



**UNA APROXIMACIÓN DEL IMPACTO DEL SISTEMA DE ASEGURAMIENTO
DE LA CALIDAD DE LA EDUCACIÓN EN ESTABLECIMIENTOS ESCOLARES
CON BAJO DESEMPEÑO.**

**Tesis para optar al grado de
Magíster en Políticas Públicas**

**Alumna: Cristina Sepúlveda
Profesor guía: Fabián Duarte**

Santiago, diciembre de 2022



UNIVERSIDAD DE CHILE
Maestría en Políticas Públicas
Facultad de Economía y Negocios
2022

Una aproximación del impacto del Sistema de Aseguramiento de la Calidad de la Educación en establecimientos escolares con bajo desempeño.

Autor: Cristina Sepúlveda Salazar

1. Introducción

La preocupación de los gobiernos por mantener y mejorar la calidad de los sistemas educativos radica en el rol clave que tiene la educación en aspectos como la cohesión social, la equidad, el empleo, la innovación y el crecimiento de los países (European Commission, 2015).

Asimismo, al ser considerada un derecho (Naciones Unidas, 1966; Unesco, 2008), se hace esencial que todas las personas puedan acceder a ella no solo en términos de equidad sino también de calidad, sin importar el grupo socioeconómico al que se pertenezca (Medina, González y Sarzoza, 2018).

Lo anterior se debe a que, en países con altos niveles de desigualdad, los sistemas educativos a menudo contribuyen a la reproducción de los patrones socioeconómicos (Delprato, Köseleci y Antequera, 2015).

En el sistema educativo chileno, la inequidad se observa en los resultados de las pruebas estandarizadas y de ingreso a las universidades, los que están directamente relacionados con el tipo de colegio al cual asiste el estudiante, lo que a su vez está fuertemente condicionado por el nivel socioeconómico de las familias (Sanhueza, Cornejo y Leyton, 2015).

En Chile, a partir del movimiento estudiantil en 2006, la llamada “revolución pingüina”, se inició una importante reforma en el sector (Bellei y Muñoz, 2021). La principal respuesta a estas demandas se materializó en la Ley General de Educación (Ley 20.370, 2009), la cual establece que es deber del Estado propender a asegurar una educación de calidad, estableciendo las condiciones necesarias para ello y verificando permanentemente su cumplimiento. Para ello, se crea el Sistema Nacional de Aseguramiento de la Calidad de la Educación o SAC (Ley 20.529, 2011).



Para algunos, el SAC no está cumpliendo su objetivo, dado que, desde 2010, el rendimiento promedio de los estudiantes se ha mantenido estable al igual que las brechas de aprendizaje.

Sin embargo, desde otra arista, los datos muestran que de las 633 escuelas clasificadas en 2016 como Insuficientes (correspondiente al 11 % de los establecimientos), solo 155 se mantuvieron en la misma categoría en el año 2019 y las restantes avanzaron a categorías de desempeño superiores. Cabe mencionar que, dentro de esas escuelas, la mayoría corresponde a establecimientos de niveles socioeconómicos bajos, municipales y con un alto índice de vulnerabilidad escolar (González, 2018).

Dado lo anterior, surge la pregunta **¿hay una mejora desde la entrada en vigor del SAC en los establecimientos escolares con bajo desempeño?**

En Chile, a pesar de la importancia de esta política, hasta el momento solo existen incipientes estudios sobre sus efectos (Falabella y de la Vega, 2014). Asimismo, la evidencia internacional también es limitada y, además, la mayoría de estos sistemas carecen de procesos para evaluar su impacto en la práctica (European Commission, 2015).

Por otra parte, esta política no ha estado exenta de críticas, y actualmente desde el Ministerio de Educación se trabaja en su reformulación (Ministerio de Educación, 2022).

En este contexto, el presente trabajo tiene por objetivo analizar los resultados de aprendizaje de los establecimientos de educación básica clasificados con bajo desempeño en 2016 y el impacto de las visitas integrales de evaluación y orientación (en adelante, visitas) en dichos resultados.

Para ello, se analizan los datos de las pruebas SIMCE de 4° básico y se utiliza una metodología cuantitativa, no experimental y longitudinal. Específicamente, para los análisis de tendencias se utiliza el método paramétrico (T-test) y para la evaluación de impacto de las visitas, el método de dobles diferencias con propensity score matching.

Identificar si las visitas son un mecanismo de apoyo efectivo, que promueve la equidad y calidad de la educación impartida a los estudiantes de niveles socioeconómicos más bajos, resulta relevante ya que incidiría en aquellos factores relacionados a las escuelas que permiten disminuir las brechas de aprendizaje.

Finalmente, evaluar el impacto del SAC en los establecimientos con bajo desempeño busca proveer de información preliminar sobre el sistema actualmente vigente, lo que puede servir de insumo para las modificaciones que se esperan realizar en esta política o para futuras investigaciones.



2. Antecedentes

2.1. Marco general

Modelo chileno

Según la ley, el SAC debe propender a asegurar la equidad, entendida como que todos los alumnos tengan las mismas oportunidades de recibir una educación de calidad. Para ello, considera un conjunto de evaluaciones, mecanismos de apoyo, fiscalización, sistemas de información y estándares para lograr un mejoramiento continuo de los aprendizajes de los estudiantes y fomentar las capacidades de los establecimientos educacionales (Ministerio de Educación, 2013).

Un elemento central del SAC corresponde al sistema de categorización del desempeño de los establecimientos que se realiza y publica cada año de manera independiente para educación básica y media.

Los establecimientos son clasificados en cuatro categorías de desempeño: Alto, Medio, Medio-Bajo e Insuficiente, según los resultados obtenidos en las evaluaciones nacionales y en otros indicadores de calidad educativa, los que son ajustados según las características de contexto y vulnerabilidad de los estudiantes.

Además, la ley estipula que un establecimiento puede perder el reconocimiento oficial si después de cuatro años se mantiene en la categoría Insuficiente y, como una medida de apoyo, contempla la realización de visitas a los establecimientos, al menos cada dos años para aquellos con desempeño Insuficiente y al menos cada cuatro años para los de desempeño Medio-bajo.

Modelos de aseguramiento de la calidad

Los enfoques de aseguramiento de la calidad se pueden diferenciar según si priorizan la rendición de cuentas y el control o la mejora y el desarrollo, y si la responsabilidad recae en organismos que son externos a las escuelas o en las propias escuelas. La elección del enfoque por cada país depende de la función social del sistema escolar y del concepto de calidad adoptado, así como por el nivel de autonomía otorgado a las escuelas (European Commission, 2015).

Dadas las características del modelo chileno se considera que el SAC consolida un modelo de *responsabilización por desempeño*, el cual pone énfasis en mecanismos de rendición de cuentas de las escuelas hacia el Estado, asociados a consecuencias y a la entrega de apoyo según sus resultados en las pruebas estandarizadas (Falabella y



de la Vega, 2016). Este tipo de enfoque se ha desarrollado en países como Colombia, Inglaterra, Estados Unidos y Nueva Zelanda.

Sin embargo, este modelo no es el único modo de rendición de cuentas, a pesar de ser una tendencia creciente. Falabella y de la Vega (2016), definen otros dos enfoques:

- a. *Responsabilización estatal*: en el cual el Estado es responsable de la oferta educativa, por lo que el foco está puesto en el control de las condiciones de entrada al sistema, presente actualmente en países como Francia, Italia y Portugal.
- b. *Responsabilización profesional*: en el cual el equipo escolar se responsabiliza por brindar aprendizajes de calidad. Pueden existir sistemas externos de evaluación y de rendición, pero estos no están asociados a políticas de competencia, recompensas o sanciones, y los resultados no son públicos. El foco está en altos estándares y barreras de entrada para la formación docente, junto a un alto estatus de la profesión y condiciones laborales atractivas. Este modelo se ha desarrollado en Finlandia, y en países como Canadá, Escocia y Bélgica, aunque en menor medida.

Modelos de gestión escolar y visitas

Los sistemas de aseguramiento se asocian regularmente a los modelos de gestión orientados a la mejora continua (Mineduc, 2013), los que se centran en factores que se desarrollan al interior de las escuelas y que inciden para lograr buenos resultados educativos y aminorar el impacto negativo de los factores socioeconómicos (Acevedo, Valenti y Aguiñaga, 2017).

Dentro de los sistemas que evalúan los procesos de gestión escolar, se pueden identificar dos grupos (Mineduc, 2013):

- a. *Sistemas que contemplan datos de las pruebas estandarizadas junto con información de visitas inspectivas*, a partir de los cuales se categoriza a los establecimientos en distintos niveles o se plantean recomendaciones. Estos resultados son públicos y originan presión para que el establecimiento mejore o mantenga sus resultados. Además, permiten focalizar el apoyo o reestructurar las escuelas. Ejemplos de estos sistemas son Inglaterra y Holanda.
- b. *Sistemas que consideran solo las observaciones de la visita de inspección*, y tienen como resultado un informe público. El informe entrega recomendaciones y, además, incide en la entrega de apoyo o número de visitas posteriores. Estos informes no dan origen a un puntaje o a una categoría de clasificación. Ejemplos de este tipo de sistema son Escocia e Irlanda.



En Chile, las visitas son de carácter indicativo y fueron concebidas como una asesoría para la gestión escolar, por tanto, formarían parte de la dimensión de apoyo y no de la rendición de cuentas.

2.2. Evidencia empírica

Evidencia internacional

Con respecto a los modelos de evaluación de los de sistemas educativos, un estudio concluye que las evaluaciones externas estandarizadas están asociadas con mejoras en el rendimiento de los estudiantes, y que el efecto es más fuerte en países con bajo rendimiento inicial. Por el contrario, sistemas que cuentan únicamente con evaluaciones y controles internos no afectan el rendimiento de los estudiantes (Bergbauer, Hanushek & Woessmann, 2018).

En relación con las políticas de responsabilización por desempeño hay evidencia heterogénea. Algunos estudios han encontrado resultados positivos en los aprendizajes de los estudiantes, aunque leves, en países como Estados Unidos, Inglaterra, India, Pakistán, Uganda y Liberia. Sin embargo, en el caso de Estados Unidos, hay estudios que no identifican una mejora (Falabella y de la Vega, 2016).

Asimismo, un análisis comparativo de los resultados en la prueba PISA de diferentes países indica que no hay evidencia para sostener que generen mayores resultados académicos. De hecho, países como Canadá, Finlandia, Singapur y Japón se destacan por obtener un alto rendimiento académico sin implementar este tipo de políticas (Falabella y de la Vega, 2016).

Por otra parte, respecto de las visitas inspectivas, Nelson y Ehren (2014) al recopilar investigaciones del tema, concluyen que medir el impacto de las inspecciones es complejo cuando está dentro de un marco de rendición de cuentas, dado el efecto de otras variables. Destacan también factores que inciden en que las inspecciones puedan tener o no un impacto, como la aceptación de los resultados, el tipo de retroalimentación, el apoyo para la mejora y el liderazgo del director.

En cuanto al impacto de las inspecciones en el rendimiento de los estudiantes, gran parte de las investigaciones muestran un efecto leve o nulo. Los casos que han encontrado resultados positivos corresponderían a Holanda, en donde mejorarían entre un 2 y un 3 % de una desviación estándar a los dos años posteriores de la inspección, y a Inglaterra, en el que se observan mejoras que se mantuvieron durante los siguientes tres años (Nelson



& Ehren, 2014).

Evidencia nacional

En Chile, existe escasa evidencia respecto de las políticas de rendición de cuentas, debido a la dificultad de hacer inferencias causales, ya sea porque existe evidencia incompleta, o bien, porque varias políticas se han implementado en forma combinada (Bellei y Muñoz, 2021).

La principal evidencia se basa en análisis de tendencias, los que muestran un estancamiento en las evaluaciones internacionales durante la última década, tanto en los resultados de los estudiantes como en la disminución de la brecha por nivel socioeconómico (Bellei y Muñoz, 2021).

En el ámbito nacional, los resultados en las pruebas SIMCE también se han mantenido estables para el mismo periodo, y solo se observa una mejora en los resultados de Lectura 4° básico y una disminución en la brecha por nivel socioeconómico en ese mismo nivel tanto en Lectura como Matemática (de 63 a 52 y de 80 a 60 puntos, respectivamente) debido al avance del grupo socioeconómico más bajo, aunque esta diferencia sigue siendo muy significativa (Agencia de la Calidad de la Educación, 2018a).

Respecto de las visitas, solo se han desarrollado estudios sobre percepciones de los actores educativos (González, 2018; Bravo, 2020), en los que se concluye que, en general, las visitas son valoradas positivamente, sin embargo, generan algunas tensiones relacionadas a los aspectos administrativos que se deben rendir.

Asimismo, los actores consideran que los informes entregados tienen acotados usos para el mejoramiento escolar y los procesos de aprendizaje, dado que las recomendaciones no son novedosas o no se traducen en pasos o acciones concretas para avanzar en este proceso. Además, perciben que esta evaluación no reconoce los esfuerzos de la gestión y liderazgo en entornos de vulnerabilidad social, por la alta incidencia de los resultados SIMCE en la categorización del desempeño.

2.3. Tensiones políticas y técnicas

Las tensiones políticas se centran, principalmente, en el modelo de responsabilización del desempeño. Quienes están a favor consideran que estas políticas permiten movilizar a los actores del sistema hacia la mejora, mediante una doble rendición de cuentas (hacia el Estado y las familias), lo que refuerza los mecanismos del aseguramiento. Por su parte, quienes están en contra, aluden a que estas políticas se enmarcan en un modelo



neoliberal y en dinámicas de mercado, con un énfasis en la competencia por la demanda y la elección de los padres, como una forma de controlar el mercado y no para hacer más eficiente el sistema público (Bellei y Muñoz, 2021).

Asimismo, al coexistir este tipo de políticas junto con reformas para fortalecer la educación pública y contener la privatización, hace que ciertos elementos se superpongan y den lugar a incongruencias, especialmente para la educación pública que es tratada simultáneamente como un servicio controlado por el Estado y, a su vez, como proveedor de mercado (Bellei y Muñoz, 2021).

Las tensiones técnicas, por su parte, se focalizan en los componentes y supuestos del modelo, entre ellos, si la calidad se puede medir y comparar, la legitimidad de los instrumentos de evaluación estandarizados, la equidad entendida como que todos deben rendir cuentas sin excusas (según los modelos de gestión escolar), el considerar los mecanismos de diferenciación, clasificación y competencia como motor de la mejora y basarse fuertemente en la motivación extrínseca (incentivos y consecuencias) para movilizar el sistema (Falabella y de la Vega, 2016).

Estas tensiones, además son reforzadas por los estudios sobre los efectos en las prácticas internas de las escuelas, que muestran, por un lado, efectos positivos como la mejora en los procesos de planificación y monitoreo de los aprendizajes, pero por otro, efectos negativos, como la desprofesionalización docente, la reducción curricular, y la mayor segregación y exclusión socioeducativa (Falabella y Opazo, 2014; Espinola y Claro, 2010; Ravitch, 2010), así como ciertas prácticas de preparación de las evaluaciones, que en algunos casos pueden lograr una posible mejora en los resultados sin que exista un aumento real en el aprendizaje de los estudiantes (mejoras “espurias”) (Shepard, Hannaway & Baker, 2009).

Las visiones más afines reconocen los efectos no deseados, sin embargo, consideran que el beneficio de la política es, en comparación, mayor, y que se pueden corregir tomando resguardos respecto de la cantidad de incentivos y la existencia de múltiples instrumentos de evaluación para reducir la posibilidad de prepararlos. Por su parte, los detractores, contraargumentan que estos efectos no son producto de la implementación, sino que responden a un diseño global del modelo. Conducen respecto de la importancia de evaluaciones externas que retroalimenten el trabajo educativo, pero acusan un desbalance en el diseño, que exagera la prescripción y la sanción por sobre el desarrollo de habilidades internas de los equipos profesionales (Falabella y de la Vega, 2016).



3. Metodología y resultados

Para responder a la pregunta si durante la implementación del SAC ha habido una mejora en los resultados de aprendizaje de los establecimientos con bajo desempeño, se realiza, en primer lugar, un análisis sobre la evolución de los resultados de aprendizaje según la primera categorización de desempeño de 2016 y, posteriormente, una evaluación de impacto de las visitas en los resultados de los establecimientos que fueron clasificados dicho año en las categorías de desempeño más bajo.

Para ello se consideran los resultados en las pruebas SIMCE de 4° básico de Lectura y Matemática según la categoría de desempeño de los establecimientos.

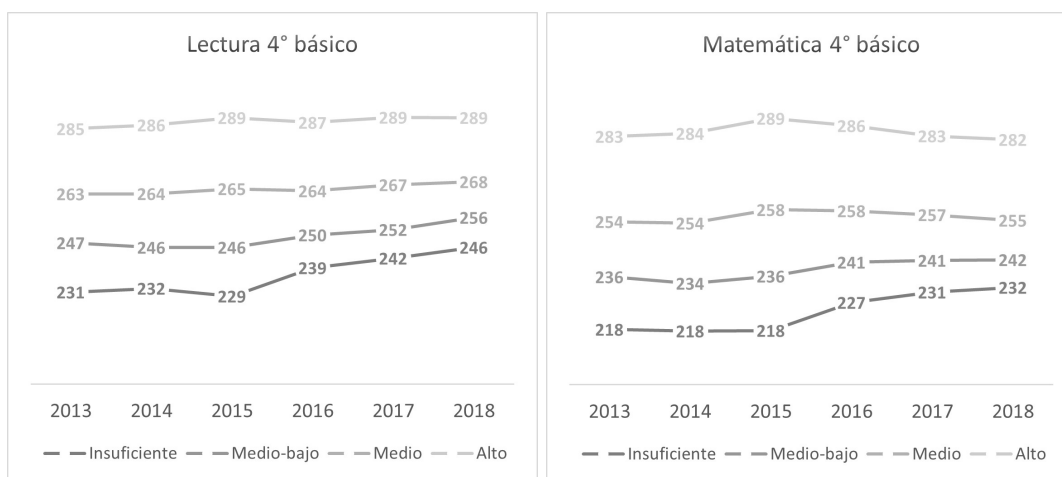
3.1. Tendencias en los resultados de aprendizaje según primera categorización de desempeño

Para realizar el análisis descriptivo de los resultados SIMCE según la clasificación de desempeño de 2016, se consideran dos variables de resultados: a) puntaje promedio SIMCE y b) distribución de estudiantes en los Estándares de Aprendizaje.

a) Puntaje promedio SIMCE

En la figura 1 se presenta la tendencia del puntaje promedio SIMCE en Lectura y Matemática 4° básico¹.

Figura 1: Tendencia puntaje promedio SIMCE según Categoría de Desempeño 2016



¹Elaboración propia a partir de bases de datos públicas de la Agencia de la Calidad de la Educación.



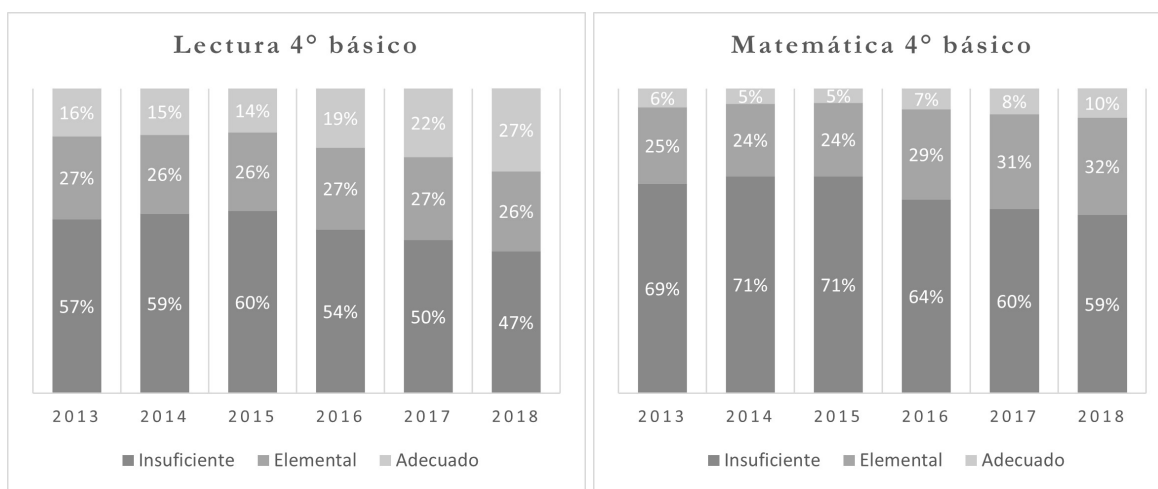
En los gráficos se observa una mejora en los establecimientos clasificados con un desempeño Insuficiente y Medio-bajo, a partir del año 2016, en ambas asignaturas.

b) Distribución de estudiantes en los Estándares de Aprendizaje

A continuación, se presentan los resultados SIMCE de los establecimientos clasificados con un desempeño Insuficiente y Medio-bajo según los Estándares de Aprendizaje, los que miden el nivel de logro de los estudiantes en estas pruebas.

En la figura 2 se presenta la evolución de la distribución de estudiantes de los establecimientos con desempeño Insuficiente en 2016² en los tres niveles de los Estándares de Aprendizajes: Insuficiente, Elemental y Adecuado.

Figura 2: Distribución de estudiantes en Niveles de Aprendizaje de establecimientos en Categoría de Desempeño Insuficiente



En los gráficos se observa una mejora desde 2016, ya que en ambas asignaturas disminuye la proporción de estudiantes con un logro Insuficiente y aumenta el porcentaje de estudiantes que alcanza el Nivel de Aprendizaje Adecuado.

Con respecto a los establecimientos que fueron clasificados en la categoría de desempeño Medio-bajo, en las dos asignaturas también hay un aumento significativo de estudiantes que alcanzan el Nivel Adecuado (9% en Lectura y 4% en Matemática) así como una disminución de aquellos que se ubican en el Nivel Insuficiente (-6% y -3%, respectivamente).

²Elaboración propia a partir de bases de datos públicas de la Agencia de la Calidad de la Educación.



3.2. Evaluación de impacto de las visitas

Para analizar si las visitas tienen relación con la mejora en los resultados de los establecimientos que en 2016 estaban en las categorías de desempeño Insuficiente y Medio-bajo, se realiza una evaluación de impacto con el método de dobles diferencias con propensity score matching (PSM).

Para establecer los grupos de tratamiento y control se identificaron a aquellos establecimientos que, siendo clasificados en 2016 con desempeño Insuficiente o Medio-bajo, recibieron visitas en 2016 y/o 2017 (grupo de tratamiento) o no recibieron ninguna visita en estos dos años (grupo de control). Además, para realizar el emparejamiento y lograr un balance adecuado entre ambos grupos se consideraron las siguientes variables: grupo socioeconómico, dependencia administrativa, región, ruralidad y resultados en las evaluaciones del año 2015³.

Por su parte, el estimador corresponde a la diferencia media entre las variables de resultado antes y después de la intervención, entre el grupo de tratamiento y de control en el soporte común, ponderada por la distribución de la probabilidad de participar en el grupo de tratados (Bernal y Peña, 2011), esto es:

$$\tau_{ATT}^{DDE} = E_{P(X)|D=1} \{E[Y_2(1) - Y_1(1)|D = 1, P(X)] - \{E[Y_2(0) - Y_1(0)|D = 0, P(X)]\}$$

Donde $E_{P(X)|D=1}$ es el valor esperado con respecto a la probabilidad de participación; el subíndice de Y indica el periodo al que corresponde la observación de la variable de resultados ($Y_1 =$ antes y $Y_2 =$ después del tratamiento); finalmente, la condición $|D$ indica si la observación corresponde a un establecimiento del grupo de tratamiento ($D = 1$) o de control ($D = 0$), (Bernal y Peña, 2011).

El estimador de dobles diferencias se calcula a partir de la diferencia media entre los resultados 2018 (último año en que se aplicaron evaluaciones) y 2015 (año previo a la primera categorización y visitas) para tres variables de resultado por asignatura:

- Puntaje promedio Simce.
- Porcentaje de estudiantes en el Nivel Insuficiente.
- Porcentaje de estudiantes en el Nivel Adecuado.

³Para mayor detalle, ver en el anexo 1 la distribución del propensity score según grupo de tratamiento y control para cada variable de resultado.



En el cuadro 1 se presenta el estimador PSM de dobles diferencias para cada una de las tres variables de resultado por asignatura.

Cuadro 1: Estimador de dobles diferencias con PSM para variables de resultado por asignatura según grupo de tratamiento y control

	Variables de resultado	Visitados		No visitados		τ_{ATT}^{DDE}	D.E.	T-stat
		N	Δ 2018-2015	N	Δ 2018-2015			
Lectura	Puntaje promedio SIMCE	712	10,4 pts.	864	12,4 pts.	-2,06 pts.	1,54	-1,33
	% estudiantes en Insuficiente	687	-7,7 %	820	-10,6 %	2,92 %	1,29	2,26 **
	% estudiantes en Adecuado	687	9,5 %	820	10,8 %	-1,33 %	1,09	-1,22
Matemática	Puntaje promedio SIMCE	735	8,1 pts.	865	12,4 pts.	-4,36 pts.	1,73	-2,52 **
	% estudiantes en Insuficiente	693	-6,8 %	822	-10,6 %	3,83 %	1,73	2,22 **
	% estudiantes en Adecuado	683	3,3 %	822	4,9 %	-1,6 %	0,94	-1,69

Como se observa en el cuadro, entre el periodo 2015-2018, ambos grupos (tratamiento y control) mejoraron sus resultados, aumentando su puntaje promedio, disminuyendo el porcentaje de estudiantes con un logro insuficiente y aumentando el porcentaje de aquellos que logran aprendizajes adecuados, tanto para Matemática como Lectura. Sin embargo, al comparar la diferencia entre grupos, se observa que estas no son significativas, y cuando las hay, los resultados son mejores para en el grupo control.

3.3. Ventajas y limitaciones

En el presente trabajo se opta por considerar los resultados de las evaluaciones de 4° básico porque se han utilizado para calcular todas las categorías de desempeño desde 2016 y a que la enseñanza básica fue el primer nivel en entrar en vigor, lo que permite tener un año más de impacto que educación media. Asimismo, el currículum para este curso no se ha modificado desde 2012, lo que permite hacer comparaciones sin las dificultades que produce un cambio curricular. Finalmente, dado el corto periodo de impacto, resulta pertinente escoger un curso inicial, donde es más probable observar algún efecto en el corto plazo que en cursos superiores, ya que, según los resultados de aprendizaje, a medida que se avanza en la trayectoria escolar se va acrecentando la brecha entre los aprendizajes esperados y los que efectivamente logran los estudiantes.



Por su parte, la selección de la metodología para la evaluación de impacto tiene la ventaja de emparejar a partir de variables observables a los participantes del grupo de tratamiento y de control con el PSM y, al usar dobles diferencias, además se puede corregir el sesgo por omisión de variables no observables, obteniendo resultados más robustos (Bernal y Peña, 2011).

Aplicar PSM con dobles diferencias responde a que las visitas no son aleatorias, ya que es la Agencia que decide cuándo y qué establecimientos visitar y, a pesar de que por recursos no es posible realizarlas en todos los establecimientos que cumplen los requisitos, es muy probable que existan diferencias sistemáticas entre el grupo de tratamiento y control que hayan incidido en una cierta priorización. Ante esto, el modelo de dobles diferencias es una forma de controlar por estas posibles diferencias preexistentes entre los dos grupos⁴ (Bernal y Peña, 2011).

Por otra parte, este estudio se basa en el supuesto que la entrada en vigor de la categoría de desempeño y las visitas fueron la única variación a partir de 2016, y las demás políticas se mantuvieron constantes, o bien, tienen una baja probabilidad de impactar en los resultados de 2018, dado que su implementación ha sido muy gradual o parcial hasta la fecha, como el caso de la ley 20.845 de Inclusión Escolar (2015) y la ley 21.040 de Nueva Educación Pública (2017).

Aun así, los resultados de este estudio deben ser considerados preliminares o como una aproximación, dados los pocos años de implementación y porque la evaluación de impacto está midiendo indistintamente el efecto a uno y dos años desde la visita, debido a que no fue posible establecer un balance adecuado entre grupo de tratamiento y control para evaluar solo el impacto luego de dos años.

Asimismo, cabe mencionar que, hay información que no se pudo considerar, como aquellos componentes del SAC que iniciaron una marcha blanca previo a su implementación, como visitas pilotos, o también herramientas que se publicaron previamente, como los distintos estándares e indicadores que están a disposición de los equipos directivos y docentes para su uso.

De igual manera, esta información resulta relevante como antecedente preliminar del SAC, aportando datos sobre su implementación. Además, los resultados de este trabajo permiten complementar los estudios existentes, así como servir de insumo para futuras investigaciones, ya sea del sistema actual o luego de incorporar modificaciones.

⁴En el anexo 2 se presenta la gráfica para cada variable de resultado que demuestra el supuesto de tendencias paralelas entre ambos grupos para que el estimador de dobles diferencias sea insesgado.



4. Análisis de resultados

Los resultados del presente estudio muestran que, en promedio, hubo una mejora en los resultados de las pruebas SIMCE de 4° básico en Lectura y Matemática, en aquellos establecimientos que en el año 2016 quedaron clasificados en la categoría de desempeño Insuficiente y Medio-bajo.

Si bien esto podría resultar contradictorio en comparación con los resultados promedio de la última década que han permanecido estables, es factible dado que las evaluaciones de desempeño también muestran que para este periodo una fracción importante de establecimientos ha empeorado significativamente sus logros, anulando los avances logrados por otros (Valenzuela, Bellei, y Allende, 2016; Bellei et al, 2020, como se citó en Bellei y Muñoz, 2021).

Además, se observa que, en cuanto a magnitud, los resultados son mejores en Lectura, sin embargo, estas diferencias se han presentado históricamente tanto en las evaluaciones nacionales como internacionales. Así, durante la década anterior hubo mejorías en esta área y no en Matemática. Del mismo modo, según los resultados de las pruebas SIMCE 2018, en 4° básico un 44,7% de los estudiantes en Lectura logra los aprendizajes adecuados, mientras que en Matemática solo un 24,6%.

La mejora en los resultados de los establecimientos con bajo desempeño también se puede relacionar con la evidencia empírica que indica que la rendición de cuentas externa aumenta el desempeño de los estudiantes, sobre todo cuando hay bajo rendimiento inicial.

A pesar de esto, no se puede asegurar que esta mejora sea sostenida en el tiempo, debido a los pocos años evaluados. Además, según estudios, los sistemas de responsabilización de desempeño pueden generar mejoras “espurias” mediante mecanismos como “enseñar para la prueba”, dada la presión que genera en los establecimientos las altas consecuencias, como la publicación de resultados o la pérdida del reconocimiento oficial, en el caso chileno.

Por su parte, la evaluación de impacto de las visitas muestran que estas tampoco se relacionan con dicha mejora, lo que se condice con la percepción que tienen los actores educativos respecto de que estas tienen usos acotados para el mejoramiento escolar, y con los resultados de la mayoría de las investigaciones internacionales que muestran un efecto leve o nulo. Considerando lo anterior, se podría inferir que las visitas en Chile se han implementado como un proceso de inspección y rendición de cuentas más que una



asesoría de apoyo para la mejora.

Sin embargo, lo anterior también puede deberse al corto periodo que se pudo medir, uno y dos años de impacto sin distinción, ya que cuando los estudios han encontrado mejorías, estas han sido a los dos y tres años posteriores a la visita.

Del mismo modo, puede haber otras herramientas incluidas en el SAC que también orienten el quehacer de los establecimientos hacia la mejora, como los distintos estándares e indicadores, dado que, como señalan los directivos en el estudio de González (2018), las recomendaciones recibidas a partir de la visita no eran novedosas, en parte, porque se referían a los Estándares Indicativos de Desempeño, ya conocidos por ellos.

Por tanto, considerando la evidencia y análisis presentados, se puede concluir que, **a pesar de la mejora estos resultados no son suficientes y aún está el desafío de seguir avanzando hacia una educación de calidad para todos.**

Según la literatura, para que el aseguramiento sea efectivo, existen factores que contribuyen a que esto se logre, como la conformación de un sistema coherente en sí y con otras políticas educativas, el desarrollo de capacidades en los actores educativos, la calidad de los maestros y el grado de autonomía otorgado a las escuelas y a las autoridades locales (European Commission, 2015).

Si analizamos estos factores, encontramos que en Chile existe un modelo de regulación híbrido, lo que lleva a un sistema contradictorio de políticas educativas (Bellei y Muñoz, 2021). Igualmente, actores técnicos y políticos, acusan un desbalance en el diseño del SAC, con un claro énfasis en la presión, en desmedro de un apoyo cercano y efectivo a los equipos escolares (Falabella y de la Vega, 2016) que permita el desarrollo de capacidades en los establecimientos.

Además, si bien las políticas de rendición de cuentas no han producido los resultados esperados, tampoco pueden explicar del todo el bajo rendimiento de la educación chilena, ya que hay otros factores adicionales como la baja calidad de la formación inicial docente, las condiciones laborales y el bajo estatus de la profesión, o bien, deficiencias institucionales, como la desarticulación o la ausencia de un nivel intermedio capaz de apoyar a las escuelas (Bellei y Muñoz, 2021).

A partir del análisis de resultados, cabe reflexionar si la articulación de las políticas educativas, el balance entre los apoyos y las exigencias, así como los mismos mecanismos de apoyo que considera el SAC, entre ellos las visitas, son suficientes y adecuados a las necesidades de los establecimientos y para lograr un mejoramiento continuo.



5. Conclusiones

El presente trabajo muestra que los resultados de los estudiantes de 4° básico de establecimientos con un desempeño Insuficiente o Medio-bajo en 2016 han presentado una mejora significativa en Lectura y Matemática, tanto en los puntajes promedio como en la movilización de estudiantes hacia mayores niveles de logro de aprendizajes evaluados por el SIMCE. Esto no se observa en establecimientos con mejor desempeño, lo que explicaría que en promedio los resultados hayan permanecido estables durante la última década.

Asimismo, se evalúa el impacto de las visitas, observándose que los establecimientos visitados no logran mejores resultados que aquellos que no las recibieron.

A partir de estos datos se puede concluir que, aunque hay un avance en los resultados de los establecimientos con bajo desempeño, no se puede asegurar que este sea sostenido en el tiempo y que las visitas integrales, como mecanismo de apoyo, no están logrando los resultados esperados, por lo que sería recomendable una revisión de su implementación, considerando los estudios de percepción de estas, así como de aquellos factores que inciden en el impacto que puedan tener, como el tipo de retroalimentación entregada.

Además se hace necesaria la articulación con otras políticas, entre ellas la de Inclusión Escolar y Nueva Educación Pública, relacionadas a factores, como la segregación escolar y el desarrollo de las capacidades a nivel local, que también influyen en los resultados de los estudiantes.

Por otra parte, y considerando el actual escenario postpandemia, resulta relevante que el SAC continúe evaluando y monitoreando los resultados de aprendizaje, sobre todo por el cierre de las escuelas y las altas tasas actuales de inasistencia, que podrían haber producido en este periodo un mayor aumento de las brechas de aprendizaje, en desmedro de aquellos estudiantes de niveles socioeconómicos más bajos.

Por ello, sería recomendable incorporar dentro del SAC indicadores que permitan ir evaluando su propio impacto, para hacer seguimiento y ajustes cuando sea pertinente, de acuerdo con las características del sistema educativo chileno.

Finalmente, es importante que el SAC se adapte a estos nuevos desafíos que le permitan garantizar calidad y equidad a todos los estudiantes, impulsar el mejoramiento continuo en los establecimientos y lograr un mayor impacto en los resultados educativos y, por ende, en el desarrollo del país.



Referencias

- [1] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2016). *Marcha blanca: Estudio de percepciones, opiniones y actitudes de sostenedores, equipos directivos y docentes sobre el Sistema de Aseguramiento de la Calidad de la Educación (SAC)*. [click](#)
- [2] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2018a). *Resultados Educativos 2018*. [click](#)
- [3] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2018b). *Hallazgos en Estudios de Calidad de la Educación (2014-2017)*. [click](#)
- [4] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2019). *Informe Técnico: Categorías de Desempeño*. [click](#)
- [5] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2021). *Visitas 2016-2019*. [Base de datos de acceso restringido, solicitada mediante Oficina de Información, Reclamos y Sugerencias, OIRS].
- [6] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2022a). *Categorías de desempeño educación básica 2016, 2017, 2018 y 2019*. [Bases de datos de acceso público]. [click](#)
- [7] AGENCIA DE LA CALIDAD DE LA EDUCACIÓN. (2022b). *SIMCE 4° básico 2013, 2014, 2015, 2016, 2017 y 2018*. [Bases de datos de acceso público]. [click](#)
- [8] ARAVENA, F. Y MADRID, R. (2021). Liderazgo en escuelas categorizadas como insuficientes en Chile: percepciones de directoras. *Revista Innovaciones Educativas*, 23(34), 113-129. [click](#)
- [9] BELLEI, C. Y MUÑOZ, G. (2021). Modelos de regulación, políticas educacionales y cambios en el sistema educativo”. Un análisis de largo plazo del caso chileno. *Journal of Educational Change*. [click](#)
- [10] BERGBAUER, A., HANUSHEK, E. & WOESSMANN, L. (2018). Testing. *National Bureau of Economic Research*, NBER Working Paper No. 24836. [click](#)
- [11] BERNAL, R. Y PEÑA, X. (2011). *Guía práctica para la evaluación de impacto*. Bogotá: Uniandes. [click](#)
- [12] BRAVO, S. (2020). Visitas de evaluación y orientación realizadas por la Agencia de Calidad de la Educación de Chile: significados otorgados por directivos de escuelas públicas. *Temps d'Educació*, (57), 267-282. [click](#)
- [13] BUCHMANN, C. & HANNUM, E. (2001). Education and stratification in developing countries. *Annual Review of Sociology*, 27(1), 77-102. [click](#)
- [14] CRISÓSTOMO, B. (2020). *Mejoramiento escolar: qué mejoran los establecimientos que avanzan hacia la mejora* (Tesis de magíster). Universidad de Chile, Santiago. [click](#)



-
- [15] DELPRATO, M., KÖSELECI, N. Y ANTEQUERA, G. (2015). Educación para todos en América Latina: evolución del impacto de la desigualdad escolar en los resultados educativos. *Revista Latinoamericana de Educación Comparada: RELEC*, 6(8), 45-75. [click](#)
- [16] ESPINOLA, V. Y CLARO, J. P. (2010). El sistema nacional de aseguramiento de la calidad: una reforma basada en estándares. En Bellei, C., Contreras, D. y Valenzuela, J. P. (Eds.) *Ecos de la revolución pingüina: Avances, debates y silencios de la reforma educacional* (pp. 51-83). Santiago, Chile: Universidad de Chile. [click](#)
- [17] EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR EDUCATION, YOUTH, SPORT AND CULTURE. (2015). *Comparative study on quality assurance in EU school education systems: policies, procedures and practices: final report*. [click](#)
- [18] FALABELLA, A. Y DE LA VEGA, L. (2016). Políticas de responsabilización por desempeño escolar: Un debate a partir de la literatura internacional y el caso chileno. *Estudios pedagógicos (Valdivia)*, 42(2), 395-413. [click](#)
- [19] FALABELLA, A. Y OPAZO, C. (2014). *Sistema de Aseguramiento de la Calidad y procesos de mejoramiento: Una mirada desde la gestión educativa INFORME FINAL*. Centro de Investigación y Desarrollo de la Educación, CIDE, Universidad Alberto Hurtado, Santiago. [click](#)
- [20] GONZÁLEZ, D. (2018). *Visitas evaluativas de la Agencia de la Calidad de la Educación* (Tesis de magíster). Proyecto FONDECYT No. 1170477. Universidad Alberto Hurtado, Santiago. [click](#)
- [21] HANUSHEK, E. & RAYMOND, M. (2004). The Effect of School Accountability Systems on the Level and Distribution of Student Achievement. *Journal of the European Economic Association*, 2(2/3), 406-415. [click](#)
- [22] LEY NO 20.370. *Diario Oficial de la República de Chile*. Santiago, Chile, 12 de septiembre de 2009. [click](#)
- [23] LEY NO 20.529. *Diario Oficial de la República de Chile*. Santiago, Chile, 27 de agosto de 2011. [click](#)
- [24] MEDINA, J., GONZÁLEZ, J. Y SARZOZA, S. (2018). Ley de aseguramiento de calidad de la educación más allá de los reportes ministeriales (2012 – 2016). *Revista de orientación educacional*, 32(61), 29-43. [click](#)
- [25] MINISTERIO DE EDUCACIÓN. (2013). *Fundamentos Estándares Indicativos de Desempeño para los Establecimientos Educacionales y sus Sostenedores*. [click](#)
- [26] MINISTERIO DE EDUCACIÓN. (2015). *Hacia un sistema completo y equilibrado de evaluación de los aprendizajes en Chile*. Informe Equipo de Tarea para la revisión del SIMCE. [click](#)
- [27] MINISTERIO DE EDUCACIÓN. (2020). *Plan de Aseguramiento de la Calidad 2020-2023*. [click](#)
-



-
- [28] MINISTERIO DE EDUCACIÓN. (2022). *Cuenta pública participativa 2022*. [click](#)
- [29] NACIONES UNIDAS. (1966). *Pacto Internacional de Derechos Económicos, Sociales y Culturales*. Asamblea General en su resolución 2200 A (XXI), 16 de diciembre. [click](#)
- [30] NELSON, R. & EHREN, M. (2014). *Review and synthesis of evidence on the (mechanisms of) impact of school inspections*. University of Twente. [click](#)
- [31] RAVITCH, D. (2010). *The death and life of the great American school system. How testing and choice are undermining education*. New York, NY: Basic Books. [click](#)
- [32] SANHUEZA, F., CORNEJO, P. Y LEYTON, J. L. (2015). La educación como agente de movilidad social. *Convergencia Educativa*, (6), 95-106. [click](#)
- [33] SHEPARD, L., HANNAWAY, J. & BAKER, E. (2009). *Standards, Assessments, and Accountability*. Education Policy White Paper, National Academy of Education, January. [click](#)
- [34] UNESCO. (2008). *Un enfoque de la educación para todos basado en los derechos humanos: marco para hacer realidad el derecho de los niños a la educación y los derechos en la educación*. [click](#)
- [35] WOESSMANN, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. *Journal of Economic Perspectives*, 30(3), 3-32. [click](#)



Anexo I. Distribución propensity score entre tratados y controles por variables de resultado

Figura 3: Variable puntaje promedio SIMCE Lectura

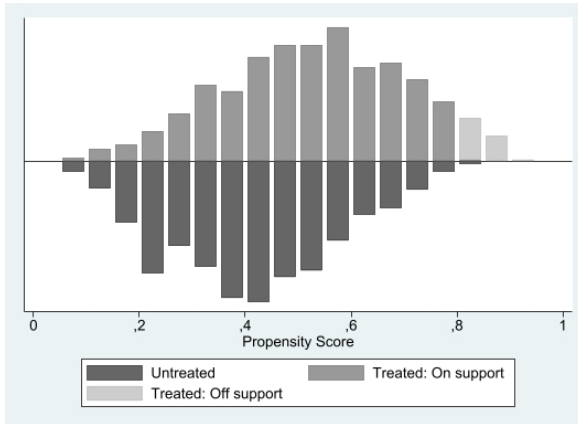


Figura 4: Variables porcentaje de estudiantes en niveles Insuficiente y Adecuado, Lectura

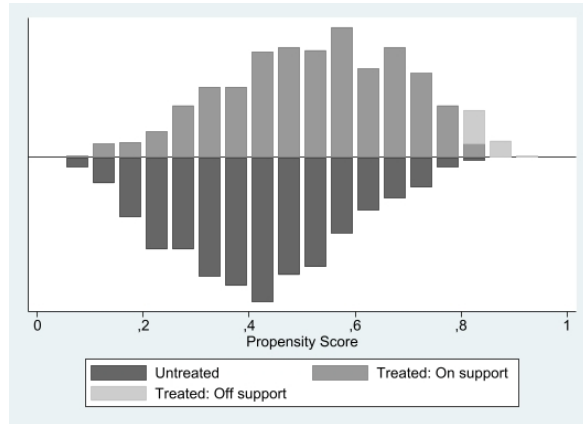


Figura 5: Variable puntaje promedio SIMCE Matemática

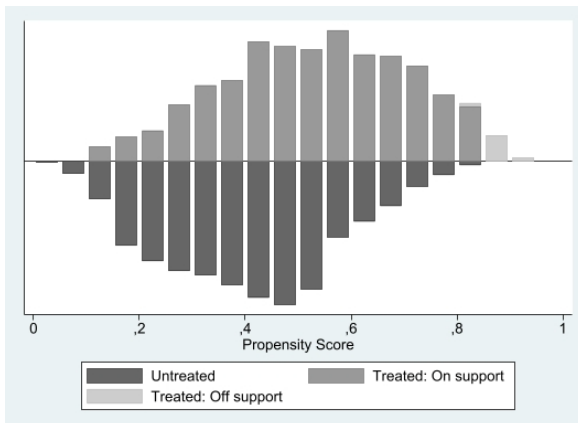
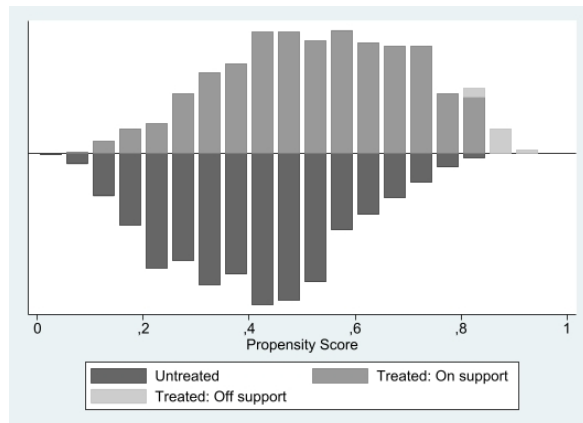


Figura 6: Variables porcentaje de estudiantes en niveles Insuficiente y Adecuado, Matemática





Anexo II. Supuesto de tendencias paralelas: evolución de variables de resultados antes y después del tratamiento entre tratados y controles

Figura 7: Evolución variables de resultado Lectura

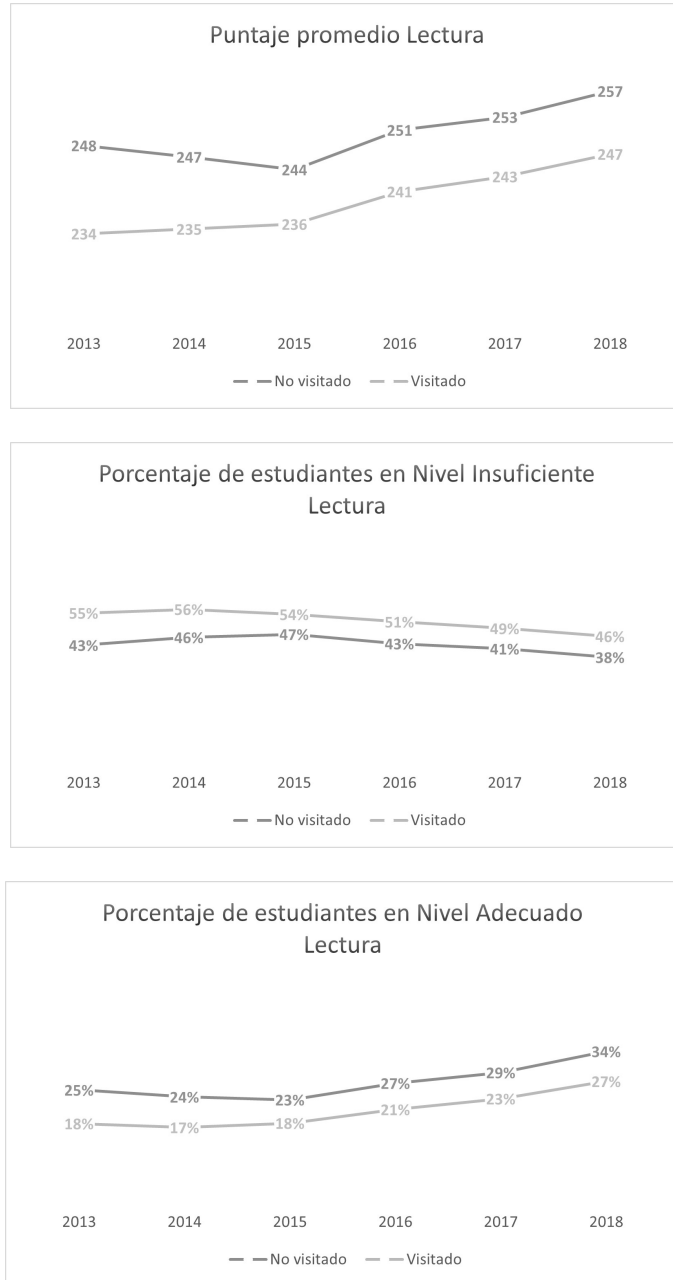




Figura 8: Evolución variables de resultado Matemática

