 controles). Las tablas C del Anexo 3 muestran el análisis de balance entre los grupos para la muestra final del estudio.

En los dos primeros paneles, la tabla C analiza variables a nivel escuela y distritales, al igual que las tablas A y B. Otra vez, de las 19 variables analizadas a nivel escuela, solo 4 muestran ligeros desbalances pero en ninguna variable de resultado de importancia como son los aprendizajes de Español y Matemáticas, ratios de aprobación o deserción del año anterior. Las variables desbalanceadas son las mismas; es decir, en tamaño de la escuela (número de alumnos y docentes), material de techos del local escolar y ruralidad.

Adicionalmente, sobre la muestra de escuelas que rindieron la ECE en 2016 (muestra final) se puede analizar balances adicionales utilizando el cuestionario a directores. En este cuestionario, se pueden analizar el balance en otros indicadores a nivel director como sexo, edad, tipo de contrato, años de experiencia y máximo nivel educativo del director. Para estas

variables tampoco se detectan desbalances (ver tercer panel de la tabla C).

Por tanto, de las 24 variables analizadas solo 4 muestran desbalance. En este caso, para corregir posibles sesgos por las diferencias en estas variables se controla por ellas en las regresiones (Glennerster y Takavarasha, 2013). La explicación y detalles se desarrollan en la siguiente sección.

4.4. Estrategia Empírica

4.4.1. Efectos sobre aprendizajes

La forma reducida para estimar el impacto del APM sobre los aprendizajes es el siguiente:

$$y_{isb1} = \beta_0 + \beta_1(\textit{Coaching})_s + X'_{isb} \Gamma + \lambda_b + \varepsilon_{isb1}$$

Donde y_i es el resultado educativo (en el follow-up) para el estudiante i , en la escuela s , en el estrato b ; $(\textit{Coaching})_s$ una variable binaria que indica que la escuela s recibió la intervención, λ_b son efectos fijos por estrato de aleatorización (25 regiones); X'_{isb0} es un vector de variables de control medidos en la línea de base; y ε_{isb1} es un término de error clusterizado a nivel escuela. De este modo, β_1 representa el efecto promedio del tratamiento del APM.

El vector de variables de control X'_{isb} incluye: i) variables a nivel estudiantes altamente correlacionada con y_{isb1} con la finalidad de incrementar el poder estadístico (Bloom et al., 2007); y ii) variables desbalanceada en la línea de base con la finalidad de corregir posibles diferencias pre-existentes (Glennerster y Takavarasha, 2013). Como se indicó en la sección 4.3.3, las variables desbalanceadas son tamaño de la escuela (número de alumnos y docentes), material de techos del local escolar y ruralidad.

Una especificación adicional es incluir los puntajes de rendimiento académicos en Español y Matemáticas del año anterior (año 2015) a nivel escuela. Esta especificación busca corregir algún posible sesgo persistente a la inclusión de covariados desbalanceados e incrementar el poder estadístico (Muralidharan, 2017; Bloom et al., 2007).

Adicionalmente, se desea tener una mirada de los efectos del APM sobre la distribución de los aprendizajes. Para tales fines, se realizan regresiones sobre los cuantiles de aprendizajes; es decir, se considera la relación entre y_{isb1} y el $(\textit{Coaching})_s$ utilizando la función condicional del cuantil q , $Q_q(y_{isb1}|\textit{Coaching})$.

Para testear los efectos heterogéneos, se estima lo siguiente:

$$y_{isb1} = \beta_0 + \beta_1(\textit{Coaching})_s + \beta_2(\textit{Coaching} \times \sigma)_j + X'_{isb0} \Gamma + \lambda_b + \varepsilon_{isb1}$$

Donde σ_j es una variable binaria que representa un sub-grupo de interés y que ahora está incluida en el vector X'_{isb0} , siendo β_2 el efecto diferencial para el sub-grupo de interés y $j \in (i, s)$ que indica que σ puede ser una variable que identifica un sub-grupo a cualquier nivel i, s .

4.4.2. Efectos sobre mecanismos

Para los mecanismos, se estiman formas reducidas del impacto del APM sobre indicadores de desempeño de los docentes, aproximado por las respuestas de los directores. Las estimaciones son realizadas a nivel escuela y está especificadas de la siguiente manera:

$$y_{sb1} = \beta_0 + \beta_3(\text{Coaching})_s + X'_{sb} \Gamma + \lambda_b + \varepsilon_{sb1}$$

Donde y_i es el desempeño del docente (en el follow-up) en la escuela s , en el estrato b ; $(\text{Coaching})_s$ una variable binaria que indica que la escuela s recibió la intervención, λ_b son efectos fijos por estrato de aleatorización (25 regiones); X'_{sb} es un vector de variables de control medidos en la línea de base; y ε_{sb} es un término de error i.i.d. De este modo, β_3 representa el efecto promedio del tratamiento del APM sobre el desempeño de los docentes.

Debido a la gran cantidad de variables analizadas (15 variables) en una misma muestra, se ajustaron los p-values utilizando las pruebas de hipótesis múltiple. Los métodos de ajuste por hipótesis múltiple utilizados son: Westfall-Young (columna 2), Bonferroni-Holm (columna 3) y Sidak-Holm (columna 4). El ajuste de Westfall-Young emplea bootstrapping para tomar en cuenta la correlación cruzada de las variables utilizadas, mientras que los ajustes de Bonferroni-Holm y Sidak-Holm asumen que las variables son independientes y, por tanto, puede resultar demasiado conservador (sobre-ajuste de p-values).

5. Resultados

Los resultados son presentados en tres subsecciones. En la primera sub-sección se analizan los efectos principales sobre aprendizajes de los estudiantes, así como los posibles efectos en los diferentes cuantiles de la distribución. Posteriormente, se ahonda un poco en el análisis para estudiar los efectos diferenciados por características de los estudiantes y de las escuelas (efectos heterogéneos).

En la segunda sub-sección, se analizan los mecanismos por los cuales el APM puede estar afectando los resultados educativos de los estudiantes. En este punto, se estudian tres hipótesis relacionadas al diseño de la intervención, así como a la evidencia de intervenciones similares.

En la tercera sub-sección, se explora la estratificación de la asignación aleatoria para estudiar efectos por regiones y la relación de los efectos encontrados con el tamaño de la intervención y la ruralidad.

5.1. Resultados educativos a nivel estudiantes

5.1.1. Aprendizajes de estudiantes

La tabla 1 muestra los efectos sobre los aprendizajes de los estudiantes en las áreas de español y matemáticas, siendo alrededor de 0.098 DE y 0.125 DE, respectivamente. Estos resultados están alrededor de la media de efectos de intervenciones similares pero se debe tener en cuenta el tiempo de maduración y escalamiento de la intervención.

Tabla 1: Efectos del APM sobre los aprendizajes

	Español			Matemáticas		
	(1)	(2)	(3)	(4)	(5)	(6)
Coaching	0.062* [0.035]	0.108*** [0.034]	0.098*** [0.032]	0.092** [0.037]	0.135*** [0.036]	0.125*** [0.034]
EF por estrato	Sí	Sí	Sí	Sí	Sí	Sí
Controles	No	Sí	Sí	No	Sí	Sí
Rezago del puntaje (escuela)	No	No	Sí	No	No	Sí
Observaciones	20,396	20,396	20,396	20,396	20,396	20,396
R-cuadrado	0.092	0.110	0.147	0.064	0.080	0.121

Nota: Errores estándar robustos en corchetes, clusterizados a nivel escuela.

Las columnas (1) y (4) muestran los efectos solo controlando por estrato (región). Las columnas (2) y (5) adicionan controles de sexo del estudiante, tamaño de la escuela (número de alumnos y docentes), material de techos del local escolar y ruralidad. Las columnas (3) y (6) adicionan el rezago del puntaje en matemáticas y español agregado a nivel escuela.

*** p<0.01, ** p<0.05, * p<0.1

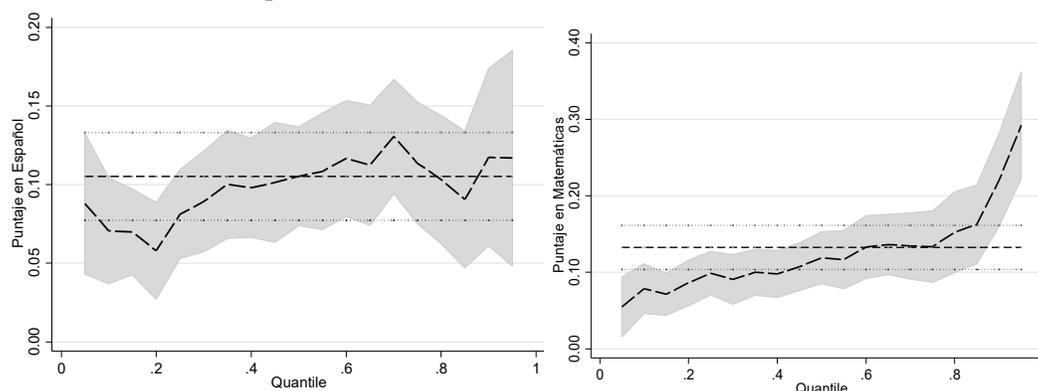
En el meta análisis realizado por Kraft et al. (2018), se muestra que las intervenciones más grandes analizadas tienen en promedio efectos menores al 0.10 DE. Aun así, las intervenciones analizadas no superan los 400 docentes intervenidos (casi el 60% no tiene ni 100 docentes intervenidos), mientras que el APM atiende aproximadamente 14,000 docentes.

Por otro lado, es importante analizar el contexto donde se desarrolla la intervención; es decir, las escuelas intervenidas son tipo multigrado y, por ello, representan mayores retos por la organización de la enseñanza en el aula (un solo docente enseña a más de un grado). El presente trabajo es el primer estudio en analizar y encontrar efectos positivos y significativos de un programa de asesoría pedagógica en este tipo de escuelas.

En la tabla 1, además se observa, que si bien se introducen algunos covariados para corregir los desbalances anteriores a la intervención (incluso los rendimientos rezagados a nivel de escuela), los resultados son bastante robustos a inclusión o exclusión de estas correcciones; es decir, los desbalances mostrados en el Anexo 3 parecen no introducir sesgos en las estimaciones.

Una pregunta de interés en los programas de asesoría es conocer si existen efectos distributivos por percentiles de aprendizajes. El gráfico 1 analiza los efectos los percentiles de la distribución de aprendizajes de Español (panel A) y Matemáticas (panel B) y los compara con el efecto promedio de tratamiento (en líneas punteadas). Para Español, los efectos por percentiles oscilan alrededor del efecto promedio pero para matemáticas se muestra una tendencia positiva en los percentiles de la distribución; es decir, los efectos son mayores en los percentiles mayores, incluso diferentes significativamente del promedio para los percentiles muy superiores (mayores a 90).

Gráfico 1: Efectos por percentiles de la distribución
Panel A: Español **Panel B: Matemáticas**

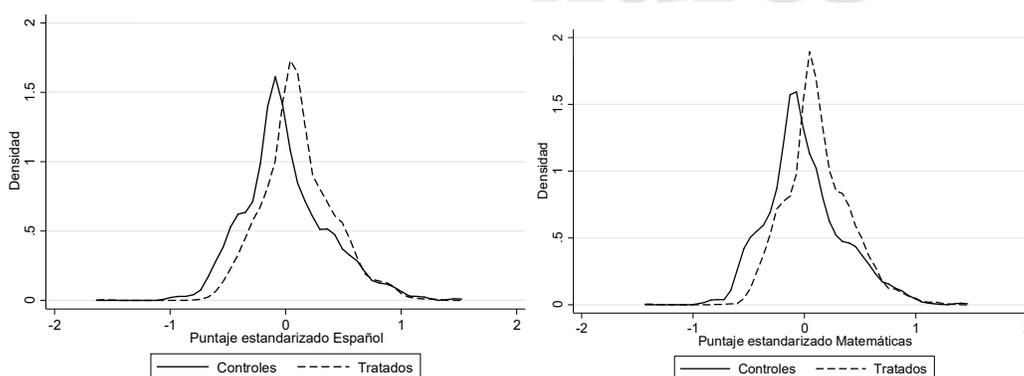


Nota: Efectos calculados por cada percentil en base a los puntajes estimados en la especificación con controles y efectos fijos por estrato (regiones). El panel A muestra los efectos por percentiles de los puntajes de los aprendizajes de Español, siendo las áreas sombreadas los intervalos de confianza de los estimadores puntuales y la línea horizontal el efecto promedio de tratamiento, mientras que el panel B muestra los efectos por percentiles puntajes de los aprendizajes de Matemáticas, bajo las mismas especificaciones.

Lo mismos resultados se observan si se analizan las densidades de los puntajes de ambos grupos tanto para Español como para Matemáticas (ver gráfico 2). En el panel A, se comparan las densidades de los puntajes de Español de los estudiantes que reciben y no reciben la intervención, y se observa que siguen distribuciones bastante similares, pero la media de los tratados (líneas punteadas) está corrida a la derecha, lo que representa el efecto de la promedio de la intervención.

En el panel B, se realiza la misma comparación pero ahora los tratados, además de tener la media corrida a la derecha, tienen mayor concentración de estudiantes con mayores puntajes lo que representa los efectos mayores en los percentiles superiores.

Gráfico 2: Densidad de puntajes de aprendizajes por grupo
Panel A: Español **Panel B: Matemáticas**



Nota: Densidades generadas en base a los puntajes estimados en la especificación con controles y efectos fijos por estrato (regiones). El panel A muestra las densidades de los puntajes de los aprendizajes de Español para los tratados (líneas punteadas) y controles (líneas sólidas), mientras que el panel B muestra los puntajes de los aprendizajes de Matemáticas, bajo las mismas especificaciones.

En resumen, el APM muestra tener efectos significativos en los aprendizajes en las dos áreas evaluadas, pero principalmente en matemáticas, tanto en el efecto promedio como en efectos en los percentiles superiores de la distribución. Los efectos promedios son menores a otras intervenciones de asesoría pedagógica (Kraft, 2018) pero en las magnitudes esperadas en la medida que se interpreten teniendo en cuenta el tiempo de maduración, escalamiento y contexto de la intervención.

5.1.2. Efectos heterogéneos

En la subsección anterior, se mostraron los efectos del APM sobre los aprendizajes de los estudiantes en las áreas de Español y Matemáticas. En la presente subsección se desea explorar si los efectos del APM son diferenciados por tipos de grupos. Para tales fines, se analiza efectos heterogéneos en el sexo de los estudiantes, sexo del director de la escuela, tipo de contratación y nivel educativo del director, así como si la escuela se encuentra ubicada en un distrito que pertenece al quintil más pobre.

La tabla 2 muestra los resultados sobre los grupos descritos. El panel A muestra los resultados sobre los resultados en Español, mientras que el panel B hace lo mismo en Matemáticas. Para cada panel se muestra solo la versión con covariados (corrigiendo por desbalance y puntajes del año anterior), regresiones con efectos fijos por estrato de aleatorización y errores estándar son robustos y clusterizados a nivel escuela.

Los resultados indican que no se disponen de efectos diferenciados por tipo de sexo del estudiante tanto en Español como Matemáticas, siendo este resultado diferente a los encontrados por Albornoz et al. (2017), quienes sí encuentran efectos diferenciados a favor de las estudiantes mujeres. En cuanto al quintil de pobreza nacional donde está ubicada la escuela, la intervención está presente, principalmente, en distritos de los dos quintiles más pobres. Para este grupo, tampoco se encuentran efectos diferenciados en las escuelas del quintil más pobre.

Sin embargo, cuando se explora el perfil del director de la escuela, sí se identifican efectos diferenciados en el sexo del director y su nivel educativo, mas no en su tipo de contratación. En el caso del sexo del director, los efectos se dan casi en la totalidad cuando el profesor es hombre. Esto puede ocurrir debido a que los asesores pedagógicos son principalmente hombres, lo que sugiere que puede existir un mayor entendimiento o afinidad entre el asesor pedagógico y el director si estos son del mismo sexo.

En el caso del nivel educativo, los efectos del APM también se dan casi en su totalidad en los directores de bajo nivel educativo, aunque solo es significativo para Matemáticas. Esto puede ocurrir porque los directivos y docentes de más bajo nivel educativos son los que mayor pueden aprovechar las asesoría pedagógicas, mientras que los directivos y docentes más capacitados pueden partir de niveles más altos de enseñanza o ser más reacios a recibir asesoramiento pedagógico dada su alta calificación (posgrado).

Tabla 2: Efectos heterogéneos del APM

Grupo	Estudiante Mujer (1)	Quintil más pobre (2)	Director hombre (3)	Director nombrado (4)	Director sin posgrado (5)
Panel A: Resultados en Español					
Coaching	0.092*** [0.035]	0.109** [0.045]	-0.021 [0.051]	0.143* [0.073]	0.021 [0.063]
Coaching x grupo	0.028 [0.029]	-0.009 [0.066]	0.174*** [0.065]	-0.048 [0.081]	0.116 [0.072]
Panel B: Resultados en Matemáticas					
Coaching	0.124*** [0.038]	0.148*** [0.048]	0.025 [0.056]	0.181** [0.077]	0.038 [0.070]
Coaching x grupo	0.018 [0.029]	-0.033 [0.069]	0.143** [0.070]	-0.058 [0.085]	0.134* [0.079]
EF por estrato	Sí	Sí	Sí	Sí	Sí
Controles	Sí	Sí	Sí	Sí	Sí
Rezago del puntaje (escuela)	Sí	Sí	Sí	Sí	Sí
Observaciones	20,396	20,396	19,652	20,174	20,142

Nota: Errores estándar robustos en corchetes, clusterizados a nivel escuela.

Todas las columnas muestran los efectos del APM controlando por estrato (región), características de los estudiantes (sexo del estudiante) y características a nivel de escuela tales como tamaño de la escuela (número de alumnos y docentes), material de techos del local escolar y ruralidad y rezago del puntaje en matemáticas y español (agregado a nivel escuela).

*** p<0.01, ** p<0.05, * p<0.1

En resumen, no se identifican efectos diferenciados por sexo de los estudiantes, pobreza donde está ubicada la escuela o por características de los directores como tipo de contratación pero sí en el sexo del director y su nivel educativo.

5.2. Mecanismos

En esta subsección se analizan los mecanismos propuestos en la subsección 3.1. Para tales fines se aprovecha el cuestionario a directores, aplicado durante el trabajo de campo de la ECE, sobre su percepción de las actividades desarrolladas por los docentes a su cargo.

5.2.1. Prácticas pedagógicas docentes

El cuestionario a directivos recoge seis indicadores relacionados a estrategias de enseñanza que el director observa en los docentes. Específicamente, el cuestionario pregunta al director su percepción de cuánto dominio tienen los docentes de su escuela sobre un conjunto de estrategias pedagógicas listadas. Las respuestas son de opción múltiple, pero para fines del presente estudio se convierten en una variable binaria que abarque las categorías de “lo domina con algo de dificultad” y “lo domina”.

La tabla 3 muestra que, de acuerdo a la percepción del director, los docentes de las escuelas que reciben APM dominan o tienen poca dificultad para dominar i) un enfoque comunicativo para la enseñanza de la escritura y la lectura (2.4 p.p.), ii) los tipos de evaluación en las diferentes áreas curriculares que enseña el docente (2.5 p.p.) y iii) estrategias para promover un ambiente de aprendizaje adecuado (2.9 p.p.). Cuando se ajusta por hipótesis múltiple, solo el indicador de estrategias para promover un ambiente de aprendizaje adecuado es significativo bajo los métodos de corrección de Westfall-Young, Bonferroni-Holm y Sidak-Holm (ver Anexo 4).

Por otro lado, los indicadores relacionados a organización y planificación no son afectados por el APM y, finalmente, tampoco el uso de tecnología de la información para la enseñanza. Esto último es esperable debido a que no es un área impulsada por la intervención por los escasos recursos tecnológicos de las escuelas multigrado. En estos casos las correcciones de hipótesis múltiples no modifica los resultados pues no son significativos.

Es necesario indicar, tal como se mencionó anteriormente, que los impactos detectados pertenecen a variables reportadas por el director de acuerdo a su percepción y a instrumentos estandarizados que midan a través de rúbricas las prácticas pedagógicas.

5.2.2. Trabajo en conjunto e involucramiento entre docentes

El cuestionario a directivos también recoge siete indicadores relacionados a actividades y frecuencia de interacción e involucramiento entre docentes. Específicamente, el cuestionario pregunta al director si los docentes de su escuela i) observan las clases de otros docentes, ii) dan retroalimentación a otros docentes sobre su forma de enseñanza, iii) trabajan en conjunto para proyectos específicos (concurso de conocimientos, proyectos de clase, etc), iv) intercambian entre ellos materiales de enseñanza, v) se involucran en debates internos sobre el desarrollo de aprendizajes de grupos específicos de estudiantes, vi) debaten sobre las prácticas de evaluación que se implementan para los estudiantes en la escuela y, vii) intercambian ideas sobre el manejo de conductas inadecuadas de los estudiantes.

Las preguntas son de opción múltiple que indica si realizan o no la actividad y la frecuencia con que lo realizan pero para fines del presente estudio se convierten en una variable binaria que indica si realizó o no las actividades referidas.

La tabla 4 muestra los impactos positivos del APM sobre todas las actividades mencionadas pero solo es significativo para las actividades de observación de clases de otros docentes (7.4 p.p.), retroalimentación a otros docentes sobre su forma de enseñanza (7.8 p.p.), trabajo en conjunto para proyectos específicos (5.5 p.p.), intercambio entre ellos de materiales de enseñanza (7.0 p.p.), involucramiento en debates internos sobre el desarrollo de aprendizajes (5.3 p.p.) y debate sobre las prácticas de evaluación que se implementan (4 p.p.).

Cuando se ajusta por hipótesis múltiple, solo los indicadores de observación de i) clases de otros docentes; ii) retroalimentación a otros docentes sobre su forma de enseñanza; iii) intercambio entre ellos de materiales de enseñanza; iv) involucramiento en debates internos sobre el desarrollo de aprendizajes, mantienen significancia al 5%, bajo los métodos de corrección de Westfall-Young, Bonferroni-Holm y Sidak-Holm (ver Anexo 4).

Tabla 3: Efectos del APM sobre prácticas pedagógicas docentes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		Estrategias para la organización de la información	Enfoque comunicativo para la enseñanza de la lectura y escritura.	Planificación curricular y diseño de sesiones de aprendizaje	Tipos de evaluación en las diferentes áreas curriculares que enseña	Uso de Tecnología de la Información y Comunicación para la enseñanza	Estrategias para promover un ambiente de aprendizaje adecuado					
Coaching	0.005 [0.012]	0.005 [0.012]	0.024** [0.010]	0.024** [0.010]	0.015 [0.009]	0.015* [0.009]	0.028** [0.011]	0.025** [0.011]	0.010 [0.018]	0.002 [0.019]	0.029*** [0.010]	0.029*** [0.010]
Media de Y (controles)	0.913	0.913	0.93	0.93	0.944	0.944	0.916	0.916	0.784	0.784	0.931	0.931
EF por estrato	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Covariados	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí
Observaciones	2,323	2,321	2,315	2,313	2,309	2,307	2,300	2,298	2,295	2,293	2,305	2,303

Nota: Errores estándar robustos en corchetes, clusterizados a nivel escuela. *** p<0.01, ** p<0.05, * p<0.1

Tabla 4: Efectos del APM sobre Trabajo en conjunto e involucramiento entre docentes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
		Observar las clases de otros docentes	Retroalimentación a otros docentes sobre su forma de enseñanza	Trabajar en conjunto para proyectos específicos	Trabajar en conjunto para proyectos específicos	Intercambiar entre ellos materiales de enseñanza	Involucrarse en debates internos sobre el desarrollo de aprendizajes	Debatir sobre las prácticas de evaluación que se implementan	Intercambiar ideas sobre el manejo de conductas inadecuadas de estudiantes					
Coaching	0.067*** [0.022]	0.074*** [0.022]	0.069*** [0.022]	0.078*** [0.022]	0.037** [0.019]	0.051*** [0.018]	0.056*** [0.017]	0.070*** [0.017]	0.044** [0.018]	0.053*** [0.018]	0.030* [0.016]	0.040** [0.016]	0.015 [0.013]	0.025* [0.013]
Media de Y (controles)	0.543	0.543	0.567	0.567	0.748	0.748	0.793	0.793	0.757	0.757	0.82	0.82	0.898	0.898
EF por estrato	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Covariados	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí
Observaciones	2,302	2,300	2,278	2,276	2,296	2,294	2,296	2,294	2,291	2,289	2,292	2,290	2,305	2,303

Nota: Errores estándar robustos en corchetes, clusterizados a nivel escuela. *** p<0.01, ** p<0.05, * p<0.1

5.2.3. Motivación de los docentes

Por último, el cuestionario a directivos también incluye preguntas relacionadas a inasistencia de docentes e inadecuada preparación de sus clases. En ambos casos los indicadores han sido convertidos en una variable binaria que indica si el evento ocurre o no.

Tabla 5: Efectos del APM sobre motivación docente

	Inasistencia por parte de los docentes		Docentes que no preparan adecuadamente sus clases	
	(1)	(2)	(3)	(4)
Coaching	-0.023 [0.018]	-0.008 [0.018]	-0.038* [0.021]	-0.016 [0.021]
Media de Y (controles)	0.219	0.219	0.403	0.403
EF por estrato	Sí	Sí	Sí	Sí
Covariados	No	Sí	No	Sí
Observations	2,314	2,312	2,318	2,316

Nota: Errores estándar robustos en corchetes, clusterizados a nivel escuela. *** p<0.01, ** p<0.05, * p<0.1

La tabla 5 muestra que el APM reduce la ocurrencia de inasistencia y poca preparación de las clases de los docentes pero no de manera estadísticamente significativa. En estos casos las correcciones de hipótesis múltiples no modifican los resultados pues no son significativos. Por tanto, los resultados de la tabla 5 concluyen que no se modifican las dimensiones de motivación docente analizadas en el presente estudio.

5.3. Efectos por región

Un análisis adicional que se puede implementar es determinar el efecto del APM por cada región explotando la asignación aleatoria estratificada. La relevancia de ello se debe a que la implementación del APM es descentralizada; es decir, si bien el APM es diseñado desde el Ministerio de Educación, los procesos de selección y contratación de los asesores pedagógicos depende cada región.

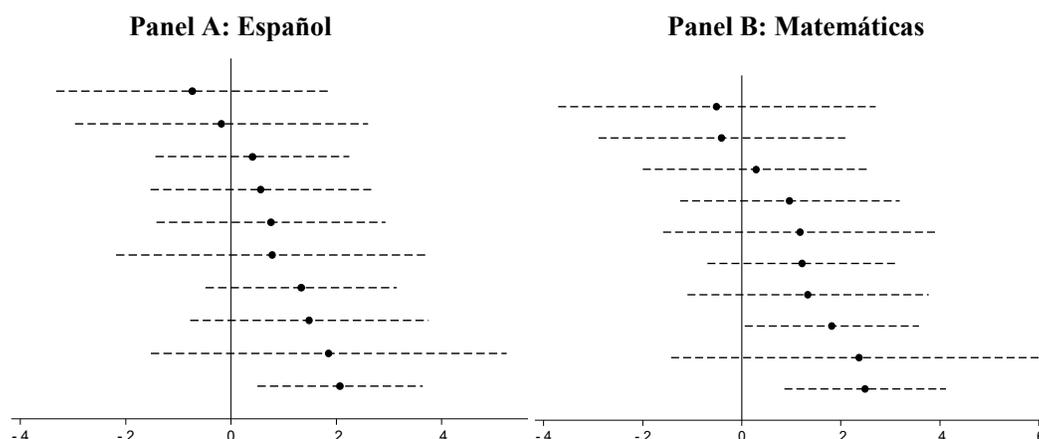
Antes de analizar estos resultados se debe considerar que una limitación que enfrenta este análisis es que algunas regiones tienen muestras muy pequeñas en alguno de los grupos de intervención (tratados o controles) lo que genera errores estándar grandes y por tanto intervalos de confianzas grandes y poder estadístico reducidos⁷.

El gráfico 3 muestra los efectos del APM por cada región, observándose una alta heterogeneidad entre ellos. Por ejemplo, la intervención muestra efectos grandes y significativos para la región Cajamarca (0.21 DE para Español y 0.25 DE para Matemáticas) pero para otras regiones, efectos pequeños y no significativos como la región Amazonas (-0.01 DE para Español y -0.05 DE para Matemáticas). Así, es posible que la intervención haya tenido mayores efectos en algunas regiones y en otras no, mientras que a nivel agregado (nacional) estos se compensen y hagan que se puedan detectar efectos de 0.098 y 0.125 DE

⁷ Los tamaños de muestra por región están en un rango entre 72 y 366 escuelas por región. El detalle de la muestra agregada y por grupo por cada región y los efectos para cada región están en la tabla A del Anexo 5.

en Español y Matemáticas, respectivamente.

Gráfico 3: Efectos del APM por región



Nota: Intervalos de confianza al 95%.

5.4. Análisis de costo efectividad

Un análisis importante se realiza a través de la determinación de costos unitarios de la intervención y la comparación de la costo-efectividad con otras intervenciones similares. La costo efectividad se calcula dividiendo los costos unitarios por estudiante entre los efectos estandarizados de la intervención. Para ello, un primer paso es analizar los costos desagregados de intervención (ver Tabla 6) y costo unitarios (ver Tabla 7). En la tabla 6 puede observarse que la mayor parte de los costos totales pertenecen a la contratación de asesores pedagógicos (72%) y en menor parte en la ejecución de los componentes de la misma (20%). Finalmente, un monto menor corresponde a la gestión de la intervención (8%) pero que por lo general no es tomado en cuenta en los pilotos mostrados en la literatura.

Tabla 6: Costos desagregados de la intervención

Componente	Costo (S/.)	Costo (US\$)	Porcentaje
Gestión de la intervención	9,162,957	2,617,988	8%
Visitas a escuelas	12,079,359	3,451,245	10%
Grupos de Inter Aprendizajes	853,500	243,857	1%
Taller Macro-regional	1,687,061	482,017	1%
Taller de actualización	9,236,295	2,638,941	8%
Contratación asesores pedagógicos	85,941,079	24,554,594	72%
TOTAL	118,960,251	33,988,643	100%

Nota: Tipo de cambio 3.5 soles por dólar

Fuente: Unidad de Planificación y Presupuesto del Ministerio de Educación

Un siguiente paso en el análisis de costo efectividad es la determinación de los costos unitarios de la intervención (ver Tabla 7). El APM tiene un costo anual por escuela de 18,576 soles (aproximadamente, 5,307 dólares) y un costo unitario anual por estudiante de 683 soles (alrededor de 195 dólares). De mismo modo, se puede analizar el costo por asesor pedagógico dividiendo el costo anual de contratación asesores por el número de ellos, siendo 45,762 soles (alrededor de 13,075) y un costo mensual por cada asesor de 3,815 soles (alrededor de 1,089 dólares).

Tabla 7: Costos unitarios de la intervención

Unidades	Unidades	Costo por unidad (S/.)	Costo por unidad (US\$)
Número de escuelas ^{1/}	6,404	18,576	5,307
Número de docentes ^{1/}	14,391	8,266	2,362
Número de estudiantes ^{1/}	174,107	683	195
Número de asesores pedagógicos ^{2/}	1,878	45,762	13,075

Nota: Tipo de cambio de 3.5 soles por dólar.

1/ Se divide el costo total de la intervención entre el número de unidades.

2/ Se divide el costo total por contratación entre el número de asesores pedagógicos.

Un tercer y último paso consiste en la comparación del costo-efectividad de la intervención con otras similares. En un intento por controlar las diferencias entre las intervenciones, se restringe los estudios aquellos que solo realicen visitas de asesoramiento presencial, que tengan información de costo de la intervención y que muestren resultados sobre las materias evaluadas en el presente estudio, reduciéndose a los estudios de Garet et al (2008) y Cilliers et al (2018). El gráfico 4 muestra el análisis comparativo con estos estudios, siendo el menos costo-efectivo el APM.

Gráfico 4: Análisis comparativo de costo-efectividad



Sin embargo, para analizar los resultados del análisis de costo efectividad se deben tener en cuenta dos aspectos:

- Mayores costos de la intervención asociados al contexto rural: El ámbito de intervención incrementa considerablemente los costos por estudiante de la implementación la intervención. En la tabla 6 se puede observar que los costos por la visitas representan el 10% del costo total de la intervención (aproximadamente 3.5 millones de dólares). Adicionalmente, la atomización de escuela repercute en el bajo número de docentes y estudiantes intervenidos por asesor pedagógico. En promedio, en el APM cada asesor pedagógico cubre durante el año escolar 92 estudiantes en 3 escuelas. En un contexto urbano, ese número de estudiantes podría cubrirse en solo dos secciones de un grado en una sola escuela.
- Menores efectos de la intervención por la escalabilidad a nivel nacional: Escalar una intervención tiene diversos retos, sobre todo en áreas rurales. En un estudio realizado por el Ministerio de Educación (MINEDU 2017) se encontró evidencia de contratación tardía de asesores pedagógicos en algunas regiones y además evidencia de reducción del perfil de los asesores pedagógicos para poder cumplir las metas de contratación. Estas limitaciones en la implementación propios de una intervención escalada (donde los procesos dependen de varios actores) repercuten en los impactos que puedan

tener la intervención; esto en comparación a la evidencia comparativa donde se pilotean intervenciones a pequeña escala y controlados (ver tabla A del anexo 6).

- Menores efectos por la falta de maduración de la intervención: Los impactos encontrados corresponden al primer año de la intervención, por lo que los resultados encontrados pueden considerarse de corto plazo y ser mayores en el mediano plazo (y por tanto más costo-efectiva). De acuerdo a la literatura (Glazerman et al, 2010), los efectos de las intervenciones de asesoría pedagógica necesitan un periodo de maduración para que los docentes mejoren sus prácticas pedagógicas.

6. Conclusiones y consideraciones finales

El presente estudio muestra los resultados de una intervención de asesoría pedagógica escalada a nivel nacional evaluada experimentalmente. Si bien existe vasta evidencia de los efectos de intervenciones de asesorías pedagógicas, a conocimiento del autor, toda la evidencia está generada en pilotos controlados o al menos en pequeña escala (Bruns et al. 2017; Albornoz, et al. 2017; Zhang, et al. 2013; Angrist y Lavy, 2001; Garet, et al, 2008; Yoshikawa et al. 2015). Así, el presente estudio brinda la primera evidencia de una intervención escalada a nivel y evaluada experimentalmente. Es importante notar que una intervención escalada a nivel nacional enfrenta diversos retos y complejidades presupuestales, logísticos y de gestión de alta relevancia para la discusión de política educativa.

Otro aspecto novedoso del presente estudio es que la intervención es implementada en escuelas multigrado. A conocimiento del autor, la evidencia de intervenciones de asesoría pedagógica para este tipo de escuelas es escasa o al menos casi inexistente. El contexto es relevante para la discusión de política educativa debido a que los docentes en escuelas multigrado enfrentan diversos retos pedagógicos para brindar una educación de calidad.

En ese mismo sentido, es necesario analizar la validez externa de los resultados. Al ser una intervención escalada, los resultados tienen validez externa en su ámbito (escuelas multigrado que reciben el APM) o incluso en escuelas rurales en contextos geográficos dispersos⁸. Sin embargo, la validez externa se reduce si se desea inferir los resultados a escuelas regulares (un docente por aula) o contextos geográficos urbanos. La validez externa también se debe analizar desde el estado del Perú, un país en desarrollo con bajos niveles de aprendizajes y calidad docentes, así como baja inversión en educación, incluso comparado con países similares o de la misma región.

Teniendo en consideración lo anterior, los resultados se pueden analizar desde la perspectiva de los diversos actores que participan de la intervención.

A nivel de estudiantes, en el primer año de implementación del APM se observan efectos sobre los aprendizajes en las áreas de Español y Matemáticas. Los efectos son similares a los encontrados en la literatura contextos similares y además se observa efectos diferenciados en los percentiles superiores en el área de matemáticas. Los resultados también indican que no identifican efectos diferenciados por sexo de los estudiantes,

⁸ En el análisis de validez externa se deben tener en cuenta que la muestra final de la evaluación difieren en tamaño, materiales e infraestructura de las escuelas del universo del APM. A pesar que no se encuentran diferencias en las variables resultado (aprendizajes y aprobación), la validez externa puede verse disminuida por estas diferencias.

pobreza donde está ubicada la escuela o por características de los directores.

A nivel de docentes, se observan algunos efectos sobre las prácticas pedagógicas de los docentes en algunas estrategias de enseñanza pero sorpresivamente no se encuentran efectos sobre indicadores de organización y planificación de las sesiones, que representan dimensiones no complicadas de mejorar y que, frecuentemente, son el punto de partida de las asesorías pedagógicas. Un punto interesante es la mejora en indicadores de interacción e involucramiento entre docentes que podrían estar asociados a los grupos de inter-aprendizajes promovidos por la intervención. Así, considerando las mejoras de las prácticas pedagógicas de los docentes, así como la interacción e involucramiento entre ellos, es posible identificar los mecanismos que podrían estar asociados a las mejoras de los aprendizajes de los estudiantes. A nivel de regiones, se observa una alta heterogeneidad de los efectos del APM que a nivel agregado se debe compensar.

Finalmente, se debe tener en cuenta que se ha evaluado el primer año de la intervención, por lo que los resultados encontrados pueden considerarse de corto plazo. De acuerdo a la literatura (Glazerman et al, 2010), los efectos de las intervenciones de asesoría pedagógica necesitan un periodo de maduración para que los docentes mejoren sus prácticas pedagógicas, por lo que se esperaría tener efectos mayores en posteriores mediciones.



7. Bibliografía

- Albornoz, Facundo; María Victoria Anauati; Melina Furman; Mariana Luzuriaga; María Eugenia Podestá and Inés Taylor (2017) Training to teach science: experimental evidence from Argentina. CREDIT Research Paper.
- Allen, Joseph; Robert Pianta; Anne Gregory; Amore Mikami and Janetta Lun (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333, 1034–1037.
- Angrist, Joshua and Victor Lavy (2001) "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools". *Journal of Labor Economics*, 19(2): 343-369.
- Blazar, David and Matthew Kraft (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis*, 37, 542–566.
- Bloom, Howard; Richburg-Hayes, Lashawn and Rebeck, Alison (2007) Using Covariates to Improve Precision for Studies That Randomize Schools to Evaluate Educational Interventions. *Educational Evaluation and Policy Analysis*, 29(1): 30-59.
- Bruns Barbara; Leandro Costa and Nina Cunha (2018). Through the looking glass: Can classroom observation and coaching improve teacher performance in Brazil? *Economics of Education Review*, 64(3): 214-250.
- Bruns, Barbara and Javier Luque (2015). *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. World Bank.
- Chetty, Raj; Friedman, John and Rockoff, Jonah (2014) Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9): 2593-2632.
- Cilliers Jacobus; Brahm Fleisch, Cas Prinsloo and Stephen Taylor (2018) How to improve teaching practice? Experimental comparison of centralized training and in-classroom coaching.
- Duflo, Esther; Rachel Glennerster and Michael Kremer (2007) "Using Randomization in Development Economics Research: A Toolkit" in T. Paul Schultz, and John Strauss (eds.) *Handbook of Development Economics*, Elsevier Science Ltd.: North Holland, Vol. 4, pp. 3895-62.
- Garet, Michael; Stephanie Cronen; Marian Eaton; Anja Kurki; Meredith Ludwig; Wehmah Jones; Howard Bloom; Fred Doolittle; Pei Zhu; Laura Szejnberg and Marsha Silverberg (2008). *The impact of two professional development interventions on early reading instruction and achievement (NCEE 2008-4030)*. Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Glazerman, Steven; Eric Isenberg; Sarah Dolfin; Martha Bleeker; Amy Johnson; Mary Grider and Matthew Jacobus (2010) *Impacts of comprehensive teacher induction: Final Results from a randomized controlled study*.
- Glennerster, Rachel and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton University Press. Retrieved from <http://www.jstor.org/stable/j.ctt4cgd52>
- Jones, Damon; David Molitor and Julian Reif (2018). *What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study.* National Bureau of Economic Research Working Paper No. 24229.

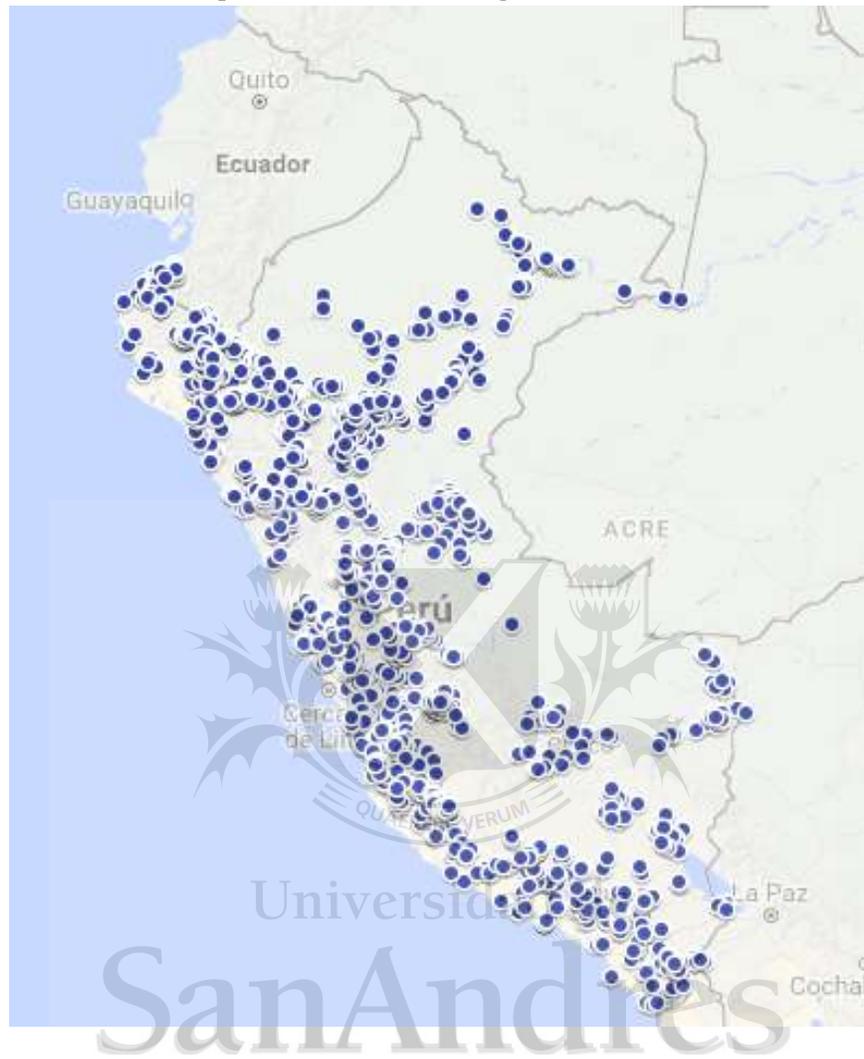
- Joyce, Bruce and Beverly Showers (1982). The coaching of teaching. *Educational Leadership*, 40(1), 4-10.
- Hanushek, Eric (2002) Teacher Quality. In Lance T. Izumi and Williamson M. Evers (ed.). pp. 1-12. Stanford: Hoover Institution Press.
- Hanushek, Eric and Steven Rivkin (2010) Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review: Papers & Proceedings* 100, 267-271.
- Kraft, Matthew; David Blazar and Dylan Hogan (2018) The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*.
- Kretlow, Allison; Nancy Cooke and Charles Wood (2011) Using In-Service and Coaching to Increase Teachers' Accurate Use of Research-Based Strategies. *Remedial and Special Education*, 33(6): 348 - 361.
- Landry, Susan; Paul Swank; Jason Anthony and Michael Assel (2011). An experimental study evaluating professional development activities within a state funded pre-kindergarten program. *Reading and Writing*, 24, 971-1010.
- McEwan, Patrick (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, 85(3), 353-394.
- Muralidharan, Karthik (2017). Field Experiments in Education in Developing Countries. *Handbook of Field Experiments*, Vol. 2. Edited by Abhijit Banerjee and Esther Duflo, Elsevier.
- Ministerio de Educación del Perú (2016). Resolución de Secretaría General N°008-2016-MINEDU. Norma que establece disposiciones para el Acompañamiento Pedagógico en la Educación Básica.
- Ministerio de Educación del Perú (2017). Informe Final. Consultoría para la evaluación del diseño e implementación de la intervención de acompañamiento pedagógico multigrado en instituciones educativas multigrado. Metis Gaia.
- Organization for Economic Cooperation and Development - OECD (2014), PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition) Student Performance in Mathematics, Reading and Science. Paris: OECD Publishing.
- Organization for Economic Cooperation and Development - OECD (2016). PISA 2015 Results (Volume 1): Excellence and Equity in Education. Paris: OECD Publishing.
- Vernon-Feagans, Lynne; Kirsten Kainz, Amy Hedrick; Mamie Ginsberg and Steve Amendum (2013) Live webcam coaching to help early elementary classroom teachers provide effective literacy instruction for Struggling Readers: The Targeted Reading Intervention. *Journal of Educational Psychology*, 105(4): 1175-1187.
- Westfall, Peter and Stanley Young (1993). Resampling-based multiple testing: Examples and methods for p-value adjustment. Vol. 279, John Wiley & Sons.
- Wolf, Sharon (2018) Impacts of Pre-Service Training and Coaching on Kindergarten Quality and Student Learning Outcomes in Ghana. *Studies in Educational Evaluation*, 59 (4): 112-123.
- Yoshikawa, H., Leyva, D., Snow, C.E., Treviño, E., Barata, M., Weiland, C., Gomez, C.J., Moreno, L., Rolla, A., D'Sa, N. y Arbour, M.C. (2015). Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Developmental Psychology*, 51(3), p.309.
- Zhang, Linxiu; Fang Lai; Xiaopeng Pang; Hongmei Yi and Scott Rozelle (2013). The impact of

teacher training on teacher and student outcomes: evidence from a randomised experiment in Beijing migrant schools. *Journal of development effectiveness*, 5(3), 339-358.



Universidad de
San Andrés

Anexo 1
Escuela que reciben APM multigrado a nivel nacional



Anexo 2

Línea de tiempo de la evaluación

La intervención inicia en febrero de 2016 con la transferencia a las regiones para la contratación de asesores pedagógicos en cada región de acuerdo al número de escuelas asignadas. El APM inicia en marzo con la contratación de los asesores, unas semanas previas al inicio de las clases en las escuelas públicas (21 de marzo).

A finales de marzo la ejecución presupuestal de la intervención era del 79% del programado; es decir, a nivel nacional solo se había contratado y realizado visitas equivalentes a 4 de cada 5 docentes. A finales de abril este ratio de ejecución presupuestal había mejorado al 89% y ya para mayo era del 100%. En julio y en agosto se realizaron los dos talleres de actualización docente aprovechando el periodo vacacional de medio año.

Finalmente, el 29 y 30 de noviembre se realizó la Evaluación Censal de Estudiantes en 2do grado de primaria y a finales de diciembre finaliza las clases escolares.



Universidad de
San Andrés

Anexo 3

Tabla A: Prueba de balance de muestra inicial

	Tratados	Controles	Nro T	Nro C	p-value
INDICADORES A NIVEL ESCUELA					
Puntaje Español 2015	533.16	534.04	1552	1070	0.73
Puntaje Matemática 2015	527.57	529.37	1552	1070	0.62
% estudiantes aprobados 2015	0.94	0.95	3796	2390	0.12
Total de estudiantes 2015	26.12	29.21	3796	2390	0.00 ***
Total de estudiantes (solo 2do grado) 2015	4.36	4.97	3796	2390	0.00 ***
Total de docentes 2015	1.74	1.94	3796	2390	0.00 ***
Ratio de estudiantes por docente 2015	14.66	14.74	3796	2390	0.72
% estudiantes con lengua nativa 2015	6.84	7.19	3796	2390	0.52
% escuelas con textos escolares 2015	0.34	0.31	3728	2351	0.12
% escuelas con cuadernos 2015	0.70	0.70	3728	2351	0.75
% escuelas con piso de material noble 2015	0.86	0.90	3796	2390	0.00 ***
% escuelas con paredes de material noble 2015	0.80	0.80	3796	2390	0.93
% escuelas con techos de material noble 2015	0.96	0.99	3796	2390	0.00 ***
% escuelas en área urbana	0.06	0.09	3796	2390	0.00 ***
% escuelas participantes AP Multigrado 2011-2015	0.32	0.31	3796	2390	0.39
% escuelas reciben otra intervención en ámbito rural	0.00	0.00	3796	2390	0.68
INDICADORES A NIVEL DISTRITO					
% hogares pobre en el distrito	63.16	62.87	3783	2378	0.41
% niños con desnutrición en el distrito 2009	27.76	27.54	3573	2249	0.43
% hogares en hacinamiento en el distrito	0.29	0.29	3748	2327	0.25
% hogares rurales en el distrito	0.79	0.78	3748	2327	0.20

Tabla B: Prueba de balance de atrición

	Tratados	Controles	Nro T	Nro C	p-value
INDICADORES A NIVEL ESCUELA					
Puntaje Español 2015	523.26	522.54	544	304	0.88
Puntaje Matemática 2015	510.66	509.31	544	304	0.85
% estudiantes aprobados 2015	0.94	0.96	2442	1381	0.11
Total de estudiantes 2015	18.31	19.42	2442	1381	0.04 **
Total de estudiantes (solo 2do grado) 2015	3.28	3.44	2442	1381	0.00 **
Total de docentes 2015	1.40	1.49	2442	1381	0.01 ***
Ratio de estudiantes por docente 2015	12.97	12.97	2442	1381	0.99
% estudiantes con lengua nativa 2015	6.42	6.73	2442	1381	0.67
% escuelas con textos escolares 2015	0.33	0.31	2389	1354	0.20
% escuelas con cuadernos 2015	0.68	0.69	2389	1354	0.76
% escuelas con piso de material noble 2015	0.83	0.87	2442	1381	0.02 **
% escuelas con paredes de material noble 2015	0.79	0.80	2442	1381	0.52
% escuelas con techos de material noble 2015	0.97	0.99	2442	1381	0.06 *
% escuelas en área urbana	0.07	0.09	2442	1381	0.00 ***
% escuelas participantes AP Multigrado 2011-2015	0.36	0.35	2442	1381	0.40

% IIEE reciben otra intervención en ámbito rural	0.00	0.00	2442	1381	0.83
INDICADORES A NIVEL DISTRITO					
% hogares pobre en el distrito	62.59	62.15	2432	1373	0.36
% niños con desnutrición en el distrito 2009	27.47	27.03	2309	1298	0.24
% hogares en hacinamiento en el distrito	0.27	0.27	2414	1345	0.77
% hogares rurales en el distrito	0.77	0.77	2414	1345	0.22

Tabla C: Prueba de balance de muestra final

	Tratados	Controles	Nro T	Nro C	p-value
INDICADORES A NIVEL ESCUELA					
Puntaje Español 2015	540.03	541.41	1008	766	0.65
Puntaje Matemática 2015	539.27	542.28	1008	766	0.48
% estudiantes aprobados 2015	0.95	0.95	1354	1009	0.69
Total de estudiantes 2015	38.91	43.47	1354	1009	0.00 ***
Total de estudiantes (solo 2do grado) 2015	6.18	7.18	1354	1009	0.00 ***
Total de docentes 2015	2.31	2.61	1354	1009	0.00 ***
Ratio de estudiantes por docente 2015	17.44	17.25	1354	1009	0.62
% estudiantes con lengua nativa 2015	7.53	7.85	1354	1009	0.67
% escuelas con textos escolares 2015	0.34	0.32	1339	997	0.35
% escuelas con cuadernos 2015	0.73	0.71	1339	997	0.28
% escuelas con piso de material noble 2015	0.92	0.94	1354	1009	0.26
% escuelas con paredes de material noble 2015	0.82	0.80	1354	1009	0.24
% escuelas con techos de material noble 2015	0.96	0.98	1354	1009	0.04 **
% escuelas en área urbana	0.05	0.08	1354	1009	0.01 **
% escuelas participantes AP Multigrado 2011-2015	0.26	0.25	1354	1009	0.51
% IIEE reciben otra intervención en ámbito rural	0.00	0.00	1354	1009	0.20
INDICADORES A NIVEL DISTRITO					
% hogares pobre en el distrito	64.04	63.94	1351	1005	0.84
% niños con desnutrición en el distrito 2009	28.22	28.30	1264	951	0.84
% hogares en hacinamiento en el distrito	0.31	0.31	1334	982	0.14
% hogares rurales en el distrito	0.80	0.80	1334	982	0.63
INDICADORES A NIVEL DIRECTIVOS					
Sexo del director (hombre=1)	0.62	0.61	1303	977	0.60
Edad del director (mayor a 45 años=1)	0.46	0.46	1343	999	0.95
Tipo de contrato (nombrado=1)	0.81	0.81	1341	995	0.90
Años de experiencia (mayor a 8 años=1)	0.38	0.36	1335	987	0.33
Máximo nivel educativo (posgrado=1)	0.26	0.27	1333	995	0.36

Tabla D: Prueba de balance entre muestra atrición y muestra final

	Tratados	Controles	Nro T	Nro C	p-value
INDICADORES A NIVEL ESCUELA					
Puntaje Español 2015	535.47	530.98	1774	848	0.12
Puntaje Matemática 2015	531.59	524.20	1774	848	0.08 *
% estudiantes aprobados 2015	0.95	0.95	2363	3823	0.15
Total de estudiantes 2015	45.29	16.55	2363	3823	0.00 ***
Total de docentes 2015	2.58	1.39	2363	3823	0.00 ***
Ratio de estudiantes por docente 2015	18.86	11.95	2363	3823	0.00 ***
% estudiantes con lengua nativa 2015	7.49	6.78	2363	3823	0.17
% escuelas con textos escolares 2015	0.33	0.31	2336	3743	0.19
% escuelas con cuadernos 2015	0.73	0.68	2336	3743	0.00 ***
% escuelas con piso de material noble 2015	0.92	0.86	2363	3823	0.00 ***
% escuelas con paredes de material noble 2015	0.82	0.79	2363	3823	0.00 ***
% escuelas con techos de material noble 2015	0.99	0.97	2363	3823	0.01 **
% escuelas en área urbana	0.13	0.05	2363	3823	0.00 ***
% escuelas participantes AP Multigrado 2011-2015	0.37	0.27	2363	3823	0.00 ***
% IIEE reciben otra intervención en ámbito rural	0.00	0.00	2363	3823	0.82
INDICADORES A NIVEL DISTRITO					
% hogares pobre en el distrito	63.73	62.47	2356	3805	0.00 ***
% niños con desnutrición en el distrito 2009	27.74	27.55	2215	3607	0.47
% hogares en hacinamiento en el distrito	0.29	0.28	2316	3759	0.24
% hogares rurales en el distrito	0.79	0.77	2316	3759	0.00 ***

Universidad de
San Andrés

Anexo 4

Ajuste por hipótesis múltiple

	Coeficiente	Error Estándar	P-value robusto	P-value Young	P-value Bonferrini	P-value Sidak
Panel A: Prácticas pedagógicas docentes						
Estrategias para la organización de la información	0.005	0.012	0.67	0.94	1.00	0.96
Enfoque comunicativo para la enseñanza de la lectura y escritura.	0.024	0.010	0.02**	0.12	0.13	0.13
Planificación curricular y diseño de sesiones de aprendizaje	0.015	0.009	0.10*	0.36	0.49	0.41
Tipos de evaluación en las diferentes áreas curriculares que enseña	0.025	0.011	0.03**	0.18	0.20	0.18
Uso de Tecnología de la Información y Comunicación para la enseñanza	0.002	0.019	0.91	0.94	1.00	0.96
Estrategias para promover un ambiente de aprendizaje adecuado	0.029	0.010	0.01***	0.06*	0.06*	0.05**
Panel B: Trabajo en conjunto e involucramiento entre docentes						
Observar las clases de otros docentes	0.074	0.022	0.00***	0.02**	0.01***	0.01***
Retroalimentación a otros docentes sobre su forma de enseñanza	0.078	0.022	0.00***	0.01***	0.01***	0.01***
Trabajar en conjunto para proyectos específicos	0.051	0.018	0.01***	0.06*	0.06*	0.06*
Intercambiar entre ellos materiales de enseñanza	0.070	0.017	0.00***	0.00***	0.00***	0.00***
Involucrarse en debates internos sobre el desarrollo de aprendizajes	0.053	0.018	0.00***	0.05**	0.04**	0.04**
Debatir sobre las prácticas de evaluación que se implementan	0.040	0.016	0.01***	0.11	0.13	0.12
Intercambiar ideas sobre el manejo de conductas inadecuadas de estudiantes	0.025	0.013	0.06*	0.30	0.36	0.31
Panel C: Motivación de los docentes						
Inasistencia por parte de los docentes	-0.008	0.018	0.65	0.94	1.00	0.96
Docentes que no preparan adecuadamente sus clases	-0.016	0.021	0.44	0.90	1.00	0.90

Anexo 5

Efectos heterogéneos por región y tamaño de muestra

En la sección 5.3 se realizó un análisis de heterogeneidad por región aprovechando el diseño aleatorio por estratos del APM. Sin embargo, una limitación que se mencionó es el reducido tamaño de muestra en algunas regiones, lo que tendría implicancias en el poder.

Para tener una idea y analizar debidamente los resultados, la tabla A detalla los tamaños de muestra agregados (por región) y desagregados (escuelas tratadas y controles) de cada región. En ella se puede observar regiones con muestra menores a 100 escuelas como Amazonas y Cusco, que tienen intervalos de confianza grandes pero también regiones que tienen más de 200 escuelas como Loreto, Huancavelica, La Libertad, San Martín y Cajamarca donde se logra identificar los principales efectos a nivel regional.

De hecho, es en estas regiones con mayor muestra, que los resultados son más estables y en todos los casos positivos, sucediendo todo lo contrario en las regiones con tamaño de muestra pequeños.

Tabla A: Efectos heterogéneos por región y tamaño de muestra

Región	Número de escuelas			Efectos en Español (DE)			Efectos en Matemáticas (DE)		
	T	C	Total	Efecto	IC inf.	IC sup.	Efecto	IC inf.	IC sup.
Amazonas	42	30	72	-0.018	-0.297	0.261	-0.050	-0.371	0.271
Cajamarca	211	155	366	0.207	0.050	0.364	0.249	0.086	0.412
Cusco	44	57	101	-0.074	-0.332	0.184	-0.040	-0.289	0.209
Huancavelica	108	151	259	0.041	-0.143	0.225	0.123	-0.069	0.315
La Libertad	123	121	244	0.133	-0.048	0.314	0.182	0.006	0.358
Lambayeque	42	56	98	0.079	-0.218	0.375	0.118	-0.159	0.394
Loreto	118	105	223	0.076	-0.142	0.293	0.030	-0.200	0.259
Piura	107	125	232	0.149	-0.078	0.376	0.134	-0.110	0.378
San Martín	35	299	334	0.057	-0.153	0.266	0.097	-0.125	0.319
Ucayali	33	60	93	0.186	-0.152	0.524	0.237	-0.143	0.617

San Andrés

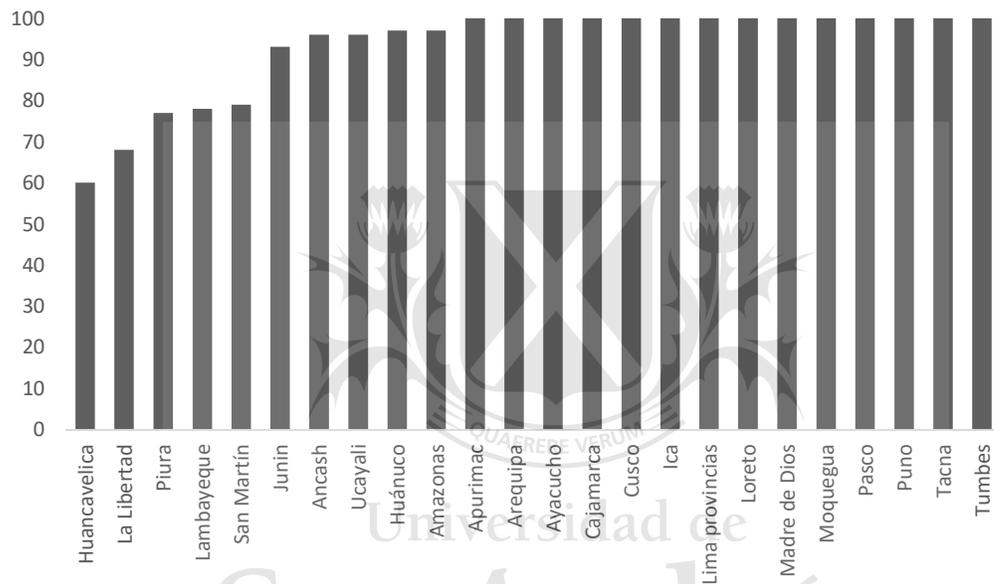
Anexo 6

Contratación de asesores pedagógicos por región

De acuerdo al estudio del Ministerio de educación (2017), el cumplimiento de metas a nivel nacional fue del 85.6%; sin embargo, al desagregar los resultados a nivel regional se observan diferencias en la implementación relacionada a la contratación de asesores pedagógicos. A mayo de 2016, del total de 24 regiones solo 14 regiones habían contratado a todos los asesores pedagógicos.

El mismo estudio detalla que para agosto de 2016, la contratación ya estaba alrededor del 99.8%, respecto al número de contrataciones de acompañantes programadas.

Gráfico A: Contratación de asesores pedagógicos por región – mayo 2016



Fuente: MINEDU (2017) – Metis Gaia