

DOCUMENTOS DE TRABAJO

Saber Investigar

Instituto Colombiano para la Evaluación de la Educación

N° 8

**¿Existe una ventaja al repetir el examen?
La relación entre la presentación del
examen de ensayo Pre Saber y los
resultados en Saber 11**

María Camila Estacio
Laura Rodríguez
Alejandro Corrales
Sergio Arango

Septiembre de 2023

Serie Documentos de Trabajo Saber Investigar

Edición No. 8
Septiembre de 2023
Edición digital
ISSN: 2954-6583

Instituto Colombiano para la Evaluación de la Educación

Oficina de Gestión de Proyectos de Investigación
Calle 26 N. 69-76, Edificio Elemento, Torre II, piso 18, Bogotá,
D. C.

Teléfono: (601) 4841410
oficinadeinvestigaciones@icfes.gov.co
<https://www.icfes.gov.co>

Autor(a/es)

María Camila Estacio
Laura Rodríguez
Alejandro Corrales
Sergio Arango

Directora General

Elizabeth Blandón Bermúdez

Jefe de Oficina De Gestión de Proyectos de Investigación (E)

Jeanit Camilo Cortés Mora

Advertencia:

El contenido de este documento es el resultado de investigaciones y obras protegidas por la legislación nacional e internacional. No se autoriza su reproducción, utilización ni explotación a NINGÚN tercero. Solo se autoriza su uso para fines exclusivamente académicos. Esta información no podrá ser alterada, modificada o enmendada.

Citar este documento en estilo APA así:

Estacio, M., Rodríguez, L., Corrales, A. y Arango, S. (2023). *¿Existe una ventaja al repetir el examen? La relación entre la presentación del examen de ensayo Pre Saber y los resultados en Saber 11.* (Documentos de Trabajo Saber Investigar No. 8). Instituto Colombiano para la Evaluación de la Educación (Icfes). <https://www.icfes.gov.co/web/guest/saber-investigar>

¿Existe una ventaja al repetir el examen? La relación entre la presentación del examen de ensayo Pre Saber y los resultados en Saber 11¹

María Camila Estacio²

Laura Rodríguez³

Alejandro Corrales⁴

Sergio Arango⁵

Resumen

En este estudio se estima el efecto de la repetición de exámenes sobre los resultados de los estudiantes en pruebas estandarizadas de alto impacto. Para hacerlo, se estima un modelo de efectos fijos a nivel de colegio y aplicación con información de los estudiantes que presentaron el examen Pre Saber, una prueba de preparación para el examen Saber 11, y datos de la población que tomó este último entre 2015 y 2018. Los resultados indican que, en promedio, presentar el examen de ensayo incrementó los puntajes en el examen Saber 11 en 0.172 desviaciones estándar (d.e.) en el puntaje global, 0.177 d.e. en matemáticas y 0.143 d.e. en lectura. Estas ganancias están explicadas, en gran parte, por una ganancia en la destreza a la hora de realizar el examen en lugar de una mejora en los aprendizajes de los estudiantes. La magnitud de las asociaciones encontradas indica que solo una pequeña parte de la brecha socioeconómica en puntajes puede explicarse por este tipo de preparación.

Palabras claves: Repetición de exámenes, preparación de exámenes, educación sombra, Saber 11, Pre Saber

¹ Las ideas, opiniones, tesis y argumentos expresados son de propiedad exclusiva de los autores y no representan el punto de vista del Icfes.

² Instituto Colombiano para la Evaluación de la Educación (Icfes); mestacio@icfes.gov.co

³ Instituto Colombiano para la Evaluación de la Educación (Icfes); lrodriguez@icfes.gov.co

⁴ Instituto Colombiano para la Evaluación de la Educación (Icfes); acorrales@icfes.gov.co

⁵ Instituto Colombiano para la Evaluación de la Educación (Icfes); saarango@icfes.gov.co



Is there an advantage to repeating the exam? The relationship between the presentation of the Pre Saber essay exam and the results in Saber 11

*María Camila Estacio
Laura Rodríguez
Alejandro Corrales
Sergio Arango*

Abstract

In this study, we provide evidence on the effect of exam retaking on students' scores in High-stakes standardized tests. To conduct this analysis, we estimated a fixed-effects model at school and application-term level using administrative data of students who took Pre Saber, a preparation test to take the Saber 11 exam, and information of the universe of students that took the latter between 2015 and 2018. The results indicate that taking Pre Saber increased Saber 11 global score in 0.172 standard deviations (s.d.), 0.177 s.d. in math and 0.143 s.d. in reading. We also provide evidence suggesting that these earnings are mostly explained by familiarity or learning-based gains instead of an increase in student knowledge. The magnitude of estimations also suggests that only a small share of the SES gap in test scores could be explained by exam retaking.

Keywords: Exam Retaking, Test Preparation, Shadow Education, Saber 11, Pre Saber



Contenido

1. Introducción.....	1
2. La efectividad de la preparación de exámenes de alto impacto	3
3. Contexto.....	6
4. Datos y métodos	8
4.1. Muestra	8
4.2. Estrategia de identificación.....	13
5. Resultados.....	16
5.1. Efectos heterogéneos	20
6. Conclusiones.....	24
7. Referencias	26
8. Anexos.....	28



1. Introducción

La efectividad de las actividades preparatorias para exámenes de “alto impacto” se ha convertido en un tema de interés para investigadores y hacedores de política pública por la amplia demanda existente de estas actividades en diferentes países del mundo (Domingue & Briggs, 2009; Cliffordson, 2004; Gómez Soler, Bernal Nisperuza, & Herrera Idárraga, 2020) y por su relevancia a la hora de explicar la brecha socioeconómica en puntajes (Appelrouth, Zabrocky, & Moore, 2017; Clotfelter & Vidgor, 2003), teniendo en cuenta que quienes acceden a estas actividades son, con mayor frecuencia, los estudiantes de entornos socioeconómicos más aventajados.

La existencia o ampliación de esta brecha por el acceso a actividades preparatorias podría tener consecuencias negativas en los estudiantes más vulnerables debido a la importancia que tienen este tipo de exámenes a la hora de determinar, entre otros, quienes acceden a un programa o institución de educación superior (Leeds, 2012; Montgomery & Lilly, 2012; Goodman, Gurantz, & Smith, 2020), quienes son beneficiarios de becas escolares y en el largo plazo, otros resultados relevantes en el mercado laboral (Chetty et al., 2011; Buchmann, Condron, & Roscigno, 2010). Estas razones han llevado a que distintos autores, la mayoría de ellos en países desarrollados, hayan cuantificado la efectividad de diferentes tipos de preparación sobre los resultados de los estudiantes en pruebas estandarizadas de alto impacto sin que hayan llegado a un consenso sobre si la preparación en sí es efectiva y tampoco sobre los tipos de preparación más efectivos.

En Latinoamérica, pocos esfuerzos se han hecho por cuantificar el efecto que tienen los distintos tipos de preparación sobre este tipo de exámenes y por indagar qué tan relevantes son a la hora de explicar la brecha socioeconómica en puntajes. Aunque el trabajo de Gómez Soler, Bernal Nisperuza, & Herrera Idárraga (2020) es uno de los pioneros en la región en estimar el efecto que tienen las actividades preparatorias sobre los resultados en el examen de salida de la educación media en Colombia, Saber 11, hasta donde se sabe, ningún trabajo en la región ha indagado por la efectividad de distintos tipos de preparación ni ha explorado

los mecanismos de transmisión de estas estrategias.

El objetivo de este trabajo es estimar la relación que tiene la repetición de exámenes, uno de los tipos de preparación que se han indagado en la literatura, sobre los resultados de los estudiantes en el examen Saber 11. Para estimar la asociación de interés, se utilizan datos administrativos del Instituto Colombiano para la Evaluación de la Educación (Icfes) de los estudiantes que presentaron el examen Pre Saber, un examen de ensayo que tiene la misma estructura del examen Saber 11 y que se presenta en las mismas condiciones, y de todos los estudiantes que presentaron este último entre 2015 y 2018. Con esta información, se estima un modelo de efectos fijos a nivel de colegio y aplicación para cuantificar la relación que existe entre presentar el examen Pre Saber y el puntaje global, el de matemáticas y el de lectura crítica estandarizados en Saber 11. Además, se indagan por los mecanismos de transmisión que explican la asociación de interés y se exploran efectos heterogéneos por sexo, nivel socioeconómico del estudiante y la naturaleza y ubicación del establecimiento educativo.

Los resultados que se reportan en este trabajo indican que presentar el examen Pre Saber está asociado a un aumento de 0.172 desviaciones estándar (d.e.) en el puntaje global, 0.177 d.e. en matemáticas y 0.142 d.e. en lectura crítica. Además, teniendo como referencia los mecanismos de transmisión que se han explorado previamente en la literatura, los resultados en este trabajo indican que las asociaciones estimadas están explicadas en su mayoría por una ganancia en la destreza a la hora de realizar el examen más que una ganancia en los aprendizajes de los estudiantes.

Los resultados para distintos subgrupos poblacionales indican que no existen diferencias relevantes en la ganancia obtenida por los estudiantes de diferentes sexos. Sin embargo, se encuentra que los estudiantes de colegios oficiales y de niveles socioeconómicos más bajos son quienes se ven más beneficiados al presentar el examen Pre Saber.

El resto del artículo está organizado de la siguiente manera: en la segunda sección, se realiza una revisión de literatura sobre trabajos relacionados con el entrenamiento para la presentación de exámenes; la tercera sección, explica los datos utilizados y la estrategia



empírica que se aborda para responder la pregunta de interés; la cuarta sección, presenta los principales resultados encontrados en la investigación; y, finalmente, en la quinta sección, se presentan las conclusiones.

2. La efectividad de la preparación de exámenes de alto impacto

En la literatura, la necesidad de prepararse para una prueba toma mayor importancia cuando dicha prueba es reconocida como un “*high-stakes test*”, concepto que Popham (1987) confiere a los exámenes que tienen consecuencias importantes sobre los examinados (promoción de grado, obtención de un diploma, entre otros) o a aquellos exámenes cuyos puntajes reflejan la calidad de la institución. De acuerdo con Messick et al. (1980), el concepto de preparación es una intervención realizada específicamente para incrementar los puntajes de una prueba de alto impacto.

Existe una amplia literatura que evalúa la efectividad de la preparación de este tipo de exámenes y explora los métodos más efectivos. Los estudios de Briggs (2001) y Buchmann, Condrón, & Roscigno (2010), por ejemplo, examinan si realizar las actividades preparatorias tienen un efecto en el resultado del SAT⁶ y exploran cuáles de las actividades son más efectivas. En el estudio de Briggs (2001), se encuentra que, aunque en general las actividades preparatorias tienen efectos positivos y estadísticamente significativos en los resultados del examen, también reconoce que puede existir una amplia varianza en la efectividad de la preparación dependiendo del tipo de actividad que se realice. Mientras utilizar videos y programas de computador no demuestra ningún aumento en el puntaje de matemáticas e incluso una ligera disminución en la prueba de razonamiento verbal, tomar un curso de preparación dado por la escuela está asociado a un aumento, en promedio, de 2 y 3 puntos respectivamente, tomar cursos ofrecidos por servicios de preparación comercial aumenta estos puntajes en 13 y 17 puntos, tener un tutor privado está asociado a aumentos en 7 y 19 puntos y utilizar libros de preparación aumenta estos puntajes en 4 y 7 puntos respectivamente, en promedio.

⁶ El examen estandarizado requerido para entrar a la universidad en los Estados Unidos

Esto ha llevado a que numerosos autores busquen identificar los tipos de preparación más efectivos. Entre estos, Briggs (2001, 2002); Domingue & Briggs (2009) se enfocan en el efecto de la preparación por parte de cursos comerciales y encuentran efectos estadísticamente significativos pero modestos en tamaño al representar mejoras aproximadas entre 3 y 20 puntos en el SAT verbal, y entre 10 y 28 puntos para el SAT matemáticas. Asimismo, en los trabajos de Montgomery & Lilly (2012) se estima que los estudiantes entrenados para el SAT a partir de este mismo método de preparación obtienen una ganancia, en promedio, de 23.5 puntos en razonamiento verbal y de 32.7 puntos en matemáticas.

Por su parte, Clotfelter & Vidgor (2001) quienes también estudian los efectos de la repetición de exámenes, encuentran que con la repetición del SAT, las personas obtienen en promedio, una mejoría de 14 puntos la cual se intensifica a medida que los individuos toman más veces el examen. Asimismo, otros como el de Cliffordson (2004), con datos de las pruebas estandarizadas en Suecia, sugieren que la repetición tiene efectos alrededor de 0.53-0.57 desviaciones estándar. En ambos casos, en sus análisis establecen que un componente relativamente pequeño de estos aumentos surge de la auto selección, entre 1.5 y 3.2 puntos, dependiendo del número de repeticiones del examen de acuerdo con el estudio de Clotfelter & Vidgor (2001), y el resto refleja un verdadero efecto atribuible a alguna combinación entre aprendizaje a lo largo del tiempo y familiaridad con la prueba, los dos mecanismos que se han usado en la literatura para explicar la efectividad de la estrategia. Uno de los estudios más recientes sobre la efectividad de la repetición de exámenes (Goodman, Gurantz, & Smith, 2020) logra estimar de manera causal que los estudiantes que repiten el SAT no solo obtienen 0.15 adicionales lo cual, posteriormente, incrementa la probabilidad de que accedan a universidades con programas de cuatro años y también de mayor calidad.

Cliffordson (2004) y Frisancho, Krishna, Lychagin, & Yavas (2016) profundizan sobre los mecanismos y ambos estudios coinciden en que la efectividad de esta estrategia puede darse o por una ganancia en la destreza necesaria para presentar la prueba, por ejemplo, al ser más eficientes al momento de enfrentarse a distintas preguntas y una asignación de tiempo óptima (Miyasaka, 2000) o por el crecimiento intelectual, es decir, el incremento de habilidades cognitivas necesarias para resolver las preguntas que se proponen en el examen. Mientras el estudio de Frisancho, Krishna, Lychagin, & Yavas (2016), que usa datos de la



prueba de entrada a la universidad en Turquía, es enfático en señalar que la repetición de exámenes es efectiva en gran medida por la ganancia en destreza para desarrollar la prueba y, en especial, para los estudiantes menos aventajados, el estudio de Cliffordson muestra que si bien la repetición de la prueba genera un incremento en el puntaje de 0.27 ,0.25 y 0.27 para las tres cohortes que tomaron el examen en 1974, 1975 y 1976, respectivamente, el incremento en el puntaje que se asociaba a crecimiento intelectual era de 0.19 desviaciones estándar.

También se han explorado diferentes efectos heterogéneos de los distintos tipos de preparación de exámenes. En términos de las pruebas en donde hay mejoras más significativas, Montgomery & Lilly (2012) encuentran que en cualquier tipo de preparación, el entrenamiento parece ser más efectivo para la competencia de matemáticas que para la verbal. Con respecto al sexo de los estudiantes, Appelrouth, Zabucky, & Moore (2017) encuentran que a pesar de que las mujeres completaron significativamente más actividades preparatorias que los hombres, los aumentos de puntaje de ambos fueron casi idénticos al ser estadísticamente iguales.

Adicionalmente, Appelrouth, Zabucky, & Moore (2017) comparan la efectividad de las estrategias de preparación en colegios públicos y privados y encuentran que, aunque los estudiantes de colegios públicos tuvieron puntajes iniciales más bajos, los aumentos en el puntaje del examen, como resultado de la preparación, también son más pequeños comparados con los colegios privados.

En el contexto Latinoamericano, pocos estudios han indagado por la efectividad de la preparación de exámenes y de los distintos tipos de estrategias. Hasta donde se sabe, el trabajo de Gómez Soler, Bernal Nisperuza, & Herrera Idárraga (2020) es pionero en la región e indaga por el efecto que tiene prepararse para el examen, independientemente del método, sobre los resultados de los estudiantes en el examen Saber 11. Asimismo, explora efectos heterogéneos y encuentra que las actividades preparatorias son menos útiles para los estudiantes de bajos ingresos. Con respecto al sexo de los estudiantes, también muestra que la preparación es más beneficiosa para las mujeres en comparación con los alumnos masculinos, ya que aumenta sus puntajes globales en aproximadamente 0,75 puntos (0,017 d.e) con respecto a los hombres.

Este estudio contribuye a la literatura existente de varias maneras. En primer lugar, usa datos de un universo de más de 2 millones de estudiantes que tomaron el examen Saber 11 entre 2015 y 2018 para estimar la asociación entre la presentación del examen de ensayo Pre Saber y los resultados en el examen Saber 11. Esto permite, a su vez, presentar evidencia en la región sobre la efectividad de la repetición de exámenes y también sobre los mecanismos que explican la efectividad de la estrategia, teniendo en cuenta que es el tipo de preparación que más se asemeja a la estructura de Pre Saber y Saber 11. Por último, permite estimar de manera preliminar qué proporción de la brecha socioeconómica en puntajes puede ser explicada por la repetición de exámenes. Esto cobra relevancia teniendo en cuenta que en estudios como en el de Goodman, Gurantz, & Smith (2020) se evidencia que la disparidad en las tasas de repetición de exámenes del SAT explica no solo una parte de la brecha socioeconómica en puntajes sino también la brecha en el acceso a programas de educación superior.

3. Contexto

En el sistema educativo colombiano, los estudiantes deben atravesar tres etapas antes de llegar a la educación superior, la cual abarca toda la educación postsecundaria formal universitaria y no universitaria. En primer lugar, deben pasar por la educación inicial que incluye servicios de atención integral a niños de hasta 6 años, en segundo lugar, la educación básica, dirigida a niños entre 6 a 14 años y finalmente, la educación media, que tiene una duración de dos años y cobija a los jóvenes de 15 y 16 años(OCDE, 2016).

Los estudiantes que están por finalizar el ciclo de educación media (grado undécimo) deben tomar un examen nacional para obtener resultados oficiales que les permita ingresar a la educación superior. Este examen, Saber 11, es un instrumento estandarizado que hace parte de los instrumentos que conforman el Sistema Nacional de Evaluación Estandarizada, por tanto, no solo apoya el proceso de selección de las Instituciones de Educación Superior (IES), sino también brinda información a los estudiantes sobre su desempeño y provee herramientas a los establecimientos educativos y los entes de control para monitorear la calidad de la educación (Icfes, 2017). Los resultados obtenidos en esta prueba también son requisitos para el alcance de becas, la consecución de créditos educativos, entre otros. Esto permite que un

alto porcentaje de los estudiantes (98,05% en el año 2016)⁷ que están por finalizar este ciclo educativo efectivamente tomen el examen.

El examen Saber 11 es diseñado, aplicado y calificado por el Instituto Colombiano para la Evaluación de la Educación (Icfes), una entidad de orden nacional e independiente del Ministerio de Educación Nacional (MEN). El examen se realiza dos veces en el año, en la aplicación del primer semestre se presentan alrededor de 20.000 estudiantes provenientes, principalmente, de colegios privados y en la aplicación del segundo semestre, más de 500.000 estudiantes de colegios oficiales y privados toman el examen, los cuales, en conjunto, representan la población que está por finalizar el ciclo de educación media.

Desde el año 2010, el Icfes puso a disposición de los estudiantes de grado décimo o demás colombianos interesados -y no inscritos en el examen Saber 11 o de validación de bachillerato- el examen Pre Saber. Este examen es una prueba de ensayo que les permite a los estudiantes conocer cómo están frente al examen Saber 11 y familiarizarse con el cuadernillo, tipos de preguntas, tiempos de respuesta, jornadas de presentación, desplazamientos y logística en general del examen Saber 11.

Los estudiantes que toman el examen Pre Saber se enfrentan a una prueba que se estructura bajo la misma metodología que el examen Saber 11 y toman el examen en los mismos días y sesiones. Asimismo, el examen Pre Saber se compone de las mismas cinco pruebas: lectura crítica, matemáticas, ciencias naturales, sociales y ciudadanas, e inglés. Aunque el examen Pre Saber tiene características similares al examen Saber 11, no lo sustituye, es decir, no arroja resultados oficiales que las IES puedan usar en sus procesos de admisión (Resolución 558, 2010).

Los estudiantes que presentan el examen Saber 11 deben pagar un precio dependiendo del colegio de donde provienen. De acuerdo con las tarifas en 2019, los estudiantes que provienen de colegios oficiales o colegios privados rango I deben pagar \$45.000 y quienes

⁷ Los cálculos se realizaron tomando como denominador al total de estudiantes de grado undécimo registrados en la base de datos del Sistema Integrado de Matrícula (SIMAT) y como numerador al total de personas que presentaron la prueba Saber 11 en calidad de estudiantes, ambos datos teniendo en cuenta a establecimientos educativos oficiales y no oficiales.

proviene de colegios privados rango II deben pagar \$68.500⁸. A diferencia del examen Saber 11, los estudiantes que desean presentar el examen de ensayo Pre Saber deben pagar \$60.000 pesos sin importar el colegio de procedencia. Cabe mencionar que, aunque el costo del examen de Pre Saber es una restricción para la preparación del examen Saber 11, especialmente para los más vulnerables, entre el 2014 y el 2018, 126.663 estudiantes de todos los estratos socioeconómicos han realizado al menos una vez esta prueba.

4. Datos y métodos

4.1. Muestra

Párrafo El Icfes posee información de todos los estudiantes que toman el examen Saber 11. Esta incluye, sus puntajes en el examen, su información personal (sexo, fecha de nacimiento, etc.), información socioeconómica, características de los colegios de donde provienen, entre otros. Igualmente, cuenta con las identificaciones de los estudiantes que tomaron el examen Pre Saber. Esto permite cruzar la información de todos los estudiantes que en algún momento tomaron el examen de ensayo Pre Saber con sus resultados posteriores en el examen Saber 11 y el resto de información recolectada.

Usando datos administrativos del Icfes, tomados de estudiantes que presentaron el examen Pre Saber entre los periodos de 2014-2 a 2018-1⁹, y de toda la población que tomó el examen Saber 11 entre los periodos desde 2015-1 hasta 2018-2¹⁰, se cruzó la información

⁸ Los rangos para los colegios privados se definen de acuerdo con el valor de la pensión de los estudiantes. Los colegios privados en el rango I son aquellos que cobran una pensión menor a \$98.000 pesos y los colegios privados rango II son los que cobran una pensión por encima de este valor.

⁹ La notación hace referencia a los periodos en los que se presenta el examen Saber 11. En este caso, 2014 hace referencia al año de presentación y el “2” hace referencia a que es la presentación que se realiza en el segundo semestre del año.

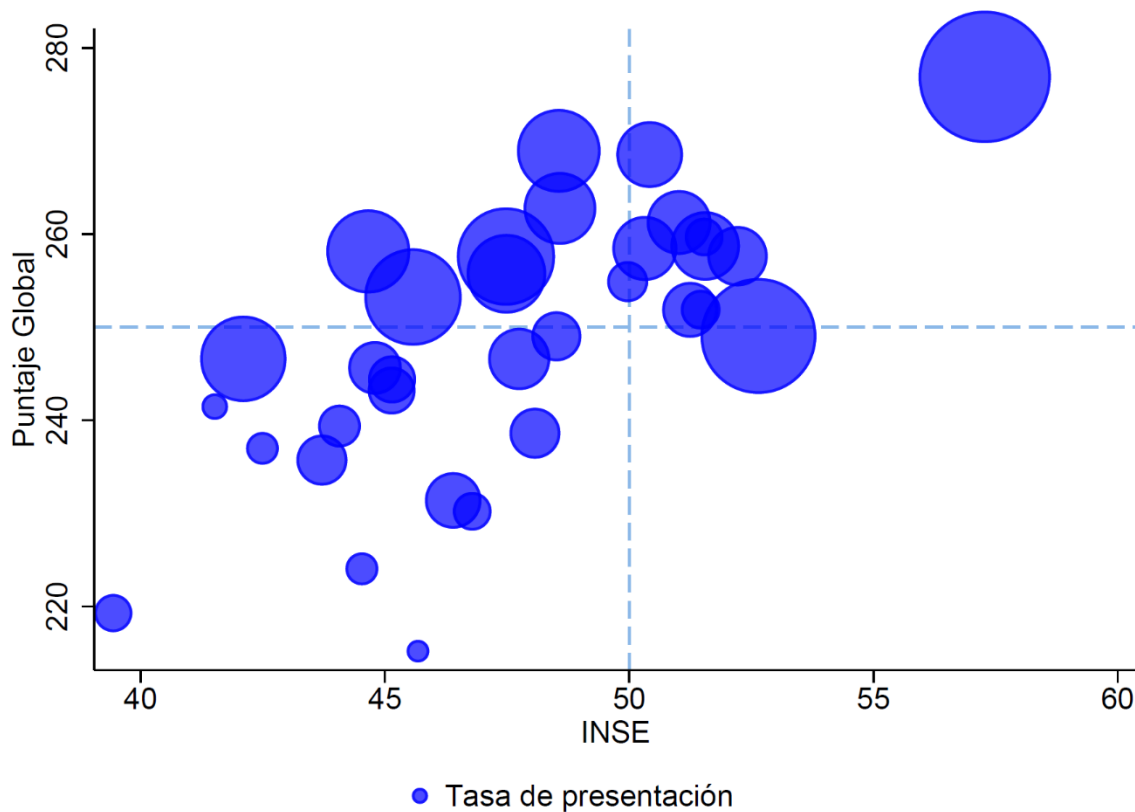
¹⁰ Se usan estos periodos debido a que el examen Saber 11 tuvo un cambio en su estructura desde el segundo semestre del 2014 con el fin de ser comparable con otras pruebas del Sistema Nacional de Evaluación Estandarizada (SNEE) (Icfes, 2013)

de los estudiantes que tomaron el examen de ensayo Pre Saber con sus posteriores resultados en el examen Saber 11 y su información sociodemográfica recopilada durante la aplicación del examen de Estado. Esto permite recolectar información de 2'405.782 estudiantes que presentaron Saber 11, entre los que se encuentran 126.663 estudiantes que previamente tomaron Pre Saber.

Debido al costo que tiene presentar el examen de ensayo Pre Saber, solo algunos estudiantes deciden tomar esta prueba de ensayo, los cuales generalmente provienen de entornos socioeconómicos más aventajados y de departamentos con menores condiciones de vulnerabilidad. El gráfico 1 muestra la relación entre el nivel socioeconómico (INSE) del estudiante promedio¹¹, a nivel departamental, la proporción de estudiantes que presentan Pre Saber sobre la población total de estudiantes que tomaron Saber 11 en el departamento (tasa de presentación) y el puntaje global promedio en el examen Saber 11 a nivel departamental. En este se puede identificar que en aquellos departamentos donde hay menor participación en la prueba Pre Saber, se presentan niveles socioeconómicos y puntajes globales más bajos que en los departamentos que tienen una mayor tasa de participación en la prueba. De modo que, el gráfico permite mostrar que la presentación del examen no es exógena, pues la mayoría de los que presentan Pre Saber están ubicados en zonas de mayor nivel socioeconómico y donde se evidencian mejores resultados en el examen Saber 11.

¹¹ Se usa información de un Índice que mide el nivel socioeconómico de los estudiantes que toman el examen Saber 11 (INSE) el cual es construido a partir de un cuestionario que es aplicado a los examinados durante la realización del examen. Para más información sobre la construcción del Índice y una descripción cualitativa de los niveles socioeconómicos que se construyen a partir de este se sugiere revisar Icfes (2019)

Gráfico 1. Participación en Pre Saber y puntaje global de Saber 11 a nivel de departamento



Fuente: elaboración propia

En la **Tabla 1** se presentan características de la población que toma Saber 11 junto con las diferencias entre los estudiantes que previamente tomaron Pre Saber y los que no. En el panel A se observan las diferencias en las medias de las características observables de los individuos y en el panel B, diferencias en las características de los establecimientos educativos de donde provienen ambos tipos de estudiantes.

Tabla 1. Diferencia de medias Pre Saber vs. No Pre Saber

Variable	Total	Pre Saber	No Pre Saber	
	(1)	(2)	(3)	(2)-(3)
Panel A: Características del estudiante				
Sexo	0.544 (0.498)	0.552 (0.497)	0.543 (0.498)	0.009*** (0.001)
Edad	17.671	17.157	17.700	- 0.543***
INSE estandarizado	-2.311	-1.274	-2.352	(0.007)
	-0.000	0.749	-0.040	0.789***
	-1.000	-1.158	(0.975)	(0.003)
Prop. madres con Ed Terciaria	0.265 (0.441)	0.535 (0.499)	0.250 (0.433)	0.285*** (0.001)
Prop. padres con Ed Terciaria	0.220	0.470	0.206	0.264***
Panel B: Características del colegio				
	(0.414)	(0.499)	(0.404)	(0.001)
Colegio oficial==1	0.714 (0.452)	0.460 (0.498)	0.728 (0.445)	- 0.267*** (0.001)
Jornada Única o completa==1	0.240 (0.427)	0.439 (0.496)	0.228 (0.420)	0.211*** (0.001)
Colegio calendario A==1	0.960 (0.195)	0.824 (0.380)	0.968 (0.176)	- 0.144*** (0.001)
Colegio en área rural==1	0.144 (0.351)	0.113 (0.317)	0.145 (0.353)	- 0.032*** (0.001)
N	2,405,782	126,663	2,279,119	

Nota: La columna (4) contiene diferencias de medias y su significancia estadística usando una prueba t asumiendo varianzas desiguales.

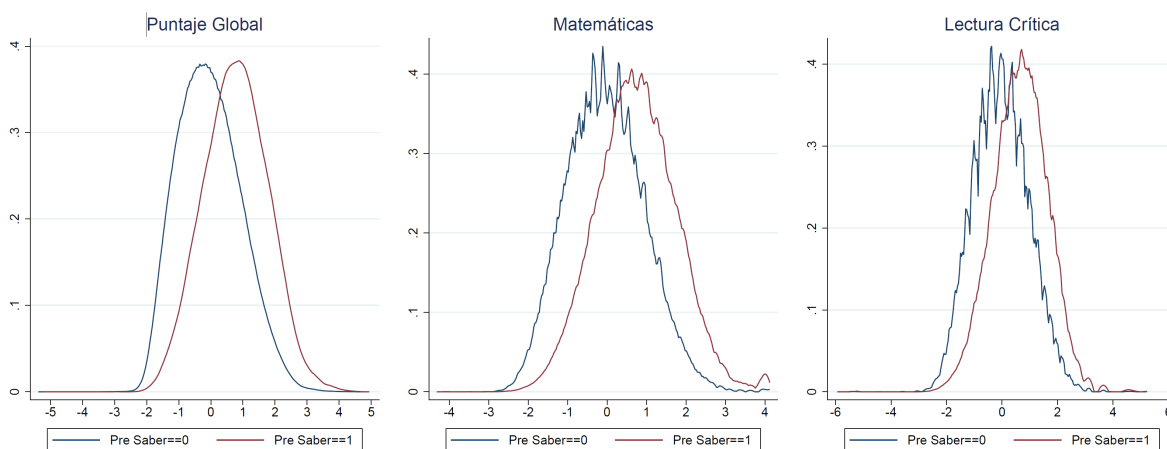
*** Significativo al 1%, ** Significativo al 5% * Significativo al 10%

En general, las personas que presentaron Pre Saber tienen un Índice de Nivel Socioeconómico (INSE) casi 0.8 desviaciones estándar más alto en comparación con los que no presentaron el examen de ensayo. Esto indica que el grupo de estudiantes que se prepara para el examen Saber 11, a través del examen Pre Saber, proviene de un entorno más favorable en términos socioeconómicos. Este resultado se evidencia también en las diferencias existentes en el nivel educativo de los padres de quienes presentan el examen de ensayo y los que no. Mientras la proporción de madres y padres con educación superior es de 53.5% y 47 %, respectivamente, para el grupo de estudiantes que presentaron el examen Pre Saber, solo el 25% de las madres y el 20% de los padres en el grupo de los estudiantes que no tomaron el examen, poseen educación terciaria.

Existen también diferencias considerables en las características de los establecimientos de donde provienen ambos grupos de estudiantes. En el panel B, se evidencia que los estudiantes que no toman el examen de ensayo provienen, con mayor frecuencia, de colegios públicos, colegios localizados en zonas rurales o de calendario A y, con menor frecuencia, provienen de colegios que tienen Jornada Única o completa.

Estas diferencias, tanto en características observables de los individuos como en las características de los establecimientos de donde provienen los estudiantes, explican una proporción importante de las diferencias que se observan en los puntajes que obtienen los estudiantes cuando toman el examen Saber 11. El **Gráfico 2** contiene las distribuciones del puntaje global y los puntajes en matemáticas y lectura crítica en el examen Saber 11, todos estandarizados, para los estudiantes que, con anterioridad, tomaron el examen de ensayo Pre Saber y los que no. En esta se evidencia que los estudiantes que toman el examen Pre Saber, tienen una distribución de puntajes a la derecha de la distribución de todos los estudiantes que no toman el examen de ensayo. En promedio, los estudiantes que toman el examen de ensayo obtienen aprox. 0.8 desviaciones estándar más en el puntaje global con respecto a los que no toman el examen Pre Saber. En matemáticas y lectura crítica, las diferencias son de 0.78 y 0.68 desviaciones estándar respectivamente.

Gráfico 2. Diferencias en los puntajes de las pruebas en el examen Saber 11 entre estudiantes que tomaron el examen de ensayo Pre Saber y los que no.



Fuente: elaboración propia

4.2. Estrategia de identificación

El objetivo de este trabajo es identificar el efecto de la presentación del examen Pre Saber sobre los resultados de los estudiantes en el examen Saber 11. Comparar los resultados de estudiantes que presentaron y no presentaron la prueba supone un reto, ya que se deben considerar características observables y no observables que difieren entre estudiantes que toman o no el examen, entre las cuales se encuentran las características no observables de las instituciones educativas a las que pertenecen.

Teniendo en cuenta lo anterior, para cuantificar la relación que existe entre presentar el examen Pre Saber (D_{ikt}) y la variable dependiente Y_{ikt} , que en este caso hace referencia al puntaje global y los puntajes de matemáticas y lectura, todos estandarizados, se estima un modelo de efectos fijos a nivel de colegio y aplicación (1). Este modelo no solo permite controlar por características observables de los estudiantes sino también aislar la influencia de cualquier otro factor no observable de los colegios que pueda estar asociado a la asociación de interés. Como consecuencia, el parámetro de interés (β_1) se estima comparando estudiantes que presentaron o no el examen de ensayo que cuentan con características observables similares y también pertenecen al mismo colegio. Por otro lado, con el fin de tener en cuenta la potencial autocorrelación de las características no observables de los

estudiantes, se usan errores estándar agrupados a nivel de colegio para realizar la inferencia estadística.

$$Y_{ikt} = \beta_0 + \beta_1 D_{ikt} + \delta_1 X_{ikt} + \gamma_t + \gamma_k + \varepsilon_{ikt} \quad (1)$$

Donde, Y_{ikt} , hace referencia al puntaje global y los puntajes de matemáticas y lectura estandarizados obtenidos en Saber 11 por el estudiante i , del colegio k en la aplicación t . D_{it} , es una variable dicótoma que toma el valor de 1 si el estudiante presentó Pre Saber y 0 en caso contrario, en la aplicación t . X_{ikt} , es un vector de variables con características sociodemográficas del estudiante i y características del establecimiento educativo k al que pertenece en la aplicación t . γ_t , representa un efecto fijo por la combinación de año- semestre de aplicación γ_k , es un vector de variables binarias por cada una de las instituciones educativas. Representa un efecto fijo por institución educativa, constante en el tiempo, el cual permite controlar por características observables y no observables de las instituciones educativas que pueden afectar los resultados de los estudiantes en la prueba Saber 11.

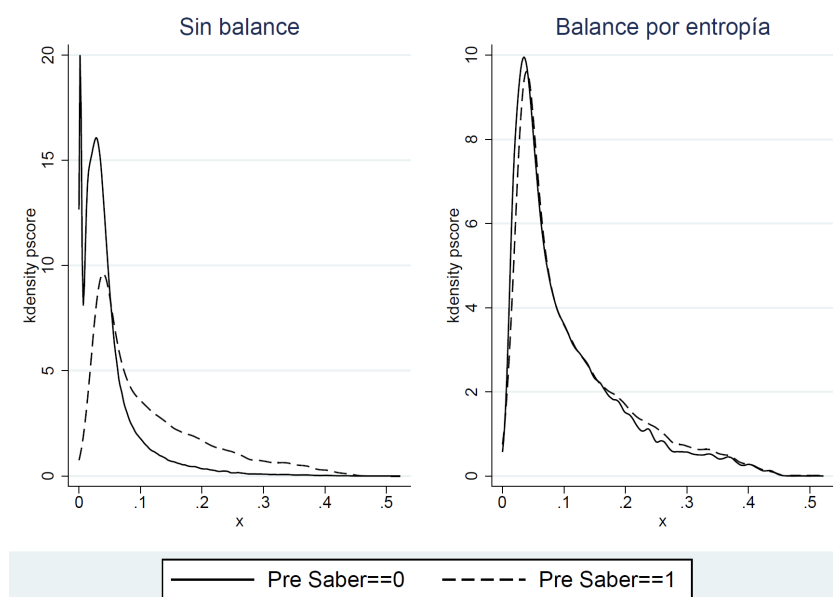
El estimador del parámetro β_1 es insesgado siempre y cuando no existan diferencias en características no observables entre los estudiantes, dentro de un mismo colegio, que tomaron o no el examen Pre Saber. En este escenario, el supuesto de identificación es difícil de cumplir debido a que probablemente quienes toman el examen tienen características no observables diferentes a los que no, incluso dentro de una misma institución. Por ejemplo, se espera que los estudiantes que tomaron el examen tengan una mayor motivación intrínseca o tengan mayores expectativas de ingresar a un programa de educación superior y, por tanto, ex ante, sean más propensos a obtener puntajes más altos en el examen Saber 11. Teniendo en cuenta lo anterior, en lugar de encontrar un estimador insesgado del efecto de interés, se busca encontrar un límite o banda superior (*Upper Bound*) del efecto de presentar Pre Saber sobre los resultados de los estudiantes en Saber 11.

Como un ejercicio de robustez, también se utilizó el método de emparejamiento porestropía propuesto por Hainmuller (2012) para estimar la asociación entre la presentación de Pre Saber y el examen Saber 11. Este método permite calcular los pesos que deben asignarse a las observaciones en el grupo control, en este caso los estudiantes que no tomaron el examen Pre Saber, con el fin de que sean comparables con las observaciones en el grupo de tratamiento en términos de sus características observables. A pesar de esta ventaja, es

relevante mencionar que este método no permite aislar las características no observables de los colegios que pueden estar asociadas al efecto de interés lo cual se logra a partir del modelo de efectos fijos a nivel de colegio presentado con anterioridad. En el gráfico 3 se presenta el soporte común antes y después del emparejamiento y se evidencia que una vez calculados los pesos para el grupo control a través del método de emparejamiento por entropía, la distribución para la probabilidad de pertenecer al grupo de tratamiento es casi la misma para los grupos de control y tratamiento.

Para la estimación de la probabilidad de pertenecer al grupo de tratamiento se usó el vector de variables X_{ikt} , el cual contiene características sociodemográficas de los estudiantes como sexo, edad, educación y ocupación de los padres, tenencia de enceres en el hogar, etc. De igual forma, se contemplan características de los establecimientos educativos a los que pertenecen los estudiantes como el tipo de jornada, naturaleza del colegio, ubicación del colegio, entre otros.

Gráfico 3. Diferencias en los puntajes de las pruebas en el examen Saber 11 entre estudiantes que tomaron el examen de ensayo Pre Saber y los que no.



Fuente: elaboración propia

5. Resultados

Los resultados presentados en la **Tabla 2** muestran la estimación de la asociación entre la presentación de Pre Saber y los resultados en Saber 11 usando el método de mínimos cuadrados ordinarios (OLS), el método de emparejamiento con balance por entropía y el modelo de efectos fijos (1). Como variables dependientes, se toman los puntajes estandarizados en matemáticas, lectura crítica y el resultado global del examen Saber 11.

Las estimaciones de la asociación de interés a partir del método de mínimos cuadrados ordinarios (OLS) y del método de emparejamiento con balance por entropía evidencian que presentar el examen Pre Saber está asociado a un aumento de 0.41 desviaciones estándar (d.e.) y 0.39 d.e. en el puntaje global del examen. En matemáticas, presentar Pre Saber está asociado a aumentos de 0.422 d.e. y 0.399 d.e. y en lectura, los aumentos son de 0.332 d.e. y 0.31 d.e. Aunque estas estimaciones aíslan, de la asociación de interés, la relación entre presentar Pre Saber y las características observables de los estudiantes y de los colegios a los que pertenecen, no permite controlar por factores no observables de los colegios.

Con el modelo de efectos fijos a nivel de colegio y aplicación, la estimación del coeficiente asociado a la variable dummy que indica si el estudiante previamente había tomado el examen Pre Saber para las tres variables dependientes indica que, en promedio, presentar el examen de ensayo resultó en un aumento en 0.172 desviaciones estándar (d.e.) en el puntaje global, 0.177 d.e. en matemáticas y 0.143 d.e. en lectura crítica. Estos coeficientes son estadísticamente significativos a un nivel de significancia del 1% una vez ajustados los errores estándar a nivel de colegio.

Los resultados presentados en la **Tabla 2** no solo son consistentes con lo encontrado previamente en la literatura sino también son similares, en magnitud, a los efectos causales encontrados por Goodman, Gurantz, & Smith (2020) quienes estiman un efecto de aprox. 0.15 d.e. La similitud entre la magnitud de las estimaciones encontradas evidencia dos puntos adicionales. Por un lado, sugiere que al usar un modelo de efectos fijos a nivel de colegio y aplicación, como el que se está usando en este trabajo, puede ser suficiente para encontrar un estimador insesgado de la relación de interés pues una vez se controla por las características observables de los estudiantes y las características observables y no observables de los colegios, la autoselección explica poco de la asociación encontrada como Clotfelter & Vidgor

(2003) señalan. Por otro lado, sugiere que una gran parte de la diferencia observada en puntajes entre los que presentan o no Pre Saber está explicada por diferencias en las características observables y no observables de los colegios en lugar de la frecuencia con la que los estudiantes acceden a los métodos de preparación para el examen.

Aunque la magnitud de la asociación encontrada en este trabajo es sustancial de acuerdo con Goodman, Gurantz, & Smith (2020), en el contexto local, explica una proporción baja de la brecha socioeconómica en puntajes. Tomando como referencia los resultados de los estudiantes en el examen Saber 11, discriminados por niveles socioeconómicos, se observa que para 2018 los estudiantes del nivel socioeconómico más bajo obtuvieron un puntaje global de -0.50 d.e. mientras que los del más alto obtuvieron, en promedio, un puntaje de 0.99 d.e. Esto implica que existe una brecha de aprox. 1.5 d.e. entre los grupos de estudiantes más y menos vulnerables y, por tanto, la repetición de exámenes, en este caso, explicaría alrededor del 11% de esta brecha.

Los resultados resumidos en la **Tabla 2** evidencian que la presentación del examen de ensayo Pre Saber está asociado a una ganancia, sobre los resultados del examen Saber 11, una vez se ha descontado la relación que existe entre la presentación del examen y las características observables de los estudiantes y las características observables y no observables de los establecimientos educativos. Esta ganancia puede darse, de acuerdo con Cliffordson (2004) debido a dos canales principales. Por un lado, el crecimiento intelectual, el cual hace referencia al incremento en las habilidades cognitivas para resolver las preguntas propuestas en el examen y por otro, la ganancia en la destreza necesaria para presentar la prueba al conocer mejor su funcionamiento y tener algunas lecciones aprendidas que influyen positivamente en su presentación.

Con el objetivo de distinguir qué proporción del efecto se debe a cada uno de los mecanismos, se toma información sobre los periodos en que tardan los estudiantes en presentar el examen Saber 11, después de presentar el examen de ensayo Pre Saber, con el fin de aislar la parte del efecto que esta explicada por la ganancia en destreza. Esto es posible bajo el supuesto de que la ganancia en puntajes que obtienen los estudiantes que presentan el examen de ensayo solo un periodo antes de rendir el examen Saber 11, se explica en su mayoría por este mecanismo.

Tabla 2. El efecto de presentar Pre-Saber sobre los resultados en Saber 11

Pre Saber==1	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica
OLS	0.414*** (0.0112)	0.422*** (0.0116)	0.332*** (0.00912)
PSM – Balance por entropía	0.390*** (0.0109)	0.399*** (0.0112)	0.314*** (0.00889)
Modelo de efectos fijos	0.172*** (0.00450)	0.177*** (0.00463)	0.143*** (0.00399)
Observaciones	2,258,277	2,258,277	2,258,277

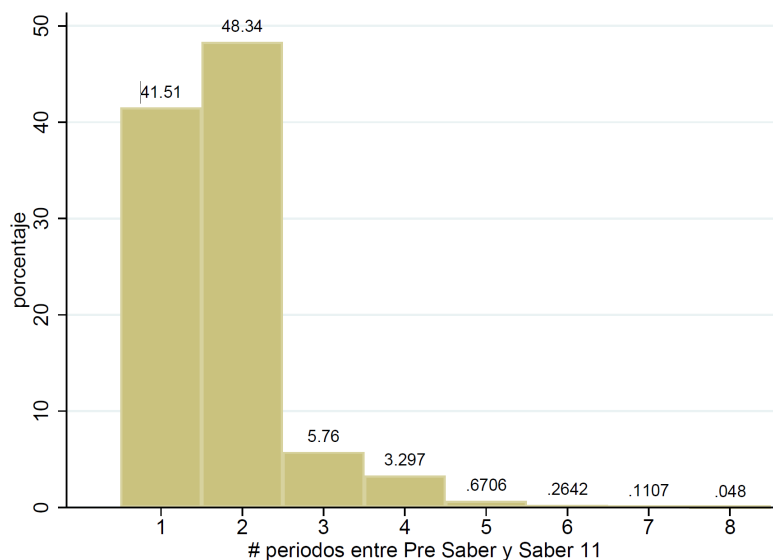
Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen.

Existen dos razones principales que justifican este supuesto. En primer lugar, los estudiantes que presentan el examen de ensayo un periodo antes de rendir el examen Saber 11 tienen solo cinco meses de preparación antes de presentar el examen de estado en condición de estudiantes y también reciben sus resultados en el examen de ensayo con menos de tres meses de anticipación. Esto deja un espacio corto de tiempo para el crecimiento intelectual antes de presentar nuevamente el examen. En segundo lugar, el examen Saber 11 evalúa no solo las competencias que son enseñadas regularmente en el ciclo de educación media, es decir los grados décimo y once, sino también evalúa las competencias que se desarrollan desde el ciclo de educación básica (Icfes, 2013), en primaria y secundaria, lo cual hace menos probable que el efecto sea explicado en su mayoría por el primero de los mecanismos.

El Gráfico 4 contiene información sobre los periodos que tardan los estudiantes en tomar Saber 11 una vez han tomado Pre Saber. En este, un periodo hace referencia a un semestre en el año, es decir, si un estudiante tomó Pre Saber en el segundo semestre de 2014 y tardó un periodo en presentar Saber 11, el estudiante presentó este último durante el primer semestre de 2015. Lo que se evidencia en la gráfica es que mientras el 41.51% de los estudiantes toman Pre Saber justo antes de tomar Saber 11, más del 50% de los estudiantes que toman Pre Saber lo hacen con 2 o más periodos de anticipación.



Gráfico 4. Porcentaje de estudiantes que toman el examen de ensayo Pre Saber uno o más periodos antes de rendir el examen Saber 11



Fuente: elaboración propia

Tabla 3. El efecto de presentar Pre-Saber un periodo antes o más

Variables	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica
Pre Saber 1 periodo antes	0.157*** (0.00559)	0.162*** (0.00573)	0.131*** (0.00516)
Pre Saber 2 periodo antes o más	0.190*** (0.00670)	0.194*** (0.00702)	0.156*** (0.00576)
Observaciones	2,258,949	2,258,949	2,258,949
R-cuadrado	0.466	0.419	0.348

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen.

Los resultados en la **Tabla 3** contienen las estimaciones para el efecto de presentar el examen de ensayo Pre Saber, un periodo antes de presentar el examen Saber 11, y el efecto para quienes lo presentan con dos o más periodos de anticipación. Los resultados de esta tabla muestran dos resultados principales. Por un lado, los estudiantes que presentan el examen de ensayo con 2 periodos de anticipación o más obtienen mayor ganancia en los puntajes del examen Saber 11 en comparación con los que toman el examen solo un periodo antes. Esto era de esperarse si quienes presentan el examen de ensayo con mayor tiempo de anticipación no se ven beneficiados solo por la ganancia en destreza sino también por el crecimiento cognitivo. Por otro lado, se evidencia que presentar el examen de ensayo Pre Saber con solo un periodo de anticipación está asociado a un incremento en el puntaje global del examen Saber 11 de 0.157 d.e. y de incrementos en las pruebas de matemáticas y lectura de 0.162 d.e. y 0.131 d.e. Estos resultados son estadísticamente significativos al 1% y al contrastarlos con los resultados de la **Tabla 2**, se evidencia que al menos el 90% del efecto asociado a la presentación del examen de ensayo puede atribuirse a la ganancia en destreza en lugar del crecimiento cognitivo.

5.1.Efectos heterogéneos

En los resultados presentados en la **Tabla 4** se exploran efectos heterogéneos por sexo de la presentación del examen de ensayo Pre Saber sobre los resultados en Saber 11. Los resultados en las columnas (1) a (3) presentan el efecto de tomar el examen de ensayo Pre Saber sobre el puntaje global del examen Saber 11 y los puntajes de matemáticas y lectura crítica para las mujeres y las columnas (4) a (5) presentan los mismos resultados para los hombres. En promedio, las mujeres que presentaron el examen Pre Saber obtuvieron 0.17 d.e. adicionales en el puntaje global del examen con respecto a las que no tomaron el examen de ensayo y obtuvieron 0.172 d.e. y 0.145 d.e. adicionales en matemáticas y lectura crítica frente al mismo grupo de comparación. Por su parte, los hombres que tomaron el examen de ensayo obtuvieron resultados similares a lo de las mujeres que también presentaron Pre Saber. Con respecto al grupo de hombres que no tomaron el examen, los efectos estimados fueron de 0.173 d.e. en el puntaje global, 0.181 d.e. en matemáticas y 0.137 d.e. en lectura crítica. Cabe

mencionar que todos los efectos estimados son estadísticamente significativos al 1%.

Tabla 4. El efecto de presentar Pre-Saber sobre los resultados en Saber 11 por sexo

Variables	Mujeres			Hombres		
	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(4) Puntaje global	(5) Matemáticas	(6) Lectura Crítica
Pre Saber==1	0.170*** (0.00498)	0.172*** (0.00510)	0.145*** (0.00463)	0.173*** (0.00529)	0.181*** (0.00549)	0.137*** (0.00474)
Observaciones	1,231,279	1,231,279	1,231,279	1,027,005	1,027,005	1,027,005
R-cuadrado	0.474	0.417	0.355	0.462	0.409	0.350

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen. La Tabla 4ª en el anexo del documento, contiene las estimaciones de la asociación de interés a partir del método de MCO y emparejamiento con balance por entropía.

Los resultados en la **Tabla 5** evidencian efectos heterogéneos por sector del establecimiento educativo. Los resultados presentados en las columnas (1) a (3) muestran los efectos de interés estimados usando únicamente la población de estudiantes en colegios oficiales mientras los resultados de las columnas (4) a (6) contienen los resultados de interés en la población de estudiantes en colegios no oficiales. Existe una diferencia relevante entre las asociaciones encontradas en estudiantes del sector oficial y el no oficial. Mientras los estudiantes del sector oficial que presentaron el examen de ensayo Pre Saber obtuvieron 0.232 d.e. en el puntaje global del examen, 0.236 d.e. en matemáticas y 0.194 d.e. en lectura por encima de quienes no tomaron el examen de ensayo, los estudiantes en el sector no oficial obtuvieron 0.107 d.e. adicionales en el puntaje global del examen, 0.11 d.e. en matemáticas y 0.087 d.e. en lectura crítica frente a los que no tomaron en el examen Pre Saber.

Los resultados de la **Tabla 6** muestran efectos heterogéneos por nivel socioeconómico¹² del estudiante. Con base en la información sociodemográfica que reportan los estudiantes durante la presentación del examen Saber 11, el Icfes clasifica a los estudiantes en cuatro

¹² Una descripción cualitativa de los cuatro niveles puede encontrarse en el Informe nacional de resultados Saber 11 2014-II - 2016-II (Icfes, 2017), además, se puede encontrar más información sobre el NSE en el Boletín 4 Saber al Detalle (Icfes, 2019).

Niveles Socioeconómicos (NSE). El Nivel Socioeconómico 1 (NSE1) hace referencia a los estudiantes más vulnerables y los estudiantes en NSE4 (NSE4) son la población menos vulnerable. Para cada uno de los cuatro niveles socioeconómicos, en esta tabla se presentan estimaciones de la asociación entre presentar Pre Saber y los resultados en el puntaje global de Saber 11 y los puntajes de matemáticas y lectura crítica.

Tabla 5. El efecto de presentar Pre-Saber sobre los resultados en Saber 11 por sector del establecimiento educativo

Variables	Colegios Oficiales			Colegios No Oficiales		
	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(4) Puntaje global	(5) Matemáticas	(6) Lectura Crítica
Pre Saber==1	0.232*** (0.00676)	0.236*** (0.00685)	0.194*** (0.00596)	0.107*** (0.00475)	0.110*** (0.00505)	0.0867*** (0.00430)
Observaciones	1,626,675	1,626,675	1,626,675	631,602	631,602	631,602
R-cuadrado	0.360	0.330	0.264	0.583	0.531	0.447

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen. La Tabla 5a en el anexo del documento, contiene las estimaciones de la asociación de interés a partir del método de MCO y emparejamiento con balance por entropía.

En general, existen diferencias considerables entre los resultados encontrados para los estudiantes de distintos niveles socioeconómicos. A primera vista se observa que entre más alto es el nivel socioeconómico del estudiante, más baja es la magnitud de la asociación entrada para la presentación del examen de ensayo. Por ejemplo, se observa que los estudiantes que presentaron el examen de ensayo Pre Saber, y que pertenecen al NSE1, obtuvieron 0.248 d.e. en el puntaje global de Saber 11, 0.242 d.e. en matemáticas y 0.212 d.e. en lectura crítica por encima de los resultados que obtuvieron quienes no presentaron el examen Pre Saber en este nivel socioeconómico. De manera análoga, para la población de estudiantes que provienen de NSE4, tomar el examen de ensayo Pre Saber representó un



aumento 0.093 d.e. en el puntaje global del examen, 0.097 d.e. en matemáticas y 0.073 d.e. en lectura crítica.

Tabla 6. El efecto de presentar Pre-Saber sobre los resultados en Saber 11 por Nivel socioeconómico del estudiante

VARIABLES	NSE 1			NSE 2			NSE 3			NSE 4		
	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica
Pre Saber	0.248*** (0.0130)	0.247*** (0.0137)	0.212*** (0.0119)	0.222*** (0.00742)	0.226*** (0.00755)	0.184*** (0.00689)	0.169*** (0.00554)	0.172*** (0.00558)	0.141*** (0.00504)	0.0929*** (0.00527)	0.0969*** (0.00570)	0.0734*** (0.00488)
Observaciones	388,625	388,625	388,625	844,025	844,025	844,025	799,592	799,592	799,592	226,035	226,035	226,035
R-cuadrado	0.335	0.311	0.230	0.337	0.315	0.238	0.349	0.324	0.246	0.433	0.399	0.292

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** p<0.01, ** p<0.05, * p<0.1. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen. La Tabla 6a en el anexo del documento, contiene las estimaciones de la asociación de interés a partir del método de MCO y emparejamiento con balance por entropía.

En la literatura también se han explorado efectos heterogéneos por sexo, nivel socioeconómico y tipo de colegio del que provienen los estudiantes. En general, los resultados encontrados previamente sugieren que el efecto de la preparación para los exámenes de alto impacto no difiere entre distintos subgrupos de la población (Appelrouth, Zabucky, & Moore, 2017). Sin embargo, los resultados presentados en las tablas 4 y 5 contrastan con lo encontrado en la literatura existente y evidencian que los estudiantes más vulnerables y que provienen de colegios oficiales obtienen un mayor beneficio de presentar el examen de ensayo Pre Saber con respecto a sus pares más aventajados en términos socioeconómicos y de colegios privados.

Una de las posibles explicaciones a este fenómeno, que se da en el contexto nacional, es el uso frecuente de métodos de preparación para el examen Saber 11 diferentes a las lecciones impartidas en los establecimientos educativos. Si bien los estudiantes de diferentes niveles socioeconómicos y diferentes tipos de colegio suelen usar estos métodos de preparación, los estudiantes que provienen de colegios privados y de entornos socioeconómicos más altos acceden a ellos con mayor frecuencia como lo establecen Gómez Soler, Bernal Nisperuza, & Herrera Idárraga (2020). Este hecho puede llevar a que las

ganancias, en términos de resultados en el examen Saber 11, sean menores para estos subgrupos siempre y cuando la mayor parte del efecto de presentar el examen de ensayo Pre Saber sea explicada por la ganancia en destreza y no por el crecimiento cognitivo de los estudiantes, pues esta ganancia en destreza es probablemente menor para los estudiantes que con anterioridad se han preparado a través de métodos que usan exámenes de ensayo similares al examen Saber 11.

6. Conclusiones

Este documento usa información del universo de estudiantes que presentaron el examen de estado Saber 11 para determinar el efecto de la repetición de exámenes, a través del examen Pre Saber, sobre los puntajes de los estudiantes en el examen de salida de la educación media. Los resultados de este trabajo evidencian que presentar el examen de ensayo Pre Saber está asociado a un aumento en el puntaje global, en matemáticas y en lectura crítica.

La magnitud de la asociación de interés varía cuando se hace análisis por subgrupos teniendo en cuenta características, tales como el sexo y nivel socioeconómico de los estudiantes o la naturaleza del colegio. A diferencia de lo encontrado en la literatura, en este trabajo los resultados difieren entre subgrupos mostrando que la ganancia más grande la obtienen los estudiantes de instituciones educativas oficiales y los estudiantes de niveles socioeconómicos más bajos.

La literatura internacional ha documentado diferentes canales a través de los cuales los diferentes tipos de preparación y en particular, la repetición de exámenes, impactan sobre los resultados en pruebas estandarizadas. Tomando información del tiempo que pasa entre la presentación de ambas pruebas como identificador de los canales, se encontró que la ganancia se da en mayor medida a través de la destreza adquirida para presentar la prueba al conocer mejor su funcionamiento y tener algunas lecciones aprendidas que influyen positivamente en su presentación, como por ejemplo el manejo eficiente del tiempo o la implementación de estrategias dependiendo del tipo de pregunta.

Finalmente, considerando que las ganancias en su mayoría se dan en términos de destreza y aunque la magnitud de la asociación encontrada en este trabajo es sustancial de acuerdo con lo encontrado en la literatura existente, la repetición de exámenes en el contexto local explica una proporción baja de la brecha socioeconómica en puntajes y por tanto, los efectos que esto

podría tener sobre otros resultados de interés que se han señalado en la literatura reciente, como el acceso a programas de educación superior de población minoritaria o vulnerable y la entrada a universidades de mayor calidad, podría ser limitado. Es necesario profundizar sobre los efectos de largo plazo de la repetición de exámenes y de la preparación de exámenes estandarizados de alto impacto.



7. Referencias

- Appelrouth, J. (2015). *Predicting Success In Preparing For High-Stakes Admissions Tests: A Moderated Mediation Analysis*. (G. S. University, Ed.) Dissertation.
- Appelrouth, J. I., Zabrocky, K. M., & Moore, D. (2017). Preparing students for college admissions tests. *Assessment in Education: Principles, Policy & Practice*, 24(1), 78-95.
- Briggs, D. (2001). The effect of admissions test preparation: Evidence from NELS-88. *Chance*, 14(1), 10-18.
- Briggs, D. (2002). SAT coaching, bias and causal inference. *Dissertation Abstracts International*, 1-178.
- Buchmann, C., Condrón, D. J., & Roscigno, V. J. (2010). Shadow Education, American Style: Test Preparation, the SAT and College Enrollment. *Social Forces*, 89(2), 435-461.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4), 1593-1660.
- Cliffordson, C. (2004). Effects of practice and intellectual growth on performance on the Swedish Scholastic Aptitude Test (SweSAT). *European Journal of Psychological Assessment*, 20(3), 192-204. doi:10.1027/1015-5759.20.3.192
- Clotfelter, C., & Vidgor, J. (2003). Retaking the SAT. *Journal of Human Resources*, 38(1), 1-33.
- Domingue, B., & Briggs, D. (2009). Using linear regression and propensity score matching to estimate the effect of coaching on the SAT. *Multiple Linear Regression Viewpoints*, 35(1), 12-29.
- Frisancho, V., Krishna, K., Lychagin, S., & Yavas, C. (2016). Better luck next time: Learning through retaking. *Journal of Economic Behavior & Organization*, 125, 120-135.
- Gómez Soler, S. C., Bernal Nisperuza, G. L., & Herrera Idárraga, P. (2020). Test Preparation and Students' Performance: The Case of the Colombian High School Exit Exam. *Cuadernos de Economía*, 39(79), 31-72.
- Goodman, J., Gurantz, O., & Smith, J. (2020). Take two! SAT retaking and college enrollment gaps. *American Economic Journal: Economic Policy*, 12(2), 115-158.
- Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1), 25-46.
- Hausknecht, J. P., Halpert, J. A., Paolo, N. T., & Gerrard, M. O. (2007). Retesting in selection: a meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology*, 92, 373.
- Icfes. (2013). *Sistema Nacional de Evaluación Estandarizada de la Educación. Alineación del Examen Saber 11*. Bogotá: Instituto Colombiano para la Evaluación de la Educación (Icfes).
- Icfes. (2017). *INFORME NACIONAL Resultados nacionales 2014-2 - 2016-2 Saber 11*. Bogotá: Instituto Colombiano para la Evaluación de la Educación (Icfes).

- Icfes. (2019). *¿Cómo se construye el índice de Nivel Socioeconómico (INSE) en el contexto de las pruebas Saber?* Bogotá: Instituto Colombiano para la Evaluación de la Educación (Icfes).
- Leeds, D. (2012). The Impact of Admissions Policies and Test Prep on SAT Scores. *Labor Lunch Seminar Series University of Michigan, University of Michigan*, 1-43.
- Messick, S., Alderman, D., Angoff, W., Jungeblut, A., Powers, D., Rock, D., . . . Stroud, T. (1980). The effectiveness of coaching for the SAT: Review and reanalysis of research from the fifties to the FTC. *ETS Research Report Series, 1*, 1980(1), ii-iii.
- Miyasaka, J. (2000). *A framework for evaluating the validity of test preparation practices*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Montgomery, P., & Lilly, J. (2012). Systematic reviews of the effects of preparatory courses on university entrance examinations in high school-age students. *International Journal of Social Welfare, 21*, 3-12.
- OCDE. (2016). *La educación en Colombia*. (M. d. Nacional, Trad.) Paris: Organización para la Cooperación y el Desarrollo Económicos (OCDE).
- Park, J., & Becks, A. (2015). Who benefits from SAT prep?: An examination of high school context and race/ethnicity. *The Review of Higher Education, (39)*(1), 1-23.
- Popham, W. J. (1987). The Merits of Measurement-Driven Instruction. *The Phi Delta Kappan, 68*(9), 679-682.
- Resolución 558. (22 de Junio de 2010). RESOLUCIÓN 558 DE 2010 ICFES. *Diario Oficial No. 47.748*. Bogotá, Colombia.

8. Anexos

Tabla 4a. La asociación entre la presentación de Pre Saber y Saber 11 por sexo

Pre Saber=1	Mujeres			Hombres		
	(1) Puntaje global	(2) Matemáticas	(3) Lectura crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura crítica
OLS	0.400*** (0.0117)	0.405*** (0.0123)	0.326*** (0.00967)	0.430*** (0.0125)	0.441*** (0.0127)	0.338*** (0.0101)
PSM - Balance por entropía	0.374*** (0.0111)	0.381*** (0.0117)	0.305*** (0.00920)	0.403*** (0.0121)	0.415*** (0.0123)	0.318*** (0.00984)
Observaciones	1,231,280	1,231,280	1,231,280	1,027,005	1,027,005	1,027,005

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** p<0.01, ** p<0.05, * p<0.1. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen.

Tabla 5a. La asociación entre la presentación de Pre Saber y Saber 11 por naturaleza del colegio

Pre Saber=1	Colegios oficiales			Colegios no oficiales		
	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica
OLS	0.442*** (0.0166)	0.455*** (0.0171)	0.357*** (0.0139)	0.358*** (0.0150)	0.363*** (0.0154)	0.287*** (0.0117)
PSM - Balance por entropía	0.438*** (0.0154)	0.451*** (0.0160)	0.352*** (0.0129)	0.329*** (0.0139)	0.335*** (0.0143)	0.263*** (0.0108)
Observaciones	1,626,684	1,626,684	1,626,684	631,597	631,597	631,597

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** p<0.01, ** p<0.05, * p<0.1. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen.

Tabla 6a. La asociación entre la presentación de Pre Saber y Saber 11 por nivel socioeconómico del estudiante

	NSE 1			NSE 2			NSE 3			NSE 4		
	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica	(1) Puntaje global	(2) Matemáticas	(3) Lectura Crítica
OLS	0.565*** (0.0199)	0.572*** (0.0211)	0.477*** (0.0174)	0.514*** (0.0188)	0.521*** (0.0191)	0.420*** (0.0160)	0.405*** (0.0129)	0.410*** (0.0132)	0.325*** (0.0104)	0.243*** (0.0142)	0.254*** (0.0151)	0.185*** (0.0110)
PSM - Balance por entropía	0.559*** (0.0190)	0.568*** (0.0202)	0.471*** (0.0166)	0.506*** (0.0176)	0.515*** (0.0180)	0.412*** (0.0150)	0.396*** (0.0124)	0.402*** (0.0127)	0.318*** (0.0101)	0.233*** (0.0138)	0.244*** (0.0147)	0.178*** (0.0107)
Observaciones	388,628	388,628	388,628	844,025	844,025	844,025	799,595	799,595	799,595	226,033	226,033	226,033

Nota: Los puntajes del examen Saber 11 están estandarizados. Los errores estándar están ajustados por clúster a nivel de colegio. *** p<0.01, ** p<0.05, * p<0.1. En todas las estimaciones se controla por el vector de características sociodemográficas de los estudiantes y características de los colegios a los que pertenecen.

