

No. **127**

Noviembre del 2025

ISSN 2215 – 7816 (En línea)

Documentos de Trabajo

Escuela de Gobierno Alberto Lleras Camargo

Evaluación de impacto de un programa
de inglés como lengua extranjera en el
contexto de una iniciativa de capacitación
laboral en Colombia

Sandra García, Andrés Ham,
Stephanie Majerowicz, Darío Maldonado.

Serie Documentos de Trabajo 2025

Edición No. 127

ISSN 2215-7816 (En línea)

Edición digital

Noviembre 2025

© 2025 Universidad de los Andes, Escuela de Gobierno Alberto Lleras Camargo

Carrera 1 No. 19 -27, Bloque Aulas

Bogotá, D.C., Colombia

Teléfono: 3394949, ext. 2073

publicaciones@uniandes.edu.co

<http://gobierno.uniandes.edu.co>

Autores

Sandra García, Andrés Ham, Stephanie Majerowicz, Darío Maldonado

Directora de la Escuela de Gobierno Alberto Lleras Camargo

María Margarita, Paca Zuleta

Coordinación editorial, Escuela de Gobierno Alberto Lleras Camargo

María Alejandra Rojas Forero

Dirección de Investigaciones, Escuela de Gobierno Alberto Lleras Camargo

Camilo Ignacio González

Diagramación de cubierta, Escuela de Gobierno Alberto Lleras Camargo

Miguel Ángel Campos Guaqueta

El contenido de la presente publicación se encuentra protegido por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por tanto su utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso, digital o en cualquier formato conocido o por conocer, se encuentran prohibidos, y solo serán lícitos en la medida en que cuente con la autorización previa y expresa por escrito del autor o titular. Las limitaciones y excepciones al Derecho de Autor solo serán aplicables en la medida en que se den dentro de los denominados Usos Honrados (Fair Use); estén previa y expresamente establecidas; no causen un grave e injustificado perjuicio a los intereses legítimos del autor o titular; y no atenten contra la normal explotación de la obra.

Evaluación de impacto de un programa de inglés como lengua extranjera en el contexto de una iniciativa de capacitación laboral en Colombia¹

Sandra García², Andrés Ham³, Stephanie Majerowicz⁴, Darío Maldonado⁵

Resumen

Este estudio examina la efectividad de un programa de aprendizaje del idioma inglés para personas de bajos ingresos en edad de trabajar en Bogotá, Colombia. Estimamos los efectos causales del programa a través de un emparejamiento y un diseño de regresión discontinua. La oferta de un cupo en el programa aumentó los puntajes de inglés en 0,2 desviaciones estándar (DE) y en 0,27 DE para los estudiantes que asistieron al menos al 20 % de las clases. El programa fue efectivo independientemente del género o el nivel socioeconómico del estudiante y no tuvo diferencias significativas de acuerdo con las características individuales. Sin embargo, los estudiantes cuyos profesores tenían una calidad de enseñanza superior al promedio tuvieron ganancias lingüísticas significativamente mayores (0,35 DE). Nuestros resultados contribuyen a la evidencia limitada sobre intervenciones efectivas que mejoran el dominio del inglés entre los adultos, particularmente aquellos de contextos vulnerables.

Palabras clave: inglés como lengua extranjera, educación no formal, educación para adultos.

Códigos JEL: I2, I21, I28

¹ Este documento presenta los resultados obtenidos en el contrato BC23-312 entre el British Council, Colombia, y la Universidad de los Andes. Los investigadores agradecen a Laura Barragán, Carolina Cruz, Adriana Giraldo y Milton Mendoza del British Council, así como a la Agencia de Educación y Ciencia de Bogotá - Atenea, por su apoyo. También agradecemos a María Paula Fajardo, Juan Camilo Quintero, Juliana Borbón y Viviana Iriarte por su excelente asistencia en la investigación en varias etapas del proyecto.

² Universidad de los Andes.

³ Universidad de los Andes.

⁴ Universidad de los Andes.

⁵ Universidad de los Andes.

Impact Evaluation of an English as a Foreign Language Program in the Context of a Job Training Initiative in Colombia⁶

Sandra García⁷, Andrés Ham⁸, Stephanie Majerowicz⁹, Darío Maldonado¹⁰

Abstract

This study examines the effectiveness of an English language learning program targeting low-income, working-age adults in Bogotá, Colombia. We estimate the program's causal effects using propensity score matching and a regression discontinuity design. Being offered a place in the program increased English test scores by 0.2 standard deviations (SD), and by 0.27 SD for participants who attended at least 20% of classes. The program proved effective regardless of participants' gender or socioeconomic status, showing no significant differences across individual characteristics. However, students taught by instructors with above-average teaching quality achieved significantly greater language gains (0.35 SD). Our findings add to the limited body of evidence on effective interventions that enhance English proficiency among adults, particularly those from vulnerable backgrounds.

Keywords: English as a Foreign Language; Non-formal Education; Adult Learning.

JEL Classification: I2; I21; I28

⁶ This document presents the results of Contract BC23-312 between the British Council Colombia and Universidad de los Andes. The researchers express their gratitude to Laura Barragán, Carolina Cruz, Adriana Giraldo, and Milton Mendoza from the British Council, as well as to Bogotá's Education and Science Agency – Atenea, for their support. We also thank María Paula Fajardo, Juan Camilo Quintero, Juliana Borbón, and Viviana Iriarte for their excellent research assistance throughout various stages of the project.

⁷ Universidad de los Andes.

⁸ Universidad de los Andes.

⁹ Universidad de los Andes.

¹⁰ Universidad de los Andes.

Tabla de contenido

1. Introducción.....	4
2. Contexto	7
2.1. Todos a la U	7
2.2. English without Borders	9
3 Datos y estrategia empírica	11
3.1. Los grupos de tratamiento y control	11
3.2. Datos	13
3.3. Estadística descriptiva de la muestra	14
3.3.1 . La muestra de individuos que presentaron la prueba EnglishScore	15
3.3.2. Tasas de asistencia	15
3.4. Metodología: estimadores de emparejamiento.....	16
3.5. Metodología alternativa: regresión discontinua	19
4. Resultados	20
4.1. Resultados principales	20
4.1.1 . El efecto del intento de tratamiento (IIT).....	20
4.1.2 . Teniendo en cuenta el cumplimiento imperfecto	21
4.2. Efectos heterogéneos	23
5. Conclusión.....	24
Referencias.....	26
Tablas y figuras.....	28
Tabla 1. Estadística descriptiva de la muestra y submuestras.....	28
Tabla 2. Estadística descriptiva de los resultados del examen EnglishScore.....	29
Tabla 3. Asociación entre el tiempo de solicitud y la asistencia	29
Tabla 4. Asignación de tratamiento de tablas de balance de PSM.....	30
Tabla 5. Estimaciones por intención de tratar en los puntajes de las pruebas.....	31
Tabla 6. Efecto de la aceptación del tratamiento en los puntajes de las pruebas	31
Tabla 7. Resultados de la estimación de RD	32
Figura 1. Gráficos de densidad para todos los partidos de PSM.....	33
Figura 2. Resultados principales	34
Figura 3. Principales efectos heterogéneos	35
Figura 4. Efectos heterogéneos: todos los gráficos.....	36
Figura 5. Gráficos de resultados de la estimación de regresión discontinua	37
Tabla A1. Tablas de balance de PSM para PSM según diferentes definiciones de tratamiento.....	38
Tabla A2. Efectos heterogéneos por características de los estudiantes	39
Tabla A3. Efectos heterogéneos por desempeño docente	40

1. Introducción

Uno de los propósitos principales de la política educativa es asegurar que los sistemas educativos preparen a las personas con las habilidades necesarias para tener éxito en el mercado laboral. Los altos costos de la educación, la incertidumbre sobre las habilidades que demandan los empleadores y las asimetrías de información con respecto a la calidad y efectividad de programas particulares justifican la intervención pública en la educación superior. Además, el rápido ritmo de cambio en las estructuras económicas de un mundo cada vez más globalizado requiere que los gobiernos alienten a los individuos a adquirir continuamente nuevas competencias buscadas por el mercado laboral. Una habilidad cada vez más demandada en los países donde la lengua materna no es el inglés es el dominio de este idioma, en la medida en que mejora la empleabilidad y las oportunidades de avance profesional en los mercados laborales internos. Las investigaciones de las últimas décadas han mostrado que en varios países, incluidos China, Polonia, India, Sudáfrica, Turquía o Ruanda, entre otros, el conocimiento del inglés por parte de los trabajadores se asocia con importantes beneficios en el mercado laboral, en particular una prima salarial significativa (Adamchik *et al.*, 2019; Azam *et al.*, 2013; Chakraborty y Bakshi, 2016; Casale y Posel, 2011; Di Paolo y Tansel, 2015; Fabo *et al.*, 2017; Garrouste, 2008; Guo y Sun, 2014; Lang y Siniver, 2009; Liwiński, 2019; Muhawenayo *et al.*, 2022; Stöhr, 2015; Wang *et al.*, 2017). Por otra parte, este idioma resulta vital en el desarrollo de varios sectores, incluidos el comercio internacional y las operaciones comerciales, así como la industria del turismo y la hospitalidad. Además, su dominio permite encontrar trabajos en el extranjero, ampliando las perspectivas de empleo y retribuyendo a los países de origen a través de remesas y otros beneficios migratorios (Agirdag, 2014; Bleakley y Chin, 2004; Chiswick, 1991).

A pesar de este papel fundamental, muchos países carecen de educación bilingüe en la educación K-12, lo que resulta en una alta proporción de adultos con poco o ningún conocimiento del inglés. Como consecuencia, se necesitan programas de enseñanza para adultos que no tuvieron la oportunidad de adquirir estas habilidades durante sus años escolares. Estos programas deben ser efectivos para desarrollar el dominio del inglés en estas poblaciones y flexibles para adaptarse a las limitaciones de horario de los trabajadores de tiempo completo, las personas con responsabilidades de cuidado y quienes viven en ciudades con largos tiempos de viaje. Este desafío se vuelve aún más difícil cuando estos programas se dirigen a adultos de bajos ingresos, quienes a menudo enfrentan mayores barreras para la participación (por

ejemplo, horas de trabajo más largas e impredecibles, mayores responsabilidades de cuidado y menos acceso a internet y dispositivos electrónicos), así como una base más débil en la educación básica y media.

Al respecto, hay muy poca evidencia sobre la efectividad de los programas de inglés como lengua extranjera (ESL son las siglas estándar provenientes del inglés) dirigidos a adultos de bajos ingresos en edad de trabajar. Las investigaciones existentes se han centrado en el inglés como segundo idioma (2L) y en establecer cómo estrategias pedagógicas específicas pueden mejorar el aprendizaje dentro de las aulas de inglés, pero no en la efectividad general de las clases sobre el dominio del inglés (Wyse *et al.* 2010 y Gao 2019). Además, entre los estudios centrados en estrategias particulares de enseñanza y aprendizaje para 2L, pocos emplean metodologías capaces de estimar los efectos causales de los programas en los resultados de interés¹¹. Finalmente, la investigación existente se concentra casi exclusivamente en la educación formal (educación secundaria o terciaria), pero ha avanzado muy poco en el contexto de la educación informal para adultos.

Borradaile *et al.* (2021) realizaron una revisión sistemática exhaustiva de las estrategias de educación de adultos, incluidas las habilidades de alfabetización y aritmética, y encontraron solo un estudio que se centra específicamente en las estrategias de adquisición del idioma inglés para adultos (Condelli *et al.*, 2010). En particular, estos autores examinan la implementación de un plan de estudios y un libro de texto diseñados para el aprendizaje de ESL, pero no encontraron efectos significativos en la adquisición del idioma. Esta falta de evidencia destaca la necesidad urgente de más investigación y desarrollo, a fin de comprender la efectividad de los programas de inglés para estudiantes adultos. Esta necesidad es aún más apremiante en una región como América Latina, que ha luchado durante décadas para mejorar sus programas de aprendizaje de inglés (Rojas y Hernández Fernández, 2018; Cronquist y Fiszbein, 2017).

Para contribuir a cerrar esta brecha, proporcionamos evidencia causal de la efectividad de un programa de aprendizaje del inglés dirigido a poblaciones vulnerables en edad laboral en Bogotá, Colombia. El programa, denominado English without Borders (EwB), fue diseñado

¹¹ La mayor parte de la literatura emplea dos estrategias diferentes: estrategias de aprendizaje de idiomas (Takeuchi, 2019) y enfoques basados en el uso (Dolgova y Tyler, 2019). Las primeras introducen a los estudiantes a estrategias cognitivas (p. ej., tomar notas o resumir textos), metacognitivas (p. ej., planificar sesiones de estudio o monitorear el progreso) o sociales (p. ej., hacer preguntas o practicar con hablantes nativos), entre otras; las segundas enfatizan el aprendizaje a través del uso frecuente del idioma.

por el British Council (BC) y se ofreció como parte del programa de formación profesional Todos a la U (TALAU), implementado por el gobierno de la ciudad de Bogotá en 2022 para mejorar las habilidades del mercado laboral de los bogotanos en sectores de alta demanda. El programa incluyó capacitación vocacional y lecciones de inglés para personas interesadas en sectores específicos, como BPO (Business Process Outsourcing), turismo y gastronomía, donde las habilidades en inglés tienen una gran demanda. Por su parte, EwB es un programa en línea, flexible, diseñado por BC para adultos que trabajan. Los participantes asisten a cinco horas de clases por semana, en línea y sincrónicas, durante seis meses, es decir, 120 horas académicas, junto con actividades complementarias opcionales como clubes de conversación. Las clases son impartidas por profesores altamente calificados, con experiencia en la enseñanza de inglés como segundo idioma, quienes emplean un enfoque pedagógico activo, basado en tareas y proyectos, y centrado en el estudiante. El programa tuvo como objetivo abordar el bajo nivel de dominio del inglés entre los adultos en Bogotá, la capital de un país donde menos del 5 % de los graduados de secundaria tienen un nivel de dominio B+ o superior en este aspecto, según el Marco Común Europeo de Referencia (MCER) (Laboratorio de Economía de la Educación, 2023).

Para determinar el efecto causal del programa EwB en el dominio del inglés, aprovechamos un cambio arbitrario en el proceso de selección de participantes. Debido a limitaciones presupuestarias, la agencia ajustó las reglas de asignación de cupos en una cohorte a mitad del proceso de solicitud. El programa decidió establecer dicho límite una vez que el proceso de inscripción estaba en marcha, y los estudiantes no lo sabían con anticipación. Así, las personas que habían presentado su solicitud antes de una fecha determinada y habían indicado interés en los cursos de inglés fueron elegibles, mientras que las que solicitaron después de esa fecha no lo fueron pero sí les podían ofrecer un cupo en el curso de formación profesional. Estas últimas personas que solicitaron el curso de inglés, pero no fueron elegidas, conforman nuestro grupo de control potencial. En este contexto usamos dos estrategias empíricas que aprovechan este cambio de regla: estimadores de emparejamiento (emparejamiento de puntaje de propensión y emparejamiento de vecino más cercano) y un diseño de regresión discontinua. Los resultados de la discontinuidad de emparejamiento y regresión son similares en magnitud; sin embargo, nuestra estrategia de regresión discontinua tiene poco poder y, por lo tanto, no es estadísticamente significativa.

Nuestros resultados muestran que el programa es efectivo para mejorar el dominio del inglés. Más específicamente, la oferta de un cupo en el curso (una estimación de intención de tratar o ITT) aumenta los puntajes de las personas en una prueba de inglés en 0,195 desviaciones estándar (DE). Sin embargo, no todas las personas a las que se les ofreció un cupo en el programa se registraron para el curso o asistieron a todas las clases. Teniendo en cuenta este cumplimiento imperfecto, encontramos que las personas que asistieron, al menos al 20 % de las clases, aumentaron su dominio del inglés en 0,3 DE. También exploramos los efectos heterogéneos del tratamiento por características del estudiante y por características del profesor asignado al estudiante, incluida una medida de la calidad de las clases que da el profesor. Aunque las estimaciones del efecto son mayores para los hombres y para las personas de bajo nivel socioeconómico (NSE), en comparación con sus pares femeninos o de alto NSE, estas diferencias no son estadísticamente significativas. Esto muestra que el programa EwB fue efectivo para mejorar el dominio del inglés para una amplia gama de personas, incluidas aquellas de los entornos socioeconómicos más bajos de nuestra muestra, quienes probablemente enfrentaron las barreras más altas para el aprendizaje. Por otro lado, encontramos que las personas a las que se les enseña, profesores que imparten clases con una calidad superior al promedio, muestran mejoras significativamente mayores en los puntajes de inglés, lo que sugiere que la calidad de la enseñanza es importante.

El resto del documento procede de la siguiente manera. La sección 2 proporciona una descripción general del programa EwB, incluidos detalles sobre el programa de formación profesional en el que está integrado. La sección 3 explica las fuentes de datos y las estrategias empíricas utilizadas, la sección 4 presenta los resultados y la sección 5 concluye.

2. Contexto

2.1. Todos a la U

TALAU es un programa de formación profesional creado en 2022 por Atenea, la Agencia de Educación Superior, Ciencia y Tecnología de la ciudad de Bogotá. Su propósito es fortalecer las habilidades de las personas en el mercado laboral en áreas de alta demanda. A partir de 2025, la agencia ha ofrecido el programa a seis cohortes diferentes. A cada cohorte del programa se le ofrecía capacitación vocacional para un conjunto diferente de sectores del mercado laboral, pero todas las cohortes compartían una estructura y criterios de selección similares. Este estudio se centra en la cuarta cohorte, que ofreció cursos de formación

profesional en los sectores de construcción, gastronomía, turismo, BPO y salud. Además de fortalecer las habilidades técnicas específicas del sector, TALAÚ buscó mejorar el dominio del inglés en todas las cohortes. En las cohortes anteriores, el curso de inglés era obligatorio para todos los estudiantes que participaban en la capacitación técnica; sin embargo, a partir de la cohorte 3, este componente se volvió opcional.

Los cursos técnicos fueron implementados por organizaciones asociadas (universidades o institutos de formación profesional); los cursos de inglés los ofreció el British Council (BC). Los socios implementadores entregaron todo el contenido del curso, pero no tenían control sobre el proceso de admisión, que fue administrado centralmente por Atenea de acuerdo con un procedimiento estandarizado¹². Los cursos podían impartirse en línea o en persona y duraban entre 40 y 300 horas de instrucción (aunque la mayoría estaban en la parte inferior del rango) en entre dos y seis meses.

Como su nombre lo indica, TALAÚ estaba abierto a un amplio espectro de la población bogotana. Los únicos requisitos de elegibilidad eran residir en Bogotá, tener un título de bachillerato y tener 18 años o más al momento de la solicitud (no había límite de edad superior). Los cupos en el programa se asignaron entre los solicitantes, utilizando un sistema de puntos y un algoritmo de selección que priorizó a las personas vulnerables y a las que ya habían participado en otros programas de la ciudad. Al presentar la solicitud, las personas completaron un formulario con su información personal, que se utilizó para asignar puntos en función de varios criterios de vulnerabilidad. Por ejemplo, se otorgaban siete puntos a mujeres, personas con discapacidad, solicitantes transgénero, pertenecientes a un grupo étnico y aquellos que habían sido víctimas de conflictos o violencia de género. Además, las personas recibieron nueve puntos por pertenecer al Grupo A del Sisbén (el grupo socioeconómico más bajo según el mecanismo de focalización de bienestar social oficial del país), siete puntos para los del Grupo B y cinco puntos para los del Grupo C. Adicional a ello, la ciudad también quería incentivar a las personas que se habían beneficiado de programas de capacitación anteriores para continuar sus trayectorias de aprendizaje, por lo que otorgó diez puntos a aquellos que habían participado o completado previamente otros programas de formación ofrecidos por el gobierno de la ciudad.

¹² Los implementadores del curso podían administrar pruebas de ubicación para determinar si los estudiantes pertenecían a cursos introductorios, intermedios o avanzados.

Todos estos criterios (excepto ser transgénero) se verifican utilizando registros oficiales del gobierno, lo que garantiza que los solicitantes no puedan falsificar su información. A través del sistema de puntos, cada solicitante recibe un puntaje total que va de 0 a 100, siendo los puntajes más altos indicadores de una mayor vulnerabilidad o exposición previa a programas gubernamentales. De este modo, el gobierno determinó la asignación de espacios para cada programa técnico disponible: el solicitante con el puntaje más alto fue asignado a su primera preferencia, luego el siguiente solicitante, y así sucesivamente, hasta que se llenaron todos los cupos. Si la primera opción de un solicitante ya estaba llena, se le asignaría su segunda o tercera opción, siempre que hubiera cupos disponibles. Si ninguno de sus programas preferidos estaba disponible, no se les ofrecía un cupo. Este proceso se repitió hasta que se llenaron todos los cupos. Si los solicitantes tenían el mismo puntaje, se usaban los siguientes criterios de desempate, en orden estricto: 1) se daba preferencia a las personas con NSE más bajo, 2) se daba preferencia al solicitante con el puntaje más alto en las categorías de poblaciones vulnerables, y 3) cuando ninguno de estos rompiera el empate, el orden de preferencia se aleatorizaba.

Los solicitantes podían seleccionar hasta tres cursos técnicos y, además, elegir si optaban por el curso EwB. También podían postularse solo al curso de inglés, pero aquellos que lo hicieron estaban sujetos a un proceso de selección diferente y no están incluidos en nuestra muestra de evaluación.

2.2. *English without Borders*

EwB es un programa diseñado e implementado por el BC en Colombia que ofrece a jóvenes y adultos la oportunidad de aprender inglés a través de cursos flexibles en línea. Los participantes se agrupan en aulas, de acuerdo con su nivel de dominio del inglés y su horario de cursos preferido en el momento de la inscripción. El diseño del curso tiene enfoques pedagógicos activos, basados en tareas y proyectos, centrados en el estudiante. Los profesores (denominados *remote teachers* o RT) facilitan el trabajo centrado en el estudiante para grupos de entre 25 y 35 estudiantes. Los RT son profesionales altamente calificados, seleccionados de acuerdo con los estándares de calidad de BC, y deben ser profesores de inglés capacitados y con experiencia en la enseñanza virtual. Además, el BC capacita a estos profesores en metodologías y prácticas pedagógicas y los monitorea constantemente a través de observaciones en el aula.

Los RT son supervisados y apoyados por *master teachers* (MT). Los MT deben tener un diploma que los acredite como profesores o seis años trabajando como docentes de inglés como lengua extranjera. También deben tener experiencia trabajando en entornos en línea, coordinando programas para profesores y capacitando a sus pares (incluida la observación y la retroalimentación en el aula). Los MT coordinan el programa, acompañan el desarrollo de las clases, lideran el programa de desarrollo profesional continuo de los RT y aseguran el cumplimiento de los estándares de calidad en la enseñanza.

La duración del curso es de 120 horas académicas, repartidas en 6 meses, con 5 horas sincrónicas de clase por semana. Además, el curso ofrece una serie de actividades voluntarias, como clubes de conversación, campamentos de entrenamiento y tutorías académicas, que brindan a los participantes oportunidades adicionales para practicar el inglés con compañeros y MT. También hay sesiones de tutoría académica en las que los RT brindan apoyo académico extracurricular. Además, el curso utiliza herramientas y recursos digitales diseñados para brindar una mayor exposición al uso del inglés a través del trabajo autónomo. Estos recursos se adaptan a las necesidades específicas de los estudiantes, quienes acceden principalmente al contenido del curso a través de sus teléfonos celulares u otros dispositivos móviles.

Los profesores evalúan el aprendizaje de los estudiantes a través de evaluaciones formativas y sumativas en la plataforma EwB. Al final del curso, los participantes con una puntuación de progreso del 60 % o superior reciben un certificado que indica que han finalizado el curso, junto con el número de horas completadas, el nombre del curso y su nivel equivalente en el Marco Común Europeo de Referencia (MCER). Los puntajes de progreso se calculan ponderando la asistencia a clase (60 %), el desempeño (20 %) y la finalización de tareas (20 %); los dos últimos son evaluados por el RT en función de rúbricas estandarizadas predefinidas. Además, los estudiantes pueden ganar puntos de progreso adicionales por participar en actividades voluntarias, como seminarios web, *global career cafes*, clubes de conversación, completar encuestas o tomar el examen EnglishScore.

El conocimiento del inglés en el curso EwB se mide utilizando la prueba EnglishScore, una prueba de clasificación para tres competencias y diez subcompetencias¹³. Fue desarrollado por el BC y el Center of Research for English Language Learning Assessment (CRELLA), de

¹³ Las tres competencias son gramática y vocabulario, expresión oral y escritura. Las subcompetencias para la gramática y el vocabulario son gramática, vocabulario, lectura y comprensión auditiva; para la expresión oral son comunicación, fluidez y pronunciación; para la escritura son lenguaje, organización y comunicación.

la Universidad de Bedfordshire¹⁴, y ha sido utilizado por muchas otras instituciones de todo el mundo (EE. UU., Perú, Colombia, Brasil y Chile).

El resultado en la prueba EnglishScore toma valores entre 0 y 599 y está alineado con el MCER; las puntuaciones entre 0 y 199 corresponden al nivel A1 o menos, entre 200 y 299 a A2, entre 300 y 400 a B1, entre 400 y 499 a B2, y de 500 o más a C1. La prueba es adaptativa, se realiza en línea y dura aproximadamente 40 minutos. El examen se presenta bajo estrictas condiciones de supervisión. Por ejemplo, la cámara debe permanecer encendida en todo momento, y se utiliza inteligencia artificial para monitorear la cara del examinado durante la prueba. Los resultados se utilizan para determinar el nivel de inglés del participante al comienzo del curso y situarlo en el curso apropiado. Al finalizar el curso se invita a los participantes a realizar la prueba final del curso de nuevo. El certificado indica el nivel de inglés del individuo en las tres competencias y diez subcompetencias, y se puede utilizar para indicar a los empleadores dichas habilidades en inglés, proporcionando a los estudiantes una motivación adicional para realizar la prueba.

3. Datos y estrategia empírica

La estimación del efecto causal del programa requiere datos de resultados tanto para las personas que se beneficiaron del programa como para un grupo comparable que no lo hizo, así como un método para comparar de manera creíble los resultados de los dos grupos. En esta sección, describimos los grupos de tratamiento y control de nuestro estudio, los datos utilizados y los dos métodos que usamos (estimadores de emparejamiento y regresión por discontinuidad).

3.1. Los grupos de tratamiento y control

Dado que un programa rara vez ocurre en el vacío, el desafío central al estimar su efecto es tener en cuenta el aprendizaje que habría ocurrido en ausencia del mismo (el contrafactual). Esto requiere medir los niveles de inglés de un grupo de control que es idéntico al grupo de tratamiento, excepto que no recibió el curso. Este es un desafío común en la investigación educativa, ya que la selección de programas a menudo está impulsada por la autoselección (lo que hace que los no participantes sean sistemáticamente diferentes ya que eligieron no

¹⁴ Puede encontrar más información sobre la prueba EnglishScore en <https://www.englishscore.com>, incluido un informe sobre cómo sus resultados se alinean con la escala de puntuación del MCER.

participar) y la evaluación generalmente está integrada dentro de los programas (lo que significa que los no participantes no son evaluados).

Esta evaluación, aprovechó la decisión arbitraria de dividir la cohorte 4 de TALAÚ en dos cohortes, según la fecha de solicitud. La oferta de cupos en las formaciones se determinó de acuerdo con el algoritmo explicado en la sección 2.1; todos los postulantes señalados por el algoritmo obtuvieron un cupo en el curso técnico de Gastronomía o BPO, pero no todos tuvieron la oferta para el curso de inglés¹⁵. La cohorte se dividió en dos, 4.1 y 4.2; los estudiantes que habían postulado antes del 27 de agosto hacían parte de la primera y los que postularon después de la segunda. En ambas cohortes había estudiantes que pidieron cupo en el curso EWB, mostrando interés por esa formación, pero solo los de la cohorte 4.1 recibieron esa oferta. La decisión de dividir la cohorte 4 en dos se tomó después de la fecha límite, no se anunció de antemano y no se había hecho nada parecido en ninguna cohorte anterior. Por lo tanto, los estudiantes no sabían que postularse antes aumentaría sus posibilidades de que se les ofreciera el curso de inglés, y quienes solicitaron participar en el curso después de la fecha límite no sabían que ya no tendrían la oportunidad de recibirlo.

Esta decisión de dividir la cohorte produjo dos grupos de estudiantes, los cuales demostraron interés en los cursos de inglés, pero solo a los del primer grupo (cohorte 4.1) se les ofreció el curso de inglés. Esto presenta una oportunidad para identificar a estudiantes que comparten características observables similares, como una motivación para aprender inglés, pero solo algunos de ellos recibieron el curso de inglés porque lo habían solicitado antes.

Encontrar un grupo de control aceptable es un desafío y medir sus resultados es aún más difícil. Por ejemplo, hay muy pocos incentivos para presentar exámenes cuando no son obligatorios y, en algunos casos, incluso contactar a los posibles participantes del grupo de control es difícil debido a las restricciones de privacidad. En ese sentido, una de las fortalezas de este estudio es la capacidad de establecer un grupo de control razonable y medir su nivel de dominio del inglés. En este caso, el BC pudo contactar a las personas del grupo de control con la ayuda de Atenea y ofrecerles incentivos monetarios (una tarjeta de autobús con \$60 000) para realizar la prueba. El hecho de que presentarla también proporcionara una certificación

¹⁵ El componente de inglés en esta cohorte se limitó a personas que habían seleccionado títulos técnicos en Gastronomía o BPO, por lo que restringimos nuestra muestra a estos estudiantes.

válida del nivel de inglés, que podría usarse al solicitar trabajo, fue un incentivo adicional para ello.

3.2. Datos

Combinamos varias fuentes de datos administrativos y de encuestas recopiladas por Atenea y BC para caracterizar a la población beneficiaria y estimar el efecto del programa en su dominio del inglés.

-Puntajes de la prueba de inglés: nuestro resultado principal es el nivel de conocimiento del inglés medido por la prueba EnglishScore, que evalúa el rendimiento en el habla, la escritura, la gramática y el vocabulario como se describió en la sección 2.2. Los estudiantes en el grupo de tratamiento que se inscriben en EwB toman la prueba EnglishScore antes de comenzar su curso para ser asignados a un nivel de inglés (aquellos que no toman la prueba son asignados al nivel de entrada y, a veces, son reasignados por el instructor en la primera semana de clase). Los estudiantes del grupo de control no fueron evaluados al inicio del estudio, pero sí fueron evaluados al final (igual que los del grupo de tratamiento). La recopilación de datos finales se llevó a cabo entre el 25 de abril y el 26 de mayo de 2024, a través de una campaña intensiva de comunicación e incentivos destinada a motivar a los estudiantes de ambos grupos a realizar la evaluación. De los 3641 estudiantes a los que se les ofreció un cupo en el programa (el grupo de tratamiento), 1030 tomaron la prueba (28 %); de los 2298 estudiantes a los que no se les ofreció un cupo (el grupo de control potencial), 757 (33%) lo hicieron.

En la tabla 2 se muestran las estadísticas descriptivas de la prueba EnglishScore para toda la muestra, desagregadas por género y NSE. Existe una amplia variación en el rendimiento de los estudiantes, con un promedio de 218 —equivalente a un puntaje en el rango inferior del nivel A2— y una DE de más de 100 puntos —más de un nivel completo en el MCER—. En promedio, las mujeres se desempeñan ligeramente peor que los hombres, mientras que existe una clara correlación entre el NSE y el puntaje en las pruebas.

-Datos administrativos de Atenea: cuando las personas solicitan el programa TALAÚ, reportan información personal que se utiliza para calcular el índice de vulnerabilidad en el que se basa la asignación. Esta información incluye edad, género, NSE medido por el Sisbén (un puntaje que mide la vulnerabilidad económica y se utiliza para orientar programas sociales), grupo étnico, además de que se registra si las personas son víctimas de conflictos o violencia de género, y si han participado en programas gubernamentales anteriores. Como se mencionó

anteriormente, los solicitantes proporcionan toda la información, que luego se verifica mediante registros gubernamentales (excepto si son transgénero).

-Información administrativa del British Council: el BC recopila datos administrativos sobre la implementación del programa. Los RT registran la asistencia de los estudiantes, ya que recibir un certificado de participación en el curso depende de cumplir con el requisito mínimo de asistencia del 60 %. Los estudiantes reciben puntos de asistencia adicionales por completar actividades extracurriculares en inglés (como participar en clubes de conversación), lo que significa que un puntaje de asistencia puede ser mayor al 100 %. Además, el BC recopila datos sobre los RT, incluidas las observaciones en el aula y las evaluaciones realizadas por los MT.

-Datos de la encuesta final para los estudiantes tratados: el BC realiza una encuesta a todos los estudiantes de EwB al final del programa, recopilando información sobre la satisfacción del curso y otras características de los estudiantes. Las preguntas indagan sobre los motivos de inasistencia, los dispositivos utilizados para las lecciones en línea y otros temas relacionados, pero esta información solo está disponible para los estudiantes del grupo de tratamiento.

3.3. Estadística descriptiva de la muestra

En la tabla 1 se muestran las características de la muestra, comparando a los estudiantes de los grupos de control y tratamiento. El grupo de tratamiento está compuesto por las 3641 personas que solicitaron tanto los cursos de BPO/Gastronomía como el curso de inglés antes del corte, mientras que el grupo de control está compuesto por las 2298 personas que solicitaron ambos cursos después del corte. La tabla muestra la media y la desviación estándar de cada variable para cada grupo, así como la diferencia de medias entre ambos.

En general, aproximadamente el 60 % de la muestra evaluada está compuesta por mujeres, el 4,5 % pertenece a un grupo étnico y el 0,05 % tiene una discapacidad. En cuanto al NSE, el 59 % se clasifica según el índice medio de pobreza proxy del Sisbén, de los cuales el 26 % se encuentra en los dos grupos más vulnerables (A o B). Además, el 4,7 % ha sido víctima del conflicto armado y el 3,2 % ha sido víctima de violencia de género.

Las columnas 2 y 3 de la tabla 1 presentan estadísticas descriptivas desagregadas si los individuos estaban en el grupo de control o de tratamiento. La columna 4 muestra que existen diferencias estadísticamente significativas en la proporción de mujeres e individuos que

pertenecen a un grupo étnico entre los estudiantes en los grupos de control y tratamiento, con un poco menos de mujeres y más estudiantes pertenecientes a grupos étnicos en el grupo de tratamiento en comparación con el grupo de control. Todas las demás variables están balanceadas. Si bien no esperamos que los dos grupos estén necesariamente balanceados, dado que el tratamiento no se asigna al azar, la falta de diferencias sistemáticas sugiere que, incluso antes del emparejamiento, los grupos no parecen ser sistemáticamente diferentes en las características observadas.

3.3.1. La muestra de individuos que presentaron la prueba EnglishScore

Desafortunadamente, no todos los estudiantes de nuestra muestra tomaron la prueba EnglishScore. Dado que nuestro análisis se limita necesariamente al grupo de individuos para los que tenemos datos de resultados, es importante evaluar hasta qué punto la muestra de estudiantes que tomaron la prueba difiere de aquellos que no lo hicieron para comprender las limitaciones de validez externa de nuestro programa.

Las columnas 5-7 de la tabla 1 comparan las características de los estudiantes que tomaron la prueba EnglishScore con las de los que no lo hicieron. Aunque esperamos que quienes toman la prueba difieran de los que no lo hacen, observamos solo ligeras diferencias entre los dos grupos. Los grupos son estadísticamente diferentes solo en el hecho de que los que no toman la prueba tienen tres puntos porcentuales más de probabilidad de pertenecer a un grupo de bajo NSE (Sisbén A o B), y aproximadamente un punto porcentual más de probabilidades de ser víctimas del conflicto o de violencia de género. Esto sugiere que, en promedio, los estudiantes que tomaron y no tomaron la prueba son similares en la mayoría de las variables que observamos, y los que no tomaron la prueba son un poco más vulnerables.

3.3.2. Tasas de asistencia

No todas las personas a las que se les ofreció un cupo en el curso de inglés decidieron inscribirse, y no todas las que se inscribieron asistieron activamente a clases. El panel B de la tabla 1 muestra las tasas promedio de asistencia a los cursos de inglés para los estudiantes de los grupos de tratamiento y control. Como es natural, estos datos solo están disponibles para los que se inscribieron en el curso EwB. Como muestra la columna 2, de las 3641 personas a las que se les ofreció el curso EwB (el grupo de tratamiento), solo se inscribieron 2149. La tasa de asistencia promedio entre los que se inscribieron fue solo del 45,6 %, lo que refleja el hecho de que algunos de los que se inscribieron nunca asistieron a una sola clase, mientras que otros abandonaron a lo largo del curso. Por otro lado, 40 estudiantes del grupo control se inscribieron

en EwB (este número se eleva a 78 si incluimos a los tratados en cohortes anteriores de TALAÚ) y su tasa de asistencia promedio fue del 40 %.

Las columnas 5-7 del panel B (tabla 1) muestran que las tasas de asistencia promedio para aquellos que tomaron el examen de salida son mucho más altas que para quienes no lo hicieron. Quienes tomaron el examen de salida tuvieron tasas de asistencia promedio 70 puntos porcentuales más altas que quienes no lo hicieron. Es importante tener esto en cuenta, ya que afecta nuestra validez externa: estamos estimando los resultados de una muestra de estudiantes con tasas de asistencia significativamente superiores a la media. Por lo tanto, los resultados que encontramos pueden no extrapolarse a los estudiantes con tasas de asistencia más bajas, que pueden diferir sistemáticamente en términos de motivación o barreras para la participación.

3.4. Metodología: estimadores de emparejamiento

Para estimar el efecto causal del programa, nos basamos principalmente en estimadores de emparejamiento, que aprovechan la decisión de dividir la cohorte en función de una fecha límite arbitraria. La intuición detrás de un estimador de emparejamiento es utilizar los datos disponibles para identificar, para cada individuo tratado, el individuo en el grupo de control que le es más similar en términos de características observables. Esto crea un grupo de control que es similar al grupo de tratamiento en características observables (y con suerte no observables), que se puede usar para estimar el aprendizaje contrafactual de inglés del grupo de tratamiento.

Los estimadores de emparejamiento se basan en el supuesto de independencia condicional: condicionada a las características en las que se realiza el emparejamiento, la asignación de tratamiento es independiente de los resultados potenciales de los individuos. En otras palabras, esto requiere asumir que los individuos en los grupos de control y tratamiento, que tienen características observables similares, también habrían tenido niveles similares de inglés en ausencia de tratamiento. En programas con autoselección, este es un supuesto fuerte y, a menudo, poco realista. Cuando el tratamiento es completamente voluntario, incluso si podemos emparejar perfectamente a los individuos en todas las características observables, el hecho de que un individuo elija el programa mientras que el otro no sugiere que pueden diferir en características importantes no observables (como la motivación para aprender inglés), lo que podría conducir a un sesgo en las estimaciones de los efectos del programa. Sin embargo, el beneficio del cambio de regla en el contexto de nuestro programa de inglés hace plausible un estimador de emparejamiento, ya que tanto los individuos de control como los de

tratamiento mostraron interés en aprender inglés. Lo que explica esto es que la razón por la que los individuos del grupo de control no participaron en el programa no se debió a una decisión personal, sino a la decisión de la agencia de no ofrecer el programa a este segundo grupo debido a restricciones presupuestarias. Podríamos suponer razonablemente que, dada la opción de tomar el programa, los del grupo de control lo habrían hecho (o al menos en la misma medida que los del tratamiento).

En este contexto, el supuesto de independencia condicional es razonable. Dado que sabemos que las características observables, como el género y el estatus socioeconómico, están asociadas con los niveles de inglés, los grupos de control y tratamiento pueden diferir en la medida en que su composición difiere en estas características. En ese sentido, la tabla 1 muestra que la composición de los grupos es ligeramente diferente en algunas características; sin embargo, una vez que corregimos este desequilibrio haciendo coincidir a los individuos en el grupo de tratamiento con el individuo de control que es más similar a ellos en características observables, sería razonable suponer que, en ausencia de tratamiento, los grupos de control y tratamiento también habrían tenido niveles similares de dominio del inglés. Por lo tanto, cualquier diferencia que observemos en la muestra emparejada podría atribuirse razonablemente al programa en sí.

Empleamos dos tipos de estimadores: emparejamiento de vecino más cercano y emparejamiento con el puntaje de propensión. El estimador de emparejamiento de vecino más cercano encuentra a los individuos del grupo de control que son más similares a los individuos del grupo de tratamiento de acuerdo con un conjunto de covariables, mientras que el estimador de emparejamiento de propensión estima la probabilidad de que una observación reciba el tratamiento dado un conjunto de covariables (la puntuación de propensión), y con esa probabilidad encuentra el grupo de individuos de control. El emparejamiento debe realizarse en las características observables utilizadas para asignar el tratamiento, por lo que emparejamos con dos conjuntos de variables: primero, emparejamos por la puntuación agregada utilizada para la asignación del tratamiento, que debería ser una estadística suficiente de todo el conjunto de características y es en sí misma una especie de puntuación de propensión, pero como verificación de robustez también coincidimos con todas las características subyacentes utilizadas para calcular la puntuación.

Una preocupación potencial con la metodología de emparejamiento es que las personas que solicitaron el programa muy temprano o muy tarde pueden diferir en características no

observables que también podrían afectar su dominio del inglés, lo que podría sesgar las estimaciones. Por ejemplo, si las personas más diligentes y motivadas solicitan el cupo antes, habría una correlación entre la diligencia y la fecha de solicitud que podría sesgar nuestras estimaciones. Podemos probar esto indirectamente utilizando nuestros datos de asistencia en la cohorte tratada. Asumiendo que el momento de la solicitud es un indicador de diligencia o motivación, esperaríamos que los estudiantes, dentro del grupo de tratamiento que solicitaron antes, tuvieran más probabilidades de tomar el curso y asistir a más clases (quizás condicionadas a características observables) que aquellos que solicitaron más cerca del límite.

Sin embargo, nuestra evidencia muestra que este no parece ser el caso. La tabla 3 presenta los resultados de una regresión de mínimos cuadrados ordinarios (MCO), que examina la asociación entre el tiempo de solicitud y la aceptación y asistencia al curso. Encontramos muy poca asociación entre el momento de la solicitud y las tasas de aceptación o asistencia al curso. Las columnas 1 y 2 presentan la relación entre el tiempo de solicitud, medido en horas, y haber tomado el curso (definido como asistencia al menos a una clase). El coeficiente se estima con precisión en 0, tanto sin controles (columna 1) como con controles (columna 2), lo que sugiere que la solicitud temprana no se asocia con un aumento en la probabilidad de que las personas se inscriban en el curso cuando se les ofrece un cupo. Las columnas 3 y 4 muestran los resultados de la asociación entre el tiempo de solicitud y las tasas de asistencia, donde la asistencia se indica como 0 para aquellos que no se inscribieron en el curso. El coeficiente es positivo y estadísticamente significativo, pero la magnitud es extremadamente pequeña: hacer la solicitud una hora más tarde se asocia con un aumento de 0,01 puntos porcentuales en el puntaje de asistencia.

Sin embargo, el hecho de que el coeficiente sea positivo sugiere que la solicitud más tarde, no antes, se asocia con una mayor asistencia, que es lo contrario de lo que hubiéramos esperado si la asistencia sirviera como un indicador de diligencia, y si los estudiantes que solicitaron antes fueran más diligentes. Esto sugeriría, en todo caso, que quienes solicitaron el cupo más tarde son más diligentes y nuestras estimaciones estarían sesgadas a la baja. Es decir, estaríamos estimando un límite inferior en los efectos de aprendizaje. Sin embargo, dada la magnitud del efecto, una conclusión más precisa puede ser que no existe una relación significativa entre las tasas de asistencia y la fecha de solicitud entre los estudiantes tratados. En cualquier caso, podemos concluir que no hay evidencia que sugiera que los estudiantes que solicitaron antes sean más diligentes, como lo indican las tasas de asistencia.

3.5. Metodología alternativa: regresión discontinua

Una solución alternativa al sesgo potencial resultante de que los estudiantes que solicitaron más tarde y más temprano tengan diferentes características no observables es implementar un diseño de regresión discontinua (RD), que compara a los estudiantes que hicieron la solicitud cerca de la fecha límite arbitraria. En general, la RD ofrece una estimación contrafactual *más limpia*, porque tiene supuestos de identificación subyacentes más débiles. Sin embargo, también es muy exigente en términos de datos cercanos al límite, lo que reduce su poder estadístico. A la luz de esta limitación, nuestra principal estrategia empírica es el estimador de emparejamiento, y presentamos el estimador de RD como una estrategia complementaria.

La intuición detrás del diseño de RD es que, incluso si los estudiantes que solicitaron muy temprano y muy tarde pueden diferir en características importantes no observadas, aquellos que solicitaron justo antes del límite deberían ser esencialmente idénticos a los que solicitaron más tarde, particularmente dado que los solicitantes no conocían la fecha límite. Por lo tanto, esperamos que, en ausencia de tratamiento, las personas cercanas al límite tengan niveles similares de inglés. Al restringir el análisis a una ventana de individuos alrededor de la fecha límite y estimar si hay un salto discontinuo en los niveles de inglés alrededor de dicha fecha, podemos estimar el efecto causal del programa en el aprendizaje.

El supuesto de identificación principal para el diseño de RD es que hay continuidad en todas las demás características a lo largo de la fecha límite, de modo que los estudiantes que solicitaron justo después del corte son un buen contrafactual para los estudiantes justo por encima del corte. El panel B de la figura 5 muestra el balance en las pocas características disponibles a través de este umbral. Además, es necesario que los estudiantes no puedan manipular el valor de la variable en ejecución alrededor del límite. En este caso, ello implicaría presentar una solicitud justo antes de la fecha límite para asegurar un cupo en el programa. Afortunadamente, la fecha límite se estableció *ex post*, nunca se había hecho antes, y los estudiantes no sabían que postularse antes de esta fecha límite aumentaría sus posibilidades de recibir un cupo en el programa de inglés. Como resultado, no creemos que pudiera haber habido manipulación en torno a esta fecha límite.

4. Resultados

4.1. Resultados principales

4.1.1. El efecto del intento de tratamiento (IIT)

La figura 2 presenta los principales resultados del programa EwB sobre el dominio del idioma (las tablas 5 y 6 presentan los mismos resultados en formato de tabla). El panel A presenta las estimaciones de intención de tratar, que muestran el efecto de que se le ofrezca un cupo en el curso de inglés como resultado de haber presentado una solicitud antes del 27 de agosto sobre el dominio del inglés, independientemente de si el estudiante finalmente se inscribió o no en el programa.

El resultado de interés en todos los modelos es la puntuación estandarizada en el examen EnglishScore administrado al final del curso tanto a los individuos de tratamiento como a los de control. La primera barra muestra los resultados de referencia de una regresión de MCO, donde a la variable dependiente se le ha ofrecido un cupo en el curso EwB. En la medida en que las personas a las que les ofrecieron un curso difieren de las que solicitaron un poco más tarde y no lo hicieron, este MCO estará sesgado.

Las otras barras muestran las estimaciones de intención de tratar, utilizando los dos estimadores de emparejamiento: las dos primeras barras muestran los resultados del estimador de emparejamiento con puntuación de propensión (PSM), y las dos últimas barras muestran los resultados del estimador de emparejamiento por vecino más cercano. En ambos casos, emparejamos dos conjuntos de covariables: 1) solo en la puntuación utilizada para asignar el programa, y 2) en el conjunto completo de características individuales utilizadas para calcular la puntuación de la asignación. La tabla 4 y la figura 1 muestran las pruebas de balance y el soporte común de las muestras brutas y emparejadas para el estimador de PSM, mostrando que logra equilibrar las muestras. Los resultados son casi idénticos en ambos estimadores de emparejamiento y similares entre los dos conjuntos de variables de emparejamiento. De acuerdo con nuestra estimación preferida, que es el estimador de PSM, usando solo la probabilidad de asignación, encontramos que haber recibido la oferta de un cupo en el curso

aumenta el puntaje de inglés en 0,195 DE. Esta ganancia equivale a 1/5 de un cambio de nivel completo del MCER (por ejemplo, pasar de A1 a A2)¹⁶.

4.1.2. Teniendo en cuenta el cumplimiento imperfecto

Aunque el hecho de que se ofrezca un cupo en el curso mejora los niveles de inglés, no todos los estudiantes a los que se les ofreció un cupo decidieron tomar el curso, y muchos de los que comenzaron lo abandonaron luego. Además, algunos individuos del grupo de control (a los que no se les ofreció un cupo en el curso) habían participado en cohortes anteriores de EwB y, por lo tanto, fueron tratados. El resultado de este cumplimiento imperfecto es que nuestras estimaciones de ITT subestiman el efecto del curso para aquellos que lo tomaron efectivamente.

El panel B de la figura 2 y la tabla 6 presentan los resultados de los estimadores que dan cuenta del cumplimiento imperfecto. Logramos esto de dos maneras: primero, usamos el mismo estimador de emparejamiento de puntaje de propensión, pero definimos los grupos de tratamiento y control como haber tomado efectivamente el curso. Para esto usamos tres definiciones de tomar el curso: 1) inscribirse en el curso, independientemente de si asistieron a alguna clase, 2) asistir al menos a una clase, o 3) asistir al menos al 20 % de las clases. La tabla A1 en el Apéndice muestra las verificaciones de equilibrio de las muestras brutas y emparejadas para cada uno de estos estimadores de PSM. Al igual que en el caso del ITT (presentado en la subsección 4.1.1.), el PSM logra equilibrar las muestras y logra relaciones de varianza cercanas a uno.

La primera barra del panel B en la figura 2 muestra nuestra estimación preferida (emparejamiento con PSM en el puntaje de asignación) de la estimación de ITT que se muestra en el panel A como referencia. Las barras 2-4 muestran los resultados del mismo estimador de emparejamiento con el puntaje de propensión, pero con el tratamiento definido como tomar el curso en cada una de las formas descritas anteriormente. Encontramos que, según el estimador de PSM, inscribirse en el curso aumentó los puntajes de inglés de los estudiantes en 0,23 DE, asistir al menos a un curso aumentó los puntajes de los estudiantes en 0,24 DE y asistir al menos al 20 % de las clases aumentó los puntajes de las pruebas de los estudiantes en 0,27 DE.

¹⁶ Como recordatorio, una DE de la variable de resultado es de 105 puntos en la puntuación estandarizada de EnglishScore, y 100 puntos es la distancia entre los diferentes niveles del MCER, por lo que 0,195 DE equivalen a una mejora en 1/5 de un nivel del MCER.

Además, ejecutamos un modelo simple de mínimos cuadrados en dos etapas (2SLS), en el que instrumentamos la aceptación del programa con la oferta del programa como resultado de presentar la solicitud antes de la fecha límite. La validez de este instrumento se basa en el supuesto de que la solicitud antes o después de la fecha límite no se correlaciona con las variables omitidas, lo cual es un supuesto exigente, pero podría ser plausible dado que la fecha límite no era conocida de antemano por los estudiantes y se decidió arbitrariamente después de que suficientes estudiantes hubieran solicitado el cupo. Además, si bien existe una preocupación potencial de que los estudiantes que solicitaron el cupo muy temprano podrían ser diferentes de los que solicitaron más tarde, de manera que también podrían afectar el aprendizaje del inglés (por ejemplo, pueden ser más diligentes), nuestras verificaciones de validez anteriores indican que este no parece ser el caso (ver tabla 3). Si bien estos resultados deben tomarse con un grano de sal, los resultados de 2SLS son muy similares a los de las estimaciones de PSM, con efectos que van desde 0,27 DE para los estudiantes que se inscribieron hasta 0,3 DE para los estudiantes que asistieron a más del 20 % de las clases.

La tabla 7 y la figura 5 presentan los resultados del diseño de RD. Desafortunadamente, debido a la baja densidad de datos alrededor de la fecha límite, el diseño de RD tiene poca potencia (panel I de la figura 5). Si bien los resultados son similares en magnitud a los encontrados en el estimador de emparejamiento, no son estadísticamente significativos. La columna 1 de la tabla 7 presenta los resultados de la primera etapa del RD, mostrando que la aplicación justo por debajo del punto de corte aumenta la probabilidad de retomar el programa EwB en 84 puntos porcentuales (valor de $p < 0,01$) (la representación gráfica de la primera etapa se puede ver en el panel III de la figura 5). La columna 2 muestra los resultados de forma reducida de que se ofrezca el programa en los puntajes de las pruebas de inglés, utilizando una regresión lineal local como es estándar en la literatura. La columna 3 presenta los resultados de 2SLS, en la que el tratamiento (definido como asistir a más del 20 % de las clases) se instrumenta con la elegibilidad inducida por el corte del programa. Todas las especificaciones de RD restringen la muestra al ancho de banda óptimo como lo sugieren Calonico y Cattaneo (2014). Los coeficientes para la forma reducida y 2SLS son similares en magnitud y consistentes con los resultados del estimador de emparejamiento; sin embargo, no son estadísticamente significativos debido a su bajo poder estadístico.

4.2. Efectos heterogéneos

A continuación, exploramos si el efecto del programa difiere según las características del estudiante o del profesor. Dado que las tasas de asistencia y las características de los profesores solo están disponibles para los estudiantes tratados, no podemos estimar los efectos heterogéneos a través de un modelo interactuado. En cambio, calculamos los efectos del tratamiento por separado para individuos con tasas de asistencia bajas y altas, así como para aquellos con baja y alta calidad de profesores. Para mantener la coherencia, realizamos un análisis de muestra dividida similar para las otras características para las que tenemos medidas para todos los estudiantes: género, NSE y sector del mercado laboral.

El panel A de la figura 3 muestra los efectos heterogéneos del tratamiento según varias características de los estudiantes, incluido género, NSE, sector de su programa de formación profesional (gastronomía o BPO) y tasas de asistencia (con una alta asistencia definida como superior a la tasa media). En esta figura, el tratamiento se define como haber asistido al menos a una clase, mientras que la figura 4 muestra el conjunto completo de resultados para las diferentes definiciones de tratamiento: las estimaciones de ITT de que se le ofreció un cupo (panel I), se inscribió en la clase (panel II), tomó al menos una clase (panel III) y tomó al menos el 20 % de las clases (panel IV). Aunque encontramos tamaños de efecto más altos para los estudiantes hombres, los de hogares de bajo NSE, los estudiantes del sector de gastronomía y los estudiantes con alta asistencia, lo que sugiere efectos diferenciales, ninguna de las diferencias estimadas es estadísticamente significativa. Estos resultados sugieren que el programa fue efectivo para aumentar los puntajes de inglés para todos los estudiantes, incluidos los de bajas condiciones socioeconómicas y las mujeres, que pueden enfrentar mayores barreras para asistir debido a las responsabilidades de cuidado.

También exploramos si las características de los profesores afectaron el aprendizaje de los estudiantes. El panel B presenta efectos de tratamiento heterogéneos, según el desempeño docente medido, utilizando los criterios de evaluación que tiene el BC para evaluar a sus docentes. La calidad de las clases de los profesores se mide a lo largo de cuatro dimensiones: 1) planificación de lecciones, 2) entrega de lecciones, 3) responsabilidad y 4) puntualidad. Construimos un agregado simple de estas cuatro medidas, al que nos referimos como la puntuación general, y comparamos los efectos del tratamiento para los estudiantes asignados a un profesor por encima de la media con los de los asignados a uno por debajo de la media. El BC asigna los estudiantes a los profesores de forma centralizada en función de la disponibilidad

de tiempo de los estudiantes y su nivel, lo que hace que sea extremadamente improbable que los estudiantes estén seleccionando profesores en particular.

Sorprendentemente, encontramos que el puntaje compuesto del profesor no se correlaciona con las diferencias en el aprendizaje de los estudiantes. Sin embargo, los estudiantes con profesores por encima de la media, en términos de calidad de la impartición de sus clases, aprenden significativamente más (en 0,35 DE) que aquellos con profesores por debajo de la media, cuyas ganancias no son estadísticamente significativas. Esto sugiere que la calidad de las clases, particularmente la calidad de la impartición de las lecciones, afecta el aprendizaje de los estudiantes. Observamos un patrón similar para los estudiantes con un profesor que tiene una calificación superior al promedio en la planificación de lecciones; sin embargo, las diferencias entre los dos grupos no son estadísticamente significativas en este caso. Curiosamente, vemos el patrón opuesto para las dos dimensiones restantes, aunque las diferencias no son estadísticamente significativas en ninguno de los casos.

5. Conclusión

Este artículo estima el efecto causal de un programa que brinda oportunidades de aprendizaje del idioma inglés a personas con vulnerabilidad económica en edad de trabajar en Bogotá, en el contexto de una iniciativa de formación profesional destinada a cerrar la brecha de habilidades existente entre la fuerza laboral y las necesidades del sector productivo. El programa es de intensidad relativamente baja: los participantes toman 5 horas de clases de inglés por semana durante 6 meses y tienen la oportunidad de participar en actividades complementarias para practicar su inglés.

Nuestros resultados muestran que el programa tiene éxito en mejorar el dominio del inglés de los participantes: el hecho de que se les ofrezca un cupo en el curso aumenta el puntaje de inglés en 0,195 DE. Además, asistir al menos al 20 % de las clases aumenta los puntajes de inglés en 0,27 DE. Una extrapolación lineal de este resultado implica que llevar al participante promedio del nivel A2 en la terminología del MCER al B1 requeriría aproximadamente 3,8 cursos de EwB. Examinamos la heterogeneidad por características de estudiantes y profesores y no encontramos diferencias significativas entre las diversas características de los estudiantes, lo que indica que una amplia gama de estudiantes se beneficia del programa. Estos hallazgos contribuyen a la evidencia limitada sobre intervenciones efectivas que mejoran el dominio del inglés entre adultos, particularmente aquellos de contextos vulnerables.

A partir de estos resultados, podemos sacar tres conclusiones. En primer lugar, es posible lograr resultados en el aprendizaje del inglés entre adultos con cursos de baja intensidad. Esta es una buena noticia, dados los desafíos que enfrentan los países en desarrollo cuando intentan mejorar las habilidades de su población activa. Sin embargo, una segunda conclusión es que esta mejora requiere una inversión monetaria significativa y esfuerzos sustanciales en el sector educativo, ya que la fuerza laboral docente de inglés de alta calidad necesaria para una actualización masiva no está disponible en la mayoría de los países en desarrollo.

La tercera conclusión derivada de nuestros hallazgos se refiere a la calidad del programa, específicamente a la calidad de los profesores. EwB es un programa de baja intensidad, pero incorpora componentes de alta calidad: brinda acceso a recursos digitales (como libros y actividades en línea) y, lo que es más importante, a profesores de alta calidad. El programa cuenta con un sólido sistema de selección, capacitación y monitoreo de profesores, lo que garantiza una instrucción de alta calidad. Nuestros resultados muestran que tener un profesor de alta calidad mejora los puntajes de inglés en 0,35 DE, lo cual es un efecto sustancial en el contexto de las intervenciones educativas (Kraft, 2020). Esto tiene implicaciones significativas para el diseño de otros programas de enseñanza de inglés: el enfoque debe estar en la calidad de los profesores.

Referencias

- Adamchik, V. A., Hyclak, T. J., Sedlak, P. y Taylor, L. W. (2019). El salario vuelve al dominio del inglés en Polonia. *Revista de Investigación Laboral*, 40, 276–295.
- Agirdag, O. (2014). Los efectos a largo plazo del bilingüismo en los hijos de la inmigración: bilingüismo estudiantil e ingresos futuros. *Revista Internacional de Educación Bilingüe y Bilingüismo*, 17, 449–464.
- Azam, M., Chin, A. y Prakash, N. (2013). Los regresos a las habilidades del idioma inglés en la India. *Desarrollo Económico y Cambio Cultural*, 61, 335–367.
- Bleakley, H. y Chin, A. (2004). Habilidades lingüísticas e ingresos: evidencia de inmigrantes infantiles. *La Revisión de Economía y Estadística*, 86, 481–496.
- Borradaile, K., Martínez, A. y Schochet, P. (2021). *Estrategias de educación de adultos: Identificación y creación de pruebas de eficacia* [Informe técnico]. Mathematica Policy Research.
- Calonico, S., Cattaneo, M. D., y Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4), 909–946.
- Casale, D. y Posel, D. (2011). Dominio del idioma inglés e ingresos en un país en desarrollo: el caso de Sudáfrica. *La Revista de Socioeconomía*, 40, 385–393.
- Chakraborty, T. y Bakshi, S. K. (2016). Prima de idioma inglés: evidencia de un experimento de política en India. *Revista de Economía de la Educación*, 50, 1–16.
- Chiswick, B. R. (1991). Hablar, leer y ganar dinero entre los inmigrantes poco calificados. *Revista de Economía Laboral*, 9, 149–170.
- Condelli, L., Cronen, A., Bos, J., Tseng, F. y Altuna, J. (2010). *El impacto de una intervención de lectura para estudiantes adultos de ESL con bajo nivel de alfabetización*. NCEE 2011-4003. Centro Nacional de Evaluación de la Educación y Asistencia Regional.
- Cronquist, K. y Fiszbein, A. (2017). *Aprendizaje del idioma inglés en América Latina*. Diálogo Interamericano.
- Di Paolo, A. y Tansel, A. (2015). Retornos a las habilidades en idiomas extranjeros en un país en desarrollo: el caso de Turquía. *La Revista de Estudios del Desarrollo*, 51, 407–421.
- Dolgova, N. y Tyler, A. (2019). Aplicaciones de los enfoques basados en el uso para la enseñanza de idiomas. En X. Gao (ed.), *Segundo manual de enseñanza del idioma inglés* (pp. 939–961). Salmer.
- Fabo, B., Beblavý, M. y Lenaerts, K. (2017). La importancia de las competencias lingüísticas en los mercados laborales de Europa Central y Oriental: evaluación basada en datos de portales de empleo en línea. *Empírica*, 44, 487–508.
- Gao, X. (2019). *Segundo manual de enseñanza del idioma inglés*. Salmer.

- Garrouste, C. (2008). Habilidades lingüísticas y rendimientos económicos. *Futuros de Políticas en la Educación*, 6, 188–201.
- Guo, Q. y Sun, W. (2014). Rendimientos económicos del dominio del inglés para graduados universitarios en China continental. *Revisión Económica de China*, 30, 290–300.
- Kraft, M. A. (2020). Interpretación de los tamaños del efecto de las intervenciones educativas. *Investigador Educativo*, 49(4), 241–253.
- Laboratorio de Economía de la Educación. (2023). *Inglés, el factor de competitividad pendiente en Colombia* [Informe de Análisis Estadístico, 69]. Autor.
- Lang, K. y Siniver, E. (2009). El regreso al inglés en un país de habla no inglesa: inmigrantes rusos e israelíes nativos en Israel. *La Revista B.E. de Análisis y Política Económica*, 9.
- Liwiński, J. (2019). La prima salarial por habilidades en idiomas extranjeros. *Empírica*, 46, 691–711.
- Muhawenayo, J., Habimana, O. y Heshmati, A. (2022). Conocimientos de idiomas extranjeros e ingresos en el mercado laboral en Ruanda. *Revista de Educación y Trabajo*, 35, 719–734.
- Rojas, J. y Hernández-Fernández, J. E. (2018). *Políticas públicas de inglés en América Latina: Buscando innovación y mejora sistémica en la enseñanza de inglés de calidad*. British Council México.
- Stöhr, T. (2015). Los retornos al uso de lenguas extranjeras específicas: evidencia de Alemania. *Revista de Comportamiento Económico y Organización*, 116, 1–20.
- Takeuchi, O. (2019). Estrategias de aprendizaje de idiomas: ideas del pasado y direcciones para el futuro. En X. Gao (ed.), *Segundo manual de enseñanza del idioma inglés* (pp. 683–702). Salmer.
- Wang, H., Smyth, R. y Cheng, Z. (2017). El Económico vuelve a dominar el inglés en China. *Revista Económica de China*, 43, 91–104.
- Wyse, D., Andrews, R. y Hoffman, J. V. (2010). *El manual internacional de Routledge sobre la enseñanza del inglés, el idioma y la alfabetización*. Routledge.

Tablas y figuras

Tabla 1. Estadística descriptiva de la muestra y submuestras

		Grupo de tratamiento			Datos de resultados		
	Total	Control	Tratamiento	Prueba t por pares	Sin examen de salida	Tomó el examen de salida	Prueba t por pares
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Datos demográficos							
Mujer	0,601 (0,490)	0,629 (0,483)	0,583 (0,493)	0,046***	0,606 (0,489)	0,587 (0,493)	0,019
Grupo étnico	0,044 (0,206)	0,034 (0,181)	0,051 (0,220)	-0,02***	0,042 (0,201)	0,049 (0,216)	-0,007
Discapacidad	0,005 (0,069)	0,005 (0,069)	0,005 (0,068)	0,000	0,005 (0,071)	0,004 (0,062)	0,001
NSE bajo (Cualquier Sisbén)	0,591 (0,492)	0,598 (0,490)	0,587 (0,493)	0,011	0,590 (0,492)	0,593 (0,491)	-0,003
NSE más bajo (Sisbén A o B)	0,262 (0,440)	0,273 (0,446)	0,255 (0,436)	0,018	0,270 (0,444)	0,245 (0,430)	0,025**
Víctima de conflicto	0,047 (0,211)	0,047 (0,212)	0,046 (0,210)	0,001	0,049 (0,216)	0,041 (0,198)	0,008
Víctima de violencia de género	0,032 (0,177)	0,034 (0,182)	0,031 (0,173)	0,003	0,035 (0,185)	0,025 (0,157)	0,010**
N	5939	2298	3641		4152	1787	
B. Datos de asistencia (solo disponible para personas tratadas)							
Asistencia	0,455 (0,497)	0,403 (0,448)	0,456 (0,498)	-0,053	0,201 (0,309)	0,909 (0,445)	-0,708***
N	2189	40	2149		1404	785	

Nota: La tabla muestra las estadísticas descriptivas de la muestra completa y una comparación de medias entre diferentes submuestras. La columna 1 presenta las estadísticas descriptivas para toda la muestra, mientras que las columnas 2 y 3 desglosan los datos por grupo de tratamiento. Las columnas 5 y 6 desglosan la muestra según si los estudiantes tomaron el examen de salida. Las DE están entre paréntesis para todas estas columnas. Finalmente, las columnas 4 y 7 muestran el coeficiente de la diferencia entre los dos grupos, con significancia estadística denotada por estrellas: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$.

Fuente: Elaboración propia.

Tabla 2. Estadística descriptiva de los resultados del examen EnglishScore

	Media	DE	p25	p50	p75	N
Todo	218,6	110,2	136	190	287	1787
Hombres	230,1	113,8	143	210	314	738
Mujeres	210,4	106,9	135	176	275	1049
Sisbén A	196,5	92,4	134	173,5	254	80
Sisbén B	193,1	102,8	125	166	230	357
Sisbén C	207,6	103,2	127	175	275	491
Sisbén D	232,2	114,5	145,5	203	286,5	132
No Sisbén	238,4	115,4	152	212	319	727

Nota: La tabla proporciona estadísticas descriptivas de los resultados del examen EnglishScore para toda la muestra evaluada al final, así como desglosadas por género y NSE, según lo medido por Sisbén. El Sisbén es una prueba de disponibilidad de recursos que mide la vulnerabilidad económica y se utiliza para orientar los programas de protección social en Colombia. Asigna a los individuos a grupos donde A es el más vulnerable, mientras que los que no están en el Sisbén son los menos vulnerables.

Fuente: Elaboración propia.

Tabla 3. Asociación entre el tiempo de solicitud y la asistencia

	Tomó curso		Asistencia	
	(1)	(2)	(3)	(4)
Tiempo de aplicación	0,0000 (0,000)	-0,0000 (0,000)	0,0001** (0,000)	0,0001* (0,000)
Controles	No	Sí	No	Sí
Media de la variable dependiente	0,531	0,531	0,286	0,286
N	3,641	3,641	3,641	3,641
R cuadrado	0,000	0,021	0,002	0,023

Nota: La tabla muestra los resultados de una regresión MCO simple entre el tiempo que cada individuo solicitó al programa (en horas), si el individuo tomó el curso en las columnas 1 y 2 (definido como asistir al menos a una clase) y las tasas de asistencia en las columnas 3 y 4. Las columnas 1 y 3 no incluyen controles, mientras que las columnas 2 y 4 incluyen un conjunto completo de controles utilizados para los estimadores coincidentes. Los errores estándar robustos se muestran entre paréntesis, y las estrellas denotan la significancia estadística: ***p<0,01, **p<0,05, *p<0,1.

Fuente: Elaboración propia.

Tabla 4. Asignación de tratamiento de tablas de balance de PSM

	Diferencias estándar		Relación de varianza	
	Crudo (1)	Emparejado (2)	Crudo (3)	Emparejado (4)
<i>A. Emparejamiento solo en la puntuación global</i>				
Puntaje global	-0,243	-0,000	0,715	0,997
<i>B. Emparejamiento en todas las características</i>				
BPO	0,011	0,012	0,998	0,998
NSE bajo (Cualquier Sisbén)	-0,139	-0,010	1,058	1,004
NSE más bajo (Sisbén A o B)	-0,191	0,018	0,800	1,022
Étnico	0,077	-0,007	1,386	0,969
Mujer	0,007	0,021	0,997	0,991
Discapacidad	0,031	-0,000	1,705	1,000
Víctima de conflicto	-0,067	0,048	0,731	1,252
Violencia de género	-0,042	0,043	0,778	1,286
Recluso	-0,041	-0,000	0,514	1,000
Reto a la U	-0,065	0,026	0,691	1,162
Inmersión en media	-0,153	0,005	0,232	1,041
SENA	-0,234	-0,023	0,155	0,853
Reto SDIS	-0,037	0,039	0,739	1,353
Parceros	-0,124	-0,006	0,525	0,966
Programa deportivo	0,012	0,019	1,366	1,999
Programa de cuidadores	-0,107	-0,020	0,230	0,770

Nota: La tabla muestra el balance posterior al emparejamiento para la intención de tratamiento con el puntaje de propensión, donde el tratamiento se define como elegible para el programa como resultado de la solicitud antes de la fecha límite. El panel A muestra el balance después de emparejar solo en la puntuación global, mientras que el panel B muestra el balance después de emparejar en todas las características disponibles. Las columnas 1 y 2 muestran las diferencias estandarizadas de las muestras brutas y emparejadas, mientras que las columnas 3 y 4 muestran la relación de varianza de las muestras brutas y emparejadas.

Fuente: Elaboración propia.

Tabla 5. Estimaciones por intención de tratar en los puntajes de las pruebas

	MCO	Estimadores de emparejamiento			
		PSM		Vecino más cercano	
		Solo puntuación	Todas las Vars.	Solo puntuación	Todas las Vars.
	(1)	(2)	(3)	(4)	(5)
Tratamiento asignado	0,218*** (0,048)	0,195*** (0,049)	0,170*** (0,051)	0,195*** (0,049)	0,180*** (0,050)
Media de los controles	-0,000	-0,000	-0,000	-0,000	-0,000
N	1,879	1,879	1,879	1,879	1,879

Nota: La tabla muestra las estimaciones del coeficiente de intención de tratar de ser elegible para el programa EwB sobre dominio del inglés, según lo medido por la prueba estandarizada EnglishScore. La columna 1 muestra los resultados de una regresión MCO ingenua para la comparación, mientras que las columnas 2-5 muestran los resultados del emparejamiento con PSM y vecino más cercano. Para ambos estimadores de emparejamiento, presentamos dos especificaciones: emparejamiento solo en el puntaje utilizado para asignar a los estudiantes al programa (columnas 2 y 4) y emparejamiento en todas las características disponibles (columnas 3 y 5). Los errores estándar robustos se muestran entre paréntesis, y las estrellas denotan significancia estadística: ***p<0,01, **p<0,05, *p<0,1.

Fuente: Elaboración propia.

Tabla 6. Efecto de la aceptación del tratamiento en los puntajes de las pruebas

	ITT	PSM			2SLS		
		Registrado	Att >0 %	Att >20 %	Registrado	Att >0 %	Att >20 %
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Elegible	0,195*** (0,049)						
Tratado		0,229*** (0,048)	0,240*** (0,048)	0,265*** (0,048)	0,270*** (0,060)	0,279*** (0,061)	0,304*** (0,067)
Media de los controles	-0,000	0,006	0,007	0,011	0,006	0,007	0,011
Estadísticas F					4,025.6	3,448.3	2,521.4
N	1,879	1,879	1,879	1,879	1,879	1,879	1,879

Nota: La tabla muestra el efecto de la aceptación del tratamiento en el dominio del inglés medido por la prueba estandarizada EnglishScore. La columna 1 muestra la estimación de ITT utilizando nuestra especificación preferida (emparejamiento con PSM solo en la puntuación) para la comparación, mientras que las columnas 2-7 muestran el efecto de la aceptación del tratamiento utilizando el emparejamiento con PSM en la aceptación (columnas 2-4) y un modelo de mínimos cuadrados de dos etapas que utiliza la oferta de un cupo como resultado de la solicitud antes de la fecha límite como instrumento para la aceptación. Para ambos estimadores, utilizamos tres definiciones de aceptación del tratamiento: 1) inscribirse en el curso, 2) asistir al menos a una clase o 3) asistir al menos al 20 % de las clases. Los errores estándar se muestran entre paréntesis y las estrellas denotan significancia estadística: ***p<0,01, **p<0,05, *p<0,1.

Fuente: Elaboración propia.

Tabla 7. Resultados de la estimación de RD

	Primera etapa (1)	Forma reducida (2)	Segunda etapa (3)
Estimación de RD	-0,840*** (0,077)	-0,353 (0,356)	0,348 (0,351)
N	1,879	1,792	1,792

Nota: La tabla muestra los resultados de los coeficientes del análisis de RD. La columna 1 muestra los resultados de la primera etapa de la probabilidad de adoptar el programa EwB en función del tiempo de aplicación (variable en ejecución). La columna 2 muestra el efecto de forma reducida de ser elegible para el inglés. Todas las columnas restringen la muestra al ancho de banda óptimo según lo definido por Calonico y Cattaneo (2014). Los errores estándar robustos se muestran entre paréntesis, y las estrellas denotan significancia estadística: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$.

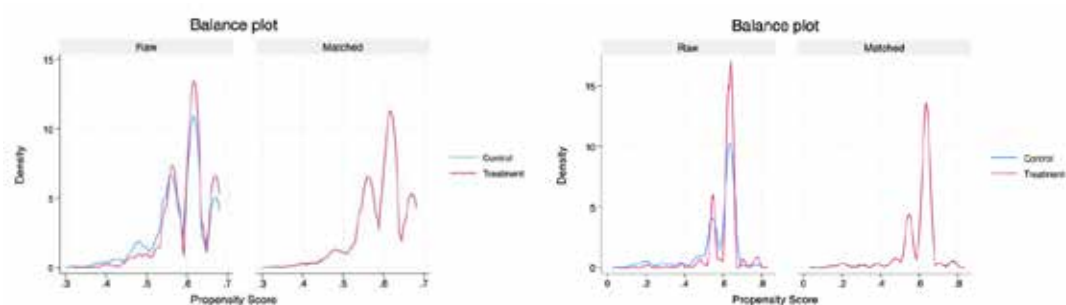
Fuente: Elaboración propia.

Figura 1. Gráficos de densidad para todos los partidos de PSM

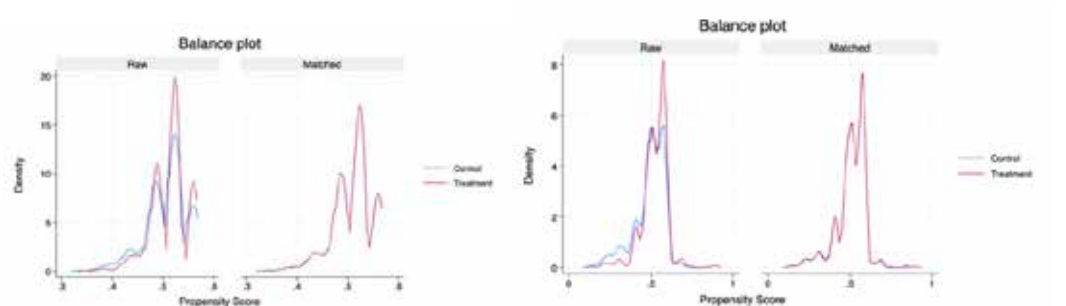
A. Puntaje global

B. Todas las variables

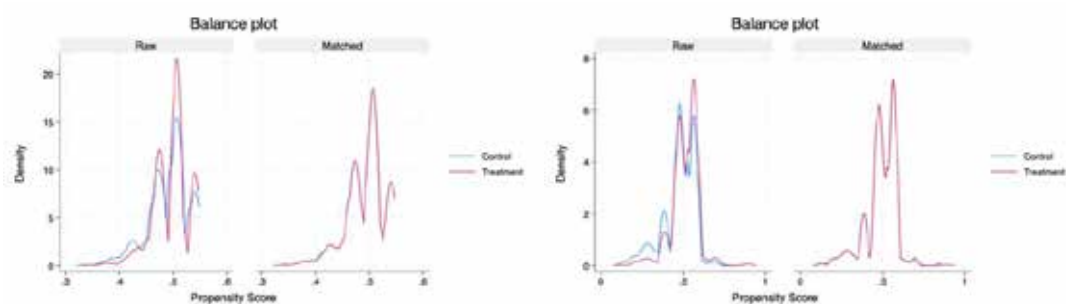
I. Asignación de tratamiento



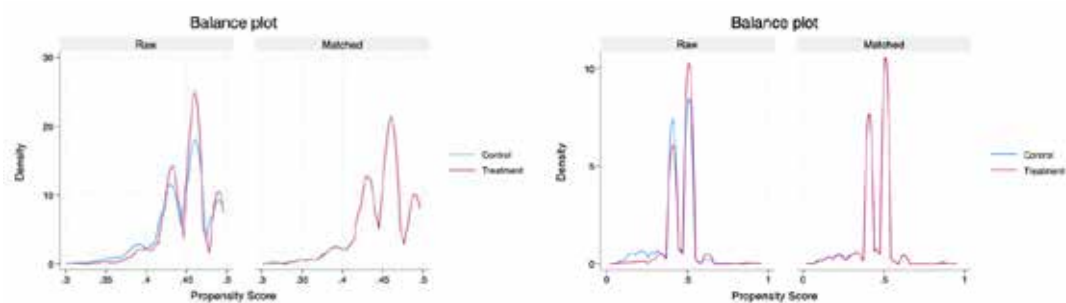
II. Aceptación del tratamiento definida como inscrito



III. La aceptación del tratamiento se define como tomar al menos una clase



IV. La aceptación del tratamiento se define como asistir al menos al 20% de las clases.

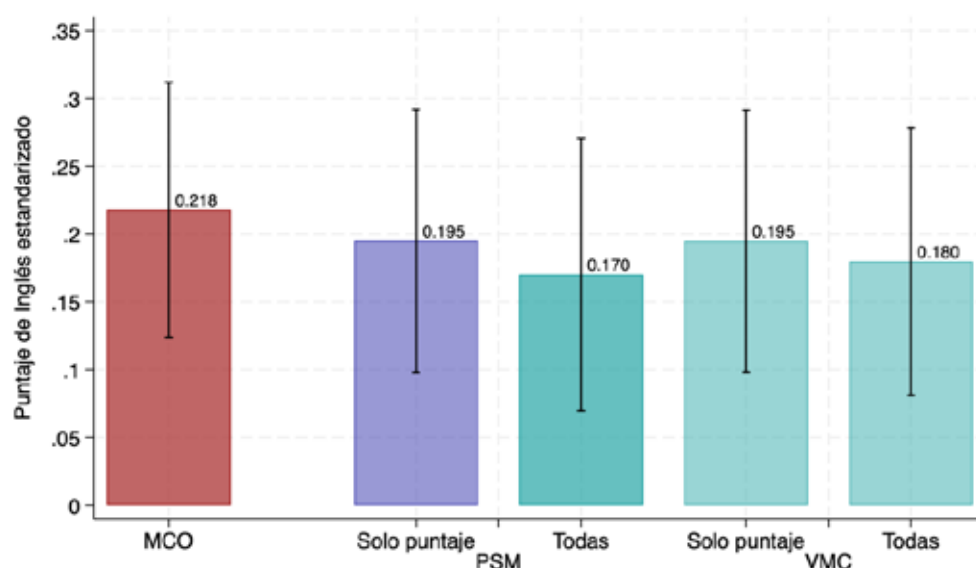


Nota: Las figuras muestran la distribución de la puntuación de propensión de los datos brutos y emparejados cuando se empareja solo en la puntuación global (columna A) o cuando se empareja en todas las variables (columna B). En el panel I, el tratamiento se define como ser elegible para el programa EwB (ITT), mientras que en los paneles II-IV, el tratamiento se define como tomar el programa de acuerdo con nuestras tres definiciones de aceptación: inscribirse en EwB, asistir al menos a una clase o asistir al menos al 20 %.

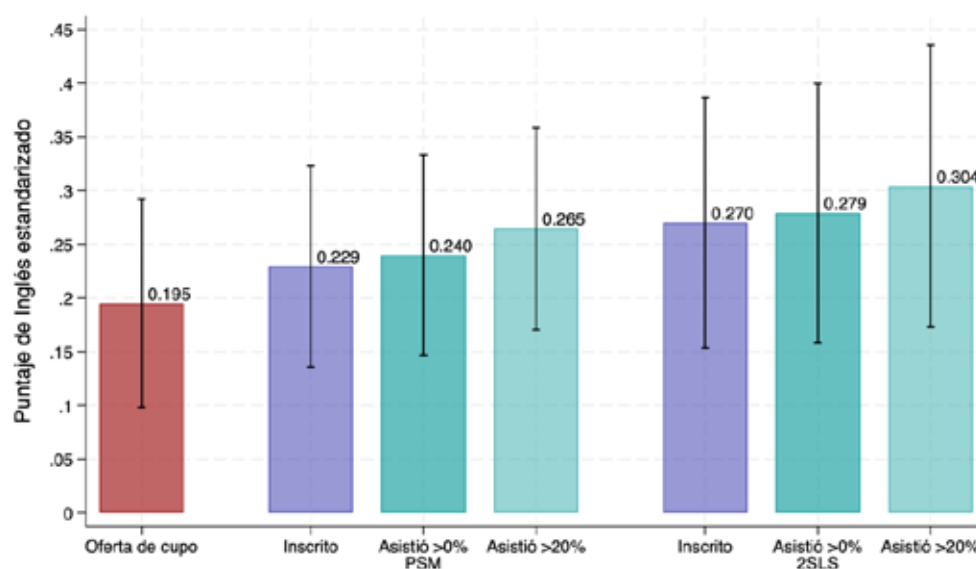
Fuente: Elaboración propia.

Figura 2. Resultados principales

B. Intención de tratar las estimaciones



B. Tener en cuenta el cumplimiento imperfecto utilizando PSM y 2SLS

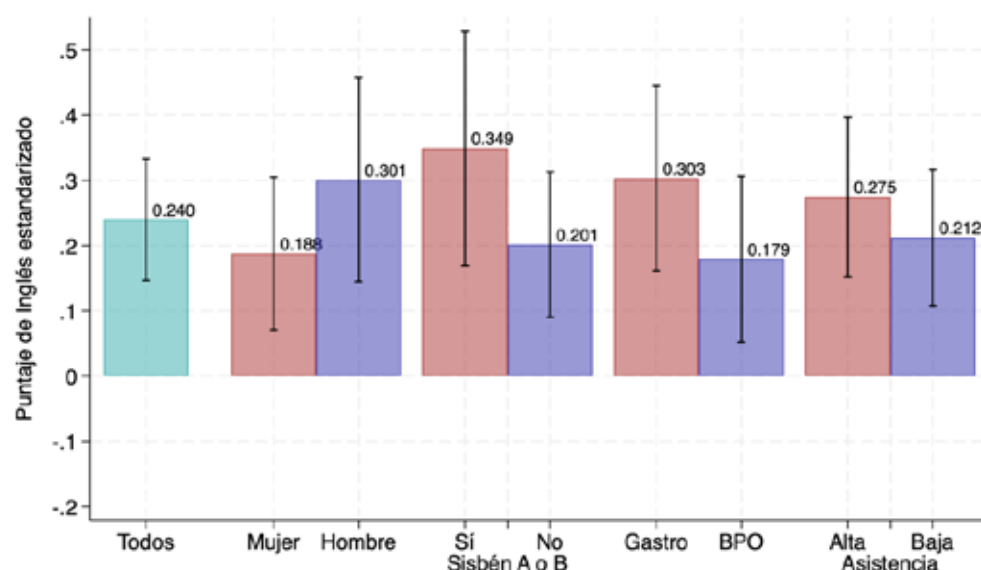


Nota: Las figuras muestran las estimaciones de coeficientes del programa EwB sobre el dominio del inglés medido por la prueba estandarizada EnglishScore. El panel A muestra la regresión de MCO ingenua en rojo, junto con los resultados del emparejamiento con PSM y vecino más cercano. Para ambos estimadores de emparejamiento, presentamos dos especificaciones: emparejar solo en el puntaje utilizado para asignar a los estudiantes al programa y emparejar en todas las características disponibles. El panel B compara la estimación de ITT de que se le ofreció un cupo utilizando nuestra especificación preferida (PSM solo en puntuación) con estimaciones del efecto de la aceptación del tratamiento utilizando PSM y IV. Usamos tres definiciones de aceptación del tratamiento: 1) inscribirse en el curso, 2) asistir al menos a una clase o 3) asistir al menos al 20 % de las clases. Para todos los coeficientes, se muestra el intervalo de confianza del 95 %.

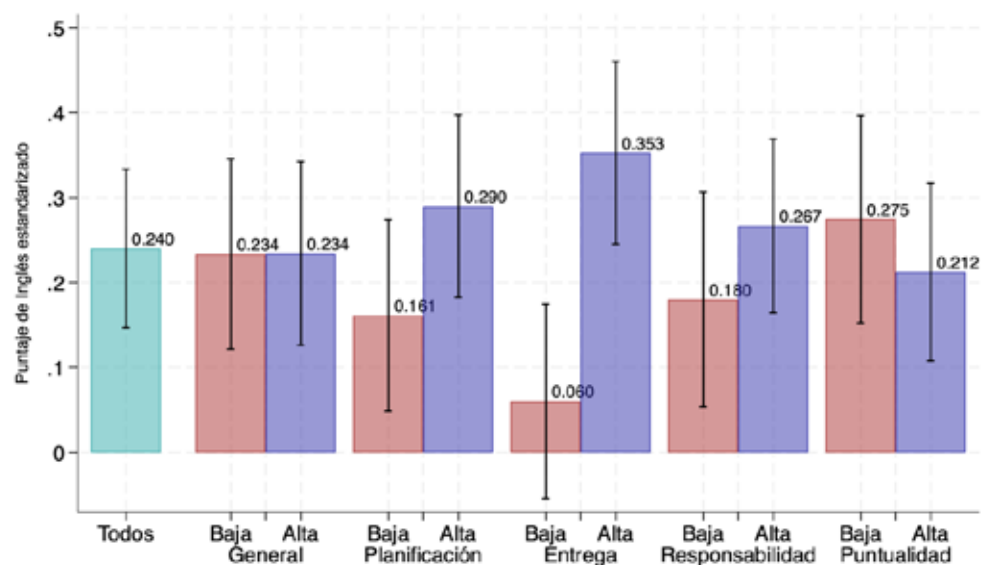
Fuente: Elaboración propia.

Figura 3. Principales efectos heterogéneos

A. Efectos heterogéneos según las características de los estudiantes



B. Efectos heterogéneos según la calidad de las clases del profesor

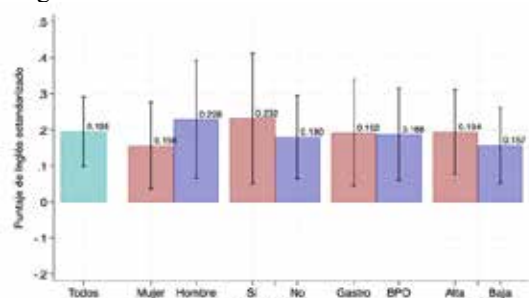


Nota: Las figuras muestran los efectos heterogéneos de la aceptación del tratamiento en los puntajes de las pruebas de inglés utilizando nuestra especificación PSM preferida (coincidiendo solo con el puntaje de la tarea). El panel A muestra efectos heterogéneos según las características de los estudiantes, incluido el género, el estado de SES, el campo de estudio y las tasas de asistencia (con una asistencia alta definida como superior a la tasa de asistencia media). El panel B muestra efectos heterogéneos del tratamiento según el desempeño docente medido en cuatro criterios de evaluación: 1) planificación de lecciones, 2) entrega de lecciones, 3) responsabilidad y 4) puntualidad, así como en un promedio simple de las cuatro dimensiones. Comparamos los efectos del tratamiento para los estudiantes asignados a un profesor por encima de la media en cada una de estas características (barra morada) con los estudiantes asignados a un profesor por debajo de la media (barra roja). En esta figura, el tratamiento se define como haber asistido al menos a una clase, mientras que la figura 4 en el Apéndice muestra el conjunto completo de resultados para las diferentes definiciones de tratamiento. Para todos los coeficientes, se muestra el intervalo de confianza del 95 %.

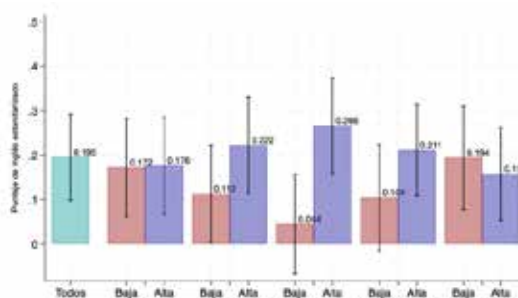
Fuente: Elaboración propia.

Figura 4. Efectos heterogéneos: todos los gráficos

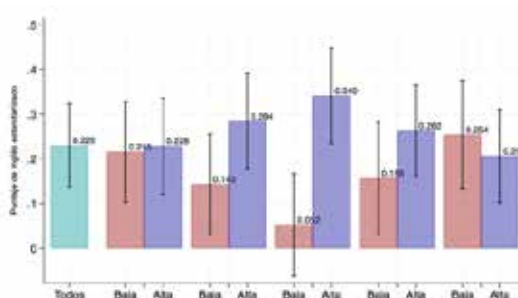
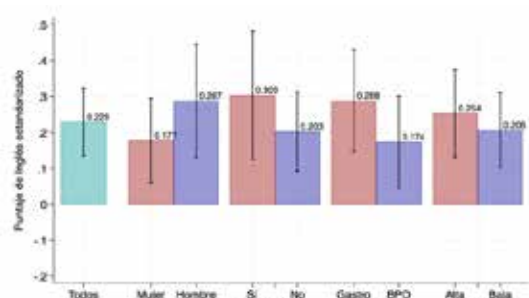
A. Características de los estudiantes
Asignación de tratamiento



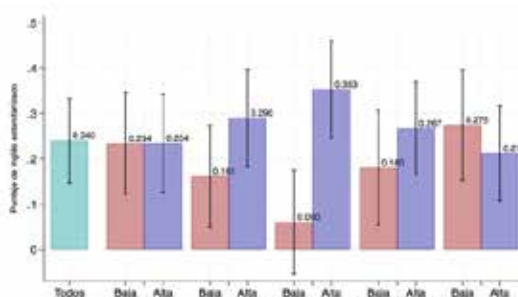
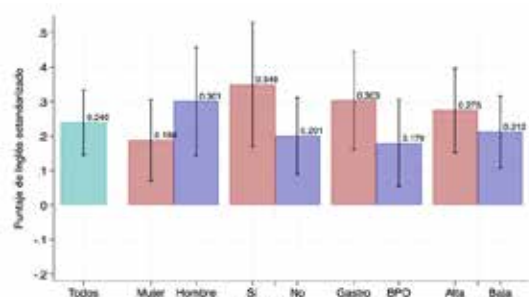
B. Calidad de las clases de los profesores



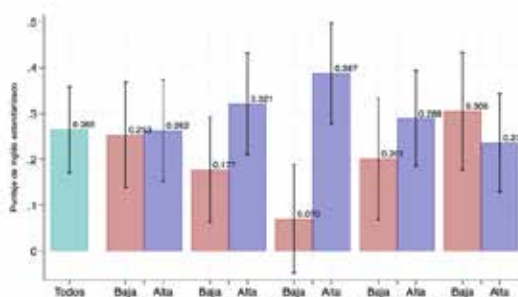
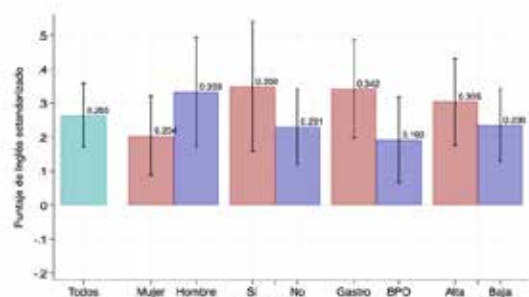
Estudiante inscrito en el curso



El estudiante asistió al menos a una clase



El estudiante asistió al menos al 20 % de las clases



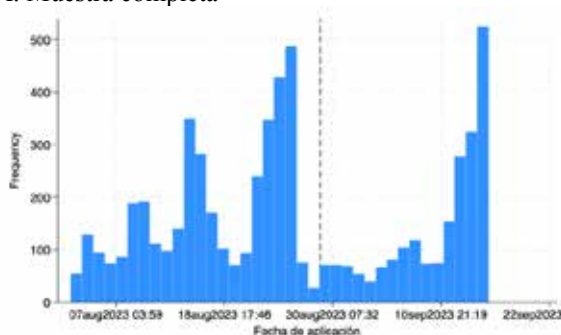
Nota: Las figuras muestran los efectos heterogéneos de la aceptación del tratamiento en los puntajes de las pruebas de inglés utilizando nuestra especificación PSM preferida (coincidiendo solo en la puntuación). La columna A muestra efectos heterogéneos por características de los estudiantes, incluido el género, el campo de estudio y las tasas de asistencia. La columna B muestran efectos heterogéneos del tratamiento según si un estudiante fue asignado a un profesor con un desempeño superior al promedio en cuatro criterios de evaluación: 1) planificación de lecciones, 2) entrega de lecciones, 3) responsabilidad y 4) puntualidad. El panel I muestra el efecto de ser asignado al tratamiento (ITT), mientras que los paneles II-IV muestran el efecto de la aceptación del tratamiento de acuerdo con nuestras tres definiciones de aceptación. Se muestra intervalo de confianza del 95 %.

Fuente: Elaboración propia.

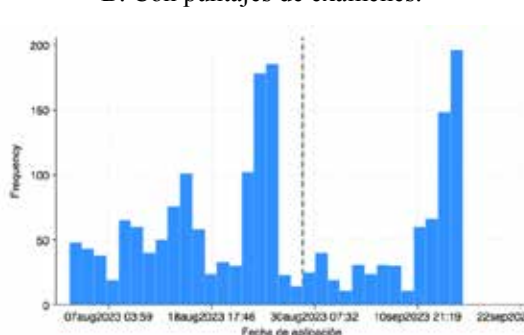
Figura 5. Gráficos de resultados de la estimación de regresión discontinua

Distribución de la variable en ejecución

A. Muestra completa

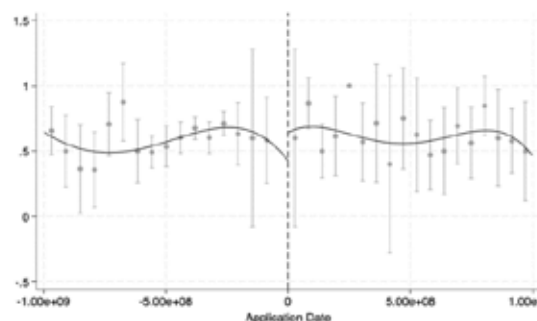


B. Con puntajes de exámenes.

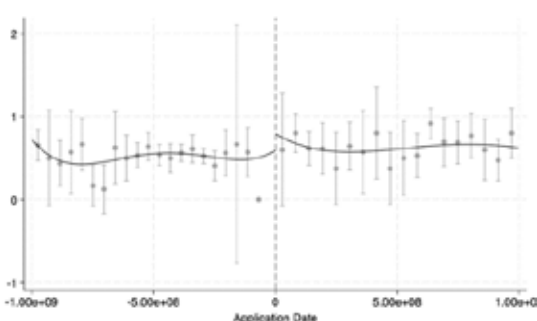


Continuidad de covariables

A. Mujer

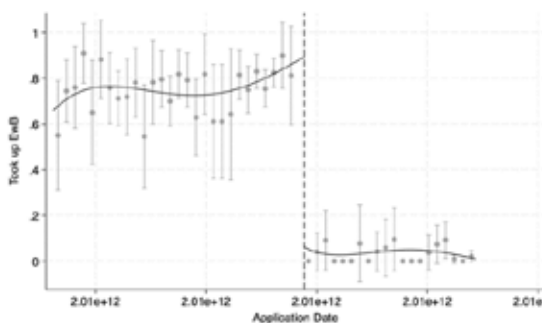


B. Bajo NSE

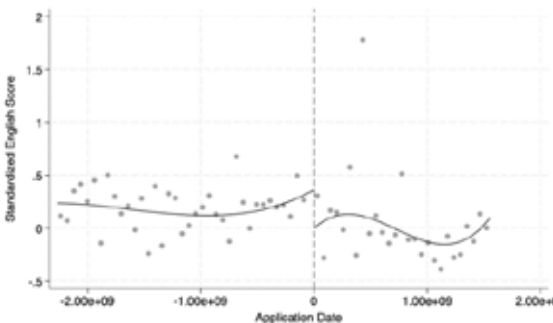


Resultados

A. Primera etapa



B. Forma reducida



Nota: Las figuras presentan los resultados de la estimación de RD. El panel I muestra la distribución de la variable de asignación para la muestra completa (figura A) y para la muestra con puntajes de prueba (figura B). El panel II muestra la continuidad de dos covariables para las que tenemos datos: porcentaje de individuos del sexo femenino (figura A) y NSE bajo, según lo indicado por si los individuos pertenecen al Sisbén (figura B). El panel III muestra los resultados gráficos de la primera etapa (figura A) y de la forma reducida en los puntajes de las pruebas de inglés (figura B). La variable de asignación es el tiempo de aplicación en milisegundos, centrado en la fecha y hora límite. Cuando es relevante, se muestran los intervalos de confianza del 95 %

Fuente: Elaboración propia.

Tabla A1. Tablas de balance de PSM para PSM según diferentes definiciones de tratamiento

	Registrarse en EwB				Asistir al menos a una clase				Asistir al menos al 20% de las clases			
	Dif. estándar		Rel. varianza		Dif. estándar		Rel. de varianza		Dif. estándar		Rel. de varianza	
	Crudo	Emparejado	Crudo	Emparejado	Crudo	Emparejado	Crudo	Emparejado	Crudo	Emparejado	Crudo	Emparejado
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(5)	(6)	(7)	(8)
<i>Panel A. Emparejamiento solo con puntaje global</i>												
Puntaje global	-0,155	-0,000	0,734	0,997	-0,143	-0,000	0,721	0,997	-0,126	-0,000	0,684	1,000
<i>Panel B. Emparejamiento con todas las características</i>												
BPO	0,047	-0,014	0,991	1,003	0,029	-0,018	0,994	1,003	0,073	-0,032	0,986	1,006
NSE bajo (Cualquier Sisbén)	-0,143	-0,040	1,057	1,015	-0,150	-0,028	1,059	1,010	-0,161	-0,038	1,061	1,013
NSE más bajo (Sisbén A o B)	-0,192	-0,016	0,793	0,981	-0,192	-0,021	0,791	0,975	-0,204	-0,017	0,776	0,979
Étnico	0,055	-0,013	1,256	0,948	0,069	0,005	1,331	1,020	0,062	-0,017	1,291	0,928
Mujer	0,101	0,000	0,961	1,000	0,124	0,004	0,952	0,998	0,184	0,000	0,926	1,000
Discapacidad	0,086	0,009	5,763	1,166	0,089	0,009	6,129	1,166	0,097	-0,000	7,330	1,000
Victima de conflicto	-0,084	0,020	0,670	1,103	-0,071	-0,012	0,710	0,941	-0,058	0,006	0,757	1,028
Violencia de género	-0,026	-0,024	0,858	0,857	-0,016	-0,034	0,911	0,816	-0,013	-0,024	0,924	0,860
Recluso	-0,020	-0,021	0,725	0,667	-0,051	0,023	0,412	1,373	-0,041	-0,027	0,493	0,626
Reto a la U	-0,032	-0,010	0,832	0,945	-0,034	0,034	0,819	1,208	-0,057	0,029	0,715	1,174
Inmersión en media	-0,099	0,020	0,402	1,208	-0,092	0,032	0,427	1,299	-0,154	-0,037	0,179	0,698
SENA	-0,162	-0,004	0,297	0,976	-0,169	0,004	0,272	1,026	-0,163	0,038	0,275	1,267
Reto SDIS	-0,085	-0,026	0,488	0,775	-0,077	-0,053	0,518	0,586	-0,076	-0,005	0,522	0,953
Parceros	-0,119	0,034	0,530	1,180	-0,132	0,012	0,489	1,065	-0,125	-0,031	0,503	0,846
Programa deportivo	0,025	0,019	1,929	1,999	0,028	0,019	2,052	1,999	0,034	0,019	2,455	1,999
Programa de cuidadores	-0,083	-0,020	0,324	0,770	-0,078	0,035	0,344	1,457	-0,094	0,030	0,248	1,413

Nota: La tabla muestra el balance posterior al emparejamiento para el estimador de emparejamiento de puntaje de propensión, de acuerdo con varias definiciones del tratamiento: las columnas 1-4 definen el tratamiento, inscribirse en EwB, las columnas 5-6 asistir al menos a una clase, y las columnas 7-8, asistir al menos al 20 % de las clases. El panel A muestra el balance después de emparejar solo en la puntuación global, mientras que el panel B muestra el balance después de emparejar en todas las características disponibles. Para cada definición, las dos primeras columnas muestran las diferencias estandarizadas de las muestras sin procesar y emparejadas, mientras que las dos últimas columnas muestran la relación de varianza de las muestras sin procesar y emparejadas.

Fuente: Elaboración propia.

Tabla A2. Efectos heterogéneos por características de los estudiantes

	Estimación global	Género		SES		Sector		Asistencia	
	(1)	Hembra (2)	Masculino (3)	Bajo (4)	Alto (5)	Gastro (6)	BPO (7)	Bajo (8)	Alto (9)
<i>A. ITT: Elegible para EwB</i>									
Elegible	0,195*** (0,049)	0,156** (0,061)	0,228*** (0,083)	0,232** (0,092)	0,180*** (0,059)	0,192** (0,076)	0,188*** (0,065)	0,110* (0,06)	0,237*** (0,056)
Media de los controles	0	-0,06	0,08	-0,2	0,08	0,02	-0,02	0	-0,01
N	1,879	1,122	757	455	1,424	851	1,028	1,333	1,469
<i>B. Aceptación: Se inscribió en EwB</i>									
Tratado	0,229*** (0,048)	0,177*** (0,06)	0,287*** (0,08)	0,303*** (0,091)	0,203*** (0,057)	0,288*** (0,073)	0,174*** (0,065)	0,146** (0,061)	0,293*** (0,056)
Media de los controles	0,01	-0,06	0,09	-0,22	0,09	0	0,01	0,01	0,01
N	1,879	1,122	757	455	1,424	851	1,028	1,333	1,469
<i>C. Aceptación: Asistió al menos una vez</i>									
Tratado	0,240*** (0,048)	0,188*** (0,06)	0,301*** (0,08)	0,349*** (0,091)	0,201*** (0,057)	0,303*** (0,073)	0,179*** (0,065)	0,166*** (0,061)	0,293*** (0,056)
Media de los controles	0,01	-0,06	0,09	-0,23	0,1	0	0,01	0,01	0,01
N	1,879	1,122	757	455	1,424	851	1,028	1,333	1,469
<i>D. Aceptación: Asistió al menos al 20% de las clases</i>									
Tratado	0,265*** (0,048)	0,204*** (0,06)	0,333*** (0,082)	0,350*** (0,097)	0,231*** (0,057)	0,342*** (0,074)	0,192*** (0,065)	0,212*** (0,069)	0,293*** (0,056)
Media de los controles	0,01	-0,06	0,1	-0,21	0,1	0	0,02	0,01	0,01
N	1,879	1,122	757	455	1,424	851	1,028	1,333	1,469

Nota: La tabla muestra los efectos heterogéneos de la aceptación del tratamiento en los puntajes de las pruebas de inglés por características de los estudiantes utilizando nuestra especificación PSM preferida (emparejamiento solo en el puntaje). La columna 1 muestra el efecto en toda la muestra, mientras que las columnas 2-9 muestran los efectos por características del estudiante, incluido el género, el estado de SES, el campo de estudio y las tasas de asistencia. Cada columna es el resultado de una regresión con una muestra dividida para el grupo respectivo. El panel A muestra el efecto de ser asignado al tratamiento (ITT), mientras que los paneles B-D muestran el efecto de la aceptación del tratamiento de acuerdo con nuestras tres definiciones de aceptación: inscribirse en EwB, asistir al menos a una clase o asistir al menos al 20% de las clases. Los errores estándar robustos se muestran entre paréntesis y las estrellas denotan significancia estadística: ***p<0,01, **p<0,05, *p<0.

Fuente: Elaboración propia.

Tabla A3. Efectos heterogéneos por desempeño docente

	Estimación global	En general		Planificación		Entrega		Responsabilidad		Puntualidad	
	(1)	Bajo (2)	Alto (3)	Bajo (4)	Alto (5)	Bajo (6)	Alto (7)	Bajo (8)	Alto (9)	Bajo (10)	Alto (11)
<i>A. ITT: Elegible para EwB</i>											
Elegible	0,195*** (0,049)	0,172** *	0,176*** (0,056)	0,112** (0,056)	0,222*** (0,056)	0,044 (0,057)	0,266*** (0,055)	0,194*** (0,060)	0,157*** (0,053)	0,227*** (0,056)	0,116** (0,055)
Media de los controles	0	-0,00	0,24	-0,30	0,43	-0,58	0,67	-0,23	0,46	0,02	0,20
N	1,879	1,439	1,451	1,367	1,523	1,321	1,569	1,357	1,533	1,481	1,409
<i>B. Aceptación: Se inscribió en EwB</i>											
Tratado	0,229*** (0,048)	0,215** *	0,228*** (0,055)	0,143** (0,057)	0,284*** (0,055)	0,052 (0,058)	0,340*** (0,055)	0,254*** (0,062)	0,206*** (0,053)	0,286*** (0,057)	0,159*** (0,055)
Media de los controles	0,01	.	.	.	.	.	.	.	.	.	.
N	1,879	1,439	1,451	1,367	1,523	1,321	1,569	1,357	1,533	1,481	1,409
<i>C. Aceptación: Asistió al menos una vez</i>											
Tratado	0,240*** (0,048)	0,234** *	0,234*** (0,055)	0,161*** (0,057)	0,290*** (0,055)	0,060 (0,058)	0,353*** (0,055)	0,275*** (0,062)	0,212*** (0,053)	0,309*** (0,057)	0,158*** (0,055)
Media de los controles	0,01	-0,07	0,24	-0,13	0,26	-0,21	0,18	-0,03	0,17	-0,02	0,41
N	1,879	1,439	1,451	1,367	1,523	1,321	1,569	1,357	1,533	1,481	1,409
<i>D. Aceptación: Asistió al menos al 20% de las clases</i>											
Tratado	0,265*** (0,048)	0,253** *	0,262*** (0,056)	0,177*** (0,059)	0,321*** (0,056)	0,070 (0,060)	0,387*** (0,056)	0,305*** (0,065)	0,236*** (0,055)	0,335*** (0,059)	0,177*** (0,057)
Media de los controles	0,01	-0,00	0,13	-0,08	0,16	-0,06	0,13	0,04	0,09	0,01	0,16
N	1,879	1,439	1,451	1,367	1,523	1,321	1,569	1,357	1,533	1,481	1,409

Nota: La tabla muestra los efectos heterogéneos de la aceptación del tratamiento en los puntajes de las pruebas de inglés por desempeño docente, utilizando nuestra especificación PSM preferida (emparejando solo en el puntaje). La columna 1 muestra el efecto en toda la muestra, mientras que las columnas 2-9 muestran los efectos según si un estudiante fue asignado a un profesor con un desempeño superior al promedio en cuatro criterios de evaluación: 1) planificación de lecciones, 2) entrega de lecciones, 3) responsabilidad y 4) puntualidad y puntaje general. Cada columna es el resultado de una regresión para la muestra del grupo respectivo. El panel A muestra el efecto de ser asignado al tratamiento (ITT), mientras que los paneles B-D muestran el efecto de la aceptación del tratamiento de acuerdo con nuestras tres definiciones de aceptación: inscribirse en EwB, asistir al menos a una clase o asistir al menos al 20 % de las clases. Los errores estándar robustos se muestran entre paréntesis y las estrellas denotan significancia estadística: ***p<0,01, **p<0,05, *p<0.

Fuente: Elaboración propia.

Documentos de trabajo es una publicación periódica de la Escuela de Gobierno Alberto Lleras Camargo de la Universidad de los Andes, que tiene como objetivo la difusión de investigaciones en curso relacionadas con asuntos públicos de diversa índole. Los trabajos que se incluyen en la serie se caracterizan por su interdisciplinariedad y la rigurosidad de su análisis, y pretenden fortalecer el diálogo entre la comunidad académica y los sectores encargados del diseño, la aplicación y la formulación de políticas públicas.

gobierno.uniandes.edu.co

     | GobiernoUAndes