

Efecto de un Programa de Fortalecimiento Educativo sobre el logro académico en estudiantes de bachillerato en las zonas rurales de Antioquia

**Curso de Evaluación de Impacto
Universidad de los Andes**

**Julián Gallego Villa
Luis Alejandro Quimbayo**

1. Introducción

Las intervenciones de política pública en la educación han demostrado tener impactos positivos en indicadores de escolaridad, aprobación y desempeño académico. Estas políticas pueden perseguir dos objetivos claros: subsidiar la demanda o fortalecer la oferta. Aunque ambos enfoques buscan mejorar los resultados educativos mediante mecanismos diferentes, el propósito de este trabajo es examinar el éxito de subsidiar la oferta en programas educativos que benefician a la población juvenil rural, específicamente en los municipios no certificados del departamento de Antioquia. Se pretende entonces comprobar si existe una relación causal entre la implementación de un programa de fortalecimiento educativo al finalizar el bachillerato y los resultados obtenidos en la prueba de Estado, Saber11.

Programas como “*Oportunidades*” en México buscaban subsidiar la demanda educativa proporcionando transferencias monetarias condicionadas a los hogares para fomentar la asistencia escolar de los niños, especialmente en las zonas rurales. Un estudio realizado por Behrman, Parker y Todd (2009) evaluó el impacto a mediano plazo de este programa, métodos de evaluación no experimentales, como el método de diferencias-en-diferencias (DID) y el emparejamiento por puntaje de propensión (PSM). Los resultados mostraron un aumento en los años de escolaridad completados, una reducción en el empleo infantil y una disminución en las tasas de migración juvenil.

Por otro lado, los programas de fortalecimiento de la oferta se enfocan en mejorar directamente los recursos y la infraestructura escolar para elevar la calidad y disponibilidad de la educación. Un ejemplo representativo de evaluación sobre este tipo de programas es el realizado por Rodríguez, Sánchez y Armenta (2010) sobre el impacto del *Proyecto de Educación Rural* (PER) en las escuelas rurales de Colombia. Este programa incluyó la implementación de modelos educativos flexibles adaptados a las necesidades de las comunidades rurales y la provisión de material didáctico especializado y capacitación docente. Utilizando varias metodologías no experimentales combinadas, el estudio encontró efectos positivos en las tasas de aprobación y desempeño académico, así como una reducción en las tasas de deserción y repetición escolar.

El propósito de este documento es evaluar el impacto de un programa de fortalecimiento educativo implementado en Antioquia. Este programa se basa en una política de subsidio a la oferta y busca mejorar los resultados de logro académico, los cuales son medidos mediante la prueba de Estado Saber11. La prueba de Estado Saber11 es una prueba de evaluación de competencias obligatoria aplicada a los estudiantes próximos a finalizar el bachillerato en Colombia. La prueba evalúa competencias en varias áreas del conocimiento, entre ellas Matemáticas, Comprensión de Textos, Ciencias Naturales y Sociales e Inglés.

A diferencia de estudios previos que han evaluado impactos en variables como la asistencia y la deserción escolar, este trabajo se centra en los resultados en puntajes académicos, ofreciendo una contribución clave a la literatura existente sobre la preparación estudiantil para la presentación de pruebas y el fortalecimiento de competencias. Estudios anteriores, como el de Rodríguez et al. (2010), han mostrado que las intervenciones a nivel escolar pueden mejorar significativamente la eficiencia educativa y la calidad, medidas estas por las tasa de aprobación y el rendimiento académico, respectivamente.

2. Descripción del programa

El Programa de Fortalecimiento Educativo de Antioquia -PFEA se implementó con el objetivo de mejorar el rendimiento académico de los estudiantes rurales en el departamento, particularmente aquellos que estaban próximos a finalizar el bachillerato. Esta iniciativa surgió como respuesta al incremento de las brechas educativas ampliadas por la pandemia en las zonas rurales del departamento, reconociendo que obtener buenos resultados en la prueba Saber11 es crucial para ampliar el acceso a la educación superior y mejorar las condiciones de empleabilidad de los egresados de la educación media.

El programa se implementó mediante un proceso de formación presencial e intensivo, con clases de 4 horas, distribuidas en 24 encuentros, abarcando un total de 92 horas de formación en 4 módulos disciplinares¹. Las clases fueron impartidas por profesores universitarios, en búsqueda de aseguramiento de la calidad y el conocimiento sobre los temas impartidos. La estructura del programa incluyó una combinación de horarios de clase y extraclase, proporcionando flexibilidad y oportunidades adicionales de aprendizaje. Además, los estudiantes recibieron materiales y las herramientas educativas necesarias para su proceso de aprendizaje y preparación.

Entre los módulos ofertados en el programa están Razonamiento Lógico, Comprensión de Textos, Inglés y Preparación para la Presentación de Pruebas. El módulo de Razonamiento Lógico se enfocó en desarrollar habilidades analíticas y de resolución de problemas matemáticos. El de Comprensión de Textos buscó mejorar la capacidad de los estudiantes para interpretar y analizar textos académicos y literarios. El módulo de Inglés se centró en fortalecer las competencias lingüísticas en inglés como segunda lengua extranjera y el módulo de Preparación para la Presentación de Pruebas buscó dotar a los estudiantes con estrategias para optimizar el uso del tiempo y los recursos disponibles para afrontar pruebas estandarizadas.

¹En el programa se ofertaron 2 módulos adicionales de 24 horas cada uno: Proyecto de Vida y Emprendimiento.

Tabla 1. Estructura temática del PFEA.

Módulo	Componente	Intensidad
Razonamiento Lógico	Formación disciplinar	72 hrs
Comprensión de Textos		
Inglés		
Preparación para pruebas	Formación transversal	24 hrs
Total		96 hrs

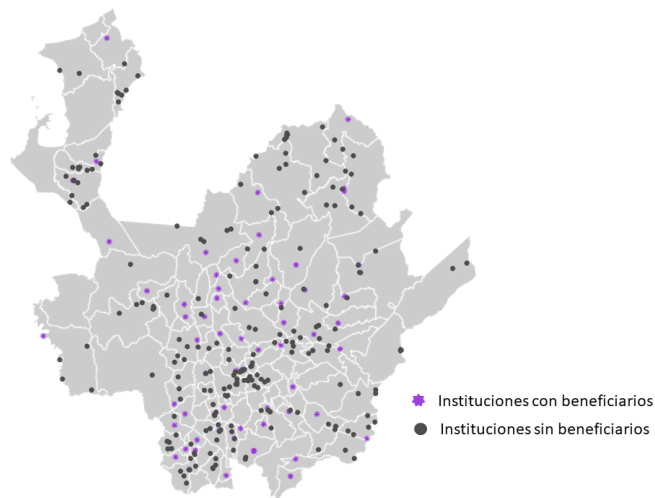
Fuente: elaboración propia

En total, el programa benefició a 3.618 estudiantes de 228 instituciones educativas gubernamentales, distribuidas en 103 municipios del departamento (Figura 1). Principalmente, los beneficiarios fueron estudiantes de grado undécimo. Sin embargo, en algunos municipios no se logró conformar grupos completos de estudiantes debido a que se asignó un cupo de 30 estudiantes por municipio, lo que llevó a la inclusión de estudiantes de varias instituciones educativas en un mismo grupo para completar el número permitido de participantes.

3. Datos

Para evaluar cómo la participación en el programa afectó los resultados académicos de los beneficiarios en la prueba Saber11, se recopiló información individual y no anonimizada del desempeño estudiantil con bases de datos del Instituto Colombiano para la Evaluación de la Educación (ICFES), entre 2019 y 2022. En la base de datos se identificó a los beneficiarios y la información individual permitió calcular medidas de desempeño académico a nivel institucional y de municipio. Estos datos se complementaron con información demográfica y socioeconómica de los municipios y las instituciones educativas, obtenidos principalmente del Departamento Nacional de Estadística -DANE, de la Policía Nacional y del Ministerio de Educación Nacional.

Figura 1. Distribución geográfica de las instituciones beneficiadas por el PFEA.



4. Estrategia empírica

En la revisión de literatura se evidenció que, en términos generales, los programas de preparación para presentar pruebas estandarizadas generan efectos positivos en los resultados académicos de los estudiantes beneficiarios. Además, se observó que se han utilizado múltiples técnicas de inferencia causal para evaluar el impacto de estos programas. Por ejemplo, estudios como el de Gómez, Bernal y Herrera (2020) emplearon métodos de comparación cruzada y emparejamiento por puntaje de propensión, mientras que Bernal y Penney (2019) utilizaron análisis de regresión para identificar los efectos heterogéneos de las intervenciones. Sin embargo, fue común el uso de metodologías basadas en diseños de diferencias-en-diferencias, como en el estudio de Londoño-Vélez et al. (2020) y Posso, Saravia y Uribe (2023) cuando se disponía de datos panel a nivel de unidad de observación con información previa y posterior a la implementación del programa. De manera similar, el estudio de Sanchez y Harnisher (2018) sobre el programa ACT Kaplan Online Prep Live (OPL) utilizó un diseño cuasi-experimental con emparejamiento por puntaje de propensión para crear grupos de tratamiento y control comparables, y análisis de regresión múltiple para evaluar el impacto del programa en los puntajes del ACT, encontrando mejoras significativas, especialmente entre estudiantes de bajos ingresos y minorías. En el mismo sentido, el estudio de Bonilla Sebá y González Borrero (2012) evaluó el impacto de varios programas de la Secretaría Distrital de Educación de Bogotá utilizando metodologías como el emparejamiento por puntaje de propensión, análisis de primeras y segundas diferencias, y regresión discontinua, demostrando mejoras significativas en la cobertura, permanencia escolar y rendimiento académico, especialmente a través de programas como subsidios condicionados, alimentación escolar y construcciones escolares.

Para el caso del PFEA, dado que la asignación al tratamiento entre instituciones educativas no fue aleatoria y no se realizó con base a una regla o criterio específico², se implementó una metodología en tres partes con el objetivo de controlar cualquier fuente de sesgo potencial que pueda surgir. A nivel de institución educativa, se estima un modelo de diferencias-en-diferencias (DID) y un modelo de diferencias-en-diferencias con soporte común (DID-SC) y a nivel de estudiante, como prueba de robustez, se estima el impacto a través del método de emparejamiento o Propensity Score Matching (PSM).

Método de diferencias en diferencias

El método de diferencias-en-diferencias – DID es un método cuasiexperimental de evaluación de impacto, que se basa en comparar las diferencias en el cambio de una variable de interés (Y_t) entre un grupo expuesto a un tratamiento ($D = 1$) y un grupo de control no expuesto ($D = 0$), antes ($t = 1$) y después ($t = 2$) de la

²En el proceso de selección de beneficiarios, inicialmente, se priorizaron las instituciones ubicadas en municipios rurales y rurales dispersos, es decir, municipios con menos de 25 mil habitantes y con una densidad poblacional inferior a 50km/h. Posteriormente, se recomendó a las administraciones municipales elegir a las instituciones educativas que mostraban mayor necesidad de mejora en los resultados de la prueba Saber11. Finalmente, el criterio para la elección de estudiantes se dejó a discreción de las directivas del establecimiento educativo.

implementación de un tratamiento, programa o política. De acuerdo con Bernal y Peña (2011), el estimador se puede escribir de la siguiente manera:

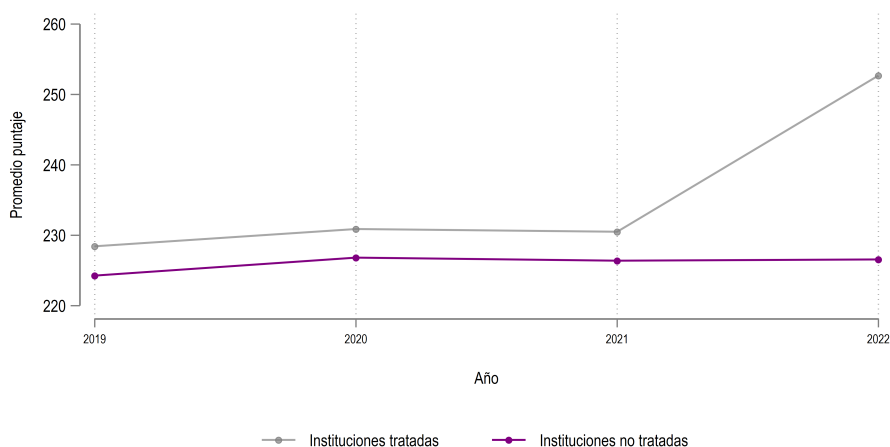
$$\hat{\tau} = [(\bar{Y}_2 | D = 1) - (\bar{Y}_2 | D = 0)] - [(\bar{Y}_1 | D = 1) - (\bar{Y}_1 | D = 0)] \quad (1)$$

donde $\bar{Y}_1 | D = 1$ es el promedio muestral de Y en el periodo antes de la intervención en el grupo de tratamiento. Por otro lado, para garantizar que el estimador $\hat{\tau}$ es insesgado y consistente, es necesario verificar el supuesto de tendencias paralelas, según el cual, en ausencia del programa, la trayectoria de la variable de interés hubiera evolucionado de manera paralela entre el grupo de tratamiento y de control a lo largo del tiempo. Esto permite asegurar que el cambio observado en la tendencia tras la implementación es atribuible a la implementación del programa y no a un cambio en la tendencia en la variable. Esta es la primera metodología utilizada en el documento.

Supuesto de tendencias paralelas

Para validar el supuesto de tendencias paralelas, se empleó un método gráfico que consistió en comparar la tendencia del promedio de los resultados generales en la prueba Saber11 en los años previos a la implementación del programa entre el grupo de tratamiento y el de control. En la Figura 2 se confirma que, si bien las instituciones educativas tratadas mostraban un desempeño superior antes de la implementación del programa en 2022, se cumple con el supuesto de tendencias paralelas. De igual forma, se verificó el supuesto con el promedio del puntaje en algunas de las pruebas específicas que componen el examen: matemáticas, lectura crítica e inglés.

Figura 2. Promedio del puntaje global en prueba Saber 11



Fuente: ICFES

Estimación del modelo

En esta metodología, el grupo de tratamiento está compuesto por las instituciones educativas con grupos completos de beneficiarios³ y se utilizó como grupo de control las instituciones educativas gubernamentales que no participaron del programa. A partir de ello, el siguiente modelo:

$$P_{emt} = \alpha_0 + \tau T_{et} + \mathbf{X}'_{mt}\beta + \gamma_e + \omega_t + e_{emt} \quad (2)$$

donde P_{emt} es el resultado de la institución educativa e , del municipio m , en el año t , en la prueba Saber11°. T_{et} es la variable indicadora del tratamiento, que toma el valor de 1 si en la institución educativa e hubo beneficiarios del programa, \mathbf{X}'_{mt} es un vector de variables de control a nivel de municipio m ; γ_e son efectos fijos a nivel de institución y ω_t de tiempo.

Método de diferencias en diferencias con soporte común

Bifulco (2002) expone que el método de diferencias-en-diferencias contribuye a mitigar el sesgo de selección si se mantienen constantes, a lo largo del tiempo, las características observables y no observables que pudieron haber influido en la selección de beneficiarios. Sin embargo, también se ha señalado que el estimador puede estar sujeto a diversas fuentes de sesgo potencial si las covariables utilizadas para inferir el resultado difieren entre el grupo de tratamiento y control (Stuart et al., 2014). Para abordar esta problemática, se adopta la metodología utilizada por Galiani et al. (2005) y Sánchez et al. (2010) como segunda estrategia.

Utilizando un modelo de respuesta binaria probit, se estima la probabilidad de participación en el programa de las instituciones educativas, considerando como variables explicativas, los resultados en la prueba Saber11 en el año anterior a la implementación del programa (2021), la condición de ruralidad del establecimiento, la proporción de maestros con estudios de formación profesoral y un índice de desempeño académico de la institución educativa construido por el ICFES con información de la prueba Saber11 del 2019, 2020 y 2021. Este método permite calcular la propensión de las instituciones a participar en el programa según las variables observables definidas, a la vez que posibilita la restricción de la muestra a aquellas instituciones con probabilidad de participación común, lo que implica características observables similares.

Soporte común

Se realizó una prueba de diferencia de medias para comprobar que las características observables de las unidades en los grupos de tratamiento y control fueran similares tanto antes y después de aplicar el soporte común. Esto asegura que cualquier diferencia en los resultados puede atribuirse al tratamiento y no a diferencias preexistentes entre los grupos.

³Con el fin de mitigar el sesgo de autoselección de estudiantes, se restringió la muestra a aquellos pertenecientes a instituciones educativas donde todos los estudiantes fueron tratados en grupos completos, es decir, a aquellas instituciones que pudieron beneficiar a todos los estudiantes de un grupo sin dejar a ninguno por fuera. Esta restricción redujo el tamaño de la muestra de estudiantes de 3.618 a 1.587.

Tabla 2. Diferencia de media en variables observables sin soporte común

Variable	Media Tratados	Media Controles	Diferencia Diff
Puntaje global (2021)	230.5 (2.09)	226.4 (1.84)	-4.11 (2.79)
Municipio rural	0.55 (0.063)	0.377 (0.035)	-0.178** (0.072)
Indice desempeño académico (ICFES)	0.639 (0.005)	0.618 (0.004)	-0.02 (0.007)
Promedio educación maestros	2.23 (0.025)	2.21 (0.019)	-0.013 (0.032)
Observaciones	110	110	110

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

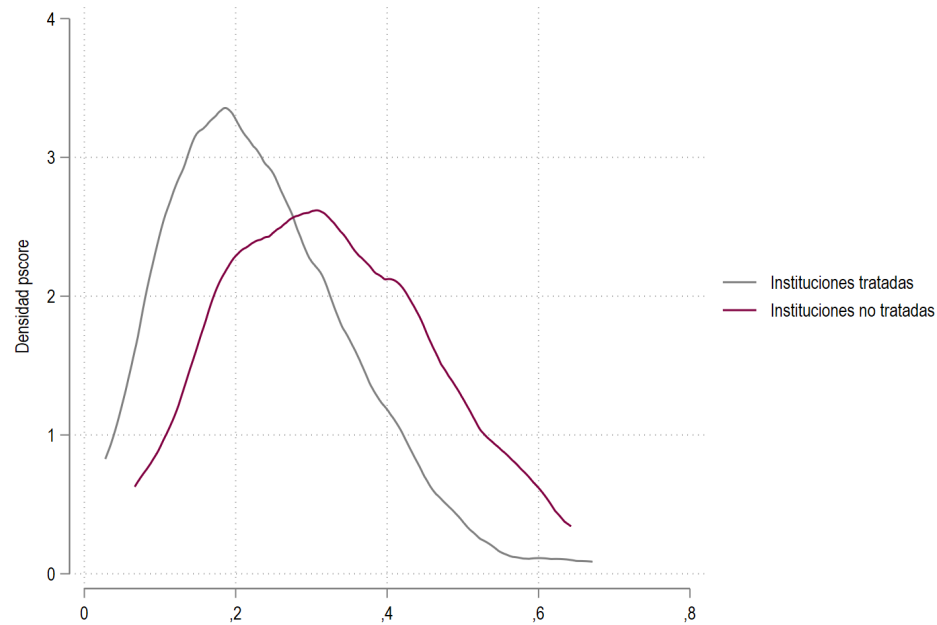
Sin reducir la muestra al soporte común, existen algunas diferencias en las características observables entre las instituciones tratadas y no tratadas. En particular, la variable de "municipio rural" muestra una diferencia significativa al 0.05, indicando que las instituciones tratadas tienden a estar ubicadas en áreas rurales con mayor frecuencia que las instituciones no tratadas. Esto podría introducir sesgo en las estimaciones del impacto del programa si no se controla adecuadamente.

Analizadas las diferencias en las variables observables, se obtuvo el puntaje de propensión o la probabilidad estimada de que una institución reciba el tratamiento dado el conjunto de variables observables indicadas. Formalmente, el puntaje de propensión es la probabilidad condicional de recibir el tratamiento dado un vector de covariables X :

$$e(X) = P(T = 1 \mid X) \quad (3)$$

donde T es una variable indicadora que toma el valor 1 si la unidad recibe el tratamiento y 0 si no lo recibe y X es un vector de características observables. Una vez que se calcularon los puntajes de propensión, se comparó la distribución de los puntajes entre el grupo de tratamiento y control y se obtuvo el soporte común, que se puede observar en la Figura 3.

Figura 3. Distribución de la propensión a ser tratado y soporte común



Determinado el soporte común y restringido las observaciones a su zona común, se calculó de nuevo la diferencia de medias para verificar si existen diferencias significativas entre los grupos de tratamiento y control. Los resultados que se muestran en la Tabla 3, indican que las diferencias en las medias de las covariables entre las instituciones tratadas y no tratadas no son estadísticamente significativas dentro del soporte común. Esto indica que las características observables están bien balanceadas entre los dos grupos, lo que sugiere que los grupos son comparables.

5. Resultados

Comprobado el supuesto de tendencias paralelas y ajustadas las unidades tratadas y de control al área de soporte común, se estimaron los modelos propuestos (Tabla 4). Los resultados de ambos modelos son consistentes y significativos para las variables evaluadas. Esto sugiere que la participación en el programa tiene un efecto positivo y significativo sobre el puntaje global de los beneficiarios, así como en las pruebas específicas de matemáticas, lectura crítica e inglés. Por otra parte, los errores estándar en el modelo DID-SC son generalmente menores que en el modelo DID, lo que sugiere una estimación más precisa al controlar el soporte común. El modelo DID-SC, que incluye la restricción del soporte común, proporciona estimaciones ligeramente diferentes pero consistentes con el modelo DID. Esto indica que el control del sesgo de selección a través del soporte común, mejora la precisión del efecto estimado.

En cuanto a la magnitud del efecto, durante el año de tratamiento, el promedio del puntaje global para el grupo de control fue de 226. Esto sugiere que el programa generó una mejora del 9.7% en este puntaje, del 11.7% en promedio para matemáticas, del 8.9% en Lectura Crítica y del 10.7% en Inglés.

Tabla 3. Diferencia de media en variables observables con soporte común

Variable	Media Tratados	Media Controles	Diferencia Diff
Puntaje global (2021)	230.50 (2.1)	232.90 (3.17)	2.42 (3.80)
Municipio rural	0.55 (0.063)	0.49 (0.073)	-0.066 (0.097)
Indice desempeño académico (ICFES)	0.639 (0.005)	0.641 (0.007)	0.002 (0.009)
Promedio educación maestros	2.23 (0.025)	2.22 (0.037)	-0.01 (0.045)
Observaciones	110	110	110

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

6. Prueba de robustez

Propensity Matching Score

El método de emparejamiento o Propensity Matching Score – PSM es un método cuasiexperimental de evaluación de impacto, que se basa en comparar las diferencias promedio en el cambio de una variable de interés (Y) entre un grupo expuesto a un tratamiento ($D = 1$) y un grupo de control no expuesto ($D = 0$), condicionando el resultado a la probabilidad de ser tratado por el programa. De acuerdo con Bernal y Peña (2011), el estimador se puede escribir de la siguiente manera:

$$\hat{\tau} = \{E[Y(1) \mid D = 1, P(X)] - E[Y(0) \mid D = 0, P(X)]\} 4 \quad (4)$$

donde $E[Y(1) \mid D = 1, P(X)]$ es el valor esperado con respecto a la probabilidad de ser tratado, dada la condición de beneficiario del programa. Para garantizar que el estimador $\hat{\tau}$ es insesgado y consistente, es necesario garantizar la condición de independencia condicional y la condición de soporte común.

Estimación

Inicialmente, por medio de un modelo probit se calculó la probabilidad de que un estudiante fuese tratado. Las variables usadas para esta estimación incluyeron el valor agregado municipal, el índice de desempeño

académico del ICFES utilizado en el modelo DID-SC y un índice multidimensional del ICFES que incorpora información personal, familiar y socioeconómica de los estudiantes. Se considera que estas variables pueden capturar características relevantes que podrían influir en la probabilidad de ser beneficiario del programa.

Seguidamente, la muestra se restringió al soporte común asegurando que las comparaciones se realizaran entre estudiantes con probabilidades similares de recibir el tratamiento. Empleando el algoritmo de agrupamiento vecinos más cercanos, se formaron parejas de estudiantes tratados y no tratados con puntajes de propensión similares. De esta manera se garantizó que cada estudiante tratado se compare con uno o más estudiantes no tratados que tengan características observables similares, lo que reduce el sesgo de selección. Por último, se procedió a la estimación del efecto del programa, evaluando la significancia de los resultados y calculando los intervalos de confianza.

Tabla 4. Resultados de las estimaciones a nivel de institución educativa

Modelos	DID	DID-SC
Variable	ITT	ITT
Puntaje global	22.19*** (2.3)	23.61** (0.665)
Matemáticas	5.28*** (0.54)	5.322*** (0.0)
Lectura crítica	4.30*** (0.45)	4.494*** (0.548)
Inglés	4.50*** (0.50)	4.405*** (0.548)
Observaciones	984	440

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

El puntaje global de los estudiantes que participaron en el PFEA aumenta en promedio 19.28 puntos en comparación con aquellos que no participaron en el programa (Tabla 5). Este efecto es estadísticamente significativo al nivel del 1%, lo que sugiere que el incremento es atribuible al programa. Este aumento representa una mejora sustancial, especialmente considerando que, en promedio, el puntaje global crece 22 puntos más con el tratamiento, equivalente a un aumento adicional del 10% respecto a la línea base. Además, los estudiantes tratados muestran un aumento significativo en las pruebas específicas: 4.31 puntos más en Matemáticas, 3.86 puntos más en Lectura crítica y 2.98 puntos más en Inglés, todos con una significancia del 1%.

7. Mecanismos

Los resultados de la evaluación del impacto del PFEA pueden explicarse por una serie de factores relacionados con el contexto socioeconómico y educativo de las instituciones rurales, así como las intervenciones específicas del programa. En primer lugar, las instituciones educativas rurales enfrentan carencias importantes en términos de acceso a bienes y servicios básicos. Esto se traduce en dificultades en los procesos de aprendizaje y la asimilación del conocimiento, ya que los estudiantes no disponen de los recursos necesarios para

un ambiente de estudio óptimo. Aunque el programa no aborda directamente estos problemas de financiación e infraestructura debido a su naturaleza, sí contribuyó a mitigar algunos de estos problemas proporcionando materiales educativos y facilitando un entorno de aprendizaje más estructurado durante la implementación de la educación de refuerzo o complementaria.

Tabla 5. Resultados de las estimaciones a nivel de individuo

Modelos	PSM
Variable	ITT
Puntaje global	19.28*** (1.32)
Matemáticas	4.31*** (0.32)
Lectura crítica	3.86*** (0.28)
Inglés	2.98*** (0.31)
Observaciones	5.176984

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

Además, la ausencia de maestros en dedicación permanentes y el insuficiente seguimiento a las recomendaciones curriculares afectan la adquisición de competencias necesarias en cada grado. El programa de fortalecimiento educativo, al ofertar clases estructuradas y la presencia de instructores cualificados, como los profesores universitarios, ayudó a subsanar este vacío, asegurando que los estudiantes reciban la instrucción necesaria para mejorar la adquisición de competencias. Por otro lado, a pesar de la firma del acuerdo de paz, algunas zonas rurales siguen siendo afectadas por conflictos armados entre disidencias y grupos de delincuencia común. Estos conflictos crean un ambiente de inseguridad que distraen a los estudiantes y afectan negativamente su desempeño académico, al obstaculizar la asistencia a clases en condiciones de normalidad. El programa, al proporcionar un entorno seguro y enfocado para el aprendizaje, ha permitido a los estudiantes concentrarse mejor en sus estudios, contribuyendo a la mejora de sus resultados.

Asimismo, el programa permitió que los estudiantes se familiaricen con el tipo de preguntas que evalúa la prueba Saber 11 y que adquieran habilidades en el manejo de recursos teóricos, prácticos y emocionales y de tiempo. Esta familiaridad es crucial, ya que reduce la ansiedad de los estudiantes y mejora su capacidad para responder eficazmente durante los exámenes, lo que se refleja en mejores puntajes. Además, para muchos estudiantes, el programa fue la primera vez que tuvieron contacto con muchas de las temáticas abordadas. La exposición a estos nuevos contenidos amplió su conocimiento y habilidades, preparándolos para la evaluación futura.

En resumen, el programa impactó a los estudiantes y a las comunidades académicas en la que se desenvuelven. La adaptación pedagógica y el enriquecimiento a través del diálogo junto con la experiencia práctica

resultaron cruciales para mantener la efectividad del programa. Aunque se encontraron dificultades logísticas y falta de recursos en las instituciones de educación, el programa demostró ser una herramienta exitosa para el fortalecimiento educativo en áreas rurales, proporcionando los medios para facilitar una educación integral y motivadora que prepara a los estudiantes para los desafíos académicos y personales.

8. Conclusiones

El Programa de Fortalecimiento Educativo de Antioquia generó un impacto positivo resultado en el desempeño académico de los estudiantes de municipios rurales. En este trabajo se utilizó metodologías robustas como el método de diferencias en diferencias (DID) con y sin soporte común y el emparejamiento por puntaje de propensión (PSM), demostrado que el programa contribuyó a mejoras significativas en los puntajes globales, así como en áreas específicas como matemáticas, lectura crítica e inglés. Estas mejoras son atribuibles a la intervención del programa, que proporcionó un entorno de aprendizaje estructurado, recursos educativos adecuados y una familiarización con el tipo de evaluación aplicada en la prueba Saber 11.

Las instituciones de educación de municipios rurales del departamento Antioquia enfrentan diversas carencias, la falta de recursos, problemas en la infraestructura e incluso el conflicto armado, y aun sí la implementación del programa ha permitido no solo mejorar el rendimiento académico medido en el puntaje de la prueba de Estado Saber 11, sino también ofrecer un espacio seguro y motivador para los estudiantes. La intervención ha facilitado que los estudiantes se familiaricen con los formatos de evaluación y desarrollen habilidades en el manejo del tiempo y los recursos, reduciendo así la ansiedad durante los exámenes. La participación de profesores universitarios y la estructura del programa han sido elementos clave en la efectividad del mismo, mostrando que, incluso en contextos desfavorables, es posible lograr resultados educativos positivos mediante intervenciones bien diseñadas y focalizadas.

9. Referencias

- Behrman, J. R., Parker, S. W., Todd, P. E. (2009). Medium-term impacts of the Oportunidades program package, including secondary and high school scholarships. *Latin American Economic Review*, 18(1), 123-155.
- Bernal, G. L., Penney, J. (2019). Scholarships and student effort: Evidence from Colombia's Ser Pilo Paga program. *Economics of Education Review*, 72, 121–130.
- Bifulco, R. (2002). Addressing Self-Selection Bias in Quasi-Experimental Evaluations of Whole-School Reform: A Comparison of Methods. *Evaluation Review*, 26(5), 545-572.
- Bonilla Sebá, E. C., González Borrero, J. I. (2012). Evaluación del impacto de algunos programas de la Secretaría Distrital de Educación en el marco de la política educativa del Plan Sectorial de Educación Bogotá: una gran escuela 2004-2008. Instituto para la Investigación Educativa y el Desarrollo Pedagógico (IDEP).
- Galiani, S., Gertler, P., Schargrodsky, E. (2005). Water for Life: The Impact of the Privatization of Water Services on Child Mortality. *Journal of Political Economy*, 113(1), 83-120.

- Gómez, S. C., Bernal, G. L., Herrera, P. (2020). Test preparation and students' performance: The case of the Colombian high school exit exam. *Cuadernos de Economía*, 39 (79), 31–72.
- Londoño-Vélez, J., Rodríguez, C., Sánchez, F. (2020). Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser Pilo Paga in Colombia. *American Economic Journal: Economic Policy*, 12 (2), 193–227.
- Posso, C., Saravia, E., Uribe, P. (2023). Acing the test: Educational effects of the SaberEs test preparation program in Colombia (No. 1237). Banco de la República, Icfes, University of California, Los Angeles, Universidad Eafit.
- Rodríguez, C., Sánchez, F., Armenta, A. (2010). Do interventions at school level improve educational outcomes? Evidence from a rural program in Colombia. *World Development*, 38(3), 415-428.
- Sanchez, E., Harnisher, J. (2018). The Impact of ACT Kaplan Online Prep Live on ACT Score Gains. *ACT Research Policy, Technical Brief*, July 2018.
- Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M., Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services and Outcomes Research Methodology*, 14(4), 166-182.
- Vera-Hernández, M. (2010). Guía práctica para la evaluación de impacto. *Desarrollo y Sociedad*, (66), 219-220.