

DOCUMENTOS DE TRABAJO

Saber Investigar

Instituto Colombiano para la Evaluación de la Educación

N° 13

**Evaluación de Impacto de la
Normatividad Específica para las
Licenciaturas (2015-2017) en los
Resultados del Examen Saber Pro**

Alejandro Farieta

Febrero de 2024

Serie Documentos de Trabajo Saber Investigar

Edición No. 13
Enero de 2024
Edición digital
ISSN: 2954-6583

Instituto Colombiano para la Evaluación de la Educación

Oficina de Gestión de Proyectos de Investigación
Calle 26 N. 69-76, Edificio Elemento, Torre II, piso 18, Bogotá, D. C.
Teléfono: (601) 4841410
proyectosinvestigacion@icfes.gov.co
<https://www.icfes.gov.co>

Autor

Alejandro Farieta

Directora General

Elizabeth Blandón Bermúdez

Jefe de Oficina De Gestión de Proyectos de Investigación

Jeanit Camilo Cortés Mora

Advertencia:

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

Citar este documento en estilo APA así:

Farieta, A. (2024). *Evaluación de Impacto de la Normatividad Específica para las Licenciaturas (2015-2017) en los Resultados del Examen Saber Pro*. (Documentos de Trabajo Saber Investigar No. 13). Instituto Colombiano para la Evaluación de la Educación (Icfes). <https://www.icfes.gov.co/web/guest/saber-investigar>



Evaluación de Impacto de la Normatividad Específica para las Licenciaturas (2015-2017) en los Resultados del Examen Saber Pro ¹

Alejandro Farieta ²

Resumen

Este artículo presenta la evaluación de impacto de las políticas para los programas de formación docente en el marco del *Plan Nacional de Desarrollo 2014-2018*, los cuales tuvieron que obtener acreditación de alta calidad, incrementar el número de créditos de prácticas y el nivel de inglés, entre otros requerimientos. Utilizando regresión multivariable, mostramos que, para el año 2021 no hay asociación entre acreditación de alta calidad en la mayoría de los módulos del Examen Saber Pro, o esta es negativa, específicamente en el puntaje global y Razonamiento Cuantitativo. En un segundo momento, usando Diferencias-en-Diferencias, mostramos cómo los cambios realizados por los programas tuvieron efectos positivos en el desempeño de los estudiantes de licenciatura con respecto a otros programas profesionales. Los resultados muestran una mejora significativa en la mayoría de los módulos del Examen Saber Pro, excepto en el módulo de Inglés. Los efectos fueron homogéneos por características de los estudiantes y las instituciones, pero heterogéneos por área de conocimiento, lo que indica que el efecto se debe principalmente a los cambios curriculares de los programas.

Palabras claves: Formación inicial de docentes; Evaluación de política educativa; acreditación de alta calidad; Saber Pro; Diferencias-en-diferencias.

¹ Las ideas, opiniones, tesis y argumentos expresados son de propiedad exclusiva del autor y no representan el punto de vista del Icfes. Esta investigación fue financiada por Icfes al ser ganadora de la Convocatoria de Estudiantes 2022. Este trabajo se ha realizado como parte de la investigación doctoral del autor, quien agradece especialmente a sus supervisores, Marcos Delprato y Mario Novelli por sus valiosos aportes y comentarios, así como a la Oficina Asesora de Gestión de Proyectos de Investigación del Icfes, quienes también hicieron aportes significativos para la mejora del presente trabajo. No obstante, cualquier error en los cálculos, o imprecisión en el análisis, es solamente responsabilidad del autor.

² Candidato a PhD en Educación y Desarrollo Internacional, Universidad de Sussex (Reino Unido). R.Farieta-Barrera@sussex.ac.uk

Impact Evaluation of the Specific Regulations for Bachelors in Education (2015-2017) on the Results of the *Saber Pro* tests

Alejandro Farieta

Abstract

This paper presents the impact evaluation of the policies for teacher education programs within the framework of the *National Development Plan 2014-2018*. The programs were compelled to obtain high-quality accreditation, increase the credits in school practices, and the English level, among other requirements. Relying on multivariable regression, it shows that, for the year 2021, there is no association between high-quality accreditation in most modules of the *Saber Pro* tests, or it is negative in the global score and quantitative reasoning. In a second step, using Differences-in-Differences, it evidences how the changes made by the programs had positive effects on the performance of students compared to other professional programs. The results show significant improvement in most modules of the *Saber Pro* test, except for the English module. The effects were homogeneous by characteristics of the students and institutions, but heterogeneous by area of knowledge, indicating that the effect is mainly due to the curricular changes of the programs.

Keywords: Initial Teacher Education; Evaluation of Educational Policy; High-Quality Accreditation; *Saber Pro* test; Differences-in-Differences.

Contenido

1	Introducción.....	4
2	Marco teórico, revisión de literatura y estado del arte	7
2.1	Políticas de programas de formación inicial de docentes	7
2.2	Acreditación de alta calidad.....	9
2.3	Factores individuales y socioeconómicos que inciden en el desempeño de los estudiantes.....	11
3	Metodología y datos	12
3.1	Muestra	13
3.2	Primera fase de análisis: interacción de acreditación de alta calidad con tipo de programa por año	18
3.3	Segunda fase de análisis: enfoque cuasiexperimental	19
3.3.1	Emparejamiento de los grupos	19
3.3.2	Estimación de DiD.....	21
4	Resultados.....	24
4.1	Desempeño de los estudiantes por tipo de programa y acreditación.....	24
4.2	Resultados del emparejamiento de los grupos.....	26
4.3	Resultados del DiD.....	30
4.3.1	Resultados del DiD por subgrupos.....	32
4.3.2	Resultados del DiD por área de la licenciatura	33
5	Conclusiones, discusión y recomendaciones	35
5.1	Acreditación de alta calidad.....	35
5.2	Efectos de los demás aspectos de las reformas a las licenciaturas.....	36
5.3	Recomendaciones	38
6	Referencias	40
	Anexos	52
	Anexo 1. Regresión multivariada con interacciones, versión completa	52
	Anexo 2. Resultados detallados del emparejamiento de los grupos	55
	Anexo 3. Balanceo de grupos por covariables. Cálculos por año.....	57
	Anexo 4. Compatibilidad del grupo tratado posterior al emparejamiento en comparación con el grupo original (0). Cálculos por año.....	59
	Anexo 5. Resultados completos del DiD, incluyendo variables de control	60
	Anexo 6. Resultados de la estimación del DiD sin balanceo de grupos	62
	Anexo 7. Comparación del DiD. Muestra balanceada y sin balancear	64
	Anexo 8. Resultados de DiD por subgrupos – diferencias no significativas	65
	Anexo 9. Resultados de DiD por área de la licenciatura – todas las áreas	69
	Anexo 10. Resultados de DiD por área de la licenciatura – diferencias no significativas	70

1 Introducción

El *Plan Nacional de Desarrollo [PND] 2014-2018* tuvo como uno de sus ejes centrales la necesidad de la mejora de la formación de maestros en Colombia. El documento conceptual que soportó el *PND 2014-2018*, emitido por el Departamento Nacional de Planeación [DNP] (2015), tituló su capítulo IV “Colombia, la más educada”. Dicho documento señalaba que a pesar de los significativos logros que se habían logrado en el país en los últimos años en términos de cobertura, había un importante reto con respecto a la calidad. Los resultados de las pruebas PISA para Colombia han sido recurrentemente bajos desde que el país inició su participación en dicha prueba internacional en el año 2006 (Organización para el Crecimiento y el Desarrollo Económico [OCDE], 2013). De acuerdo con dicho documento, “Solo en matemáticas, estos resultados indican un retraso de casi 3 años de escolaridad de los niños colombianos de 15 años frente a sus pares en países de la OCDE” (DNP, 2015, p. 78). Como el mismo documento señala, hay una importante correlación entre el desempeño de los estudiantes en pruebas internacionales como PISA y el crecimiento económico de los países, tal como lo prueban varios estudios internacionales (Hanusheck, 2011; Hanusheck *et al.*, 2018; RAND Corporation, 2013). De acuerdo con un estudio financiado por la Fundación Compartir (García, *et ál.*, 2014), citado en el mencionado documento conceptual del *PND 2014-2018*, los colegios con mejor y más alta formación tienen docentes con desempeños significativamente mucho más altos que colegios en donde los profesores tienen niveles más bajos de formación. Este mismo estudio mostraba que de los programas de formación de docentes, los puntajes más altos se encuentran en las universidades que cuentan con acreditación de alta calidad, y los programas con este mismo reconocimiento también mostraban desempeños más altos.

En el *PND 2014-2018*, establecido por el Congreso de la República [CRC] mediante la *Ley 1753 de 2015* (CRC, 2015, jun. 9, art. 222) se obligó a la obtención de acreditación de alta calidad a los programas de formación de docentes —también conocidos como *Licenciaturas*, como lo establece el art. 25 de la *Ley 30 de 1992* (CRC, 1992, dic. 28)— en un plazo no superior a dos años. De acuerdo con las condiciones mínimas para obtención de acreditación vigentes en aquel momento, establecidas por el Ministerio de Educación [MEN] a través del Consejo Nacional de Acreditación [CNA], un programa puede

postularse a obtener dicho reconocimiento si cumple con al menos cuatro cohortes de egresados (MEN/CNA, 2013). Por ello, los programas que aún no cumplieran con el requisito de las cuatro cohortes tendrían dos años a partir del cumplimiento dicho requisito. Adicionalmente, y como complemento del requisito de acreditación, la Presidencia de la República [PRC] promulgó el *Decreto 2450 de 2015*, el cual reglamentaba las condiciones de calidad para otorgamiento y renovación de renovación de registro calificado de los programas de formación de docentes (PRC, 2015, dic. 17) y adicionalmente, el Ministerio de Educación Nacional (MEN) promulgó la *Resolución 2041 de 2016*, que establecía las características específicas de calidad de los programas de licenciatura para la obtención, renovación o modificación de registro calificado (MEN, 2016, feb. 3). Estas nuevas condiciones de calidad para estos programas los obligó a aumentar sus prácticas pedagógicas a 50 créditos; para los programas a distancia, al menos 40 créditos deberían ser presenciales. Adicionalmente, las licenciaturas en inglés deberían cumplir con un mínimo de C1 en el Marco Común Europeo de Referencia para lenguas, y el resto de las licenciaturas deberían obtener por lo menos un nivel de A2, que incrementaría a B1 luego de tres años de vigencia de la norma. Al año siguiente el MEN, en la *Resolución 18583 de 2017*, redujo el número de créditos obligatorios de prácticas a 40, y el de créditos presenciales en programas a distancia a 20 (MEN, 2017, sep. 15).

De acuerdo con lo expresado por la entonces viceministra de Educación Superior, imponer acreditación a las licenciaturas “llevará al cierre de al menos la mitad de los programas de licenciaturas y por eso invito a ser valientes. Quedarán otros 200, los de mejor calidad” (Ariza, 2018). Cerca de 120 programas fueron cerrados entre 2017 y 2018 (Arias, *et al.*, 2018). Para prevenir el cierre de 106 licenciaturas que atendían regiones priorizadas por el PND, en 2017 se sancionó el Decreto el *Decreto 892* (PRC, 2017, mayo 28), lo que evitó la pérdida de cupo de cerca de 13.000 estudiantes, como el mismo decreto señala. No obstante, entre 2017 y 2019, 101 licenciaturas obtuvieron acreditación de alta calidad por primera vez.

El posterior *Plan de Desarrollo 2018-2022*, o Ley 1955 de 2019 (CRC, 2019, mayo 25) derogó la acreditación obligatoria para las licenciaturas, y el Decreto 1330 (PRC, 2019, jul. 25) cambió la normatividad para los programas de pregrado a nivel nacional, eliminando la normatividad específica para dichos programas de formación de docentes. Sin embargo, más allá de las razones políticas que llevaron a derogar el requisito de acreditación para

las licenciaturas como su regulación específica, vale la pena preguntarse si estas medidas tuvieron algún impacto en el desempeño de los estudiantes de licenciatura, tal como tanto el documento conceptual del *PND 2014-2018* presuponía (DNP, 2015). Aunque la acreditación de alta calidad se asocia normalmente con mejores desempeños en exámenes de estado (García, *et al.*, 2014; Bayona-Rodríguez, *et al.*, 2018; Cayón, *et al.*, 2020), los estudios normalmente se centran en la acreditación institucional, pero hay muy poca literatura dedicada a estudiar el impacto de la acreditación de programas. De hecho, el estudio realizado por García *et al.* (2014, citado en DNP, 2015) no incluyó entre sus variables la acreditación de programas, sino solamente la acreditación institucional. Algunos estudios muestran que no necesariamente se da esta asociación (Farieta-Barrera, 2020; 2024) o que, en algunos módulos como Inglés, no hay un aumento significativo en los últimos 10 años a pesar del incremento de las instituciones y los programas con acreditación de alta calidad (Benavides, 2021).

De 2017 a 2021 los puntajes de los estudiantes en dichos programas no parecen estar mejorando significativamente, ni estarse nivelando con los puntajes de estudiantes de los demás programas de pregrado, tal como lo muestran los informes presentados por el ICFES desde 2017 (ICFES, 2018, 2020, 2021a). Así, la pregunta que nos queremos hacer es ¿qué impacto tuvo en el desempeño de los estudiantes el requisito de acreditación y las demás reformas curriculares que hicieron los programas para ajustarse a la normatividad específica para las licenciaturas?

La pregunta supondrá dos objetivos y tipos de análisis ligeramente diferentes. El primero tiene que ver con cómo se comportó el puntaje de los estudiantes de licenciatura en función del requisito de acreditación. Dado que se trató del componente más agresivo de la política específica para los programas de licenciatura, era de esperarse que tuviera incidencia en el desempeño de los programas. Sin embargo, como veremos, esto no sucedió exactamente así. En un segundo momento, evaluaremos cuál fue el efecto que tuvieron estas políticas para los programas de licenciatura. Dado que el efecto de los cambios de los currículos y de los programas tomó bastante tiempo, y los primeros graduados de las cohortes reformadas tendrían que estarse graduando cuatro años después de las reformas a los programas, evaluaremos si para el año 2021 hay impacto en los desempeños de los estudiantes.

En la primera parte emplearemos metodología de regresión multivariable, examinando la interacción entre el tipo de programa (licenciatura/profesional) y el tipo de registro (registro calificado/acreditación de alta calidad) a lo largo de los años 2012 a 2021. Para la segunda parte, emplearemos un modelo de diferencias-en-diferencias [DiD], utilizando como grupo de control a los demás programas profesionales. Dadas las sustanciales diferencias que hay entre estudiantes de licenciatura con respecto a los estudiantes de los demás programas profesionales, tanto en términos de antecedentes académicos como de condiciones socioeconómicas, combinaremos el DiD con otras técnicas de emparejamiento —*matching*—, más específicamente con emparejamiento exacto para las variables categóricas, y emparejamiento por índice de propensión —*propensity score matching*—.

2 Marco teórico, revisión de literatura y estado del arte

2.1 Políticas de programas de formación inicial de docentes

Los sistemas educativos en América Latina han sido denominados por Hanusheck & Woessman “The Latin American Growth Puzzle”: han crecido significativamente en los últimos años, pero ello no ha tenido el efecto esperado en las economías nacionales como sí ha pasado en otros países con economías similares, como las del sudeste asiático (Hanusheck & Woessman, 2012). Una serie de estudios y de reportes de política pública de diferentes organizaciones han evidenciado que una de las explicaciones para esto puede estar en la selección, formación y remuneración de los futuros docentes (Bruns & Luque, 2015; OREALC/UNESCO, 2013, 2016; OCDE/CEPAL/CAF, 2014, RAND Corporation, 2013).

De acuerdo con Cochran-Smith, desde finales de los años 90 las políticas de formación de docentes de muchos países se han convertido en lo que ella denominó “A public policy problem”: tratar de identificar qué elementos se pueden manipular desde la política pública de formación de maestros con miras a mejorar la calidad de la educación (Cochran-Smith, 2005; 2021; Furlong *et al.*, 2008; Tatto & Menter, 2019). El razonamiento que gira bajo este supuesto es que al mejorar los programas de formación de docentes (a nivel de pregrado) habrá un impacto futuro en los docentes cuando estén en ejercicio, lo que a su vez tendrá un impacto en los desempeños de sus futuros estudiantes y en

consecuencia en el desarrollo y la economía a nivel nacional (Barber & Mourshed, 2007; National Commission on Teaching & America's Future, 1997; OCDE, 2005; Rockoff, 2004; Tucker, 2011).

Estudios muestran cómo el desempeño de los estudiantes depende principalmente de la formación y calidad de los docentes (Hanusheck & Woessman, 2015). Igualmente, es cada vez más aceptada la tesis de que, aparte de las condiciones individuales y socioeconómicas, el principal factor de aprendizaje de los estudiantes en nivel básico y secundario es la calidad de los docentes (Burroughs *et al.*, 2019; Coenen *et al.*, 2018). Algunos muestran que profesores mejor capacitados y con más experiencia tienen estudiantes con mayores oportunidades de ingresar a la universidad, obtener mejores trabajos, vivir en áreas de mayores ingresos, estar menos expuestos a embarazos juveniles y tener mejores planes de pensión o retiro al final de sus vidas (Chetty *et al.*, 2014).

Sin embargo, a pesar del acuerdo que hay sobre la importancia de los docentes en la educación básica y media, hay considerables discrepancias sobre cuáles son las características de los profesores con mayor efectividad en el aprendizaje. En una revisión de literatura, Burroughs *et al.* afirman hay evidencia mixta acerca de la incidencia de factores como el nivel de formación de los docentes, o el nivel de formación específica en el área de educación, o la experiencia de los docentes, o el conocimiento profesional de los docentes, las certificaciones que posean u otras oportunidades de aprendizaje (Burroughs *et al.*, 2019). En otra revisión de literatura, Bardach & Klassen encuentran que las correlaciones entre el nivel cognitivo de los docentes y el desempeño de los estudiantes es negativa o nula, aunque muestran también que las heterogeneidades de unos y otros hacen difícil establecer este tipo de correlaciones (Bardach & Klassen, 2020).

La medida del gobierno nacional de exigir acreditación de alta calidad a las licenciaturas convirtió este reconocimiento en un modelo de rendición de cuentas (*accountability*), con vistas a mejorar la calidad de la formación de docentes (Valderrama-Leongómez *et al.*, 2019; Verger, 2013). La rendición de cuentas ha sido una de las principales estrategias a nivel mundial para mejorar los programas de formación de docentes (Cochran-Smith, 2021), aunque en general con resultados no siempre satisfactorios. Como afirman Hanusheck & Woessman, las medidas de rendición de cuentas pueden ser obstáculos sin sentido que refuerzan la burocracia de la formación

docente en lugar de monitorear el impacto de los maestros en el rendimiento de los estudiantes (Hanusheck & Woessman, 2015).

2.2 Acreditación de alta calidad

En los últimos 30 años, más de 100 países alrededor del mundo han creado instituciones o mecanismos para el aseguramiento de la calidad de la educación superior (Garwe, 2019). En Colombia, el Sistema de Aseguramiento de la Calidad de la Educación Superior [SACES], a cargo del Ministerio de Educación Superior [MEN], se encarga de la evaluación de las condiciones de calidad tanto de los programas como de las instituciones. Sin embargo, desde 1992 se creó el Consejo Nacional de Acreditación [CNA] el cual tiene por objetivo reconocer a las instituciones como a los programas que ostenten los más altos criterios de calidad (MEN/CNA, 2020). Colombia fue uno de los primeros países en el mundo en tener un sistema de acreditación de alta calidad para sus instituciones y sus programas (MEN/CNA, 2013; Cayón *et al.*, 2020). Se trata de un proceso voluntario al que se pueden someter tanto instituciones como programas de pregrado o posgrado, de manera que se les reconozca públicamente sus altos estándares y sea un elemento diferenciador para futuros estudiantes para identificar los programas con más altos niveles de calidad (Duque, 2021), y surgió como respuesta al incremento de instituciones y programas que se abrieron en los años 90s, principalmente de universidades privadas y cuya calidad se consideró dudosa; fenómeno que no solo se dio en Colombia, sino que se replicó en muchas partes de América Latina (Fernández, 2003).

La acreditación de alta calidad en Colombia se basa en un modelo de evaluación por pares, y tienen tres momentos. El primero, que consta de la autoevaluación del programa o la institución. El segundo, en el cual el CNA envía a las universidades y los programas una comisión de “pares académicos” para que evalúen unas condiciones específicas de los programas o las instituciones. El último paso lo hace el CNA con base en los informes de los pares académicos y las evidencias que recojan en su visita (MEN/CNA, 2013). Los lineamientos de acreditación mencionan específicamente, en el *Factor 4: procesos académicos; característica 16: integralidad del currículo*, que la evaluación del programa analice no solamente el puntaje de los estudiantes en el examen *Saber Pro*, sino que además se haga un estudio del valor agregado del programa con respecto al puntaje

obtenido por los estudiantes en el examen *Saber 11^o*, que se realiza antes de ingresar a la universidad (MEN/CNA, 2013, p. 38).

Los sistemas de acreditación por lo general siguen unos estándares específicos que se aplican de igual manera o de manera contextual (Haakstad, 2001). Hay estudios a nivel internacional que muestran que las instituciones acreditadas tienen una mayor producción científica (Nigsch & Schenker-Wicki, 2013). No obstante, estudios a nivel nacional muestran que el efecto, por el contrario, ha sido una reducción de la producción investigativa de los docentes (Bayona-Rodríguez, *et al.*, 2018). Se ha argumentado que los procesos de acreditación mejoran la cultura organizacional de las instituciones y la eficiencia en sus procesos, así como la innovación y la consecución de metas a mediano y largo plazo (Lejeune & Vas, 2009; Romero, 2008), aunque también se evidencia en la literatura que en países como Hong Kong, la burocracia asociada a la acreditación y el excesivo monitoreo de parte del Estado ha conllevado a que se disminuya el tiempo y la inversión en las actividades centrales de las universidades, a saber, docencia e investigación (Tam, 1999). Como todo modelo de evaluación por pares, es muy demandante en términos de tiempo y de recursos (Verger, 2013), y está sujeto a diversos problemas como conflictos de intereses de parte de los evaluadores, como se ha evidenciado particularmente en el modelo de acreditación chileno (Barroilhet, *et al.*, 2022), aunque hay estudios que muestran que en Catar los sistemas de acreditación han tenido un impacto en la reducción de la corrupción al interior de las universidades (Mattar, 2022).

Para el caso colombiano, en general se evidencia que la acreditación institucional está asociada con el mejor desempeño de los estudiantes (García *et al.*, 2014; Bayona-Rodríguez *et al.*, 2018; Timarán-Pereira *et al.*, 2016). Sin embargo, en el caso de la acreditación de programas, su asociación con el desempeño de los estudiantes en el examen *Saber Pro* no es consistente y la evidencia al respecto es mixta. Por una parte, en programas como economía se ha encontrado evidencia a favor de los programas acreditados en el desempeño de los estudiantes (Sarmiento *et al.*, Van, 2015); lo mismo sucede con los programas de ingeniería en la región Caribe colombiana (Romero *et al.*, 2020), sin embargo, en las licenciaturas en filosofía la acreditación de los programas no ha resultado siendo un determinante de desempeño (Farieta-Barrera, 2020; 2024). Otros estudios no toman como variable de interés la acreditación del programa, sino el

porcentaje de programas acreditados por institución, independientemente de si esta está acreditada o no, con resultados positivos (Rodríguez, Ariza, & Ramos, 2014). Otros estudios muestran que los programas más antiguos y consolidados no tienen diferencias significativas de desempeño en contraste con programas más nuevos o recientemente creados (Camacho, Messina, Uribe, 2016). Un estudio de Barrera-Osorio & Bayona-Rodríguez (2019) mostró que estudiar en una universidad prestigiosa tiene efectos en la probabilidad de obtener un mejor empleo y de tener mayores ingresos, pero no en la mejora del desempeño de los estudiantes, creando un efecto de señuelo, aunque sin mayor efectividad que otras universidades en el desarrollo de capital humano. Otros estudios muestran que la acreditación de alta calidad está más ligada al nivel socioeconómico de sus estudiantes, lo que ha constituido un mecanismo para incrementar la desigualdad (Cuenca, 2016). Por último, un reporte de la OCDE afirma que el requisito de acreditación obligatoria para las licenciaturas puede ser inconveniente debido a lo demandante que resulta para las instituciones (Radinger *et al.*, 2018).

2.3 Factores individuales y socioeconómicos que inciden en el desempeño de los estudiantes

Uno de los principales determinantes de desempeño de los estudiantes a nivel universitario en Colombia es su desempeño previo al ingreso a la universidad (Castro *et al.*, 2018; Camacho *et al.*, 2016; García *et al.*, 2014). Aparte de este, el nivel socioeconómico individual de los estudiantes en Colombia es uno de los factores que más incide en su desempeño académico (Avendaño *et al.*, 2016; García *et al.*, 2015; OECD, 2016; 2019) así como otros factores estrechamente relacionados, como la educación de los padres, las condiciones de la vivienda y el estrato socioeconómico (Castro *et al.*, 2018; Cuenca, 2016; García *et al.*, 2014; Melguizo *et al.*, 2016; Rodríguez, Ariza & Ramos, 2014; Rojas, 2019; Timarán-Pereira *et al.*, 2016). Adicionalmente, también las desigualdades entre las regiones y departamentos en Colombia tienen incidencia en el desempeño de los estudiantes (Arias-Velandia *et al.*, 2021; Fajardo *et al.*, 2021), las cuales a su vez hacen que en las regiones más pobres del país los profesores que son de estas mismas regiones terminen teniendo habilidades académicas menores a los de otras regiones (Arias *et al.*, 2018; Bonilla-Mejía *et al.*, 2018). En particular las regiones más pobres tienen más estudiantes en programas a distancia, en donde también se han encontrado diferencias

significativas de desempeño en comparación con estudiantes que cursan sus estudios de manera presencial (Pineda & Celis, 2018; Rodríguez, Gómez & Ariza, 2014, Timarán-Pereira *et al.*, 2016). De igual manera, existen importantes brechas de género que afectan en particular a las mujeres (Abadía & Bernal, 2017; Rodríguez, Ariza & Ramos, 2014); hay evidencia de un mejor desempeño de las mujeres en lectura, aunque más bajo en matemáticas y ciencias (OECD, 2016, 2019), aunque otros estudios muestran que el desempeño de las mujeres solo es mejor en escritura, en particular en los programas a distancia (Arias-Velandia *et al.*, 2018).

3 Metodología y datos

Como señalamos anteriormente, presentaremos dos análisis para evaluar cómo las políticas específicas para los programas de formación de docentes tuvieron efectos en el desempeño de los estudiantes por medio de un enfoque cuasiexperimental. Para poder establecer una línea de comparación y controlar por otros tipos de programas, examinaremos los resultados de Saber Pro de los años 2012 a 2021, y analizaremos los cambios que ha habido de 2017 en adelante que puedan ser atribuidos a la normatividad específica para las licenciaturas, utilizando como grupo de control a los estudiantes de los demás programas profesionales. En una primera parte, examinamos por medio de regresión multivariada el comportamiento de los programas acreditados frente a los no acreditados, controlando por año y tipo de programa, dado que la acreditación de alta calidad fue el principal requerimiento para las licenciaturas. En la segunda parte del análisis, examinamos si posteriormente a la proclamación de la normatividad hay algún efecto que se pueda atribuir a la normatividad específica para las licenciaturas. Esto lo haremos por medio de análisis de diferencias-en-diferencias junto con la combinación de técnicas de emparejamiento —*matching*—, combinando emparejamiento exacto para variables categóricas y emparejamiento por índice de propensión —*propensity score*— para las demás variables. Esto permitirá comparar el grupo tratado (programas de licenciatura) con pares equivalentes o lo más similares posibles en el grupo de control (otros programas profesionales). Los análisis se realizan por cada una de las competencias genéricas del examen Saber Pro, para determinar si los efectos son homogéneos para

todas las competencias o si hay unas que son más afectadas que otras. A continuación, describiremos en detalle la muestra y la estrategia de análisis

3.1 Muestra

Para el análisis utilizamos los resultados de Saber Pro de los años 2012 a 2021 tanto de estudiantes de licenciatura (grupo tratado) como de estudiantes profesionales (grupo control). Como principal variable de control del desempeño de los estudiantes, tenemos los resultados del Examen Saber 11º de los estudiantes. Dados los cambios que ha habido en el Examen Saber 11º a partir de 2014, utilizaremos como variable de control el percentil de los estudiantes en dicho examen, y emparejaremos estudiantes por año de presentación de Saber 11º. Para tener una mejor comparabilidad, hemos restringido la muestra a los estudiantes que presentaron Saber 11º entre 2 y 10 años antes del Examen Saber Pro. La Tabla 1 describe la población escogida por año de presentación de ambos exámenes.

Tabla 1. Muestra por año de presentación de exámenes Saber 11º y Saber Pro

Periodo Saber 11º	Año de presentación Saber Pro										Total	
	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021		
2006-1	2,374	2,173	1,195	873	743							7,358
2006-2	20,244	20,401	11,995	8,441	7,880							68,961
2007-1	1,877	1,786	927	581	540	421						6,132
2007-2	23,491	29,763	17,019	11,527	9,993	8,065						99,858
2008-1	2,565	5,650	2,927	1,753	1,375	1,073	803					16,146
2008-2	4,289	34,084	24,474	16,222	12,849	9,916	7,519					109,353
2009-1	88	1,798	2,104	1,267	1,044	844	618	549				8,312
2009-2	357	8,886	34,006	24,445	17,564	12,948	9,669	8,342				116,217
2010-1	29	448	3,221	2,830	1,976	1,056	666	563	351			11,140
2010-2	409	851	11,822	40,264	33,500	24,161	16,953	14,559	10,433			152,952
2011-1		32	532	2,913	3,213	1,949	1,019	719	499	321		11,197
2011-2		426	936	11,542	47,468	34,282	23,404	18,741	12,594	10,019		159,412
2012-1			31	368	3,221	3,234	1,794	1,102	682	468		10,900
2012-2			392	826	13,969	49,089	33,674	25,425	16,932	13,018		153,325
2013-1				20	419	3,408	3,039	1,833	995	627		10,341
2013-2				305	1,099	14,794	47,264	36,545	23,395	16,953		140,355
2014-1					26	431	2,999	3,206	1,757	905		9,324
2014-2					479	1,121	14,828	50,721	33,223	22,597		122,969
2015-1						23	412	3,151	3,125	1,606		8,317
2015-2						402	947	16,557	49,434	32,828		100,168
2016-1							14	377	2,995	2,540		5,926
2016-2							330	835	15,678	47,905		64,748
2017-1								11	382	2,744		3,137
2017-2								294	902	14,349		15,545
2018-1									6	429		435
2018-2									125	861		986
2019-1										7		7
2019-2										87		87
Total	55,723	106,298	111,581	124,177	157,358	167,217	165,952	183,530	173,508	168,264		1,413,608

Fuente: elaboración propia con datos del ICFES.

La distribución de la muestra por tipo de programa se presenta en Tabla 2. Dado que en 2016 la escala de evaluación fue modificada, estandarizaremos los puntajes y presentaremos los puntajes en desviaciones estándar (σ). Los estudiantes de los programas de licenciatura han obtenido puntajes sistemáticamente más bajos en todos los módulos del examen Saber Pro. Las mayores diferencias se presentan en el módulo de Razonamiento Cuantitativo, mientras que las menores en Comunicación Escrita. El Examen Saber Pro cambió la escala de evaluación en 2016, como se evidencia en la Tabla 2. Para permitir la comparabilidad en el análisis, los resultados se estandarizaron por año.

Tabla 2. Muestra por año y media de puntajes de Saber Pro por tipo de programa

Año	Saber Pro Global				Razonamiento Cuantitativo				Lectura Crítica			
	Otro profesional		Licenciatura		Otro profesional		Licenciatura		Otro profesional		Licenciatura	
	N	μ (σ)	N	μ (σ)	N	μ (σ)	N	μ (σ)	N	μ (σ)	N	μ (σ)
2012	51,569	10.386 (0.877)	4,154	9.986*** (0.848)	51,562	10.352 (1.075)	4,154	9.710*** (0.857)	51,559	10.348 (0.962)	4,154	9.991*** (1.005)
2013	97,385	10.310 (0.835)	8,913	9.954*** (0.807)	97,382	10.325 (1.101)	8,912	9.670*** (0.857)	97,378	10.352 (0.993)	8,911	10.087*** (1.033)
2014	101,173	10.248 (0.844)	10,408	9.872*** (0.819)	101,163	10.263 (1.084)	10,406	9.629*** (0.878)	101,161	10.291 (1.037)	10,407	9.957*** (1.065)
2015	112,977	10.417 (0.870)	11,200	10.110*** (0.832)	112,948	10.425 (1.172)	11,198	9.754*** (0.979)	112,952	10.659 (0.989)	11,200	10.405*** (0.992)
2016	142,318	154.760 (22.799)	15,040	145.272*** (22.690)	142,314	156.310 (30.054)	15,038	139.044*** (26.743)	142,311	155.167 (29.944)	15,036	146.160*** (30.704)
2017	150,821	151.985 (24.470)	16,396	142.031*** (24.366)	150,810	154.790 (30.854)	16,393	136.798*** (27.459)	150,801	155.135 (31.928)	16,395	147.807*** (32.345)
2018	149,918	152.190 (23.702)	16,034	141.975*** (23.733)	149,895	155.551 (30.610)	16,026	137.251*** (28.030)	149,895	154.942 (30.777)	16,030	146.614*** (31.302)
2019	166,577	151.129 (23.614)	16,953	142.477*** (22.993)	166,553	152.023 (31.889)	16,948	134.845*** (27.470)	166,551	153.463 (31.431)	16,950	148.369*** (30.895)
2020	159,448	153.951 (25.944)	14,060	145.483*** (25.481)	159,442	154.347 (32.118)	14,059	137.790*** (28.151)	159,358	156.216 (29.768)	14,047	149.571*** (29.838)
2021	154,652	149.632 (26.617)	13,612	141.690*** (25.194)	154,583	149.712 (33.750)	13,603	132.431*** (28.914)	154,285	151.035 (30.952)	13,576	144.221*** (29.964)

	Competencias Ciudadanas				Inglés				Comunicación Escrita			
	Otro profesional		Licenciatura		Otro profesional		Licenciatura		Otro profesional		Licenciatura	
	N	μ (σ)	N	μ (σ)	N	M (σ)	N	μ (σ)	N	μ (σ)	N	μ (σ)
2012	51,551	10.269 (0.978)	4,154	9.846*** (0.963)	51,496	10.789 (1.452)	4,133	10.286*** (1.374)	50,597	10.392 (1.183)	4,081	10.328*** (1.227)
2013	97,358	10.101 (1.045)	8,904	9.721*** (0.991)	97,266	10.689 (1.448)	8,878	10.287*** (1.392)	96,892	10.150 (0.992)	8,863	10.115*** (1.030)
2014	101,143	10.015 (1.003)	10,403	9.623*** (0.965)	101,048	10.660 (1.458)	10,378	10.240*** (1.363)	100,357	10.113 (0.999)	10,325	10.028*** (1.046)
2015	112,928	10.291 (0.978)	11,194	9.985*** (0.955)	112,784	10.734 (1.443)	11,156	10.456*** (1.404)	111,806	10.109 (1.004)	11,139	10.050*** (1.073)
2016	142,258	154.267 (30.280)	15,028	144.540*** (30.084)	142,058	156.789 (31.630)	14,969	147.959*** (31.388)	140,836	151.934 (30.546)	14,949	150.218*** (32.158)
2017	150,718	145.832 (33.042)	16,371	136.103*** (32.622)	150,622	156.819 (31.152)	16,333	147.205*** (31.281)	144,161	153.146 (32.887)	15,581	150.280*** (33.369)
2018	149,804	146.342 (33.287)	16,004	135.708*** (31.975)	149,619	157.450 (31.441)	15,948	148.907*** (31.647)	145,692	151.406 (31.288)	15,370	148.707*** (31.974)
2019	166,464	146.619 (33.595)	16,930	136.760*** (32.247)	166,237	158.386 (32.231)	16,883	151.111*** (32.423)	161,662	150.023 (27.897)	16,237	148.358*** (27.941)
2020	159,272	156.386 (31.768)	14,038	147.233*** (31.094)	158,874	162.798 (32.584)	13,971	155.164*** (33.125)	154,291	145.565 (33.388)	13,466	145.164*** (33.560)
2021	153,941	148.171 (34.342)	13,529	138.728*** (32.128)	153,406	161.220 (31.236)	13,472	153.722*** (31.165)	151,203	143.633 (34.647)	13,303	145.540*** (35.559)

Nota: μ : promedio; σ : desviación estándar. Resultados de la prueba T en asteriscos. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Fuente: elaboración propia con datos del ICFES.

Los programas a los cuales pertenecen los estudiantes se presentan en la Tabla 3. Durante los años 2012 a 2021, 839 programas obtuvieron acreditación de alta calidad, de los cuales 183 son licenciaturas. De estas, 128 obtuvieron dicho reconocimiento durante la vigencia de la normatividad surgida del *PND 2014-2018*. Valga la pena mencionar que hubo programas que cambiaron de denominación para poderse ajustar a la normatividad (MEN, 2016, feb. 3; 2017, sept. 15), o incluso hubo algunos que decidieron presentar un nuevo registro calificado para no tener que presentarse al proceso de acreditación de alta calidad y así evitar el cierre del programa (Arias *et al.*, 2018).

Tabla 3. Estudiantes tipo de programa y año de acreditación

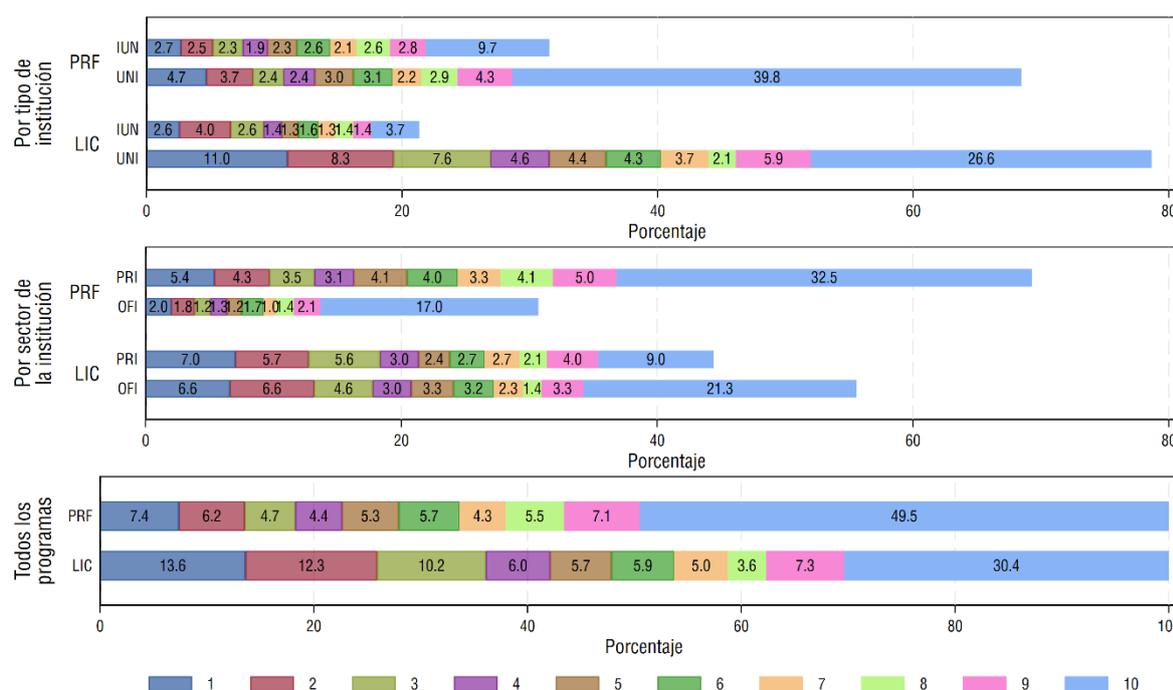
Año de acreditación	Otros profesionales		Licenciatura		Total	
	Programas	Estudiantes	Programas	Estudiantes	Programas	Estudiantes
Antes de 2011	547	338,461	71	21,580	618	360,041
2012	104	55,444	23	6,552	127	61,996
2013	48	19,402	7	1,827	55	21,229
2014	52	27,607	11	2,444	63	30,051
2015	60	29,920	11	2,751	71	32,671
2016	83	40,030	23	5,400	106	45,430
2017	61	31,393	95	18,553	156	49,946
2018	80	29,086	5	546	85	29,632
2019	89	35,681	1	106	90	35,787
2020	55	21,638	4	560	59	22,198
2021	24	6,585	3	477	27	7,062
Total acrad. entre 2012 y 2021	1,203	635,247	254	60,796	1,457	696,043
Sin acreditación	2,682	651,591	444	65,974	3,126	717,565
Total	3,885	1,286,838	698	126,770	4,583	1,413,608

Fuente: elaboración propia con datos del ICFES y SNIES (<https://hecaa.mineducacion.gov.co/consultaspublicas/programas>).

En la Figura 1 presentamos el porcentaje de programas por tipo, y el número de años que el programa incluye estudiantes en la presentación del examen Saber Pro. El total de años del estudio es 10 (2012 a 2021), pero como muestra la Figura 1, solamente el 30 por ciento de las licenciaturas tuvieron estudiantes presentando el examen a lo largo de todo el periodo de estudio. Contrasta con lo que sucede con los programas profesionales, en donde cerca del 50% de los programas tuvieron estudiantes presentando el examen a lo largo del mismo periodo. Por otra parte, el 36 por ciento de los programas reporta estudiantes durante uno (13.6%), dos (12.3%) o tres años (10.2%). Esto evidencia un significativo cambio en la oferta de programas de licenciatura. En parte debido a la normatividad para las licenciaturas, algunas de estas cambiaron su código SNIES por lo que se pudo haber perdido la trazabilidad del programa y no es comparable el puntaje de los estudiantes previo a la promulgación de nuevas condiciones de calidad para las

licenciaturas (Rodríguez-Ávila *et al.*, 2018). Adicionalmente, algunas universidades, sobre todo privadas y con acreditación, decidieron abrir programas de licenciatura como consecuencia de la nueva normatividad de las licenciaturas (Arias, *et al.*, 2018). Ello hace que no sea posible utilizar en el análisis solamente los programas que se mantuvieron durante todo el periodo estudiado, pues quedarían por fuera cerca del 70% de los programas. Los cambios se dieron mayormente en las instituciones privadas, en donde solamente el 9% de los programas se mantuvo a lo largo del periodo, mientras que en las públicas fue el 21%. De igual manera, las instituciones universitarias presentaron significativamente menos programas durante el total del periodo (3.7%) que las universidades (26.6%). Hemos decidido incluir en el estudio la totalidad de los programas, porque la política contemplaba justamente el cierre de los programas con el peor desempeño (Ariza, 2018).

Figura 1. Número de años en que un programa reporta estudiantes que presentan Examen Saber Pro.



Nota: PRF: otros programas profesionales. LIC: Licenciaturas. IUN: Institución Universitaria; UNI: Unieversidad. PRI: Institución de carácter privado. OFI: Institución oficial

Las estadísticas descriptivas de las variables de control se presentan en la Tabla 4. Hay un mayor porcentaje de mujeres en las licenciaturas, y la edad promedio es mayor. Así mismo, los antecedentes académicos de los estudiantes (percentil en el Examen Saber

11^o) son menores. Adicionalmente, los estudiantes de los programas de licenciaturas provienen de contextos socioeconómicos más desfavorables en términos de estrato socioeconómico, educación de los padres, así como del Índice de Pobreza Multidimensional del departamento del que provienen [IPM], y del Índice de Necesidades Básicas Insatisfechas del municipio en el que viven [NBI].

Tabla 4. Estadísticas descriptivas y test comparativos de las variables de control

Variable	Otro profesional			Licenciatura			Diferencia
	N	μ o %	σ	N	μ o %	σ	
Mujer (*)	757,936	58.90%		85,733	67.63%		8.73%***
Estrato (*)							
Rural o sin estrato	3,277	0.26%		933	0.76%		0.49%***
Estrato 1	163,822	13.02%		36,072	29.20%		16.18%***
Estrato 2	393,782	31.30%		50,707	41.05%		9.75%***
Estrato 3	429,458	34.13%		29,967	24.26%		-9.87%***
Estrato 4	166,590	13.24%		4,446	3.60%		-9.64%***
Estrato 5	65,015	5.17%		1,004	0.81%		-4.35%***
Estrato 6	36,235	2.88%		392	0.32%		-2.56%***
Trabaja actualmente (*)	805,347	64.05%		86,550	70.07%		6.02%***
Edad	1,286,816	23.252	2.071	126,763	23.780	2.719	0.528***
IPM	1,286,838	0.200	0.037	126,770	0.209	0.038	0.009***
NBI	1,286,838	7.988	8.604	126,770	11.671	14.253	3.683***
Educación del padre	1,264,262	10.781	5.214	124,157	8.281	4.984	-2.500***
Educación de la madre	1,262,707	11.562	4.569	124,005	9.266	4.628	-2.296***
Valor de la matrícula	1,276,568	4.219	3.105	126,258	1.523	1.417	-2.695***
Percentil en Saber 11 ^o	1,286,832	71.534	24.421	126,768	60.233	26.610	-11.302***
Programa a distancia (*)	101,089	7.86%		26,400	20.83%		12.97%***
institución acreditada (*)	556,036	43.21%		51,567	40.68%		-2.53%***
Institución pública (*)	399,318	31.03%		82,375	64.98%		33.95%***
Universidad (*)	930,198	72.29%		97,250	76.71%		4.43%***

Nota: μ: promedio; σ: desviación estándar. Las variables categóricas, indicadas con (*) se evalúan con prueba de χ^2 (chi cuadrado). Las variables continuas se evalúan con prueba t. * p < 0.05, ** p < 0.01, *** p < 0.001.

No fueron incluidas la variable de región o departamento, debido a que estudiantes del mismo programa pueden vivir en ciudades o regiones distintas, sobre todo en el caso de los programas a distancia. En su lugar, y por principio de parsimonia, hemos añadido las dos variables socioeconómicas de las regiones: IPM y NBI, los cuales se han asociado con el desempeño de los estudiantes y actualmente se utilizan como uno de los criterios para la asignación de recursos para educación básica por municipio (Barrera-Osorio *et al.*, 2012; Fajardo *et al.*, 2021). Los valores para estas variables han sido tomadas de los informes emitidos del censo nacional del año 2018 (Departamento Administrativo Nacional de Estadística [DANE], 2018).

En términos de los tipos de instituciones y programas, hay un porcentaje significativamente mayor de estudiantes de licenciaturas en programas a distancia. De igual manera, la población de estos programas se encuentra en mayor proporción en

universidades públicas. Por otra parte, hay proporcionalmente menos estudiantes en universidades con acreditación de alta calidad. En Colombia hay dos tipos de instituciones de educación superior que ofrecen títulos profesionales: las instituciones universitarias o escuelas tecnológicas, y las universidades. El reconocimiento como universidad es otorgado por el MEN en función de una mayor cobertura de áreas de conocimiento, número de programas, enfoque en investigación y en posgrados. Los estudiantes de licenciatura atienden mayoritariamente universidades.

3.2 Primera fase de análisis: interacción de acreditación de alta calidad con tipo de programa por año

En la primera parte se usará análisis de regresión multivariado, examinando la interacción entre el tipo de programa (profesional u otra licenciatura), el reconocimiento ante el MEN (registro calificado o acreditación de alta calidad) y el año en que el estudiante presentó el Examen Saber Pro. Esto permitirá examinar los cambios en el tiempo que ha tenido la asociación entre la acreditación de alta calidad y el desempeño de los estudiantes. El análisis de regresión variable será estimado usando el siguiente modelo (Osborne, 2017; Weisberg, 2014):

$$y = (\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \beta_{123}x_1x_2x_3 + \tilde{x}\tilde{\beta} + \epsilon)(1)$$

Donde y es el puntaje obtenido por el estudiante en cada uno de los módulos de Saber Pro y en el puntaje global del Examen. β_0 es el intercepto del puntaje para el total de estudiantes. El valor de x_1 es 1 si se trata de una licenciatura. x_2 es 1 si es un programa acreditado, y x_3 es una variable categórica indicando el año de presentación del examen. β_{123} es nuestro coeficiente principal de interés, el cual estima la interacción de las variables x_1 , x_2 y x_3 . $\tilde{x}\tilde{\beta}$ denota las variables de control, y ϵ es el error estándar. Los errores estándar se calculan de manera robusta, en donde la varianza de la observación i se calcula mediante $\hat{\sigma}_i^2 = \{n/(n - k)\}u_i^2$ en donde n es el número total de observaciones, k es el número de regresores, u_i es el residual estimado, y $n/(n - k)$ se incluye para mejorar las estimaciones con muestras pequeñas (Imbens & Kolesár, 2016). Los cálculos se realizaron utilizando el software Stata 18®.

3.3 Segunda fase de análisis: enfoque cuasiexperimental

El impacto de las políticas específicas de formación de docentes que se estimará utilizando una aproximación cuasiexperimental, combinando la técnica de Diferencias en Diferencias [DiD] (Donald & Lang, 2007; Lechner, 2011) con emparejamiento (*matching*) entre grupo tratado (licenciaturas) y grupo control (otros programas profesionales). El DiD nos permitirá establecer si la tendencia del desempeño de los estudiantes de licenciatura es diferente antes y después de la implementación de estas políticas, tomando como grupo de control a los estudiantes de otros programas profesionales, los cuales no sufrieron dichos cambios en su normatividad. Utilizar *matching* nos permitirá evaluar la compatibilidad entre ambos grupos, dada la heterogeneidad que existe entre ellos (cf. Tabla 4). Usaremos como periodo pre-tratamiento los resultados de los estudiantes de los años 2012 a 2016. El requisito de acreditación dio dos años a los programas para obtener dicho reconocimiento (CRC, 2015, jun. 9, art. 222), y la normatividad se inauguró en el año 2015 pero quedó finalmente establecida en el año 2016 (MEN, 2016, feb. 3). Por ello, es posible hallar efectos en el año 2017. Sin embargo, dada la adaptación curricular que debieron hacer los programas, es posible que estos efectos no se presenten en dicho año sino en años posteriores, dado que los planes curriculares toman de 4 a 5 años en implementarse.

3.3.1 Emparejamiento de los grupos

Dada la heterogeneidad que se presenta entre los estudiantes de acuerdo con el tipo de programa (cf. Tabla 4), estimamos la probabilidad de los estudiantes de matricularse a un programa de licenciatura, de manera que se comparen estudiantes con similares características socioeconómicas, y se descarten observaciones que no tengan una contraparte similar tanto en el grupo tratado (licenciaturas) como en el grupo control (otros profesionales), de manera que se mitigue el sesgo por selección de programa (Stuart *et al.*, 2014). Esto lo haremos verificando lo que Rosenbaum & Rubin denominan el supuesto de independencia condicional (*conditional independence assumption* [CIA]), que implica que las variables de resultado para los grupos tratado (Y_0) y control (Y_1) son independientes de la escogencia del tipo de programa. Esto quiere decir que el efecto promedio del tratamiento (*average treatment effect* [ATE]) es independiente del programa escogido cuando se condiciona por un conjunto de covariables observadas (X).

En otras palabras, CIA: $(Y_i, Y_0) \perp\!\!\!\perp D \mid X$. Este cálculo se hará utilizando la técnica de emparejamiento por índice de propensión —*Propensity Score Matching* [PSM]—, combinada con ajuste de regresión (*regression adjustment*), y para las variables categóricas, se empleará emparejamiento exacto (*exact matching*). Ello permitirá, además, balancear las covariables de manera que los grupos de comparación sean lo más homogéneos posible, no solo en términos del índice de propensión sino también de las demás covariables del modelo.

El índice de propensión —*propensity score* [PS]— es la probabilidad condicional de que el estudiante haya escogido un programa de licenciatura en función de las covariables (X), y se calcula mediante la siguiente ecuación (Rosenbaum & Rubin, 1983):

$$p(X) = Pr\{T = 1|X\} = E\{T|X\} \quad (3)$$

Para determinar la ponderación del emparejamiento (*matching weight*), utilizamos el cálculo de índice de propensión por kernel, en donde el peso del emparejamiento es definido así:

$$W_{ij} = \frac{K(P_j - P_i)/a}{\sum_{k \in D_k = 0} G(P_k - P_i)/a} \quad (4)$$

Donde K is la función de densidad de kernel de Epanechnikov: $K(x) = \frac{3}{4}(1 - x^2)$, para $|x| < 1$, a es el parámetro de ancho de banda, P_j es el PS para la observación j en el grupo control, P_i denota el PS de la observación i en el grupo tratado, y $P_j - P_i$ representa la distancia entre PSs. El ancho de banda se calcula por validación ponderada cruzada con respecto a Y (Frölich, 2004, 2005; Galdo *et al.*, 2008). Las ponderaciones del emparejamiento son normalizadas, de manera que, para cada estudiante en un programa de licenciatura, $\sum_{j \in D=0} W_{ij} = 1$. Se aplica adicionalmente balanceo de entropía para conseguir el balanceo de las covariables de los grupos, ajustando las diferencias en desviación estándar, varianza y asimetría, y así reducir la dependencia del modelo del ATE (Galdo, *et al.*, 2008; Hainmueller 2012). Adicionalmente, como se mencionó anteriormente, el PS se recalculará por medio de ajuste de regresión por covariables, para añadir doble robustez al modelo y eliminar el sesgo de selección (Huber *et al.*, 2012; Jann, 2017a; King & Nielsen, 2019; Rubin & Thomas, 2000).

$$Y_i = (\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \epsilon_i) \quad (6)$$

Las covariables X empleadas son las siguientes: género, edad, IPM, NBI, educación de los padres, estrato socioeconómico, condición laboral, matrícula de la universidad, y percentil en el Examen Saber 11º. Adicionalmente, se hará emparejamiento exacto por variables categóricas, que incluyen género, estrato socioeconómico y condición laboral como condiciones individuales. También incluiremos el año de presentación del Examen Saber 11º, debido a los problemas de comparabilidad de este examen a lo largo de los años. Incluiremos también emparejamiento exacto por las variables de la institución a la cual pertenece el estudiante (pública/privada; institución universitaria/universidad) en función de las diferencias que estas presentan en los análisis descriptivos. El ATT lo estimaremos año por año, para evaluar las tendencias y así poder utilizar la ponderación del emparejamiento (*matching weight*) para el DiD que presentaremos a continuación. La estimación la realizamos con el software Stata 18[®] y para el cálculo del PS utilizamos el comando *kmatch* (Jann, 2017b).

3.3.2 Estimación de DiD

El modelo DiD que emplearemos (Donald & Lang, 2007) es el siguiente:

$$DiD = (\bar{y}_{p=1,t=1} - \bar{y}_{p=1,t=0}) - (\bar{y}_{p=0,t=1} - \bar{y}_{p=0,t=0}) \quad (2)$$

Donde \bar{y} es el valor promedio de la variable resultado a nivel del estudiante, p es el tipo de programa, donde 1 es el grupo tratado (licenciatura) y 0 es el grupo control (otros programas profesionales). el subíndice t denota el tiempo, donde 0 es antes de la intervención, y 1 si es posterior a ella. El resultado de la estimación es el *efecto del tratamiento en el grupo tratado (average treatment effect on treated [ATET])*. \bar{y} se estima utilizando mínimos cuadrados ordinarios (OLS) con base en el siguiente modelo:

$$y_{ipt} = \gamma_p + \gamma_t + X_{ipt}\beta + D_{pt}\delta + \epsilon_{ipt}$$

$$y_{ipt} = DiD_{ipt} + \epsilon_{ipt} \quad (3)$$

En donde i es el estudiante, p el programa al que pertenece, y t es tiempo del tratamiento (antes/después). La variable γ_p denota los efectos fijos de programa, la

variable γ_t los efectos fijos de tiempo; X_{ipt} son las covariables del modelo y D_{pt} es el tratamiento que varía de acuerdo con el programa y el tiempo, que es igual a 1 si el estudiante pertenece al programa tratado p (licenciatura) en el año t , y 0 en caso contrario. ϵ_{ipt} es el error estándar. El coeficiente δ es el ATET. Con ello controlamos por diferencias sistemáticas entre el periodo anterior y posterior que son comunes a ambos grupos (γ_t) y, adicionalmente, controlamos por diferencias sistemáticas entre el grupo tratado y el grupo control que son constantes a lo largo del tiempo (X_{ipt}). Adicionalmente, los efectos fijos de programa (γ_p) capturan la gestión institucional al interior de los programas, así como el hecho de que cada programa tenga una estructura curricular y unos mismos objetivos de aprendizaje, lo que hace que las observaciones no sean independientes, sino que estén agrupadas en clústeres, que en este caso son los programas a los que pertenecen los estudiantes. De igual manera, el cálculo de los errores estándar se hará por clúster (Cameron & Miller, 2015), en este caso, por programa.

Verificaremos dos supuestos fundamentales de todo modelo DiD. El primero es el de compatibilidad entre el grupo tratado y el grupo control, el cual evaluaremos balanceando ambos grupos por medio de emparejamiento (Fredriksson & Olivera, 2019; Lindner & McConnell, 2019), como hemos explicado en la sección anterior. El segundo supuesto es el de tendencias paralelas, el cual supone que, en ausencia del tratamiento, el grupo tratado habría seguido la misma tendencia en comparación con el grupo de control. Por ello, usaremos como periodo pre-tratamiento los resultados de los estudiantes de los años 2012 a 2016. Dado que el requisito de acreditación dio dos años a los programas para obtener dicho reconocimiento (CRC, 2015, jun. 9, art. 222), y la normatividad se inauguró en el año 2015 pero quedó finalmente establecida en el año 2016 (MEN, 2016, feb. 3), es posible hallar efectos en el año 2017. Sin embargo, dada la adaptación curricular que debieron hacer los programas, es posible que estos efectos no se presenten en dicho año sino en años posteriores. Los planes curriculares toman de 4 a 5 años para implementarse completamente para los estudiantes nuevos, por lo que su efecto puede ser progresivo dependiendo de la estrategia de transición que haya tomado cada programa. Por ello, verificamos el supuesto de tendencias paralelas para los años 2012 a 2016 y evaluaremos el ATET para los años 2017 a 2021.

El supuesto de tendencias paralelas se evalúa utilizando el siguiente modelo:

$$y_{ipt} = DiD_{ipt} + w_i d_{t,0} t \zeta_1 + w_i d_{t,1} t \zeta_2 + \epsilon_{ist} \quad (4)$$

En donde la variable $d_{t,0}$ ($d_t = 0$) denota los periodos previos al tratamiento, y la variable $d_{t,1}$ ($d_t = 1$) los periodos posteriores al tratamiento. La variable $w_i = 1$ si el individuo pertenece al grupo tratado y $w_i = 0$ en caso contrario. El coeficiente ζ_1 captura las diferencias de las pendientes entre el grupo tratado y el grupo control en los años previos al tratamiento, mientras que ζ_2 captura las diferencias en las pendientes de los grupos posteriormente al tratamiento. Si ζ_1 es 0, las tendencias lineales del resultado son paralelas previamente al tratamiento (años 2012 a 2016). Lo mismo sucede con ζ_2 , pero el resultado es irrelevante para probar tendencias paralelas. Utilizaremos la prueba de Wald para evaluar si $\zeta_1 = 0$. Por ello, la hipótesis nula de la prueba de tendencias paralelas es que las tendencias lineales son paralelas.

En los resultados presentaremos la estimación del ATET sin emparejamiento y con emparejamiento. Por medio de PSM no es posible hacer test de tendencias paralelas, dado que la estimación se hace año por año, y además se trata de una prueba no paramétrica, por lo que sus resultados no pueden considerarse concluyentes. Adicionalmente, para el PSM la estimación del error estándar se hace de manera robusta por individuo, mientras que para el DiD se hace por clúster —*i.e.*, por programa—, lo que arroja intervalos de confianza distintos. Al final presentaremos los resultados de la estimación del ATET combinando el modelo de DiD pero utilizando la muestra balanceada por medio del PS y utilizando la ponderación del emparejamiento (Lindner & McConnell, 2019).

Por último, para identificar qué características de los estudiantes, programas o instituciones favorecieron mayores efectos, estimaremos el ATET por subgrupos y compararemos los resultados utilizando test de Wald, para determinar si las diferencias son significativas entre subgrupos. Para las variables dicotómicas emplearemos cada categoría (género, acreditación institucional y de programa, tipo y sector de la institución, modalidad del programa, estudiante trabaja). Para las variables continuas, dividiremos los grupos en niveles alto y bajo: percentil en Saber 11º ($\leq 65 <$), educación total de los padres (≤ 16 años entre padre y madre $<$), NBI ($\leq 6.2 <$), IPM ($\leq 0.206 <$), estrato ($\leq 2 <$), valor matrícula universidad (< 1 salario mínimo legal vigente \leq). Adicionalmente, dividiremos la muestra por el número de años en que el programa presentó estudiantes al Examen Saber

Pro (cf. Figura 1) para determinar qué tanto efecto causó el cambio en la oferta de los programas. Y para finalizar, dividiremos las licenciaturas de acuerdo con las áreas de conocimiento establecidas por la normatividad (MEN, 2016, feb. 3, 2017 sept. 15), para determinar si hubo diferencias de efecto de acuerdo con el tipo de programa. En este caso, evaluaremos la diferencia entre el ATET por área y el ATET general.

4 Resultados

4.1 Desempeño de los estudiantes por tipo de programa y acreditación

A continuación, presentamos el resultado de las estimaciones del modelo de interacción entre tipo de programa y acreditación de alta calidad por año de presentación del examen, para cada uno de los módulos de Saber Pro:

Tabla 5. Resultados de la regresión multinivel, con interacciones por año, tipo de programa y acreditación de alta calidad (versión abreviada)

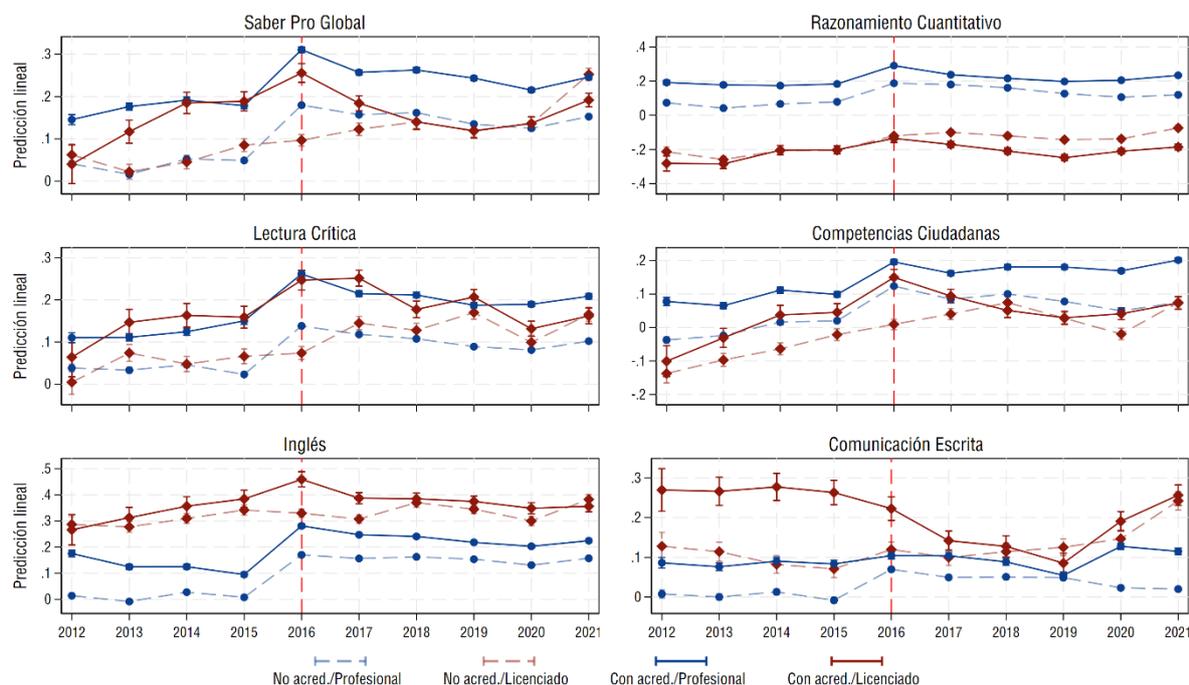
	Puntaje global	Razonamiento Cuantitativo	Lectura Crítica	Competencias Ciudadanas	Inglés	Comunicación Escrita
Año=2013 x Lic. x Prog. acreditado=1	0.0611	0.0253	0.0075	0.0553	0.0842*	0.0127
Año=2014 x Lic. x Prog. acreditado=1	0.1276***	0.0815*	0.0495	0.0841*	0.1312**	0.0546
Año=2015 x Lic. x Prog. acreditado=1	0.1006**	0.0847**	-0.0217	0.0664*	0.1371***	0.0374
Año=2016 x Lic. x Prog. acreditado=1	0.1547***	0.0684*	0.0623	0.1462***	0.2006***	0.0053
Año=2017 x Lic. x Prog. acreditado=1	0.0882**	0.0582	0.0225	0.0547	0.1706***	-0.0752*
Año=2018 x Lic. x Prog. acreditado=1	0.0261	0.0425	-0.0419	-0.0258	0.1180**	-0.0876*
Año=2019 x Lic. x Prog. acreditado=1	0.0171	0.0108	-0.0489	-0.0233	0.1469***	-0.1088**
Año=2020 x Lic. x Prog. acreditado=1	0.0368	0.0158	-0.0638*	0.0194	0.1583***	-0.1234**
Año=2021 x Lic. x Prog. acreditado=1	-0.0275	-0.0391	-0.0979**	-0.0500	0.0879*	-0.1440***
Covariables						
Mujer	-0.0594***	-0.3627***	0.0378***	0.0121***	-0.0928***	0.1551***
Edad	-0.0276***	-0.0261***	-0.0225***	-0.0155***	-0.0203***	-0.0162***
IPM	-2.7429***	-2.1739***	-1.9887***	-2.4815***	-2.1720***	-1.5724***
NBI	-0.0024***	-0.0030***	-0.0027***	-0.0016***	-0.0003**	-0.0024***
Ed. Padre	0.0056***	0.0008***	0.0039***	0.0018***	0.0128***	0.0014***
Ed. Madre	0.0092***	0.0039***	0.0069***	0.0068***	0.0155***	0.0017***
Estrato	0.0584***	0.0126***	0.0181***	0.0120***	0.1599***	0.0127***
Trabaja	-0.0363***	0.0150***	-0.0514***	-0.0368***	-0.0622***	-0.0051**
Valor matrícula universidad (SMLV)	0.0203***	0.0088***	0.0048***	0.0055***	0.0421***	0.0156***
Percentil Saber 11º	0.0199***	0.0175***	0.0189***	0.0176***	0.0143***	0.0082***
Institución pública	0.1811***	0.2441***	0.1367***	0.1189***	0.1199***	0.0815***
Universidad	0.0737***	0.0474***	0.0818***	0.0786***	0.0315***	0.0570***
Intercepto	-0.6995***	-0.1963***	-0.6993***	-0.7023***	-1.0647***	-0.2024***
N	1,364,528	1,364,410	1,364,378	1,363,995	1,363,038	1,336,368
Log likelihood	-1,460,182	-1,611,975	-1,626,238	-1,688,599	-1,569,262	-1,839,427
R ² ajustado	0.4997***	0.3897***	0.3648***	0.3137***	0.442***	0.0989***

Nota: Lic.: programa de licenciatura. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Se llevó a cabo la prueba Wald para evaluar si el modelo con interacciones tiene mayor poder explicativo (R^2) que el modelo sin interacciones. Los resultados en la fila de R^2 . Cálculos propios del autor con base en datos del Icfes. La versión completa se presenta en el

Anexos

Anexo 1.

Figura 2. Marginales predictivos para la interacción año por tipo de programa y acreditación de alta calidad. Intervalo de confianza 95%



Como se observa en los resultados (Tabla 5, Figura 2), la tendencia de los puntajes de los estudiantes de programas de licenciatura acreditados es significativamente diferente de los puntajes de los estudiantes de otros programas acreditados en todos los módulos del examen. Esto evidencia un cambio significativo de la asociación entre la acreditación del programa y el desempeño de los estudiantes. Los cambios son más notorios justo después de 2016, año en el que empezó a regir la normatividad específica para las licenciaturas. Si bien en el puntaje global se muestra un descenso del desempeño de las licenciaturas entre los años 2018 a 2020, luego hay un aumento significativo con respecto a los programas profesionales, lo que parece indicar que, de haber un efecto en las políticas, este se presenta de manera gradual en el tiempo, y no parece ser igual en cada una de las competencias. Es notorio que los puntajes aumentan considerablemente más para las licenciaturas que no contaban con acreditación de alta calidad, lo que se evidencia

por el hecho de que el desempeño de los estudiantes de licenciatura no se correlaciona con la acreditación en los últimos años. Esto sucede tanto en el puntaje global, en donde en el último año los estudiantes de licenciaturas no acreditadas superan significativamente los puntajes de los estudiantes de licenciaturas acreditadas. Lo mismo sucede en Razonamiento Cuantitativo a partir del año 2017. En las demás competencias, en el año 2021, no se observan diferencias significativas en función de acreditación para las licenciaturas, mientras que en los demás programas profesionales sí hay una diferencia significativa de desempeño.

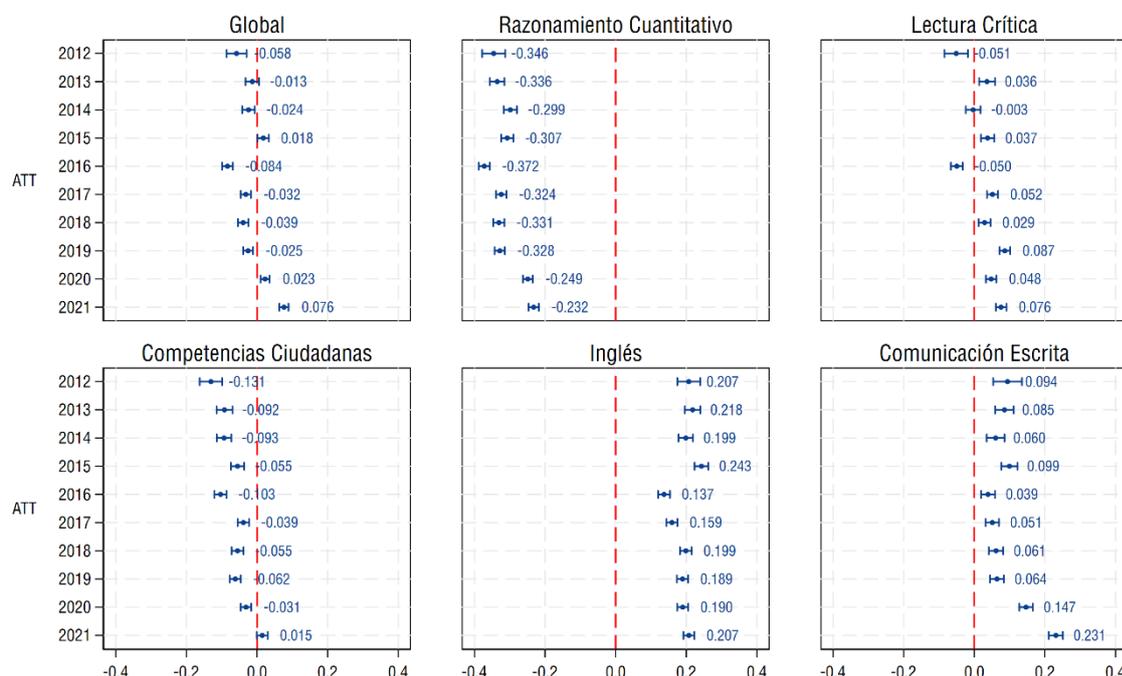
La diferencia más constante entre las licenciaturas y los demás programas profesionales se presentan en el módulo de Razonamiento Cuantitativo, inferior para las licenciaturas, pero contraria en el módulo de Inglés, en donde es mayor para estos programas. En ambos casos la acreditación se comporta de manera diferente al resto de los programas profesionales, en donde los programas sin acreditación obtienen puntajes más altos en ambos módulos. En el módulo de Comunicación Escrita los estudiantes de licenciatura también han tenido un mejor desempeño histórico, si bien en el año 2012 había una diferencia significativa entre licenciaturas con y sin acreditación, esta diferencia ha ido desapareciendo. En Competencias Ciudadanas, la acreditación fue un factor diferencial, pero en los últimos años la brecha se ha cerrado, si bien se mantiene una diferencia importante con respecto a otros programas profesionales acreditados. Las variables de control se comportaron de acuerdo como indica la literatura; sin embargo, valga la pena señalar que las variables IPM y NBI resultaron ser significativas, dando respaldo a la hipótesis que planteamos al comienzo que estas variables pueden explicar las diferencias de desempeño entre estudiantes. En términos de las instituciones, fue positivo en el desempeño de los estudiantes pertenecer a institución pública y universidad, que pertenecer a institución privada e institución universitaria o escuela tecnológica.

4.2 Resultados del emparejamiento de los grupos

A continuación, presentamos el resultado del emparejamiento de los grupos junto con el cálculo del ATT de los años 2012 a 2021 (Figura 3). En este caso, el ATT es la diferencia de puntaje entre los estudiantes de licenciatura con respecto a los programas profesionales. Dado que los resultados han sido estandarizados, los valores se presentan en desviaciones

estándar $[\sigma]$. En el Anexo 2 presentamos los resultados detallados, incluyendo el tamaño de los grupos antes y después del emparejamiento, así como el ancho de banda (*bandwidth*) y la pérdida de balance entre los grupos (*loss*). La corrección del sesgo se calcula con base en el efecto previo al emparejamiento (*Naive Average Treatment [NATE]*).

Figura 3. ATT para licenciaturas por módulo en comparación con otros programas



Se aprecia una mejora en el puntaje global del examen Saber Pro para los estudiantes de licenciatura, especialmente en los últimos dos años. En el módulo de Razonamiento Cuantitativo hay sistemáticamente una diferencia negativa, que llega a ser hasta de 0.37 σ , aunque también parece haber una reducción de la brecha en los últimos años. En el módulo de lectura parece también haber mejoría en los últimos años, sobre todo desde 2017, al igual que sucede con el módulo de Competencias Ciudadanas. Aunque el módulo de Inglés ha sido históricamente positivo para las licenciaturas, no parece haber una mejora significativa. Por último, en el módulo de Comunicación Escrita, que históricamente ha sido positivo para las licenciaturas, parece también haber una mejora en los últimos años con respecto a los años anteriores.

Figura 4. Balanceo de grupos por índice de propensión

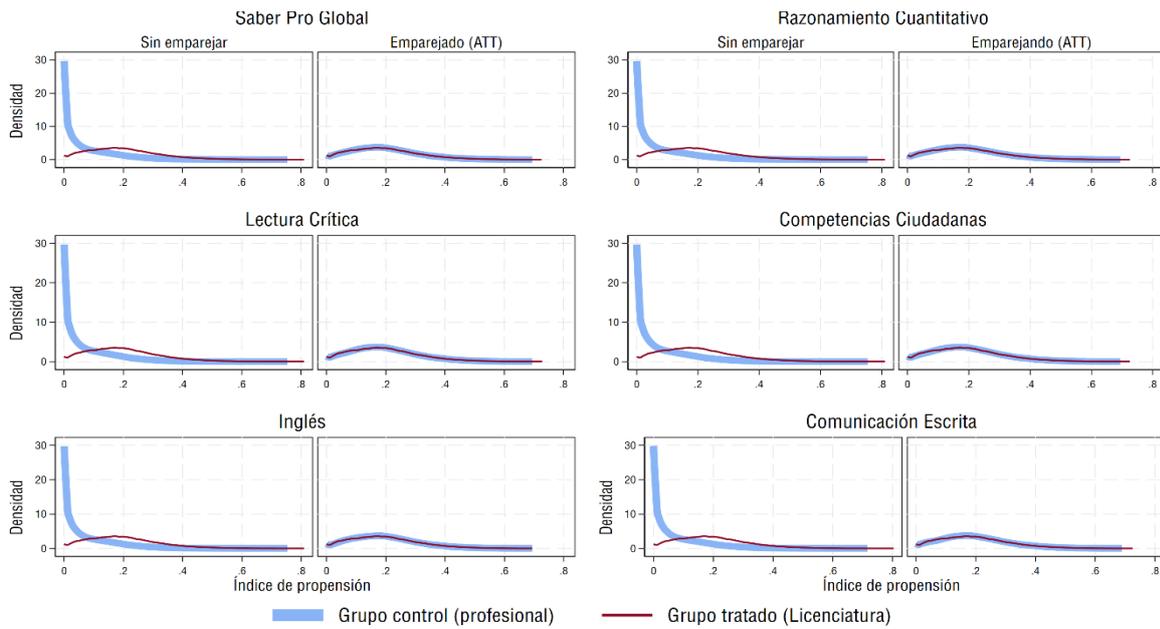
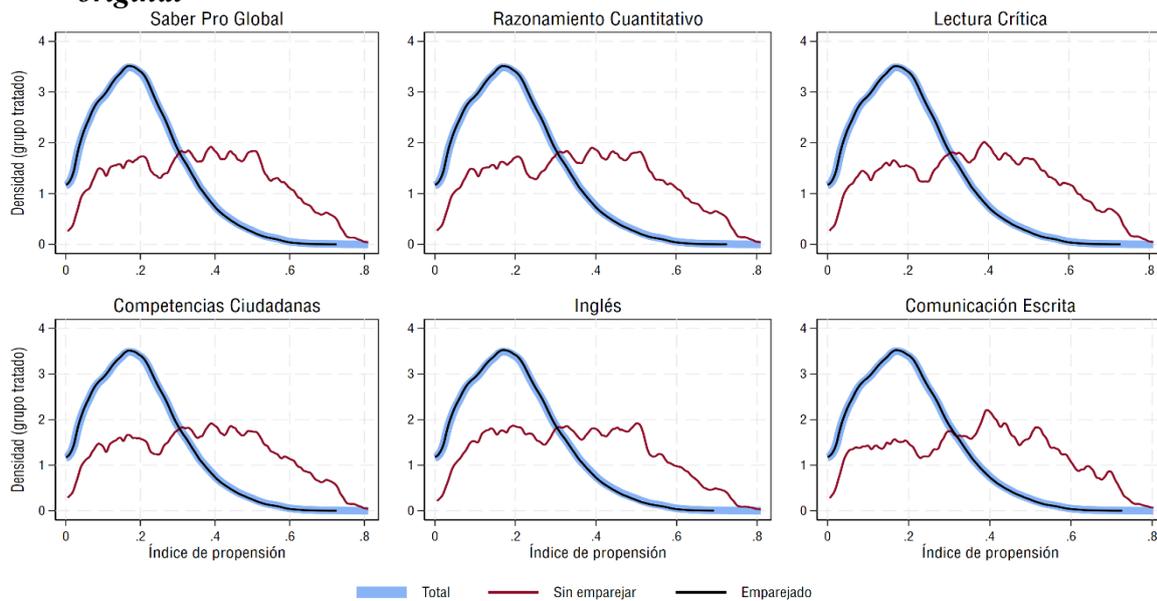
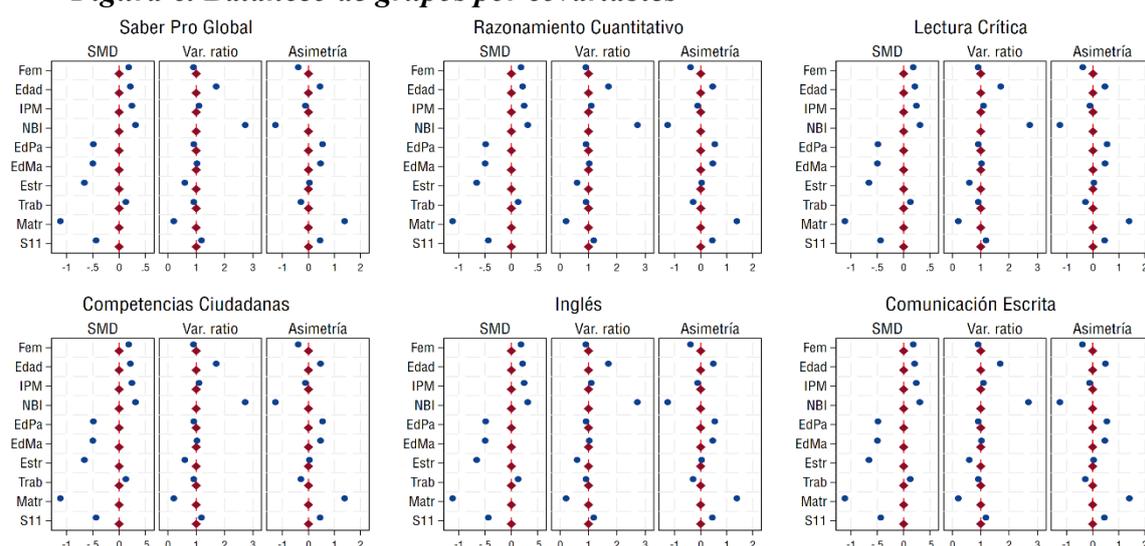


Figura 5. Compatibilidad de índice de propensión entre el grupo emparejado y el original



En términos de la calidad del emparejamiento, como se evidencia en la Figura 4, ha habido una corrección significativa del grupo de control debido a las considerables diferencias que tenía con el grupo tratado en función del índice de propensión. Ello permite mitigar el sesgo de selección, es decir, una vez hecho el emparejamiento, hay una probabilidad similar de que los estudiantes hayan escogido un programa u otro. También ha habido una corrección significativa del sesgo en función de las covariables del modelo.

Figura 6. Balanceo de grupos por covariables

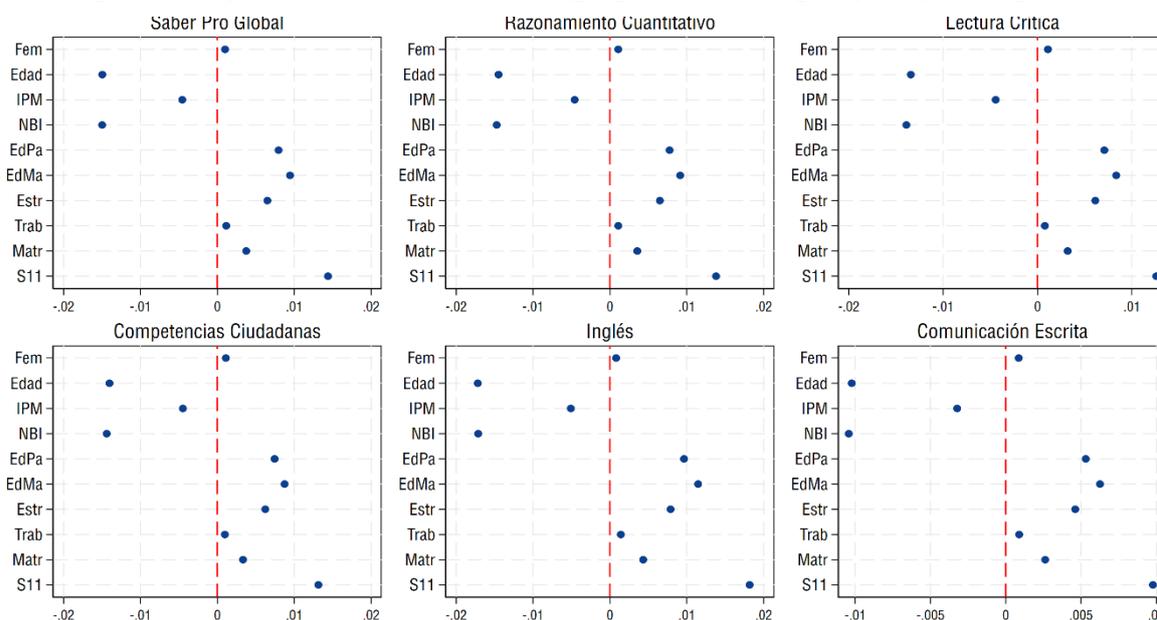


Nota: SMD: Diferencia de media estándar; Fem: Mujer; IPM: Índice de Pobreza Multidimensional por Región; NBI: Índice de Necesidades Básicas Insatisfechas por Municipio; EdPa: Educación del padre; EdMa: Educación de la madre; Estr: Estrato socioeconómico; Trab: Estudiante trabaja; Matr: Valor matrícula universidad; S11: Percentil en Saber 11

La Figura 6 muestra que ha habido una corrección considerable del balanceo de los grupos en función de las covariables, mitigando significativamente el sesgo de selección. Los grupos emparejados no presentan diferencias significativas en diferencia de la media estándar (SMD), la cual es casi nula, a comparación con los grupos sin emparejar, en donde la SMD era considerablemente más alta. La ratio de la varianza ha sido corregida a valores muy cercanos a 1, y la diferencia en asimetría (*skewness*) de las covariables también ha sido corregida a valores muy cercanos a cero. La pérdida de balance total (*loss*), o sesgo remanente, es muy cercana a cero ($2.24e-12 - 8.87e-17$) (Anexo 2, *Loss*). En el Anexo 3 presentamos la comparación del SMD, ratio de la varianza y asimetría del grupo control y el grupo tratado antes y después del emparejamiento, calculados año a año. Dado que los cálculos del ATT se hicieron año por año, también utilizamos los balanceos de los grupos

con respecto a las covariables de acuerdo con el año del examen. También se comprobó la compatibilidad del grupo emparejado con respecto al grupo tratado original, de manera que no haya cambios significativos debido al emparejamiento. Como lo muestra la Figura 5, no hay cambios significativos entre el grupo tratado posterior al emparejamiento en función del índice de propensión. De igual manera, tampoco hay cambios significativos del grupo tratado posterior al emparejamiento en función de las covariables (Figura 7), pues todas presentan valores inferiores a 0.1, el umbral establecido por la literatura (Nguyen *et al.*, 2017). En el Anexo 4 mostramos con detalle la compatibilidad entre el grupo emparejado y el grupo original de acuerdo con las covariables, mostrando los resultados año por año. Como mencionamos anteriormente, en la estimación por *matching* no podemos comprobar las tendencias lineales previamente a la promulgación de la normatividad específica para las licenciaturas, por lo que contrastaremos estos resultados con los obtenidos por DiD.

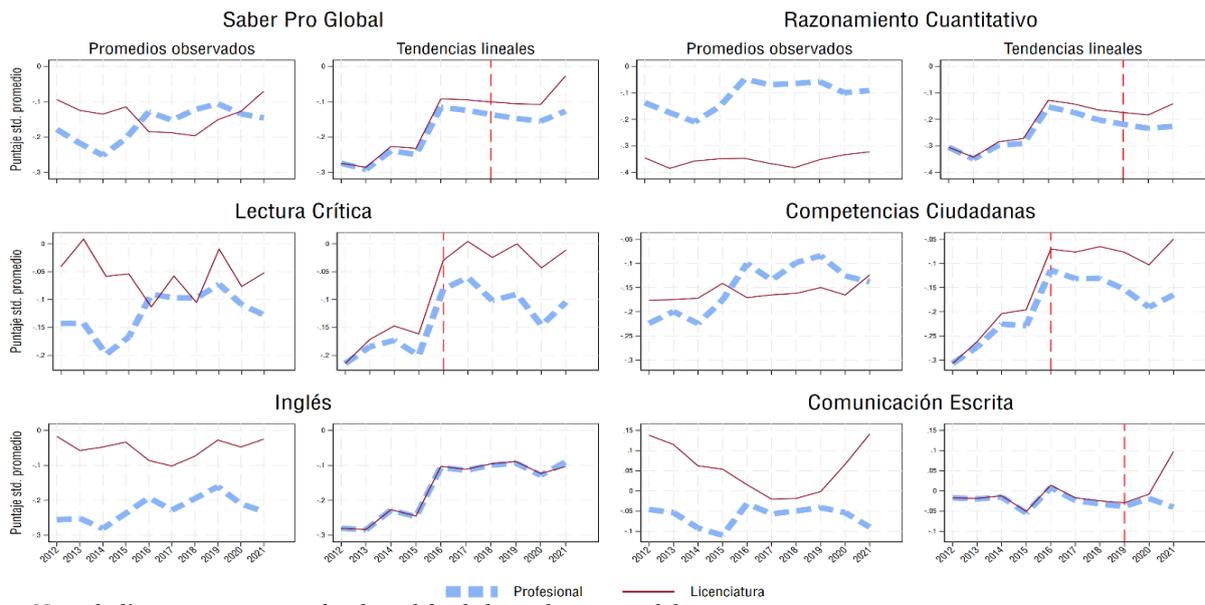
Figura 7. Diferencia de SMD entre el grupo tratado emparejado y el original



Nota: **Fem:** Mujer; **IPM:** Índice de Pobreza Multidimensional por Región; **NBI:** Índice de Necesidades Básicas Insatisfechas por Municipio; **EdPa:** Educación del padre; **EdMa:** Educación de la madre; **Estr:** Estrato socioeconómico; **Trab:** Estudiante trabaja; **Matr:** Valor matrícula universidad; **S11:** Percentil en Saber 11

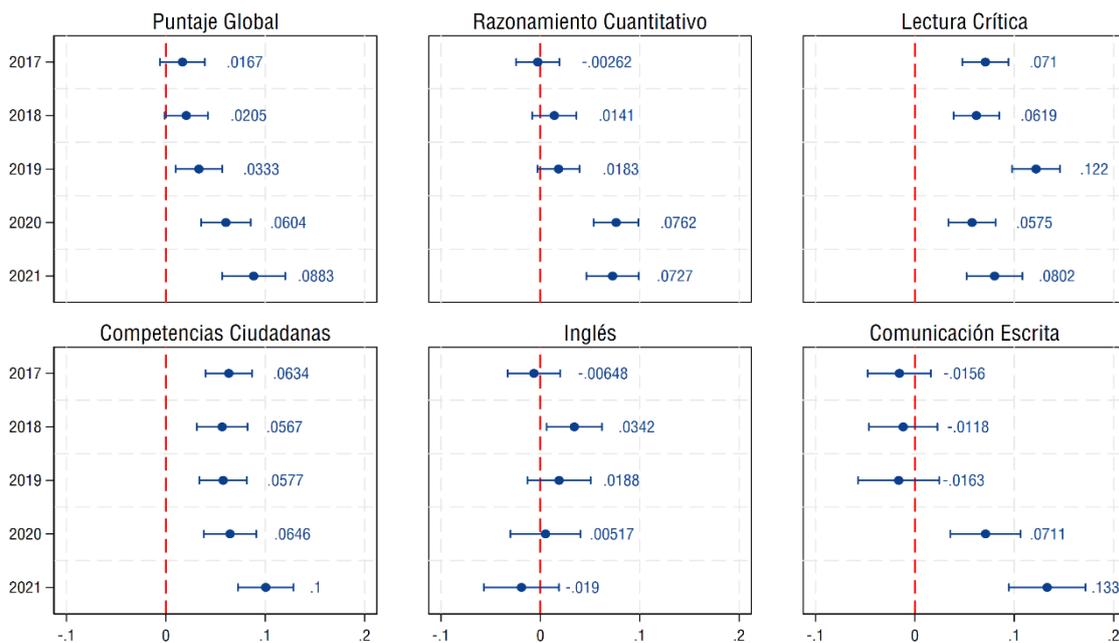
4.3 Resultados del DiD

Figura 8. Tendencias lineales por módulo del examen



Nota: la línea roja transversal indica el fin de la tendencia paralela.

Figura 9. Resultados de la estimación del modelo DiD por año y módulo



Lo primero que vale la pena destacar es que la estimación de DiD supera la prueba de tendencias paralelas para todos los módulos. En el puntaje global, Razonamiento Cuantitativo, Lectura Crítica y Competencias Ciudadanas al 95% ($p > 0.05$); en Comunicación Escrita al 99% ($p > 0.01$) y en Inglés al 99.9% ($p > 0.001$) (cf. Figura 8, Anexo 5, *Ptrends*). Esto quiere decir que, al balancear los grupos por índice de propensión y covariables, se evidencia que antes de 2016 los puntajes de estudiantes de licenciatura seguían la misma tendencia que los demás programas profesionales, tal como se observa en la Figura 8. Los resultados del ATET indican que, para el módulo de Lectura Crítica se encuentra un efecto desde el año 2017 que se sostiene en el tiempo hasta 2021 y varía de 0.057 a 0.122 σ . Lo mismo sucede con el módulo de Competencias Ciudadanas, con el mayor efecto en el último año (0.1 σ). En el puntaje global, el efecto surge en 2019 y aumenta con el tiempo. En Razonamiento Cuantitativo y Comunicación Escrita, el efecto surge en 2020. En esta última crece considerablemente para el último año, siendo el mayor efecto por módulo (0.133 σ). El módulo de Inglés, aunque pasa por muy poco la prueba de tendencias lineales solo presenta efectos en 2018 con bajo nivel de significancia (0.34 σ , $p < .05$). Es importante mencionar que los resultados son similares a las que se presentan en la estimación del ATT por medio de *matching* (Figura 3, Anexo 2), en donde se evidenciaban aumentos de los puntajes de los estudiantes de licenciatura en los últimos años en el puntaje global, en Razonamiento Cuantitativo, Competencias Ciudadanas y Comunicación Escrita. Los resultados completos de la estimación, incluyendo las variables de control y el tamaño de los grupos, los presentamos en el Anexo 5. Los resultados de la estimación del DiD sin balanceo la presentamos en el Anexo 6 y la comparación de los resultados con balanceo y sin balanceo los presentamos en el Anexo 7.

4.3.1 Resultados del DiD por subgrupos

En general, hubo muy pocas diferencias de ATET por subgrupos, y las encontradas, tuvieron bajos niveles de significancia. Como mostramos en la Tabla 6, la mayor diferencia se presenta entre estudiantes de bajo percentil de Saber 11^o en comparación con los de alto percentil en el último año en el resultado general y en el módulo de Inglés. Esto indica que los efectos se presentaron en mayor medida en estudiantes con antecedentes académicos bajos. Hay también un efecto mayor para estudiantes que trabajan tanto en el módulo de Competencias Ciudadanas como de Comunicación Escrita en los últimos años. Por último, en el año 2017 hubo una diferencia significativa por estrato

socioeconómico, pero dicha diferencia desapareció en los siguientes años, lo que indica una nivelación de los resultados por estrato socioeconómico. Los resultados por subgrupos cuyas diferencias no fueron significativas los presentamos en el Anexo 8. No hubo diferencias significativas por las demás covariables del modelo, lo que indica un efecto moderadamente homogéneo. Valga la pena señalar que no hubo diferencias por la antigüedad del programa, por lo que los efectos no se pueden atribuir solamente al cambio en la oferta de programas, sino también a los cambios curriculares de estos.

Tabla 6. Resultados del DiD por subgrupos - diferencias significativas

Resultados por nivel de Percentil en Saber 11^o												
	Puntaje Global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Bajo (≤ 65)	Alto (> 65)	Bajo (≤ 65)	Alto (> 65)	Bajo (≤ 65)	Alto (> 65)	Bajo (≤ 65)	Alto (> 65)	Bajo (≤ 65)	Alto (> 65)	Bajo (≤ 65)	Alto (> 65)
ATET												
2017	0.0326*	0.0090	-0.0130	0.0141	0.0646***	0.0848***	0.0759***	0.0588***	0.0298	-0.0363	-0.0053	-0.0245
2018	0.0254	0.0148	-0.0120	0.0358**	0.0652***	0.0581***	0.0627***	0.0524**	0.0684***	-0.0003	-0.0177	-0.0122
2019	0.0691***	-0.0079	0.0196	0.0071	0.1423***	0.0935***	0.0667***	0.0430**	0.0629***	-0.0294	0.0062	-0.0412
2020	0.0893***	0.0247	0.0583***	0.0793***	0.0941***	0.0151	0.0804***	0.0399**	0.0418*	-0.0371	0.0730**	0.0588**
2021	0.1542***	0.0298	0.0868***	0.0543***	0.1194***	0.0424**	0.1319***	0.0721***	0.0515*	-0.0860**	0.1591***	0.1061***
Test												
2017		0.0675		0.1282		0.0639		0.0441		0.5390		0.1219
2018		0.0141		0.3052		0.0061		0.0161		0.6404		0.0143
2019		0.7250		0.0225		0.2771		0.1001		1.1202		1.0037
2020		0.4303		0.0596		0.7115		0.1958		0.7870		0.0412
2021		1.7103*		0.1333		0.7770		0.5102		1.5943*		0.5126
Resultados por condición laboral												
	No trabaja	Trabaja	No trabaja	Trabaja	No trabaja	Trabaja	No trabaja	Trabaja	No trabaja	Trabaja	No trabaja	Trabaja
ATET												
2017	0.0355*	0.0073	0.0142	-0.0091	0.0846***	0.0640***	0.0465**	0.0717***	-0.0111	-0.0071	0.0078	-0.0294
2018	-0.0089	0.0281*	0.0239	0.0144	0.0373	0.0689***	0.0034	0.0773***	-0.0083	0.0363*	-0.0233	-0.0137
2019	0.0364	0.0320*	0.0216	0.0180	0.0865***	0.1281***	0.0315	0.0681***	0.0219	0.0123	0.0305	-0.0293
2020	0.0453*	0.0627***	0.0735***	0.0789***	0.0306	0.0626***	0.0303	0.0809***	-0.0368	0.0058	0.0784*	0.0657***
2021	0.0770**	0.0917***	0.0636*	0.0784***	0.0738**	0.0807***	0.0409	0.1209***	-0.0460	-0.0193	0.1904***	0.1232***
Test												
2017		0.1884		0.1752		0.1042		0.1829		0.0031		0.7507
2018		0.4033		0.0273		0.2887		2.1766*		0.4840		0.0532
2019		0.0077		0.0054		0.6528		0.5570		0.0256		2.2836*
2020		0.1314		0.0132		0.4528		1.1657		0.5024		0.0755
2021		0.0956		0.1104		0.0213		2.7294*		0.1801		1.4441*
Resultados por Estrato socioeconómico												
	Bajo (≤ 2)	Alto (> 2)	Bajo (≤ 2)	Alto (> 2)	Bajo (≤ 2)	Alto (> 2)	Bajo (≤ 2)	Alto (> 2)	Bajo (≤ 2)	Alto (> 2)	Bajo (≤ 2)	Alto (> 2)
ATET												
2017	-0.0003	0.0526**	-0.0091	0.0094	0.0540***	0.1088***	0.0366**	0.1287***	-0.0093	-0.0056	-0.0216	-0.0057
2018	0.0161	0.0342*	0.0095	0.0290	0.0516***	0.0867***	0.0376*	0.1068***	0.0398**	0.0163	-0.0110	-0.0122
2019	0.0327*	0.0234	0.0178	0.0131	0.1178***	0.1216***	0.0454**	0.0784***	0.0270	-0.0117	-0.0256	0.0030
2020	0.0658***	0.0399*	0.0786***	0.0734***	0.0598***	0.0455*	0.0695***	0.0497*	0.0187	-0.0448	0.0641**	0.0863**
2021	0.0952***	0.0680***	0.0771***	0.0554**	0.0873***	0.0570**	0.0947***	0.1205***	-0.0114	-0.0459	0.1372***	0.1282***
Test												
2017		0.3464		0.0547		0.4199		1.6488*		0.0013		0.0960
2018		0.0522		0.0617		0.2214		1.2795		0.0721		0.0006
2019		0.0142		0.0037		0.0026		0.2608		0.1739		0.5256
2020		0.1236		0.0059		0.0440		0.0902		0.4200		0.1816
2021		0.1529		0.0876		0.2307		0.1867		0.1197		0.0230

Nota: * p < 0.05, ** p < 0.01, *** p < 0.001. ATET ajustado por covariables, efectos de grupo y año de presentación del examen. Errores estándar ajustados por programa.

4.3.2 Resultados del DiD por área de la licenciatura

Hubo algunas diferencias significativas por áreas del conocimiento de las licenciaturas, como mostramos en la Tabla 7. Las licenciaturas en ciencias religiosas presentaban puntajes negativos a 2020, pero en 2021 fueron positivos y superiores al resto de

licenciaturas. Estos mismos programas en 2018 tenían resultados negativos en Lectura Crítica, pero posteriormente se emparejaron con los demás programas. Lo mismo sucedió con las licenciaturas en sociales, historia y geografía, así como áreas artística y cultural, que presentaban resultados negativos en los años 2017 a 2019 para el módulo de Comunicación Escrita, pero a partir de 2020 se emparejaron con los demás programas. Las licenciaturas en educación básica primaria, especial, comunitaria y psicopedagogía presentaron efectos en Lectura Crítica desde 2019. Las licenciaturas en filosofía presentaron puntajes superiores en Inglés en 2020 pero la diferencia no se mantuvo. Adicionalmente, los programas de matemáticas y física presentaron mejoras mayores al resto de los programas en Razonamiento Cuantitativo. En el Anexo 9 presentamos los resultados por cada una de las áreas cuyas diferencias no fueron significativas con respecto al ATET general para todos los programas incluidos en la Figura 9, y en el Anexo 10 presentamos la representación gráfica para todas las áreas, con diferencias significativas y no significativas.

Tabla 7. ATET por áreas y test de Wald con respecto al ATET general – diferencias significativas

Módulo	Religiosa y teología						Filosofía, ética y valores humanos		Matemáticas y física	
	Puntaje global		Lectura Crítica		Competencias Ciudadanas		Inglés		Razonamiento Cuantitativo	
	ATET	Wald	ATET	Wald	ATET	Wald	ATET	Wald	ATET	Wald
2017	0.0031	0.0171	-0.1156	1.6368	0.1266	0.2834	0.1216*	1.661	0.0232	0.0868
2018	-0.2005*	2.4442	-0.2382***	4.4464*	0.0557	0.0001	0.0552	0.0569	0.1753***	3.9340*
2019	-0.082	1.2632	-0.0721	2.5889	-0.0953	1.6063	0.1357*	1.0819	0.1494***	2.5655
2020	-0.2118	4.5631*	-0.1967	2.2565	-0.1315	3.4286	0.2360***	4.8716*	0.1010*	0.0933
2021	0.2911*	4.3355*	0.1354	0.1183	0.3996**	11.3319***	0.1374*	1.8555	0.2279***	3.434*

Módulo	Preescolar, primaria, especial, comunitaria, psicopedagogía		Biología, química, naturales y ambientales		Sociales, historia, geografía		Artística y cultural	
	Comunicación Escrita							
	ATET	Wald	ATET	Wald	ATET	Wald	ATET	Wald
2017	0.0067	0.3633	-0.0611	0.6046	-0.1829*	3.7707	-0.1134*	4.4285*
2018	0.0291	1.201	-0.0074	0.0087	-0.2154*	6.2896*	-0.1153**	6.4263*
2019	0.0599*	4.0719*	-0.0971	2.5355	-0.2135*	5.2654*	-0.1351*	8.3380**
2020	0.0899**	0.1754	-0.1082	3.9249*	-0.06	2.044	-0.0487	3.3477
2021	0.1709***	0.5898	0.0273	1.9243	0.0323	1.1752	0.116	0.0912

Nota: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATET ajustado por covariables, efectos de grupo y año de presentación del examen. Errores estándar ajustados por programa.

5 Conclusiones, discusión y recomendaciones

5.1 Acreditación de alta calidad

Los resultados muestran que la acreditación de alta calidad no necesariamente indica un mejor desempeño en los programas de licenciaturas. En el puntaje global del examen Saber Pro y en el módulo de Razonamiento Cuantitativo, los puntajes promedio son superiores para estudiantes que pertenecen a programas sin acreditación de alta calidad, mientras que en los demás módulos no hay diferencias significativas en función de la acreditación. Esto concuerda con estudios previos según los cuales, en las licenciaturas en filosofía, la acreditación no está asociada con mejor desempeño de los estudiantes o en algunos casos está asociada de manera negativa (Farieta, 2020; 2024). La obligatoriedad de la acreditación de alta calidad para las licenciaturas pudo haber cambiado la manera en que esta se comportaba, pues para obtener el reconocimiento, es posible que los programas se hayan enfocado en otros factores decisivos de la acreditación, como mejorar la planta profesoral, la investigación o la proyección social. Habría que esperar unos años para examinar de nuevo si la acreditación en las licenciaturas, sin el requisito de obligatoriedad, está asociada con mejor desempeño de los estudiantes.

Es difícil sacar conclusiones sobre la acreditación de los programas en general, debido a que el grupo de comparación fue conformado por todos los demás programas profesionales. Pero habría que analizar con más detalle si la asociación entre acreditación y mejor desempeño de los estudiantes se da en función del tipo del programa y el área de conocimiento a la que corresponde. Dado que no era este el objetivo del estudio, no podemos concluir nada al respecto, pero es una importante línea de investigación para el futuro, y es posible que la asociación se presente en algunas áreas y otras no, tal como reporta la literatura, según la cual hay asociación en áreas como economía (Sarmiento *et al.*, 2015) o ingeniería (Romero *et al.*, 2020), pero no en las licenciaturas en filosofía (Farieta, 2020; 2024). Lo que sí se puede concluir con claridad es que la acreditación de alta calidad no mejora automáticamente el desempeño de los estudiantes, e imponerla como obligatoria resulta excesivo e ineficiente. Así lo preveía ya un informe realizado por la OCDE, quien advertía que podría ser demasiado demandante para las universidades (Radinger *et al.*, 2018). Al ser el desempeño de los estudiantes un factor para obtener acreditación (MEN/CNA, 2013), lo normal es que la relación causal fuese a la inversa: una

vez obtenida una mejora en el desempeño de los estudiantes, se obtendría la acreditación de alta calidad. Sin embargo, dado la acreditación y el mejor desempeño de los estudiantes son procesos independientes, no es necesario que exista correlación entre ambos.

También vale tener en cuenta que la literatura señala que la acreditación es cada vez más burocrática y en algunos casos esto podría estar incidiendo negativamente en el desempeño de los estudiantes, dado que se invierten más recursos en gestión y menos en docencia e investigación (Arias, *et al.*, 2018; Hanusheck & Woessman, 2015). Esto se debe al alto costo en términos de tiempo y de recursos que supone un proceso de evaluación por pares como la acreditación de alta calidad (Verger, 2013). La simplificación del proceso de acreditación y la focalización de más recursos de los programas en docencia podría tener una incidencia importante en el desempeño académico, pero hace falta mayor investigación en dicha área. Por otra parte, la acreditación del programa no implica necesariamente que el desempeño de sus estudiantes sea sostenidamente bueno a través del tiempo, como sucede, por ejemplo, con la investigación. Como reporta la literatura, las universidades acreditadas tienden a disminuir su producción investigativa una vez obtenida la acreditación (Bayona-Rodríguez, *et al.*, 2018), lo que a su vez puede tener un efecto negativo en el desempeño de los estudiantes (Portocarrero-Sierra *et al.*, 2020).

5.2 Efectos de los demás aspectos de las reformas a las licenciaturas

Como mostramos en la sección anterior, aparte de la acreditación, los demás cambios curriculares que llevaron a cabo las licenciaturas parecen haber tenido efectos positivos en el desempeño de sus estudiantes. Esto sucedió en casi todos los módulos del Examen, excepto en Inglés, a pesar de que este fue uno de los requerimientos principales de la nueva normatividad. Tal como lo refiere la literatura, a pesar de muchos esfuerzos, las universidades se han visto estancadas en los últimos 10 años en el mejoramiento de los niveles de inglés de sus estudiantes (Benavides, 2021). Es difícil dar una explicación unívoca para ello, pero como lo refiere un informe sobre el modelo de valor agregado del ICFES (2021b), el módulo de Inglés en Saber Pro es el que más se relaciona con el nivel socioeconómico, muy por encima de los demás módulos. Adicionalmente, la enseñanza de una segunda lengua puede resultar sumamente costosa en términos de tiempo y de recursos tanto para las universidades como para los estudiantes. Como lo expresó el en

su momento rector de la Universidad Pedagógica Nacional, “una lengua no se aprende por decreto” (Atehortúa, 2017, citado en Rodríguez-Ávila *et al.*, 2021), quien además evidenció que lograr un nivel A2 para un estudiante sin ningún conocimiento de la lengua tomaría al menos dos años con cursos de cinco horas semanales; algo que irrumpiría por completo el currículo de las universidades.

El efecto fue bastante homogéneo, aunque vale la pena resaltar el mayor efecto en estudiantes cuyos antecedentes académicos —percentil en Saber 11º— eran inferiores. Se evidencia entonces que no solo hubo una mejora general para los estudiantes de los programas de licenciatura, sino que el mayor efecto se dio en quienes más lo requerirían, aportando así a que los programas aporten a la equidad en términos del desempeño de sus estudiantes. Esto lo evidencia también en los resultados por estrato socioeconómico en el módulo de Competencias Ciudadanas, en donde se evidenciaba una brecha de desempeño por estrato en 2017 pero posteriormente esta brecha se cerró.

Sin embargo, también vale la pena señalar cómo el efecto fue diferenciado por tipo de licenciatura, en donde hubo áreas que se nivelaron con las demás (sociales y artística) pero a su vez hubo áreas que aumentaron con respecto al resto, como fue el caso de las licenciaturas en matemáticas y física, quienes mejoraron justamente en Razonamiento Cuantitativo. Esto evidencia que los cambios en los currículos de estas licenciaturas habrían fortalecido un aspecto que de antemano se consideraría fuerte. Vale la pena señalar que la mayor brecha entre las licenciaturas se encuentra precisamente en el módulo de Razonamiento Cuantitativo (García *et al.*, 2014; Farieta, 2022), y aunque los programas en matemática y física hayan mejorado, los demás se encuentran en deuda en este aspecto. Hubo una mejora similar en los programas de ciencias religiosas y teología en el módulo de Competencias Ciudadanas, también posiblemente asociada con cambios en los planes de estudios. Es importante señalar que las reformas impuestas a las licenciaturas no precisaron requisitos específicos para los planes de estudios más allá del incremento en créditos de prácticas. Esto hizo que algunos programas tuvieran que reducir sus créditos disciplinares (Arias *et al.*, 2018; Farieta *et al.*, 2024; Valderrama-Leongómez *et al.*, 2019), por lo que el efecto pudo haber sido mayor si se hubiera hecho más énfasis en los contenidos disciplinares, como sucede con las licenciaturas en filosofía (Farieta, 2022). Sin embargo, falta todavía indagar con más detalle cómo el aumento en las prácticas pedagógicas pudo haber tenido efecto en el desempeño en Saber Pro, el cual

se preveía en el módulo específico en educación (Farieta, 2020), pero no en los módulos genéricos.

El requisito del número de créditos de prácticas parece, entonces, haber incidido de manera heterogénea los programas. Había licenciaturas que, previamente a la normatividad, no tenían créditos de prácticas o tenían un número muy bajo (Farieta, 2018). Sin embargo, el incremento obligatorio a 50 créditos (MEN, 2016, feb. 3) —que luego se redujo a 40 (MEN, 2017 sept. 15)— parece ser arbitrario y no parece funcionar de igual manera en todos los programas. Tal vez la decisión no tenga que imponerse por ley en los programas, sino que sea parte del funcionamiento de las evaluaciones realizadas por la sala de educación de CONACES para el otorgamiento y renovación del registro calificado. También porque es importante reconocer que las prácticas no pueden ser iguales en todas las licenciaturas, sino que hay que entender sus matices no solo disciplinares sino también contextuales (Rodríguez-Ávila *et al.*, 2021).

Aunque el éxito de la reforma parece estar centrado en la gestión de los programas, queda mucha investigación por hacer sobre cómo esta incide en el desempeño de los estudiantes. Aunque es claro que la investigación tiene un efecto positivo en el desempeño de los estudiantes (Portocarrero-Sierra *et al.*, 2020), todavía no hay evidencia concluyente sobre cómo las condiciones contractuales de la planta docente y su formación influyen en el desempeño de los estudiantes. También parece necesario estudiar con más detalle por qué estudiantes con mayores niveles socioeconómicos o con mejores antecedentes académicos (Saber 11^º) no optan por estudiar una licenciatura, y qué se debería hacer al respecto, tal como discutiremos con detalle en la siguiente subsección.

5.3 Recomendaciones

Como hemos mostrado, la acreditación no tuvo efecto en la mejora del desempeño de los estudiantes, por lo que es claro que la principal recomendación es no implementar medidas similares en el futuro para la mejora de la calidad académica. Dada la precipitación con la que se exigió a las licenciaturas la obtención de acreditación de alta calidad, tal vez no fue posible hacer un rastreo de los programas con mejores desempeños en el Examen Saber Pro, para que fueran estos quienes obtuvieran acreditación de alta calidad. Sin embargo, una mejor interacción entre el CNA y el ICFES podría impactar

significativamente la manera en que los resultados del Examen Saber Pro inciden en la acreditación. A pesar de que el ICFES emite anualmente informes de valor agregado para cada los programas de pregrado, el hecho de que no haya asociación entre acreditación y desempeño puede ser evidencia de que los programas no hacen uso de estos estudios de valor agregado para sus autoevaluaciones, o que es un criterio poco significativo para obtener acreditación de alta calidad. Si bien la acreditación no puede basarse solamente en el desempeño de los estudiantes, un informe de valor agregado específico para el CNA podría servir para recomendar o destacar el desempeño de los programas con mayor valor agregado. Igualmente, podría servir también para señalar qué programas con acreditación de alta calidad están teniendo descenso en su desempeño. Esto advertiría al programa para que incluya la mejora del desempeño académico de sus estudiantes como un punto central en sus planes de mejora. De esta manera se podría evitar también que la acreditación de alta calidad genere mayor desigualdad (Cuenca, 2016), pues la acreditación tendería a beneficiarse menos del nivel socioeconómico de sus estudiantes y sus antecedentes académicos, y se enfocaría más en el valor agregado del programa y su aporte a la nivelación académica.

Adicionalmente, es claro que imponer un nivel de inglés como requisito para las licenciaturas no ha funcionado, y se deberían buscar otro tipo de estrategias, sobre todo para permitir una mayor articulación del requisito de segunda lengua con los planes de estudios de los programas. Es preciso tener en cuenta que las condiciones socioeconómicas afectan el nivel de inglés de los estudiantes en mayor medida que en otras competencias (ICFES, 2021b), por lo que es no solo una muestra de desigualdad en el sistema, sino que puede ser la causa de su incremento. Estudiantes con mejor conocimiento de inglés no solo tienen oportunidad de ingresar a mejores posgrados, sino que también pueden acceder más fácilmente a becas en el exterior y otro tipo de ayudas internacionales. Es claro que el problema de la segunda lengua no se puede resolver solamente desde el interior del programa, sino que requiere acciones tanto institucionales como incluso interinstitucionales y de política pública, sobre todo para estudiantes de bajos niveles socioeconómicos. Por ejemplo, una mayor cooperación o convenios específicos entre las instituciones públicas que ofrecen cursos de idiomas y las universidades, así como una mayor interacción de la segunda lengua dentro del currículo.

Por otro lado, a pesar del enorme esfuerzo hecho tanto por programas como por los miembros de CONACES y del CNA, persiste una brecha importante de desempeño entre estudiantes de licenciatura y estudiantes de los demás programas, tanto en desempeño como en condiciones socioeconómicas. Ahora bien, los informes en los que se basó la normatividad a las licenciaturas señalaban tres aspectos importantes en la mejora de la planta docente: selección, formación y remuneración (Bruns & Luque, 2014; García *et al.*, 2014; OCDE, 2005; OREALC/UNESCO, 2016). La política se ha centrado solamente en la formación, pero, aunque sea exitosa, sus limitaciones muestran que tal vez es hora de centrarse en los otros dos aspectos, pues las instituciones por sí solas no pueden solucionar un problema estructural de esta magnitud. Los salarios de los docentes, así como la carga laboral y la percepción negativa de la profesión docente, entre otros aspectos, pueden estar fomentando que no ingresen los mejores estudiantes. Este aspecto no solo requiere mayor investigación sino también una intervención más focalizada desde la política pública, no solo para incentivar a los mejores estudiantes a ingresar a un programa de licenciatura, sino también a hacer de la profesión docente una ocupación más deseada no solo en términos económicos sino también sociales y de calidad de vida. Entre otras, ese era el éxito que se atribuía a otros países en el mundo (García *et al.*, 2014) y no solamente a los programas de formación docente.

6 Referencias

- Abadía, L. K., & Bernal, G. (2017). A Widening Gap? A Gender-Based Analysis of Performance on the Colombian High School Exit Examination. *Revista de Economía Del Rosario*, 20(1), 5–31. <https://doi.org/10.12804/revistas.urosario.edu.co/economia/a.6144>
- Arias G., D. H., Díaz F., O. C., Garzón B., I., León P., A. C., Rodríguez A., S. P., & Valbuena U., E. O. (2018) *Entre las exigencias de calidad y las condiciones de desigualdad: Formación inicial de profesores en Colombia*. Bogotá: Universidad Pedagógica Nacional. <http://repositorio.pedagogica.edu.co/handle/20.500.12209/7826>
- Arias-Velandia, N., Rincón-Báez, W. U., & Cruz-Pulido, J. M. (2018). Desempeño de mujeres y hombres en educación superior presencial, virtual y a distancia en Colombia. *Panorama*, 12(22), 57–69. <https://doi.org/10.15765/pnrm.v12i22.1142>

- Arias-Velandia, N., Rincón-Báez, W. U., & Cruz-Pulido, J. M. (2021). Diferencia de logro geolocalizado en educación presencial y a distancia en Colombia. *Revista Electrónica de Investigación Educativa*, 23, 1–22. <https://doi.org/10.24320/redie.2021.23.e18.3711>
- Ariza, N. (2018). *Educación superior para todos los jóvenes colombianos*. Center for Latin American Studies. University of California, Berkeley. <https://clasberkeley.wpcomstaging.com/2018/09/01/building-better-higher-education-in-colombia/>
- Avendaño, R., Barrera-Osorio, F., Nieto-Parra, S. & Vever, F. (2016). *Understanding Student Performance Beyond Traditional Factors: Evidence from PISA*. OECD Development Centre Working Papers 331. Paris: OECD Publishing. <https://doi.org/10.1787/18151949>
- Barber, M., & Mourshed, M. 2007. *How the World's Best-Performing School Systems Come Out on Top*. Washington, D.C.: McKinsey. <https://www.mckinsey.com/industries/education/our-insights/how-the-worlds-best-performing-school-systems-come-out-on-top>
- Bardach, L., & Klassen, R. M. (2020). Smart teachers, successful students? A systematic review of the literature on teachers' cognitive abilities and teacher effectiveness. *Educational Research Review*, 30(100312). <https://doi.org/10.1016/j.edurev.2020.100312>
- Barrera-Osorio, F.; Maldonado, D., & Rodríguez, C. (2012). *Calidad de la educación básica y media en Colombia: diagnóstico y propuestas*. Documentos CEDE 41. Bogotá: Universidad de los Andes. https://economia.uniandes.edu.co/components/com_booklibrary/ebooks/dcede2012-41.pdf
- Barrera-Osorio, F., & Bayona-Rodríguez, H. (2019). Signaling or better human capital: Evidence from Colombia. *Economics of Education Review*, 70, 20–34. <https://doi.org/10.1016/j.econedurev.2019.02.006>
- Barroilhet, A., Ortiz, R., Quiroga, B. F., & Silva, M. (2022). Exploring Conflict of Interest in University Accreditation in Chile. *Higher Education Policy*, 35(2), 479–497. <https://doi.org/10.1057/s41307-020-00217-7>

- Bayona-Rodríguez, H., Bedoya, J., Pulido, X., & Sánchez, F. (2018). Efectos (Algunos No Esperados) De La Acreditación De Alta Calidad En Las Universidades Colombianas (SSRN Scholarly Paper Núm. 3215465). <https://doi.org/10.2139/ssrn.3215465>
- Benavides, J. E. (2021). Level of English in Colombian Higher Education: A Decade of Stagnation. *Profile: Issues in Teachers' Professional Development*, 23(1), 57–73. <https://doi.org/10.15446/profile.v23n1.83135>
- Bonilla-Mejía, L., Londoño-Ortega, E., Cardona-Sosa, L., & Trujillo-Escalante, L. D. (2018). *¿Quiénes son los docentes en Colombia? Características generales y brechas regionales*. Documentos de Trabajo sobre Economía Regional y Urbana 276. Bogotá: Banco de la República. <https://doi.org/10.32468/dtseru.276>
- Bruns, B., & Luque, J. (2015). *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. Washington, D.C.: World Bank. <https://doi.org/10.1596/978-1-4648-0151-8>
- Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., & Schmidt, W. (2019). A Review of the Literature on Teacher Effectiveness and Student Outcomes. En N. Burroughs, J. Gardner, Y. Lee, S. Guo, I. Touitou, K. Jansen, & W. Schmidt (Eds.), *Teaching for Excellence and Equity: Analyzing Teacher Characteristics, Behaviors and Student Outcomes with TIMSS* (pp. 7–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-16151-4_2
- Camacho, A., Messina, J., & Uribe, J. P. (2016). *The Expansion of Higher Education in Colombia: Bad Students or Bad Programs?* Discussion Paper IDB-DP-452. Washington: InterAmerican Development Bank, Department of Research and Chief Economist. <https://doi.org/10.18235/0000303>
- Cameron, A. C., and D. L. Miller. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50, 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- Castro, M., Ruiz, J., & Guzmán, F. (2018). Cruce de las pruebas nacionales Saber 11º y Saber Pro en Antioquia, Colombia: una aproximación desde la regresión geográficamente ponderada (GWR). *Revista Colombiana de Educación*, (74), 63–79. <https://doi.org/10.17227/rce.num74-6898>
- Cayón, E., Correa, J. S., & Sarmiento, J. A. (2020). Does Quality Make a Difference for Higher Education Graduates in Colombia? Measuring Differences in Accredited and Non-accredited Institutions with Similar Financial Characteristics. *WSEAS*

- Transactions on Business and Economics*, 17, 859–868.
<https://doi.org/10.37394/23207.2020.17.84>
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633–2679. <https://doi.org/10.1257/aer.104.9.2633>
- Cochran-Smith, M. (2005). The new teacher education: For better or for worse? *Educational Researcher*, 34(7), 3–17.
<https://doi.org/10.3102/0013189X034007003>
- Cochran-Smith, M. (2021). Rethinking teacher education: The trouble with accountability. *Oxford Review of Education*, 47(1), 8–24.
<https://doi.org/10.1080/03054985.2020.1842181>
- Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., & Klaveren, C. (2018). Teacher Characteristics and their Effects on Student Test Scores: A Systematic Review. *Journal of Economic Surveys*, 32, 848–877.
<https://doi.org/10.1111/joes.12210>
- CRC. (1992, dic. 28) Ley 30 de 1992, Por la cual se organiza el servicio público de la Educación Superior. *Diario oficial*, 40700.
http://www.mineducacion.gov.co/1621/articulos-104539_archivo_pdf.pdf
- CRC. (2015, jun. 9). Ley 1753. “Plan Nacional de Desarrollo 2014-2018”. *Diario oficial*, 151(49538).
http://www.secretariassenado.gov.co/senado/basedoc/ley_1753_2015.html
- CRC. (2019, mayo 25) Ley 1955. “Plan Nacional de Desarrollo 2018-2022”. *Diario Oficial*, 155(50964).
http://www.secretariassenado.gov.co/senado/basedoc/ley_1955_2019.html
- Cuenca, Andrea. (2016). Desigualdad de oportunidades en Colombia: impacto del origen social sobre el desempeño académico y los ingresos de graduados universitarios. *Estudios pedagógicos*, 42(2), 69–93. <https://doi.org/10.4067/S0718-07052016000200005>
- DNP. (2015). *Plan Nacional de Desarrollo 2014-2018: todos por un nuevo país* (Vols. 1-2). Bogotá: DNP. <https://www.dnp.gov.co/Plan-Nacional-de-Desarrollo/Paginas/Que-es-el-Plan-Nacional-de-Desarrollo.aspx>

- DANE. (2018). Censo Nacional de Población y Vivienda 2018. Microdatos. <https://microdatos.dane.gov.co/index.php/catalog/643>
- Donald, S. G., and K. Lang. (2007). Inference with difference-in-differences and other panel data. *Review of Economics and Statistics*, 89(2): 221–233. <https://doi.org/10.1162/rest.89.2.221>
- Duque, J. F. (2021). A comparative analysis of the Chilean and Colombian systems of quality assurance in higher education. *Higher Education*, 82, 669–683. <https://doi.org/10.1007/s10734-020-00633-z>
- Fajardo, E., Beleño Montagut, L., & Romero, H. (2021). Incidencia de los factores socioeconómicos en la calidad de la educación media regional en Colombia. *Interciencia*, 46(3), 118–125. https://www.interciencia.net/wp-content/uploads/2021/04/04_6787_Com_Fajardo_v46n3_8.pdf
- Farieta Barrera, R. A. (2018) Análisis comparativo de los planes de estudio de los programas de pregrado en filosofía en Colombia, segunda parte: contenidos. *Folios*, (48), 167–191. DOI: <https://doi.org/10.17227/folios.48-8142>
- Farieta-Barrera, A. (2020). Initial Education of Philosophy Teachers in Colombia: Association between new Public Policy Requirements and National Standardized Tests. *Australian Journal of Teacher Education*, 45(6). <https://doi.org/10.14221/ajte.2020v45n6.3>
- Farieta, A. (2022) La incidencia de la filosofía en la lectura crítica y otras competencias genéricas en los programas de formación de docentes en Colombia. En W. Herrera (ed.), *Actualidad y Defensa de la Filosofía* (pp. 241-290). Bogotá, Universidad del Rosario – Universitaria Agustiniiana. <https://doi.org/10.12804/urosario9789587848182>
- Farieta, A. (2024). The Effect of Philosophy on Critical Reading: Evidence from Initial Teacher Education in Colombia. *International Journal of Educational Development*, 104, 102974. <https://doi.org/10.1016/j.ijedudev.2023.102974>
- Farieta, A.; Gómez, M. & Almeida, S. (2015). Análisis comparativo de los planes de estudio de filosofía en Colombia, parte 1: generalidades y flexibilidad. *Folios* (42), 87–104. <https://doi.org/10.17227/01234870.42folios87.104>
- Farieta, A., & Murillo, L. A. (2024). Efectos de las reformas a las políticas de formación de docentes en programas de filosofía en Colombia. *Perfiles Educativos*, 46(184), 53-73. <https://doi.org/10.22201/iisue.24486167e.2024.184.61186>

- Fernández L., N. (2003). Higher Education, Quality Evaluation and Accreditation in Latin America and MERCOSUR. *European Journal of Education*, 38(3), 253. <https://doi.org/10.1111/1467-3435.00145>
- Fredriksson, A., & Oliveira, G. M. de. (2019). Impact evaluation using Difference-in-Differences. *RAUSP Management Journal*, 54(4), 519–532. <https://doi.org/10.1108/RAUSP-05-2019-0112>
- Frölich, M. (2004). Finite-sample properties of propensity-score matching and weighting estimators. *The Review of Economics and Statistics*, 86(1), 77–90. <https://www.jstor.org/stable/3211661>
- Frölich, M. (2005). Matching estimators and optimal bandwidth choice. *Statistics and Computing*, 15:197–215. <https://doi.org/10.1007/s11222-005-1309-6>
- Furlong, J., Cochran-Smith, M., & Brennan, M. (2008). Politics and policy in teacher education: International perspectives. *Teachers and Teaching*, 14(4), 265–269. <https://doi.org/10.1080/13540600802037660>
- Galdo, J. C., Smith, J., & Black, D. (2008). Bandwidth selection and the estimation of treatment effects with unbalanced data. *Annales d'Économie et de Statistique*, 91/92, 89–216. <https://doi.org/10.2307/27917245>
- García J., S., Maldonado C., D., Perry R., G., Rodríguez O., C., & Saavedra C., J. E. (2014). *Tras la excelencia docente: cómo mejorar la educación de todos los colombianos*. Bogotá: Fundación Compartir, Punto Aparte. <https://www.fundacioncompartir.org/pdf/Tras%20la%20excelencia%20docente%20-%20estudio%20final.pdf>
- García J., S., Rodríguez O., C, Sánchez T., F & Bedoya O., J. (2015). La lotería de la cuna: la movilidad social a través de la educación en los municipios de Colombia. Universidad de los Andes, Facultad de Economía, CEDE. <https://repositorio.uniandes.edu.co/handle/1992/8605>
- Garwe, E. C. (2019). Quality assurance agencies: Creating a conducive environment for academic integrity. *South African Journal of Science*, 115(11–12), 1–7. <https://dx.doi.org/10.17159/sajs.2019/6231>
- Haakstad, J. (2001). Accreditation: The new quality assurance formula? Some reflections as Norway is about to reform its quality assurance system. *Quality in Higher Education*, 7(1), 77-82.

- Hainmueller, J. (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis*, 20(1), 25–46. <https://doi.org/10.1093/pan/mpr025>
- Hanusheck, E. A. (2011). The Economic Value of Higher Teacher Quality. *Economics of Education Review*, 30(3), 466-479. <https://doi.org/10.1016/j.econedurev.2010.12.006>
- Hanusheck, E. A., Piopiunik, M., & Wiederhold, S. (2018). The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance. *Journal of Human Resources*. Advance online publication. <https://doi.org/10.3368/jhr.55.1.0317.8619R1>
- Hanusheck, E. A., & Woessman, L. (2012). Schooling, Educational Achievement and the Latin American Growth Puzzle. *Journal of Development Economics*, 99(2), 497–512. <https://doi.org/10.1016/j.jdeveco.2012.06.004>
- Hanushek, R., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economies of growth*. MIT Press.
- Huber, M., M. Lechner, C. Wunsch. 2013. The performance of estimators based on the propensity score. *Journal of Econometrics*, 175(1), 1–21. <https://doi.org/10.1016/j.jeconom.2012.11.006>
- ICFES. (2018). *Saber Pro. Informe Nacional de Resultados 2016-2017*. Bogotá: ICFES. <https://www2.icfes.gov.co/documents/39286/1652832/Saber+Pro-informe+nacional+de+resultados+2016-2017+%282018%29.pdf>
- ICFES. (2020). *Informe Nacional Saber Pro 2016-2019*. Bogotá: ICFES. <https://www.icfes.gov.co/documents/39286/7239514/Informe+nacional+de+resultados+Saber+Pro+2016-2019.pdf>
- ICFES. (2021a). *Informe Nacional de Resultados del Examen Saber Pro 2020*. Vol. 1. Bogotá: ICFES. https://www.icfes.gov.co/documents/39286/1689945/23_12_21_INFORME_PR_O+%281%29.pdf
- ICFES. (2021b). *Medición de los efectos de la educación superior en Colombia sobre el aprendizaje estudiantil. Informe técnico*. Bogotá: ICFES. <https://www.icfes.gov.co/documents/39286/10060622/Informe+técnico+%281%29.pdf>

- Imbens, G. W., and M. Kolesár. (2016). Robust standard errors in small samples: Some practical advice. *Review of Economics and Statistics*, 98, 701–712. https://doi.org/10.1162/REST_a_00552
- Jann, B. (2017a). *Why propensity scores should be used for matching*. German Stata Users' Group Meetings 2017 01. Stata Users Group. https://www.stata.com/meeting/germany17/slides/Germany17_Jann.pdf
- Jann, B. (2017b). KMATCH: Stata module for multivariate-distance and propensity-score matching, including entropy balancing, inverse probability weighting, (coarsened) exact matching, and regression adjustment. *Statistical Software Components S458346*, Boston College Department of Economics, revised 19 Sep 2020. <https://ideas.repec.org/c/boc/bocode/s458346.html>
- King, G., & Nielsen, R. (2019). Why Propensity Scores Should Not Be Used for Matching. *Political Analysis*, 27(4), 435–454. <https://doi.org/10.1017/pan.2019.11>
- Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3): 165–224. <http://doi.org/10.1561/08000000014>
- Lejeune, C. y Vas, A. (2009). Organizational culture and effectiveness in business schools: A test of the accreditation impact. *Journal of Management Development*, 28(8), 728–741. <https://doi.org/10.1108/02621710910985504>
- Lindner, S., & McConnell, K. J. (2019). Difference-in-differences and matching on outcomes: A tale of two unobservables. *Health Services and Outcomes Research Methodology*, 19(2), 127–144. <https://doi.org/10.1007/s10742-018-0189-0>
- Mattar, M.Y. (2022). Combating Academic Corruption and Enhancing Academic Integrity through International Accreditation Standards: The Model of Qatar University. *Journal of Academic Ethics*, 20, 119–146. <https://doi/10.1007/s10805-021-09392-7>
- Melguizo, T., Sanchez, F., & Velasco, T. (2016). Credit for Low-Income Students and Access to and Academic Performance in Higher Education in Colombia. *World Development*, 80(C), 61–77. <https://doi.org/10.1016/j.worlddev.2015.11.018>
- MEN. (2016, feb. 3). Resolución 2041. *Diario Oficial*, 152(49776). https://www.mineducacion.gov.co/1759/articles-356982_recurso_1.pdf

- MEN. (2017, sep. 15). Resolución 18583. *Diario oficial*, 153(50357).
http://www.imprenta.gov.co/diariop/diario2.indice?v_num=50.357
- MEN/CNA. (2013). *Lineamientos para la acreditación de programas de pregrado*. Bogotá: MEN/CNA. https://www.mineducacion.gov.co/1621/articulos-342684_recurso_1.pdf
- MEN/CNA. (2020). *Lineamientos y aspectos por evaluar para la acreditación en alta calidad de programas académicos*. https://www.cna.gov.co/1779/articulos-404750_norma.pdf
- National Commission on Teaching & America's Future. (1997). *Doing what matters most: Investing in quality teaching*. National Commission on Teaching and America's Future.
- Nguyen, T.-L., Collins, G. S., Spence, J., Daurès, J.-P., Devereaux, P. J., Landais, P., & Le Manach, Y. (2017). Double-adjustment in propensity score matching analysis. *BMC Medical Research Methodology*, 17(1), 78. <https://doi.org/10.1186/s12874-017-0338-0>
- Nigsch, S. y Schenker-Wicki, A. (2013). Shaping performance: do international accreditations and quality management really help? *Journal of Higher Education and Management*, 35(6): 668-681. <https://doi.org/10.1080/1360080X.2013.844669>
- OCDE. (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Paris: OECD. <https://www.oecd.org/education/school/34990905.pdf>
- OCDE. (2013). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science* (Vol. I). Paris: OECD. <https://doi.org/10.1787/9789264201118-en>
- OCDE. (2016). *PISA 2015 Country Note for Colombia*. Paris: OECD. <https://www.oecd.org/pisa/PISA-2015-Colombia.pdf>
- OCDE. (2019). *PISA 2018 Country Note for Colombia*. Paris: OECD. https://www.oecd.org/pisa/publications/PISA2018_CN_COL.pdf
- OCDE/CEPAL/CAF/. (2014). *Latin American Economic Outlook 2015: Education, Skills and Innovation for Development*. Paris: OECD. <https://doi.org/10.1787/leo-2015-en>

- OREALC/UNESCO. (2013). *Antecedentes y criterios para la Elaboración de Políticas Docentes en América Latina y el Caribe*. Santiago de Chile: OREALC/UNESCO.
<http://unesdoc.unesco.org/images/0022/002232/223249S.pdf>
- OREALC/UNESCO. (2016). *Perspectivas sobre políticas docentes en América Latina y el Caribe. Aprendizajes de la Estrategia Regional de Docentes de la OREALC/UNESCO 2011-2016*. Santiago de Chile: OREALC/UNESCO.
<http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/Perspectivas-Políticas-Docentes-LAC.pdf>.
- Osborne, J. (2017). *Regression & linear modeling*. SAGE Publications, Inc,
<https://doi.org/10.4135/9781071802724>
- Pineda, P., & Celis, J. (2018). The Maelstrom of Online Programs in Colombian Teacher Education. *Education Policy Analysis Archives*, 26(114).
<https://doi.org/10.14507/epaa.26.3873>.
- Portocarrero-Sierra, L., Restrepo-Morales, J. A., Arias-Calderón, J. E. (2020). Impact evaluation of high quality accreditation in Colombian higher education institutions. *Formación Universitaria*, 13(6), 37–50.
<https://doi.org/10.4067/S0718-50062020000600037>
- PRC. (2015, dic. 17). Decreto 2450. *Diario oficial*, 49729.
https://www.mineducacion.gov.co/1759/articles-357048_recurso_1.pdf.
- PRC. (2017, mayo 28). Decreto 892. *Diario oficial*, 50247.
<http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=69320>.
- PRC. (2019, jul. 25). Decreto 1330. *Diario oficial*, 51025.
https://www.mineducacion.gov.co/1780/articles-387348_archivo_pdf.pdf.
- RAND Corporation. (2013). *Teachers Matter: Understanding Teachers' Impact on Student Achievement*. <http://www.rand.org/education/projects/measuring-teacher-effectiveness/teachers-matter.html>
- Radinger, T., Echazarra, A., Guerrero, G., & Valenzuela, J. P. (2018), *OECD Reviews of School Resources: Colombia 2018*. Paris: OECD Publishing.
<https://doi.org/10.1787/9789264303751-en>
- Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *The American Economic Review*, 94(2), 247–252.
<https://doi.org/10.1257/0002828041302244>.

- Rodríguez A., G., Ariza D., M., & Ramos R., J. L. (2014). Calidad institucional y rendimiento académico: El caso de las universidades del Caribe colombiano. *Perfiles educativos*, 36(143), 10–29. [https://doi.org/10.1016/S0185-2698\(14\)70607-5](https://doi.org/10.1016/S0185-2698(14)70607-5)
- Rodríguez A., G., Gómez L., V., & Ariza D., M. (2014). Calidad de la educación superior a distancia y virtual: un análisis de desempeño académico en Colombia. *Investigación & Desarrollo*, 22(1), 58–99. <https://doi.org/10.14482/indes.22.1.6079>
- Rodríguez-Ávila, S. P., Díaz-Flórez, O. C., & Arias-Gómez, D. H. (2021). Los efectos de las políticas de calidad en las licenciaturas en Colombia: Balance y alternativas. *Revista Colombiana de Educación*, 81, 35–59. <https://doi.org/10.17227/rce.num81-10688>
- Rojas O., D. E. (2019). Relaciones de Clase en el Sistema Universitario y su Efecto sobre el Rendimiento Académico: El Caso de Bogotá. *Multidisciplinary Journal of Educational Research*, 9(1), 1–24. <https://doi.org/10.17583/remie.2019.3999>
- Romero, E.J. (2008). AACSB accreditation: Addressing faculty concerns. *Academy of Management Learning & Education*, 7(2), 245–255. <https://doi.org/10.5465/amle.2008.32712622>
- Romero, V., Seguratorres, D., & Narvárez, M. (2020). Efectividad en las acreditaciones de calidad en programas de ingeniería industrial de la región caribe colombiana según desempeño en pruebas Saber Pro. *Formación Universitaria*, 13(6), 37–50. <https://doi.org/10.4067/S0718-50062020000600037>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.2307/2335942>
- Rubin, D. & Thomas, N. (2000). Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates. *Journal of the American Statistical Association*, 95(450), 573–585. <https://doi.org/10.1080/01621459.2000.10474233>
- Sarmiento E., J. A., Silva A., A. C., & van Gameren, E. (2015). Quality differences of higher education and its determinants in a less-developed country. *Journal of Higher Education Policy and Management*, 37(2), 204–221. <https://doi.org/10.1080/1360080X.2015.1019118>.

- Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M., & Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health services & outcomes research methodology*, 14(4), 166–182. <https://doi.org/10.1007/s10742-014-0123-z>
- Tam, M. (1999). Quality Assurance Policies in Higher Education in Hong Kong. *Journal of Higher Education Policy and Management*, 21(2), 215-226. <https://doi.org/10.1080/1360080990210208>
- Tatto, M., & Menter, I. (2019). The significance of teacher education. In M. Tatto & I. Menter (Eds.), *Knowledge, policy and practice in teacher education: A cross-national study* (pp. 9–17). Bloomsbury Academic.
- Timarán-Pereira, S.R., Hernández-Arteaga, I., Caicedo-Zambrano, S.J., Hidalgo-Troya, A., & Alvarado-Pérez, J.C. (2016). *Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional*. Bogotá: Universidad Cooperativa de Colombia. <https://doi.org/10.16925/9789587600490>
- Tucker, M. S. (Ed.). (2011). *Surpassing Shanghai: An agenda for American education built on the world's leading systems*. Harvard Education Press.
- Valderrama-Leongómez, M., Murillo, L. A., Farieta-Barrera, A. y Vela, M. (2019). Licenciaturas en filosofía: disputas y articulaciones entre lo disciplinar y lo pedagógico. En: M. L. Cárdenas-Beltrán, C. R. Cáceda y L. A. Murillo (eds.). *Formación de docentes en universidades latinoamericanas* (pp. 71–112). Bogotá, D. C.: Editorial Uniagustiniana. <https://doi.org/10.28970/9789585498273.02>.
- Verger, A. (2013) Políticas de mercado, Estado y universidad. Hacia una conceptualización y explicación del fenómeno de la mercantilización de la Educación Superior. *Revista de Educación*, (360), 268–291. <https://sede.educacion.gob.es/publiventa/detalle.action?cod=15529>
- Weisberg, S. (2014). *Applied Linear Regression*. 4th ed. Hoboken, NJ: Wiley.

Anexos

Anexo 1. Regresión multivariada con interacciones, versión completa

	Puntaje global	Razonamiento Cuantitativo	Lectura Crítica	Competencias Ciudadanas	Inglés	Comunicación Escrita
Mujer	-0.0594*** (0.0013)	-0.3627*** (0.0014)	0.0378*** (0.0014)	0.0121*** (0.0015)	-0.0928*** (0.0014)	0.1551*** (0.0017)
Edad	-0.0276*** (0.0003)	-0.0261*** (0.0003)	-0.0225*** (0.0003)	-0.0155*** (0.0004)	-0.0203*** (0.0003)	-0.0162*** (0.0004)
IPM	-2.7429*** (0.0229)	-2.1739*** (0.0254)	-1.9887*** (0.0259)	-2.4815*** (0.0269)	-2.1720*** (0.0248)	-1.5724*** (0.0314)
NBI	-0.0024*** (0.0001)	-0.0030*** (0.0001)	-0.0027*** (0.0001)	-0.0016*** (0.0001)	-0.0003** (0.0001)	-0.0024*** (0.0001)
Ed. Padre	0.0056*** (0.0001)	0.0008*** (0.0002)	0.0039*** (0.0002)	0.0018*** (0.0002)	0.0128*** (0.0002)	0.0014*** (0.0002)
Ed. Madre	0.0092*** (0.0002)	0.0039*** (0.0002)	0.0069*** (0.0002)	0.0068*** (0.0002)	0.0155*** (0.0002)	0.0017*** (0.0002)
Estrato	0.0584*** (0.0007)	0.0126*** (0.0008)	0.0181*** (0.0008)	0.0120*** (0.0008)	0.1599*** (0.0008)	0.0127*** (0.0009)
Trabaja	-0.0363*** (0.0014)	0.0150*** (0.0015)	-0.0514*** (0.0015)	-0.0368*** (0.0016)	-0.0622*** (0.0015)	-0.0051** (0.0019)
Valor matrícula universidad (SMLV)	0.0203*** (0.0003)	0.0088*** (0.0004)	0.0048*** (0.0004)	0.0055*** (0.0004)	0.0421*** (0.0004)	0.0156*** (0.0004)
Percentil Saber 11 ^o	0.0199*** (0.0000)	0.0175*** (0.0000)	0.0189*** (0.0000)	0.0176*** (0.0000)	0.0143*** (0.0000)	0.0082*** (0.0000)
Inst. pública	0.1811*** (0.0019)	0.2441*** (0.0021)	0.1367*** (0.0021)	0.1189*** (0.0022)	0.1199*** (0.0020)	0.0815*** (0.0025)
Universidad	0.0737*** (0.0016)	0.0474*** (0.0018)	0.0818*** (0.0018)	0.0786*** (0.0019)	0.0315*** (0.0017)	0.0570*** (0.0022)
Año=2013	-0.0249*** (0.0049)	-0.0317*** (0.0055)	-0.0050 (0.0056)	0.0133* (0.0058)	-0.0224*** (0.0052)	-0.0077 (0.0066)
Año=2014	0.0114* (0.0049)	-0.0072 (0.0055)	0.0076 (0.0055)	0.0535*** (0.0057)	0.0134** (0.0051)	0.0051 (0.0065)
Año=2015	0.0078 (0.0048)	0.0053 (0.0053)	-0.0152** (0.0054)	0.0574*** (0.0056)	-0.0062 (0.0050)	-0.0159* (0.0064)
Año=2016	0.1384*** (0.0047)	0.1141*** (0.0053)	0.1000*** (0.0053)	0.1610*** (0.0055)	0.1563*** (0.0049)	0.0619*** (0.0063)
Año=2017	0.1159*** (0.0047)	0.1063*** (0.0052)	0.0800*** (0.0053)	0.1218*** (0.0054)	0.1422*** (0.0048)	0.0417*** (0.0062)
Año=2018	0.1206*** (0.0047)	0.0877*** (0.0053)	0.0695*** (0.0053)	0.1374*** (0.0055)	0.1482*** (0.0049)	0.0429*** (0.0063)
Año=2019	0.0937*** (0.0047)	0.0534*** (0.0052)	0.0508*** (0.0053)	0.1150*** (0.0055)	0.1395*** (0.0048)	0.0410*** (0.0063)
Año=2020	0.0835*** (0.0047)	0.0327*** (0.0052)	0.0428*** (0.0053)	0.0871*** (0.0054)	0.1170*** (0.0049)	0.0154* (0.0063)
Año=2021	0.1114*** (0.0047)	0.0463*** (0.0053)	0.0639*** (0.0053)	0.1121*** (0.0055)	0.1431*** (0.0049)	0.0122 (0.0063)
Licenciatura	0.0211 (0.0135)	-0.2867*** (0.0145)	-0.0338* (0.0155)	-0.1002*** (0.0150)	0.2725*** (0.0158)	0.1201*** (0.0187)
Año=2013 x Licenciatura	-0.0150 (0.0164)	-0.0141 (0.0175)	0.0745*** (0.0188)	0.0271 (0.0182)	0.0131 (0.0192)	-0.0060 (0.0226)
Año=2014 x Licenciatura	-0.0284 (0.0160)	0.0139 (0.0169)	0.0358 (0.0183)	0.0202 (0.0177)	0.0092 (0.0186)	-0.0508* (0.0221)
Año=2015 x Licenciatura	0.0151	0.0036	0.0768***	0.0587***	0.0613**	-0.0410

	(0.0158)	(0.0168)	(0.0180)	(0.0177)	(0.0188)	(0.0220)
Año=2016 x Licenciatura	-0.1041*** (0.0154)	-0.0208 (0.0164)	-0.0307 (0.0176)	-0.0139 (0.0172)	-0.1134*** (0.0179)	-0.0701*** (0.0212)
Año=2017 x Licenciatura	-0.0556*** (0.0156)	0.0071 (0.0166)	0.0606*** (0.0177)	0.0559** (0.0174)	-0.1216*** (0.0179)	-0.0700** (0.0214)
Año=2018 x Licenciatura	-0.0428** (0.0157)	0.0054 (0.0168)	0.0539** (0.0178)	0.0744*** (0.0175)	-0.0649*** (0.0183)	-0.0566** (0.0218)
Año=2019 x Licenciatura	-0.0363* (0.0155)	0.0170 (0.0166)	0.1147*** (0.0176)	0.0502** (0.0173)	-0.0812*** (0.0181)	-0.0436* (0.0216)
Año=2020 x Licenciatura	-0.0099 (0.0156)	0.0423* (0.0168)	0.0520** (0.0178)	0.0309 (0.0175)	-0.1036*** (0.0183)	0.0034 (0.0219)
Año=2021 x Licenciatura	0.0786*** (0.0156)	0.0929*** (0.0168)	0.0975*** (0.0178)	0.1001*** (0.0175)	-0.0472* (0.0185)	0.1020*** (0.0222)
Prog. acreditado=1	0.1043*** (0.0072)	0.1187*** (0.0081)	0.0720*** (0.0075)	0.1151*** (0.0078)	0.1612*** (0.0077)	0.0785*** (0.0086)
Año=2013 x Prog. acreditado=1	0.0558*** (0.0088)	0.0174 (0.0100)	0.0059 (0.0092)	-0.0260** (0.0096)	-0.0285** (0.0094)	-0.0023 (0.0106)
Año=2014 x Prog. acreditado=1	0.0345*** (0.0087)	-0.0114 (0.0099)	0.0066 (0.0092)	-0.0196* (0.0096)	-0.0637*** (0.0094)	-0.0009 (0.0105)
Año=2015 x Prog. acreditado=1	0.0252** (0.0085)	-0.0144 (0.0095)	0.0552*** (0.0090)	-0.0364*** (0.0093)	-0.0742*** (0.0092)	0.0132 (0.0103)
Año=2016 x Prog. acreditado=1	0.0266** (0.0083)	-0.0157 (0.0093)	0.0514*** (0.0087)	-0.0429*** (0.0091)	-0.0506*** (0.0088)	-0.0439*** (0.0101)
Año=2017 x Prog. acreditado=1	-0.0045 (0.0082)	-0.0608*** (0.0092)	0.0247** (0.0086)	-0.0376*** (0.0090)	-0.0703*** (0.0086)	-0.0240* (0.0101)
Año=2018 x Prog. acreditado=1	-0.0034 (0.0082)	-0.0638*** (0.0092)	0.0319*** (0.0087)	-0.0345*** (0.0091)	-0.0828*** (0.0086)	-0.0405*** (0.0102)
Año=2019 x Prog. acreditado=1	0.0041 (0.0080)	-0.0478*** (0.0091)	0.0263** (0.0085)	-0.0122 (0.0089)	-0.0968*** (0.0085)	-0.0727*** (0.0100)
Año=2020 x Prog. acreditado=1	-0.0137 (0.0080)	-0.0199* (0.0090)	0.0365*** (0.0084)	0.0044 (0.0087)	-0.0894*** (0.0085)	0.0259** (0.0099)
Año=2021 x Prog. acreditado=1	-0.0110 (0.0080)	-0.0045 (0.0091)	0.0346*** (0.0085)	0.0117 (0.0088)	-0.0940*** (0.0086)	0.0167 (0.0100)
Licenciatura x Prog. acreditado=1	-0.1265*** (0.0276)	-0.1866*** (0.0278)	-0.0125 (0.0290)	-0.0784** (0.0285)	-0.1817*** (0.0340)	0.0631 (0.0338)
Año=2013 x Licenciatura x Prog. acreditado=1	0.0611 (0.0325)	0.0253 (0.0331)	0.0075 (0.0349)	0.0553 (0.0340)	0.0842* (0.0411)	0.0127 (0.0407)
Año=2014 x Licenciatura x Prog. acreditado=1	0.1276*** (0.0318)	0.0815* (0.0326)	0.0495 (0.0340)	0.0841* (0.0337)	0.1312** (0.0403)	0.0546 (0.0401)
Año=2015 x Licenciatura x Prog. acreditado=1	0.1006** (0.0312)	0.0847** (0.0318)	-0.0217 (0.0334)	0.0664* (0.0331)	0.1371*** (0.0395)	0.0374 (0.0393)
Año=2016 x Licenciatura x Prog. acreditado=1	0.1547*** (0.0309)	0.0684* (0.0315)	0.0623 (0.0326)	0.1462*** (0.0324)	0.2006*** (0.0382)	0.0053 (0.0386)
Año=2017 x Licenciatura x Prog. acreditado=1	0.0882** (0.0302)	0.0582 (0.0307)	0.0225 (0.0319)	0.0547 (0.0318)	0.1706*** (0.0368)	-0.0752* (0.0378)
Año=2018 x Licenciatura x Prog. acreditado=1	0.0261 (0.0303)	0.0425 (0.0308)	-0.0419 (0.0321)	-0.0258 (0.0320)	0.1180** (0.0371)	-0.0876* (0.0384)
Año=2019 x Licenciatura x Prog. acreditado=1	0.0171 (0.0300)	0.0108 (0.0305)	-0.0489 (0.0317)	-0.0233 (0.0317)	0.1469*** (0.0368)	-0.1088** (0.0379)
Año=2020 x Licenciatura x Prog. acreditado=1	0.0368 (0.0299)	0.0158 (0.0304)	-0.0638* (0.0318)	0.0194 (0.0314)	0.1583*** (0.0370)	-0.1234** (0.0380)
Año=2021 x Licenciatura x Prog. acreditado=1	-0.0275 (0.0299)	-0.0391 (0.0307)	-0.0979** (0.0319)	-0.0500 (0.0316)	0.0879* (0.0370)	-0.1440*** (0.0384)
Intercepto	-0.6995*** (0.0105)	-0.1963*** (0.0112)	-0.6993*** (0.0115)	-0.7023*** (0.0119)	-1.0647*** (0.0107)	-0.2024*** (0.0137)
N	1,364,528	1,364,410	1,364,378	1,363,995	1,363,038	1,336,368

Log likelihood	-1,460,182	-1,611,975	-1,626,238	-1,688,599	-1,569,262	-1,839,427
R ² ajustado	0.4997***	0.3897***	0.3648***	0.3137***	0.442***6	0.0989***

Nota: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Se llevó a cabo la prueba Wald para evaluar si el modelo con interacciones tiene mayor poder explicativo (R^2) que el modelo sin interacciones. Los resultados en la fila de R^2 . Cálculos propios del autor con base en datos del ICFES.

Anexo 2. Resultados detallados del emparejamiento de los grupos

		Puntaje Global									
		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ATT		-0.058***	-0.013	-0.024**	0.018*	-0.084***	-0.032***	-0.039***	-0.025***	0.023***	0.076***
NATE		-0.457***	-0.434***	-0.457***	-0.373***	-0.410***	-0.404***	-0.402***	-0.350***	-0.319***	-0.292***
N											
Emparejado											
Sí		3,805	8,350	9,811	10,518	14,459	15,844	13,918	15,128	13,248	12,517
No		298	425	498	533	547	527	459	443	264	390
Total		4,103	8,775	10,309	11,051	15,006	16,371	14,377	15,571	13,512	12,907
Controles											
Usado		41,957	82,118	84,766	94,800	119,985	131,645	119,148	136,537	135,146	123,129
No usado		9,160	14,386	15,344	16,749	21,155	18,859	17,740	16,982	19,452	23,488
Total		51,117	96,504	100,110	111,549	141,140	150,504	136,888	153,519	154,598	146,617
Bandwidth		0.116	0.133	0.123	0.086	0.087	0.101	0.104	0.107	0.089	0.067
Loss		1.86e-15	1.03e-14	4.21e-16	5.96e-14	1.27e-15	2.38e-16	3.06e-16	3.19e-16	4.76e-14	5.71e-16

		Razonamiento Cuantitativo									
		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ATT		-0.346***	-0.336***	-0.299***	-0.307***	-0.372***	-0.324***	-0.331***	-0.328***	-0.249***	-0.232***
NATE		-0.626***	-0.624***	-0.606***	-0.583***	-0.571***	-0.582***	-0.577***	-0.534***	-0.516***	-0.515***
N											
Emparejado											
Sí		3,867	8,413	9,890	10,577	14,593	15,841	13,904	15,172	13,246	12,519
No		236	361	417	472	411	527	470	396	266	386
Total		4,103	8,774	10,307	11,049	15,004	16,368	14,374	15,568	13,512	12,905
Controles											
Usado		42,355	82,470	85,294	95,953	122,445	131,647	118,965	137,171	135,068	123,231
No usado		8,755	14,031	14,806	15,567	18,691	18,846	17,919	16,337	19,527	23,368
Total		51,110	96,501	100,100	111,520	141,136	150,493	136,884	153,508	154,595	146,599
Bandwidth		0.181	0.167	0.157	0.111	0.136	0.101	0.101	0.126	0.088	0.069
Loss		2.52e-15	2.13e-14	7.47e-15	1.87e-15	2.24e-12	1.11e-15	7.28e-16	9.89e-16	3.97e-14	5.19e-16

		Lectura Crítica									
		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ATT		-0.051**	0.036**	-0.003	0.037***	-0.050***	0.052***	0.029***	0.087***	0.048***	0.076***
NATE		-0.360***	-0.264***	-0.319***	-0.264***	-0.295***	-0.229***	-0.251***	-0.152***	-0.223***	-0.220***
N											
Emparejado											
Sí		3,891	8,439	9,958	10,643	14,557	15,870	13,955	15,128	13,274	12,544
No		212	334	350	408	445	500	421	442	237	361
Total		4,103	8,773	10,308	11,051	15,002	16,370	14,376	15,570	13,511	12,905
Controles											
Usado		42,419	82,629	85,602	96,633	122,003	132,054	119,646	136,570	135,769	123,901
No usado		8,688	13,868	14,496	14,892	19,130	18,430	17,238	16,938	18,823	22,680
Total		51,107	96,497	100,098	111,525	141,133	150,484	136,884	153,508	154,592	146,581
Bandwidth		0.207	0.196	0.196	0.134	0.124	0.109	0.114	0.108	0.102	0.077
Loss		2.32e-15	2.04e-16	5.39e-16	1.26e-15	4.56e-13	4.08e-16	8.87e-17	6.56e-16	1.01e-14	2.54e-16

		Competencias Ciudadanas									
		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ATT		-0.131***	-0.092***	-0.093***	-0.055***	-0.103***	-0.039***	-0.055***	-0.062***	-0.031***	0.015
NATE		-0.428***	-0.363***	-0.388***	-0.317***	-0.319***	-0.297***	-0.304***	-0.288***	-0.288***	-0.276***
N											
Emparejado											
Sí		3,817	8,420	9,955	10,677	14,563	15,890	13,925	15,133	13,302	12,482
No		286	346	350	368	431	456	446	435	208	414
Total		4,103	8,766	10,305	11,045	14,994	16,346	14,371	15,568	13,510	12,896
Controles											
Usado		42,029	82,521	85,587	96,950	122,155	132,493	119,334	136,655	136,335	122,401
No usado		9,070	13,957	14,493	14,551	18,925	17,910	17,528	16,827	18,254	24,116
Total		51,099	96,478	100,080	111,501	141,080	150,403	136,862	153,482	154,589	146,517
Bandwidth		0.123	0.179	0.196	0.159	0.130	0.123	0.108	0.111	0.116	0.062
Loss		1.93e-15	2.54e-13	5.41e-16	2.29e-15	8.66e-13	5.03e-16	2.59e-16	4.70e-16	1.71e-15	5.93e-16

		Inglés									
		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ATT		0.207***	0.218***	0.199***	0.243***	0.137***	0.159***	0.199***	0.189***	0.190***	0.207***
NATE		-0.371***	-0.297***	-0.302***	-0.197***	-0.279***	-0.311***	-0.256***	-0.216***	-0.233***	-0.238***
N											
Emparejado											
Sí		3,772	8,181	9,686	10,464	14,433	15,812	13,896	15,112	13,178	12,519
No		311	560	594	543	502	497	469	445	329	384
Total		4,083	8,741	10,280	11,007	14,935	16,309	14,365	15,557	13,507	12,903
Controles											
Usado		41,694	80,884	83,727	94,228	120,832	131,867	118,920	136,423	132,790	123,226
No usado		9,350	15,502	16,261	17,130	20,052	18,440	17,929	17,022	21,779	23,295

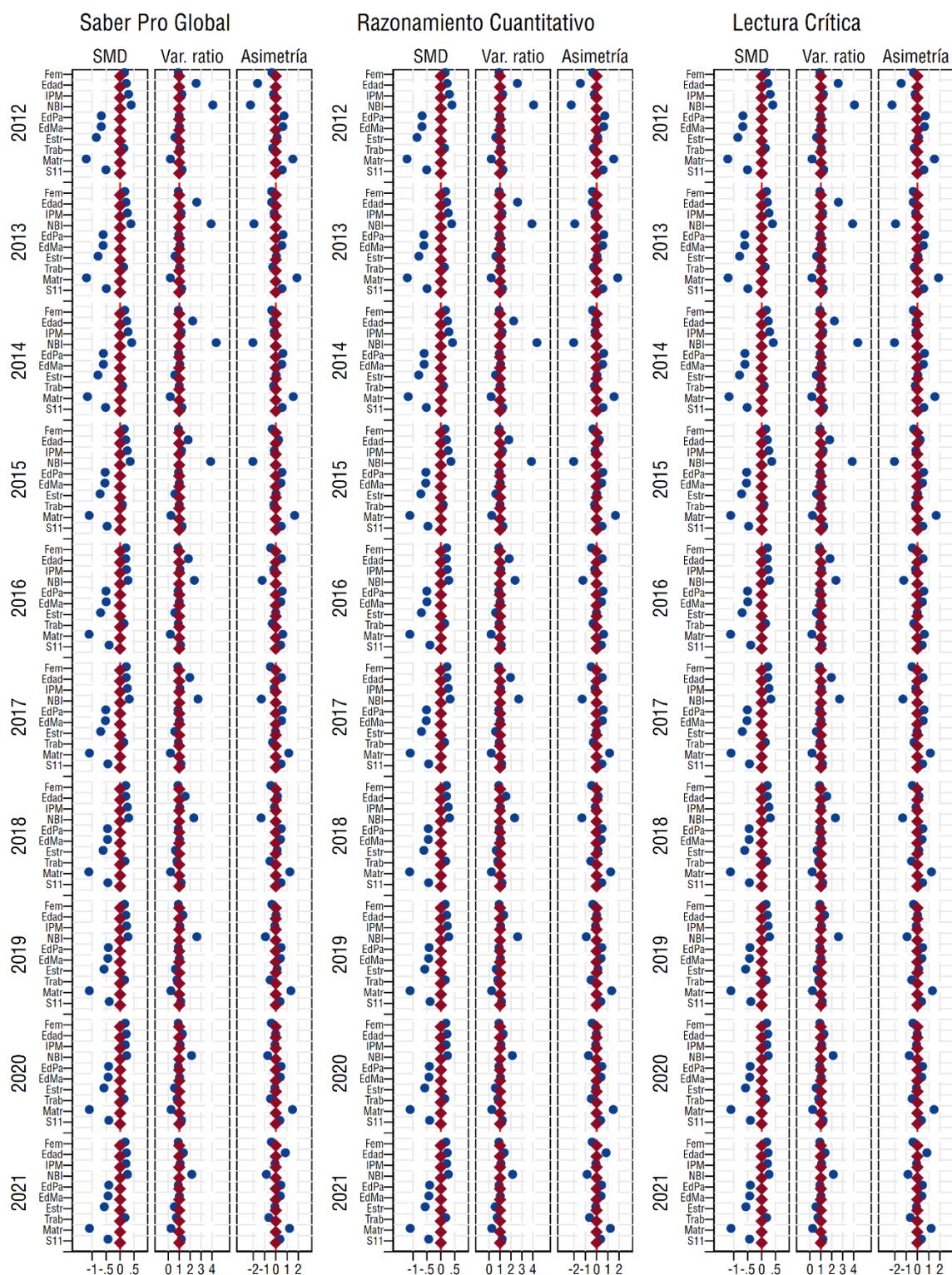
Total	51,044	96,386	99,988	111,358	140,884	150,307	136,849	153,445	154,569	146,521
Bandwidth	0.105	0.083	0.093	0.081	0.101	0.109	0.101	0.106	0.066	0.069
Loss	8.99e-14	2.78e-14	9.13e-16	1.20e-14	8.66e-15	7.33e-16	6.12e-16	2.14e-16	1.32e-15	3.05e-16

Comunicación Escrita

	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
ATT	0.094***	0.085***	0.060***	0.099***	0.039***	0.051***	0.061***	0.064***	0.147***	0.231***
NATE	-0.051**	-0.038**	-0.088***	-0.059***	-0.056***	-0.088***	-0.078***	-0.057***	-0.011	0.058***
N										
Emparejado										
Sí	3,824	8,473	9,996	10,729	14,646	15,249	13,538	14,605	12,800	12,286
No	209	253	239	277	290	310	283	333	145	336
Total	4,033	8,726	10,235	11,006	14,936	15,559	13,821	14,938	12,945	12,622
Controles										
Usado	41,384	82,643	85,735	97,425	123,274	127,839	117,688	133,652	132,555	121,726
No usado	8,773	13,492	13,998	13,753	17,420	16,156	15,713	15,590	17,030	21,701
Total	50,157	96,135	99,733	111,178	140,694	143,995	133,401	149,242	149,585	143,427
Bandwidth	0.211	0.344	0.368	0.284	0.262	0.213	0.167	0.148	0.163	0.085
Loss	6.30e-15	1.40e-15	2.77e-15	4.91e-14	6.95e-13	1.90e-15	2.52e-16	7.21e-16	1.76e-15	5.23e-16

Nota: * p < 0.05, ** p < 0.01, *** p < 0.001. ATT: tratamiento promedio en el grupo tratado. NATE: tratamiento promedio previo al emparejamiento (con sesgo).

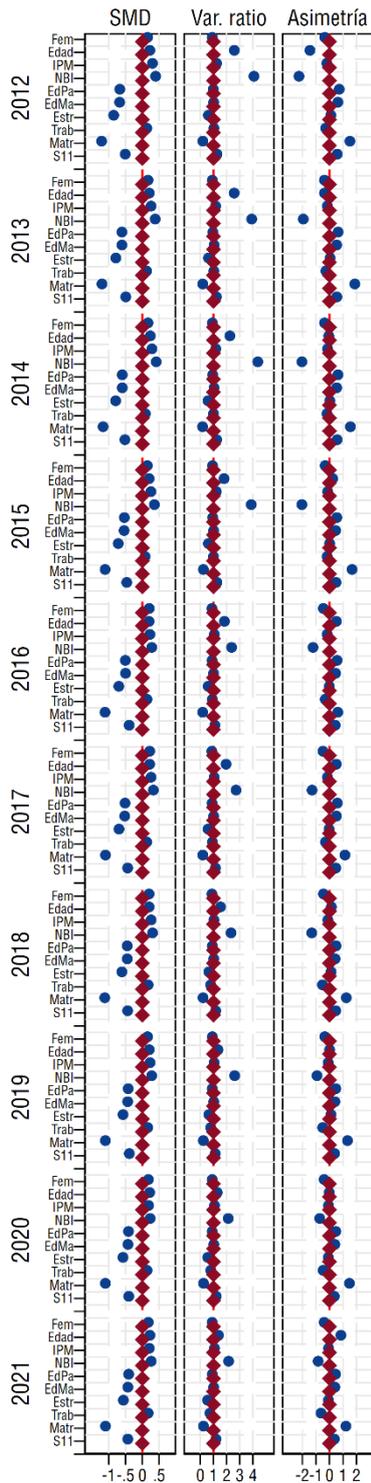
Anexo 3. Balanceo de grupos por covariables. Cálculos por año



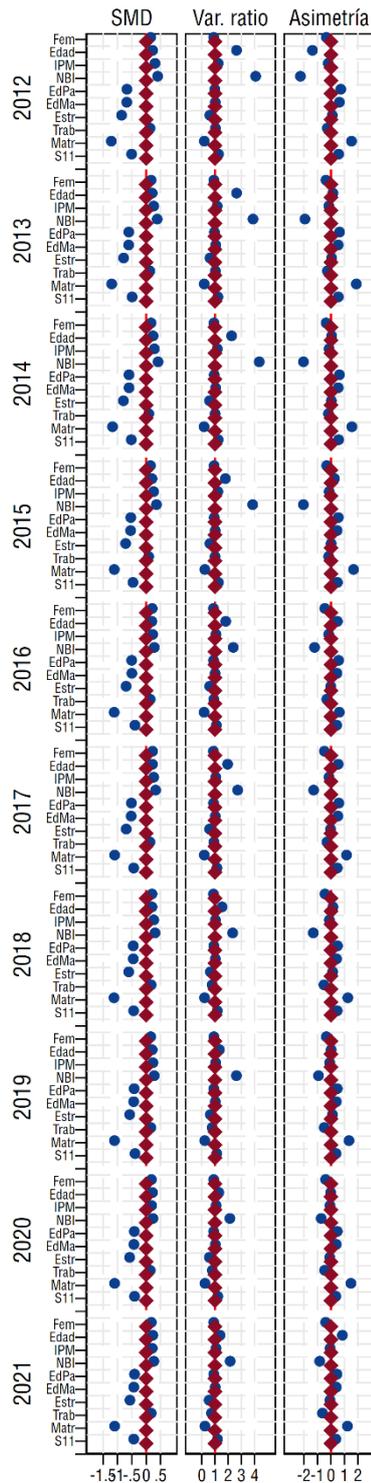
Nota: SMD: Diferencia de media estándar; Fem: Mujer; IPM: Índice de Pobreza Multidimensional por Región; NBI: Índice de Necesidades Básicas Insatisfechas por Municipio; MUN: Tipo de municipio; EdPa: Educación del padre; EdMa: Educación de la madre; Estr: Estrato socioeconómico; Trab: Estudiante trabaja; Matr: Valor matrícula universidad; S11: Percentil en Saber 11

● Sin emparejar
◆ Emparejado

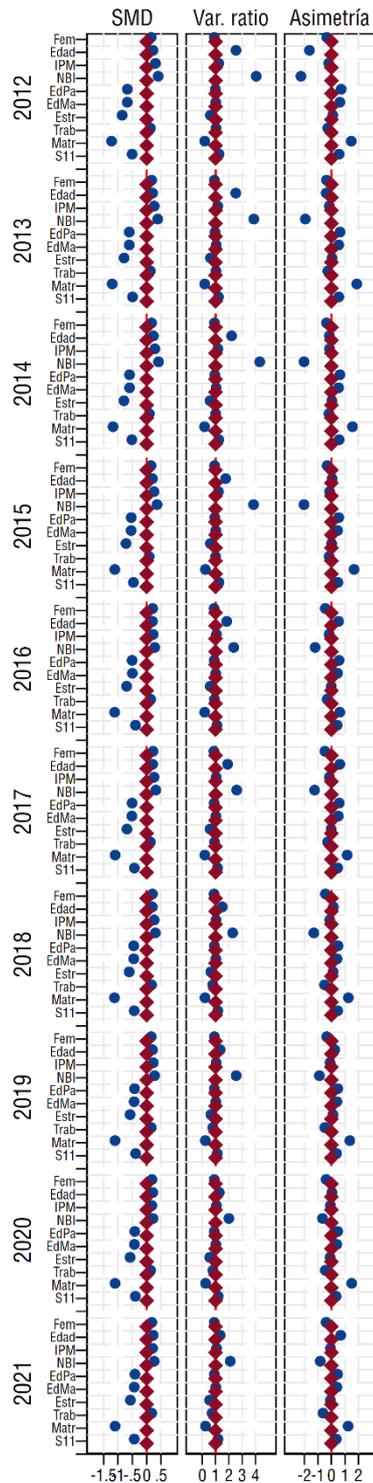
Competencias Ciudadanas



Inglés



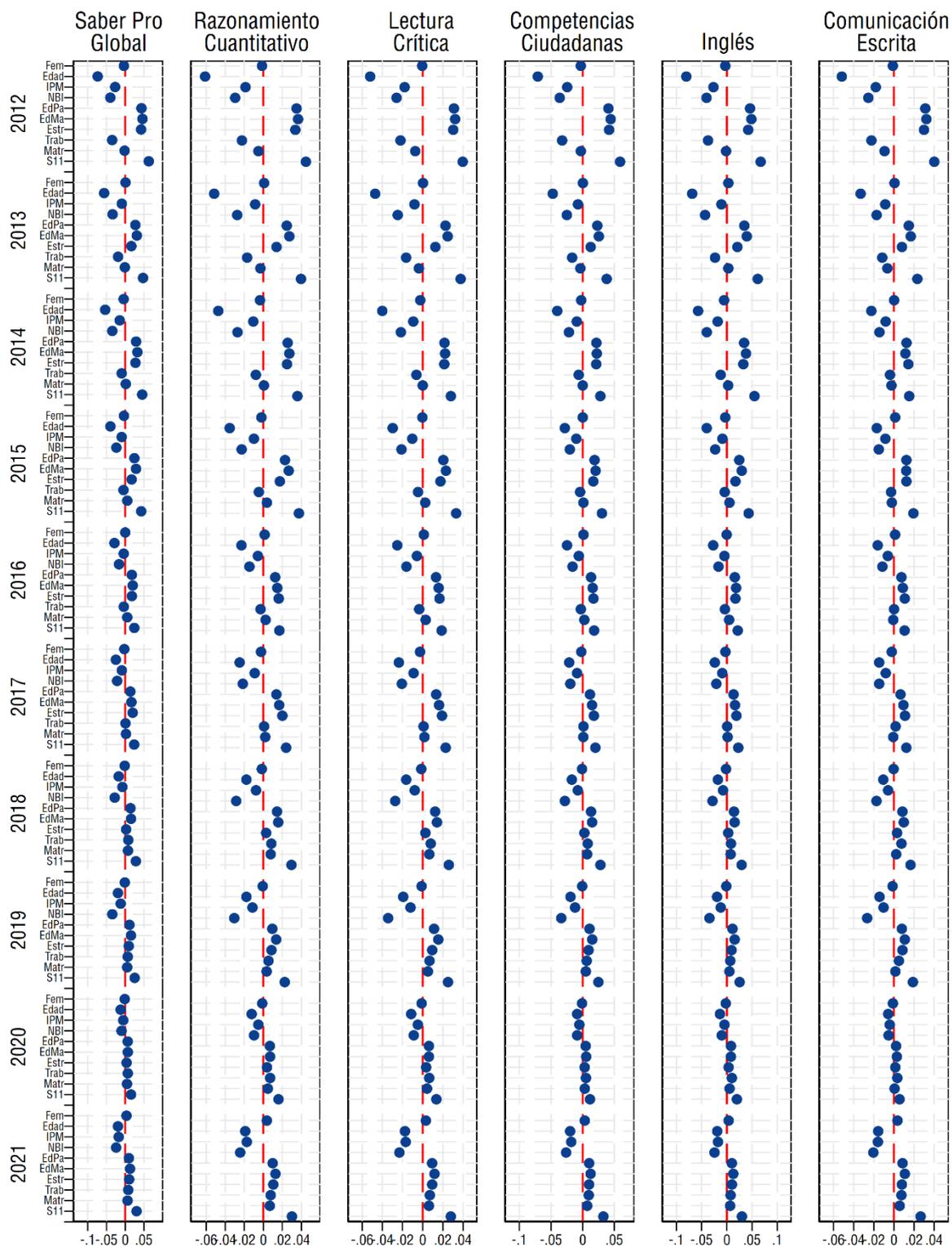
Comunicación Escrita



Nota: **SMD**: Diferencia de media estándar; **Fem**: Mujer; **IPM**: Índice de Pobreza Multidimensional por Región; **NBI**: Índice de Necesidades Básicas Insatisfechas por Municipio; **MUN**: Tipo de municipio; **EdPa**: Educación del padre; **EdMa**: Educación de la madre; **Estr**: Estrato socioeconómico; **Trab**: Estudiante trabaja; **Matr**: Valor matrícula universidad; **S11**: Percentil en Saber 11

● Sin emparejar
◆ Emparejado

Anexo 4. Compatibilidad del grupo tratado posterior al emparejamiento en comparación con el grupo original (0). Cálculos por año



Nota: Fem: Mujer; IPM: Índice de Pobreza Multidimensional por Región; NBI: Índice de Necesidades Básicas Insatisfechas por Municipio; MUN: Tipo de municipio; EdPa: Educación del padre; EdMa: Educación de la madre; Estr: Estrato socioeconómico; Trab: Estudiante trabaja; Matr: Valor matrícula universidad; S11: Percentil en Saber 11

Anexo 5. Resultados completos del DiD, incluyendo variables de control

Puntaje global					
	2017	2018	2019	2020	2021
ATET	0.0167	0.0205	0.0333**	0.0604***	0.0883***
Controles					
Fem	-0.0549***	-0.0539***	-0.0467***	-0.0537***	-0.0505***
Edad	-0.0295***	-0.0289***	-0.0293***	-0.0292***	-0.0297***
IPM	-0.9061***	-0.9482***	-1.1198***	-0.9357***	-0.8231***
NBI	-0.0036***	-0.0035***	-0.0031***	-0.0030***	-0.0030***
EdPa	0.0035***	0.0030***	0.0031***	0.0028***	0.0030***
EdMa	0.0068***	0.0074***	0.0067***	0.0062***	0.0059***
Estr	0.0394***	0.0333***	0.0319***	0.0359***	0.0333***
Trab	0.0236***	0.0286***	0.0228***	0.0193***	0.0203***
Matr	0.0233***	0.0238***	0.0264***	0.0257***	0.0231***
S11	0.0134***	0.0137***	0.0140***	0.0146***	0.0144***
Intercepto	-0.3367***	-0.3517***	-0.3236***	-0.3962***	-0.3902***
Nº programas					
Control	3,366	3,463	3,624	3,770	3,864
Tratados	408	412	439	457	469
Total	3,774	3,875	4,063	4,227	4,333
Ptrends	2.7126	3.1113	2.6386	4.2343*	3.0479
Razonamiento Cuantitativo					
ATET	-0.0026	0.0183	0.0087	0.0762***	0.0727***
Controles					
Fem	-0.2367***	-0.2429***	-0.2302***	-0.2329***	-0.2272***
Edad	-0.0275***	-0.0274***	-0.0267***	-0.0276***	-0.0285***
IPM	-0.9744***	-0.9767***	-1.0553***	-0.9039***	-0.9431***
NBI	-0.0022***	-0.0021***	-0.0018***	-0.0017***	-0.0019***
EdPa	0.0009*	0.0011*	0.0007	0.0006	0.0009*
EdMa	0.0057***	0.0052***	0.0044***	0.0047***	0.0043***
Estr	0.0135***	0.0112***	0.0103***	0.0122***	0.0110***
Trab	0.0538***	0.0561***	0.0516***	0.0479***	0.0483***
Matr	0.0244***	0.0244***	0.0263***	0.0272***	0.0249***
S11	0.0106***	0.0109***	0.0112***	0.0118***	0.0116***
Intercepto	-0.0600	-0.0733	-0.0842	-0.1376**	-0.0935*
Nº programas					
Control	3,369	3,468	3,627	3,775	3,868
Tratados	408	412	439	457	470
Total	3,777	3,880	4,066	4,232	4,338
Ptrends	1.4391	0.8539	0.5658	0.2140	0.2639
Lectura Crítica					
ATET	0.0710***	0.0619***	0.1219***	0.0575***	0.0802***
Controles					
Fem	-0.0031	-0.0037	0.0147**	0.0046	0.0118*
Edad	-0.0232***	-0.0233***	-0.0239***	-0.0244***	-0.0247***
IPM	-0.8594***	-0.9764***	-1.1125***	-0.9479***	-0.9630***
NBI	-0.0033***	-0.0030***	-0.0028***	-0.0031***	-0.0028***
EdPa	0.0029***	0.0028***	0.0025***	0.0021***	0.0024***
EdMa	0.0055***	0.0053***	0.0046***	0.0048***	0.0042***
Estr	0.0255***	0.0201***	0.0213***	0.0208***	0.0200***
Trab	0.0102*	0.0155**	0.0123**	0.0075	0.0082
Matr	0.0158***	0.0167***	0.0166***	0.0182***	0.0146***
S11	0.0136***	0.0140***	0.0142***	0.0148***	0.0148***
Intercepto	-0.4172***	-0.4089***	-0.3854***	-0.4349***	-0.4214***
Nº programas					
Control	3,370	3,470	3,628	3,776	3,870
Tratados	408	412	439	457	470
Total	3,778	3,882	4,067	4,233	4,340
Ptrends	1.3065	2.0500	1.5102	2.5873	1.5672
Competencias Ciudadanas					
ATET	0.0634***	0.0567***	0.0577***	0.0646***	0.1005***

Controles					
Fem	-0.0104*	-0.0057	-0.0038	-0.0115*	-0.0060
Edad	-0.0184***	-0.0183***	-0.0184***	-0.0172***	-0.0182***
IPM	-1.0847***	-1.0435***	-1.1742***	-1.2241***	-1.1369***
NBI	-0.0023***	-0.0024***	-0.0023***	-0.0020***	-0.0019***
EdPa	0.0008	0.0003	0.0002	-0.0003	0.0004
EdMa	0.0040***	0.0050***	0.0045***	0.0040***	0.0036***
Estr	0.0198***	0.0126***	0.0101**	0.0152***	0.0142***
Trab	0.0185***	0.0189***	0.0138**	0.0100*	0.0132**
Matr	0.0173***	0.0162***	0.0192***	0.0175***	0.0154***
S11	0.0130***	0.0131***	0.0135***	0.0147***	0.0143***
Intercepto	-0.5145***	-0.5236***	-0.5089***	-0.5988***	-0.5739***
Nº programas					
Control	3,367	3,468	3,626	3,775	3,867
Tratados	408	412	439	457	470
Total	3,775	3,880	4,065	4,232	4,337
Ptrends	2.8501	1.2891	2.8712	0.5184	0.6409

Inglés

ATET	-0.0065	0.0342*	0.0188	0.0052	-0.0190
Controles					
Fem	-0.0828***	-0.0815***	-0.0790***	-0.0725***	-0.0775***
Edad	-0.0177***	-0.0164***	-0.0171***	-0.0160***	-0.0163***
IPM	-0.5402***	-0.6585***	-0.6903***	-0.5046***	-0.4278**
NBI	-0.0028***	-0.0028***	-0.0027***	-0.0025***	-0.0030***
EdPa	0.0070***	0.0066***	0.0068***	0.0066***	0.0065***
EdMa	0.0091***	0.0098***	0.0098***	0.0094***	0.0096***
Estr	0.0825***	0.0761***	0.0766***	0.0825***	0.0794***
Trab	-0.0210***	-0.0174***	-0.0220***	-0.0251***	-0.0262***
Matr	0.0250***	0.0244***	0.0231***	0.0237***	0.0235***
S11	0.0072***	0.0073***	0.0075***	0.0079***	0.0078***
Intercepto	-0.4206***	-0.4164***	-0.4067***	-0.5082***	-0.4976***
Nº programas					
Control	3,366	3,465	3,625	3,771	3,866
Tratados	408	412	439	457	469
Total	3,774	3,877	4,064	4,228	4,335
Ptrends	9.0485**	7.8472**	8.5621**	9.3780**	9.1552**

Comunicación Escrita

ATET	-0.0156	-0.0118	-0.0163	0.0711***	0.1331***
Controles					
Fem	0.1050***	0.1108***	0.1074***	0.0914***	0.0921***
Edad	-0.0176***	-0.0177***	-0.0176***	-0.0189***	-0.0189***
IPM	0.4530	0.5452*	0.4453	0.2633	0.6068**
NBI	-0.0030***	-0.0035***	-0.0029***	-0.0028***	-0.0031***
EdPa	0.0007	0.0002	0.0009	0.0008	0.0005
EdMa	0.0011	0.0017**	0.0011	0.0009	0.0008
Estr	0.0076*	0.0051	0.0014	0.0058	0.0009
Trab	0.0430***	0.0501***	0.0468***	0.0446***	0.0443***
Matr	0.0045	0.0053	0.0099**	0.0096**	0.0074*
S11	0.0063***	0.0064***	0.0063***	0.0069***	0.0070***
Intercepto	-0.1908**	-0.2069**	-0.1711**	-0.1504*	-0.2055**
Nº programas					
Control	3,369	3,468	3,624	3,772	3,869
Tratados	408	412	439	456	470
Total	3,777	3,880	4,063	4,228	4,339
Ptrends	4.8674*	4.4160*	5.0198*	4.5292*	2.6476

Nota: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATET ajustado por covariables, efectos de grupo y año de presentación del examen Saber Pro. Errores estándares ajustados por programa. Fem: Mujer; IPM: Índice de Pobreza Multidimensional por Región; NBI: Índice de Necesidades Básicas Insatisfechas por Municipio; MUN: Tipo de municipio; EdPa: Educación del padre; EdMa: Educación de la madre; Estr: Estrato socioeconómico; Trab: Estudiante trabaja; Matr: Valor matrícula universidad; S11: Percentil en Saber 11º; Ptrends: prueba de tendencias paralelas (H_0 = las tendencias son paralelas).

Anexo 6. Resultados de la estimación del DiD sin balanceo de grupos

Puntaje global					
	2017	2018	2019	2020	2021
ATET	0.0244*	0.0094	0.0087	0.0368**	0.0594***
Controles					
Fem	-0.0531***	-0.0548***	-0.0496***	-0.0558***	-0.0515***
Edad	-0.0342***	-0.0339***	-0.0345***	-0.0348***	-0.0361***
IPM	-0.9394***	-1.0197***	-1.1167***	-1.0957***	-1.0199***
NBI	-0.0041***	-0.0041***	-0.0039***	-0.0033***	-0.0032***
EdPa	0.0051***	0.0049***	0.0047***	0.0044***	0.0046***
EdMa	0.0076***	0.0081***	0.0077***	0.0072***	0.0069***
Estr	0.0286***	0.0274***	0.0253***	0.0270***	0.0269***
Trab	0.0196***	0.0185***	0.0163***	0.0100***	0.0122***
Matr	0.0130***	0.0132***	0.0135***	0.0139***	0.0135***
S11	0.0156***	0.0159***	0.0162***	0.0167***	0.0166***
Intercepto	-0.2764***	-0.2894***	-0.2712***	-0.3010***	-0.2809***
Nº programas					
Control	3,392	3,494	3,651	3,803	3,915
Tratados	408	412	441	457	470
Total	3,800	3,906	4,092	4,260	4,385
Ptrends	2.5901	3.3825	3.4143	4.2285*	3.2583
Razonamiento Cuantitativo					
ATET	0.0137	0.0147	0.0121	0.0310***	0.0252*
Controles					
Fem	-0.2769***	-0.2831***	-0.2708***	-0.2705***	-0.2639***
Edad	-0.0339***	-0.0328***	-0.0329***	-0.0347***	-0.0363***
IPM	-0.8651***	-0.9075***	-0.9060***	-0.8439***	-0.7028***
NBI	-0.0027***	-0.0026***	-0.0024***	-0.0023***	-0.0022***
EdPa	0.0015***	0.0013***	0.0011***	0.0012***	0.0011***
EdMa	0.0051***	0.0050***	0.0045***	0.0046***	0.0044***
Estr	0.0035*	0.0024	-0.0002	0.0004	0.0008
Trab	0.0574***	0.0558***	0.0545***	0.0470***	0.0485***
Matr	0.0164***	0.0162***	0.0162***	0.0175***	0.0170***
S11	0.0129***	0.0132***	0.0134***	0.0141***	0.0139***
Intercepto	0.1290***	0.0996***	0.0906**	0.0658*	0.0815**
Nº programas					
Control	3,392	3,494	3,651	3,803	3,915
Tratados	408	412	441	457	470
Total	3,800	3,906	4,092	4,260	4,385
Ptrends	0.4818	0.9757	0.9660	1.0088	0.7332
Lectura Crítica					
ATET	0.0811***	0.0526***	0.1057***	0.0223*	0.0433***
Controles					
Fem	0.0006	-0.0021	0.0175***	0.0036	0.0139***
Edad	-0.0282***	-0.0288***	-0.0293***	-0.0302***	-0.0318***
IPM	-0.8431***	-0.9301***	-0.9821***	-0.9131***	-0.8048***
NBI	-0.0040***	-0.0038***	-0.0038***	-0.0033***	-0.0031***
EdPa	0.0035***	0.0032***	0.0029***	0.0027***	0.0028***
EdMa	0.0054***	0.0056***	0.0050***	0.0049***	0.0044***
Estr	-0.0014	-0.0027	-0.0052**	-0.0046**	-0.0043*
Trab	0.0023	0.0026	0.0008	-0.0058**	-0.0041
Matr	0.0064***	0.0068***	0.0070***	0.0076***	0.0071***
S11	0.0153***	0.0156***	0.0160***	0.0165***	0.0165***
Intercepto	-0.2940***	-0.2833***	-0.2799***	-0.3089***	-0.2940***
Nº programas					
Control	3,392	3,494	3,651	3,803	3,915
Tratados	408	412	441	457	470
Total	3,800	3,906	4,092	4,260	4,385
Ptrends	1.9823	2.7720	2.6720	3.8048	2.8917
Competencias Ciudadanas					
ATET	0.0629***	0.0462***	0.0297**	0.0214	0.0406**

Controles					
Fem	-0.0170***	-0.0131***	-0.0105***	-0.0223***	-0.0141***
Edad	-0.0243***	-0.0248***	-0.0255***	-0.0240***	-0.0260***
IPM	-1.1883***	-1.1345***	-1.3075***	-1.1580***	-1.0902***
NBI	-0.0029***	-0.0029***	-0.0026***	-0.0025***	-0.0022***
EdPa	0.0013***	0.0012***	0.0009***	0.0006*	0.0008**
EdMa	0.0045***	0.0053***	0.0050***	0.0043***	0.0042***
Estr	-0.0087***	-0.0102***	-0.0112***	-0.0112***	-0.0101***
Trab	0.0194***	0.0166***	0.0142***	0.0087***	0.0084***
Matr	0.0092***	0.0094***	0.0098***	0.0095***	0.0100***
S11	0.0146***	0.0148***	0.0151***	0.0162***	0.0159***
Intercepto	-0.2928***	-0.3141***	-0.2840***	-0.4048***	-0.3678***
Nº programas					
Control	3,391	3,494	3,651	3,803	3,915
Tratados	408	412	441	457	470
Total	3,799	3,906	4,092	4,260	4,385
Ptrends	10.2104**	6.2337*	7.8664**	5.0802*	5.0121*

Inglés

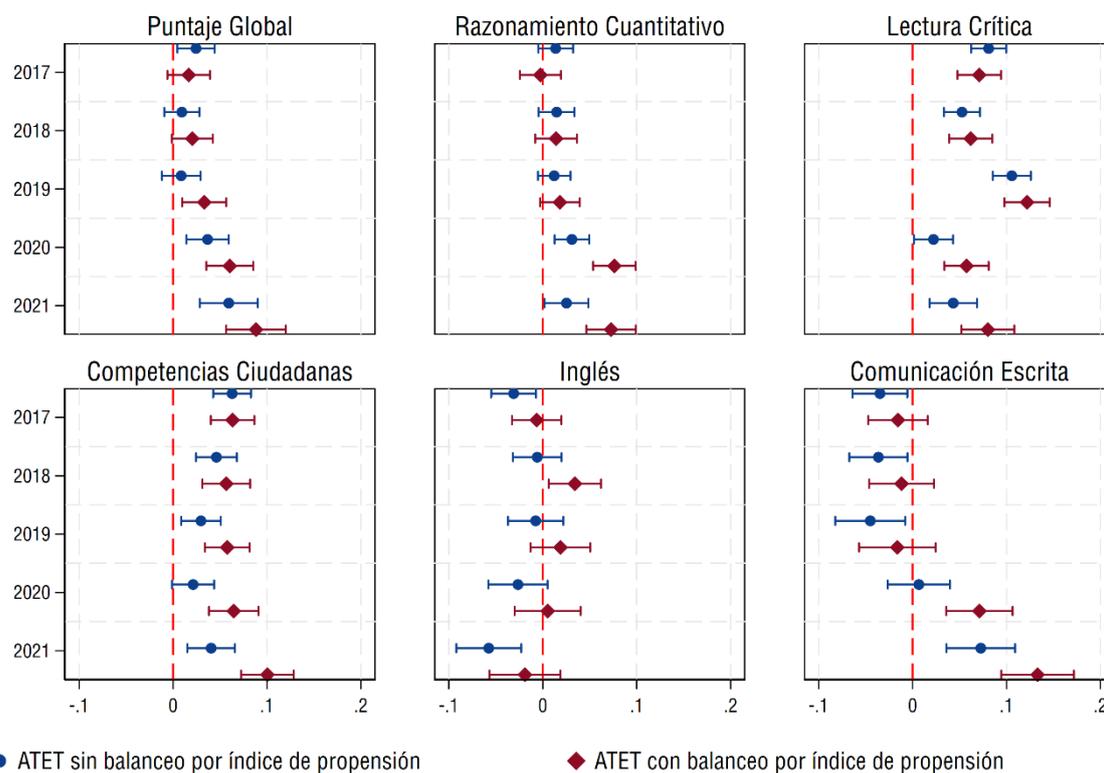
ATET	-0.0311*	-0.0059	-0.0076	-0.0264	-0.0576**
Controles					
Fem	-0.0779***	-0.0760***	-0.0763***	-0.0680***	-0.0739***
Edad	-0.0187***	-0.0183***	-0.0186***	-0.0187***	-0.0193***
IPM	-0.9845***	-1.0724***	-1.1143***	-1.1194***	-1.0769***
NBI	-0.0028***	-0.0029***	-0.0029***	-0.0027***	-0.0029***
EdPa	0.0105***	0.0106***	0.0105***	0.0100***	0.0102***
EdMa	0.0118***	0.0121***	0.0122***	0.0121***	0.0121***
Estr	0.1092***	0.1079***	0.1066***	0.1123***	0.1127***
Trab	-0.0330***	-0.0353***	-0.0362***	-0.0415***	-0.0405***
Matr	0.0126***	0.0130***	0.0127***	0.0131***	0.0128***
S11	0.0103***	0.0104***	0.0106***	0.0109***	0.0109***
Intercepto	-0.5296***	-0.5254***	-0.5202***	-0.5558***	-0.5495***
Nº programas					
Control	3,391	3,494	3,651	3,803	3,915
Tratados	408	412	441	457	470
Total	3,799	3,906	4,092	4,260	4,385
Ptrends	7.9980**	8.5436**	8.0645**	8.5802**	8.3045**

Comunicación Escrita

ATET	-0.0347*	-0.0364*	-0.0450*	0.0067	0.0726***
Controles					
Fem	0.1538***	0.1556***	0.1474***	0.1326***	0.1324***
Edad	-0.0189***	-0.0187***	-0.0183***	-0.0210***	-0.0209***
IPM	0.4155**	0.3558**	0.4456***	0.1534	0.1561
NBI	-0.0031***	-0.0031***	-0.0029***	-0.0027***	-0.0027***
EdPa	0.0011***	0.0007*	0.0010***	0.0009***	0.0010***
EdMa	0.0014***	0.0018***	0.0014***	0.0010**	0.0008**
Estr	-0.0094***	-0.0097***	-0.0104***	-0.0091***	-0.0113***
Trab	0.0374***	0.0384***	0.0374***	0.0325***	0.0365***
Matr	0.0028**	0.0028**	0.0033***	0.0044***	0.0039***
S11	0.0067***	0.0067***	0.0065***	0.0073***	0.0073***
Intercepto	-0.1798***	-0.1749***	-0.1802***	-0.1097***	-0.1098***
Nº progs					
Control	3,389	3,491	3,646	3,798	3,914
Tratados	408	412	441	456	470
Total	3,797	3,903	4,087	4,254	4,384
Ptrends	7.4738**	9.2885**	10.5400**	8.4020**	6.3823*

Nota: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATET ajustado por covariables, efectos de grupo y año de presentación del examen. Errores estándares ajustados por programa. Fem: Mujer; IPM: Índice de Pobreza Multidimensional por Región; NBI: Índice de Necesidades Básicas Insatisfechas por Municipio; MUN: Tipo de municipio; EdPa: Educación del padre; EdMa: Educación de la madre; Estr: Estrato socioeconómico; Trab: Estudiante trabaja; Matr: Valor matrícula universidad; S11: Percentil en Saber 11º; Ptrends: prueba de tendencias paralelas (H0 = las tendencias son paralelas).

Anexo 7. Comparación del DiD. Muestra balanceada y sin balancear



Anexo 8. Resultados de DiD por subgrupos – diferencias no significativas

Resultados por género

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Masculino	Femenino	Masculino	Femenino	Masculino	Femenino	Masculino	Femenino	Masculino	Femenino	Masculino	Femenino
ATET												
2017	0.0158	0.0168	-0.0043	-0.0013	0.0624**	0.0743***	0.0604**	0.0642***	-0.0138	-0.0040	-0.0113	-0.0136
2018	0.0205	0.0209	0.0407*	0.0011	0.0498*	0.0681***	0.0457*	0.0618***	0.0110	0.0436**	-0.0143	-0.0107
2019	0.0239	0.0362**	-0.0010	0.0285*	0.1058**	0.1291***	0.0708**	0.0497***	-0.0025	0.0276	-0.0099	-0.0203
2020	0.0589**	0.0619***	0.0828**	0.0740***	0.0185	0.0772***	0.0686**	0.0621***	0.0062	0.0036	0.0592*	0.0770***
2021	0.0700**	0.0985***	0.0589**	0.0810***	0.0538**	0.0937***	0.0914**	0.1021***	-0.0166	-0.0200	0.0917**	0.1544***
Wald												
2017	0.0001		0.0015		0.0184		0.0022		0.0103		0.0019	
2018	0.0000		0.2282		0.0427		0.0489		0.1314		0.0057	
2019	0.0166		0.1212		0.0662		0.0681		0.1033		0.0649	
2020	0.0011		0.0120		0.5307		0.0063		0.0008		0.0910	
2021	0.1129		0.0714		0.2481		0.0201		0.0014		1.0596	

Resultados por nivel de educación de los padres

Nivel bajo: 16 años o menos de educación de los padres, entre ambos

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto
ATET												
2017	0.0123	0.0186	-0.0090	0.0009	0.0724**	0.0689***	0.0613**	0.0629***	-0.0020	-0.0120	-0.0267	-0.0064
2018	0.0122	0.0274*	-0.0083	0.0327*	0.0546**	0.0676***	0.0514**	0.0613***	0.0415**	0.0259	-0.0286	0.0055
2019	0.0317*	0.0316*	0.0084	0.0248	0.1186**	0.1211***	0.0573**	0.0540***	0.0287	0.0079	-0.0407	0.0074
2020	0.0656**	0.0541***	0.0808**	0.0699***	0.0657**	0.0471**	0.0654**	0.0584***	0.0142	-0.0009	0.0612*	0.0805***
2021	0.1176**	0.0653***	0.0784**	0.0683***	0.1027**	0.0593***	0.1068**	0.0943***	-0.0024	-0.0300	0.1640**	0.1084***
Wald												
2017	0.0034		0.0109		0.0012		0.0003		0.0101		0.1401	
2018	0.0331		0.2324		0.0231		0.0231		0.0431		0.5056	
2019	0.0000		0.0366		0.0011		0.0021		0.0608		1.0528	
2020	0.0186		0.0163		0.0610		0.0075		0.0326		0.1207	
2021	0.5286		0.0171		0.3345		0.0330		0.1039		0.8421	

Resultados por nivel de valor matrícula universidad

Nivel bajo: 1 salario mínimo legal vigente o menos.

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto
ATET												
2017	0.0016	0.0312*	0.0210	-0.0173	0.0603**	0.0791***	0.0513**	0.0778***	-0.0201	0.0057	-0.0388	0.0060
2018	0.0231	0.0228	0.0384*	-0.0042	0.0661**	0.0616***	0.0534*	0.0662***	0.0243	0.0434**	-0.0103	-0.0052
2019	0.0326	0.0401**	0.0413*	-0.0008	0.1240**	0.1216***	0.0563**	0.0683***	0.0217	0.0225	-0.0360	0.0095
2020	0.0583**	0.0661***	0.0961**	0.0612***	0.0499*	0.0628***	0.0495*	0.0729***	0.0117	-0.0006	0.0365	0.1095***
2021	0.0780**	0.1027***	0.0817**	0.0687***	0.0794**	0.0836***	0.0753**	0.1252***	-0.0285	-0.0109	0.1178**	0.1588***
Wald												
2017	0.0311		0.0525		0.0155		0.0287		0.0134		0.1681	
2018	0.0000		0.0588		0.0008		0.0066		0.0073		0.0027	
2019	0.0017		0.0480		0.0002		0.0059		0.0000		0.1972	
2020	0.0017		0.0340		0.0064		0.0185		0.0027		0.2760	
2021	0.0207		0.0053		0.0008		0.1085		0.0052		0.0761	

Resultados por nivel de Índice de Pobreza Multidimensional

Nivel bajo: NBI inferior o igual a 6.2

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto
ATET												
2017	0.0307	0.0025	0.0174	-0.0224	0.0722** *	0.0667***	0.0936** *	0.0296	-0.0122	-0.0004	-0.0259	-0.0089
2018	0.0281	0.0126	0.0261	-0.0007	0.0614** *	0.0642***	0.0735** *	0.0361	0.0287	0.0413*	-0.0097	-0.0135
2019	0.0406*	0.0256	0.0168	0.0194	0.1289** *	0.1125***	0.0735** *	0.0388*	0.0203	0.0198	-0.0039	-0.0332
2020	0.0590** *	0.0618***	0.0793** *	0.0754***	0.0532** *	0.0616***	0.0691** *	0.0616***	0.0012	0.0159	0.0590*	0.0812**
2021	0.0799** *	0.0899***	0.0524**	0.0893***	0.0700** *	0.0872***	0.1151** *	0.0848***	-0.0140	-0.0233	0.1100** *	0.1460***
Wal d												
2017		0.0786		0.1954		0.0036		0.6044		0.0070		0.0646
2018		0.0340		0.1033		0.0013		0.2935		0.0086		0.0046
2019		0.0318		0.0010		0.0450		0.2576		0.0000		0.2776
2020		0.0013		0.0025		0.0151		0.0119		0.0117		0.1028
2021		0.0123		0.1632		0.0437		0.1533		0.0036		0.2069

Resultados por nivel de Índice de Necesidades Básicas Insatisfechas

Nivel bajo: IPM inferior o igual a 0.206

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto	Bajo	Alto
ATET												
2017	0.0205	0.0082	0.0159	-0.0257	0.0731** *	0.0651***	0.0974** *	0.0204	-0.0149	-0.0059	-0.0438*	0.0134
2018	0.0194	0.0183	0.0139	0.0085	0.0583** *	0.0642***	0.0776** *	0.0289	0.0227	0.0407*	-0.0230	0.0012
2019	0.0266	0.0360*	0.0031	0.0320*	0.1202** *	0.1219***	0.0693** *	0.0389*	0.0131	0.0200	-0.0163	-0.0157
2020	0.0508**	0.0619***	0.0731** *	0.0731***	0.0513**	0.0575***	0.0646** *	0.0533**	-0.0037	0.0116	0.0560*	0.0856***
2021	0.0767** *	0.0955***	0.0458**	0.0994***	0.0804** *	0.0783***	0.1112** *	0.0842***	-0.0196	-0.0241	0.1109** *	0.1491***
Wal d												
2017		0.0210		0.2632		0.0108		1.2157		0.0049		0.9398
2018		0.0002		0.0046		0.0064		0.5721		0.0197		0.2100
2019		0.0121		0.1193		0.0004		0.2009		0.0023		0.0001
2020		0.0206		0.0000		0.0086		0.0269		0.0146		0.2130
2021		0.0506		0.3858		0.0007		0.1389		0.0010		0.2651

Resultados por sector de la institución

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Privado	Oficial	Privado	Oficial	Privado	Oficial	Privado	Oficial	Privado	Oficial	Privado	Oficial
ATET												
2017	0.0437**	0.0020	-0.0151	0.0030	0.0824** *	0.0652***	0.0753** *	0.0566***	0.0197	-0.0194	0.0419	-0.0488*
2018	0.0331*	0.0140	0.0013	0.0203	0.0712** *	0.0583***	0.0688** *	0.0510**	0.0481*	0.0276	-0.0026	-0.0218
2019	0.0457**	0.0259	-0.0127	0.0317*	0.1241** *	0.1198***	0.0519**	0.0606***	0.0379	0.0090	0.0259	-0.0417
2020	0.0827** *	0.0485**	0.0559**	0.0851***	0.0781** *	0.0462**	0.0951** *	0.0514***	0.0148	-0.0006	0.1201** *	0.0483*
2021	0.1107** *	0.0764***	0.0688**	0.0730***	0.0857** *	0.0729***	0.1436** *	0.0773***	0.0076	-0.0338	0.1398** *	0.1305***
Wal d												
2017		0.0791		0.0197		0.0169		0.0243		0.0398		1.0227
2018		0.0186		0.0195		0.0102		0.0274		0.0116		0.0576
2019		0.0233		0.1281		0.0014		0.0071		0.0217		0.8559
2020		0.0667		0.0579		0.0813		0.1473		0.0060		0.4863
2021		0.0726		0.0012		0.0133		0.3816		0.0373		0.0076

Resultados por tipo de institución

	Puntaje global	Razonamiento Cuantitativo	Lectura Crítica	Competencias Ciudadanas	Inglés	Comunicación Escrita
--	----------------	---------------------------	-----------------	-------------------------	--------	----------------------

	Inst. Univ.	Universida d										
ATET												
2017	0.0529*	0.0061	-0.0070	-0.0007	0.0812**	0.0684***	0.0888**	0.0555***	0.0230	-0.0165	0.0323	-0.0298
2018	0.0508*	0.0124	0.0024	0.0182	0.0765**	0.0588***	0.0887**	0.0477**	0.0620**	0.0247	0.0025	-0.0138
2019	0.0774**	0.0218	0.0100	0.0211	0.1217**	0.1223***	0.0822**	0.0520***	0.0670*	0.0036	0.0398	-0.0270
2020	0.1290**	0.0424**	0.0845**	0.0738***	0.1074**	0.0424**	0.1177**	0.0457**	0.0319	-0.0074	0.1281**	0.0594**
2021	0.1807**	0.0683***	0.0949**	0.0686***	0.1245**	0.0677***	0.1809**	0.0778***	0.0359	-0.0367	0.1930**	0.1225***
Wal d												
2017	0.1598		0.0039		0.0180		0.1176		0.0696		0.6596	
2018	0.1192		0.0187		0.0331		0.2155		0.0686		0.0710	
2019	0.2326		0.0110		0.0000		0.0870		0.1756		1.2047	
2020	0.4842		0.0086		0.4051		0.4213		0.0668		0.6183	
2021	0.7134		0.0416		0.3004		0.7560		0.1285		0.5165	

Resultados por acreditación institucional

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	No acred.	Acreditada	No acred.	Acreditada	No acred.	Acreditada	No acred.	Acreditada	No acred.	Acreditada	No acred.	Acreditada
ATET												
2017	0.0132	0.0225	-0.0354*	0.0530**	0.0690**	0.0902***	0.0477**	0.0862***	0.0036	-0.0358	0.0148	-0.0424
2018	0.0264	0.0177	0.0056	0.0382*	0.0531**	0.0743***	0.0433*	0.0725**	0.0532**	-0.0132	-0.0015	0.0005
2019	0.0598**	0.0085	0.0211	0.0237	0.1318**	0.1240***	0.0593**	0.0607**	0.0452	-0.0316	0.0121	-0.0281
2020	0.0798**	0.0478*	0.0713**	0.0867***	0.0712**	0.0518*	0.0703**	0.0542*	0.0212	-0.0376	0.0949**	0.0670*
2021	0.1252**	0.0749**	0.0917**	0.0583**	0.1041**	0.0843***	0.1405**	0.0723**	0.0106	-0.0475	0.1183**	0.1460***
Wal d												
2017	0.0051		0.5807		0.0310		0.1325		0.0584		0.4989	
2018	0.0043		0.0602		0.0284		0.0846		0.1582		0.0007	
2019	0.1220		0.0003		0.0032		0.0001		0.1474		0.2436	
2020	0.0445		0.0111		0.0212		0.0153		0.0785		0.0613	
2021	0.0953		0.0409		0.0180		0.2293		0.0577		0.0539	

Resultados por acreditación del programa

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	No acred.	Acreditado	No acred.	Acreditado	No acred.	Acreditado	No acred.	Acreditado	No acred.	Acreditado	No acred.	Acreditado
ATET												
2017	0.0052	0.0364	-0.0347*	0.0644***	0.0622**	0.0867***	0.0541**	0.0911***	-0.0154	-0.0346	0.0074	-0.0420
2018	0.0281	-0.0007	0.0021	0.0516**	0.0722**	0.0358	0.0717**	0.0363	0.0423*	-0.0269	-0.0068	-0.0112
2019	0.0407*	0.0082	0.0044	0.0396*	0.1219**	0.1194***	0.0615**	0.0531*	0.0435	-0.0541	-0.0060	-0.0179
2020	0.0798**	0.0306	0.0625**	0.0835***	0.0746**	0.0285	0.0676**	0.0365	0.0377	-0.0531	0.0921**	0.0538
2021	0.1045**	0.0679**	0.0750**	0.0798***	0.0881**	0.0688**	0.1078**	0.0736**	-0.0019	-0.0585	0.1277**	0.1315***
Wal d												
2017	0.0589		0.6542		0.0424		0.1152		0.0148		0.4146	
2018	0.0460		0.1266		0.0829		0.1128		0.1886		0.0035	
2019	0.0661		0.0763		0.0004		0.0061		0.3182		0.0311	
2020	0.1473		0.0294		0.1594		0.0743		0.2781		0.1608	
2021	0.0754		0.0012		0.0244		0.0894		0.0795		0.0014	

Resultados por modalidad del programa

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	Distancia	Presencial	Distancia	Presencial	Distancia	Presencial	Distancia	Presencial	Distancia	Presencial	Distancia	Presencial
ATET												
2017	0.0121	0.0182	-0.0462	0.0095	0.0617*	0.0740***	0.0539*	0.0648***	0.0193	-0.0154	-0.0076	-0.0156
2018	0.0160	0.0228	-0.0310	0.0258*	0.0706**	0.0592***	0.0718**	0.0539***	0.0462*	0.0290	-0.0271	-0.0007
2019	0.0245	0.0362**	0.0058	0.0228	0.1252**	0.1225***	0.0315	0.0645***	0.0158	0.0154	-0.0152	-0.0082

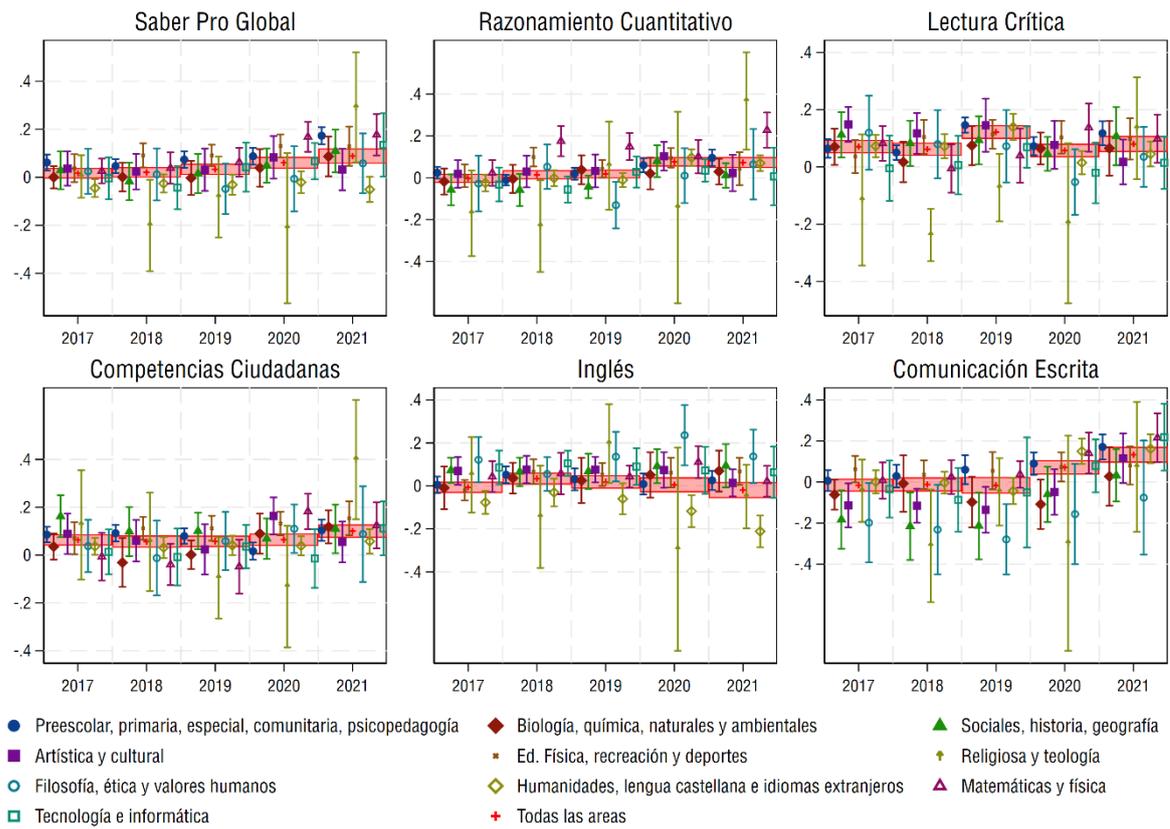
2020	0.0978**	0.0516***	0.0605*	0.0795***	0.1208** *	0.0433***	0.1009**	0.0534***	0.0342	-0.0049	0.0647	0.0739***
2021	0.1708** *	0.0718***	0.1067** *	0.0668***	0.1547** *	0.0652***	0.1667** *	0.0853***	0.0214	-0.0334	0.1477** *	0.1324***
Wal d												
2017	0.0007		0.0486		0.0031		0.0023		0.0195		0.0031	
2018	0.0007		0.0294		0.0021		0.0049		0.0077		0.0285	
2019	0.0018		0.0025		0.0001		0.0140		0.0000		0.0030	
2020	0.0305		0.0039		0.0929		0.0275		0.0197		0.0026	
2021	0.1357		0.0113		0.1218		0.0869		0.0273		0.0069	

Resultados por programa que presenta estudiantes de 2012 a 2021

	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita	
	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí
	ATET											
2017	0.0170	0.0164	-0.0034	-0.0018	0.0766** *	0.0676***	0.0596**	0.0651***	0.0075	-0.0127	-0.0174	-0.0146
2018	0.0295	0.0171	0.0286	0.0081	0.0756** *	0.0553***	0.0774** *	0.0473**	0.0595**	0.0234	-0.0309	-0.0019
2019	0.0509*	0.0255	0.0332	0.0124	0.1246** *	0.1203***	0.0559**	0.0581***	0.0489*	0.0072	-0.0042	-0.0231
2020	0.0719**	0.0552***	0.0894** *	0.0704***	0.0917** *	0.0440**	0.1040** *	0.0495***	0.0036	0.0066	0.0748*	0.0698***
2021	0.1070** *	0.0802***	0.0927** *	0.0640***	0.0964** *	0.0734***	0.1296** *	0.0887***	-0.0063	-0.0245	0.1411** *	0.1309***
Wal d												
2017	0.0000		0.0001		0.0024		0.0013		0.0048		0.0006	
2018	0.0035		0.0133		0.0110		0.0419		0.0157		0.0692	
2019	0.0151		0.0170		0.0006		0.0002		0.0191		0.0371	
2020	0.0066		0.0142		0.0750		0.0972		0.0001		0.0012	
2021	0.0185		0.0300		0.0178		0.0706		0.0027		0.0045	

Nota: * p < 0.05, ** p < 0.01, *** p < 0.001. ATET ajustado por covariables, efectos de grupo y año de presentación del examen. Errores estándar ajustados por programa.

Anexo 9. Resultados de DiD por área de la licenciatura – todas las áreas



Anexo 10. Resultados de DiD por área de la licenciatura – diferencias no significativas

Biología, química, naturales y ambientales													
	Puntaje global		Razonamiento Cuantitativo		Lectura Crítica		Competencias Ciudadanas		Inglés		Comunicación Escrita		
	ATET	Wald	ATET	Wald	ATET	Wald	ATET	Wald	ATET	Wald	ATET	Wald	
2017	0.0001	0.0325	-0.0177	0.0327	0.0705*	0.0000	0.0355	0.1256	-0.0090	0.0007	-0.0611	0.6046	
2018	0.0008	0.0649	-0.0051	0.0611	0.0173	0.3945	-0.0309	1.9406	0.0362	0.0005	-0.0074	0.0087	
2019	-0.0030	0.1736	0.0385	0.0625	0.0749*	0.3820	0.0019	0.6319	0.0252	0.0040	-0.0971	2.5355	
2020	0.0390	0.0530	0.0203	0.4957	0.0646*	0.0078	0.0897*	0.0953	0.0513	0.2100	-0.1082	3.9249*	
2021	0.0869*	0.0004	0.0293	0.3390	0.0651	0.0495	0.1180** *	0.0573	0.0695	0.7512	0.0273	1.9243	
Sociales, historia, geografía													
2017	0.0289	0.0186	-0.0574	0.3103	0.1122**	0.1908	0.1625** *	1.0348	0.0732*	0.764	-0.1829*	3.7707	
2018	-0.0175	0.2104	-0.0578	0.5599	0.0827*	0.0527	0.0982	0.2167	0.0655	0.123	-0.2154*	6.2896*	
2019	0.0159	0.042	-0.0434	0.3852	0.0934*	0.0903	0.1008**	0.2176	0.0677	0.2656	-0.2135*	5.2654*	
2020	0.0488	0.0205	0.0792*	0.001	0.0456	0.0176	0.0687	0.0021	0.0915*	0.7952	-0.06	2.044	
2021	0.1092*	0.064	0.0112	0.3846	0.1075*	0.0928	0.1098*	0.0099	0.0935	0.9529	0.0323	1.1752	
Artística y cultural													
2017	0.0362	0.068	0.0186	0.0712	0.1486** *	1.272	0.0891*	0.1529	0.0696*	0.6653	-0.1134*	4.4285*	
2018	0.0236	0.0015	0.0298	0.0314	0.1173**	0.5404	0.0602	0.0025	0.0764*	0.2234	-	6.4263*	
2019	0.0316	0.0005	0.0328	0.0335	0.1457**	0.1034	0.024	0.2586	0.0758*	0.3808	-0.1351*	8.3380* *	
2020	0.0829	0.0814	0.1027**	0.0855	0.0767	0.0681	0.1636** *	1.8305	0.0742	0.4876	-0.0487	3.3477	
2021	0.0324	0.6211	0.0231	0.3037	0.0177	0.7721	0.0549	0.4851	0.0141	0.1085	0.116	0.0912	
Ed. Física, recreación y deportes													
2017	0.0379	0.0561	0.0092	0.0181	0.0367	0.1997	0.0711*	0.0098	-0.0198	0.0240	0.0629	2.5949	
2018	0.0914** *	0.5784	0.0992** *	0.8075	0.1049** *	0.2729	0.1125** *	0.4963	0.0664*	0.1394	0.0363	0.9904	
2019	0.0913** *	0.4370	0.0426*	0.0682	0.1141** *	0.0105	0.1123** *	0.4386	0.0529	0.1458	0.0551	3.5162	
2020	0.1304** *	0.6436	0.0952** *	0.0438	0.1036** *	0.4015	0.1321** *	0.6445	0.0619	0.3757	0.0737*	0.0022	
2021	0.1296**	0.1981	0.0438	0.0891	0.0980**	0.0474	0.1545** *	0.3614	0.0678*	0.7471	0.0806	0.7973	
Religiosa y teología													
2017	0.0031	0.0171	-0.1699	0.5552	-0.1156	1.6368	0.1266	0.2834	0.0516	0.0719	-0.0425	0.0243	
2018	-0.2005*	2.4442	-0.2309*	3.4783	0.2382** *	4.4464*	0.0557	0.0001	-0.1439	0.6469	-0.3108*	2.1683	
2019	-0.082	1.2632	0.0577	0.0579	-0.0721	2.5889	-0.0953	1.6063	0.1968*	0.7746	-0.0476	0.0472	
2020	-0.2118	4.5631*	-0.142	1.5113	-0.1967	2.2565	-0.1315	3.4286	-0.2943	2.8259	-0.2978	1.1723	
2021	0.2911*	4.3355*	0.3665**	1.7846	0.1354	0.1183	0.3996** *	11.3319**	-0.0494	0.0253	0.0741	0.0353	
Filosofía, ética y valores humanos													
2017	0.0247	0.0069	-0.0277	0.0722	0.1193	0.1293	0.0380	0.0861	0.1216*	1.6610	-0.1981*	1.7691	
2018	0.0118	0.0068	0.0526	0.1854	0.0795	0.0174	-0.0121	0.5133	0.0552	0.0569	-0.2318*	2.0074	
2019	-0.0485	0.5438	-0.1307*	3.0243	0.0724	0.1244	0.0589	0.0001	0.1357*	1.0819	-	3.5632	
2020	-0.0064	0.3942	0.0102	0.5212	-0.0526	0.6700	0.1103*	0.2284	0.2360** *	4.8716 *	-0.1562	1.7209	
2021	0.0580	0.0726	0.0644	0.0066	0.0354	0.0753	0.0880	0.0137	0.1374*	1.8555	-0.0760	1.0435	
Humanidades, lengua castellana e idiomas extranjeros													
2017	-0.0449*	0.7430	-0.0248	0.1238	0.0730** *	0.0011	0.0369*	0.2236	-	0.0765**	0.5236	0.0011	0.1666
2018	-0.0259	0.4411	-0.0018	0.0568	0.0722** *	0.0260	0.0315	0.2460	-0.0302	0.4110	-0.0032	0.0444	
2019	-0.0308	0.9095	-0.0107	0.2309	0.1395** *	0.0865	0.0407*	0.1063	-0.0598	0.5581	-0.0438	0.4894	
2020	-0.0207	1.4875	0.0964** *	0.1046	0.0149	0.5756	0.0388	0.2340	-	0.1171**	1.1372	0.1497** *	2.9059

2021	-0.0505	3.5363	0.0699** *	0.0018	0.0452*	0.3676	0.0580*	0.6412	-	0.2114** *	1.6839	0.1620** *	0.3153
Matemáticas y física													
2017	0.0264	0.016	0.0232	0.0868	0.0822**	0.0256	-0.0064	0.8432	0.0434	0.3375	0.0074	0.247	
2018	0.0383	0.0633	0.1753** *	3.9340*	-0.0056	0.9753	-0.039	2.3262	0.0577	0.0741	0.0245	0.5026	
2019	0.0626*	0.1706	0.1494** *	2.5655	0.0396	1.3316	-0.048	1.9595	0.0432	0.0814	0.0373	1.5626	
2020	0.1683** *	2.5294	0.1010*	0.0933	0.1373**	1.6781	0.1826** *	2.7841	0.1107**	1.4462	0.1408**	1.2229	
2021	0.1770** *	1.8179	0.2279** *	3.434*	0.0985*	0.0883	0.1244*	0.1337	0.0216	0.1881	0.2159** *	2.1271	
Tecnología e informática													
2017	-0.0040	0.0383	-0.0319	0.0698	-0.0044	0.7915	0.0138	0.2198	0.0860*	0.7005	-0.0343	0.0883	
2018	-0.0429	0.4673	-0.0561	0.4110	0.0064	0.5148	-0.0079	0.8791	0.1056** *	0.4421	-0.0863	1.9768	
2019	0.0424	0.0137	0.0279	0.0079	0.0694	0.4769	0.0359	0.0839	0.0902*	0.5962	-0.0507	0.3879	
2020	0.0669	0.0068	0.0348	0.1931	-0.0206	1.2964	-0.0141	1.2625	0.0723	0.3867	0.0794	0.0190	
2021	0.1348*	0.4336	0.0066	0.3566	0.0147	1.0420	0.1120	0.0337	0.0636	0.7149	0.2186**	1.2711	
Preescolar, primaria, especial, comunitaria, psicopedagogía													
2017	0.0615** *	0.2973	0.0242	0.0928	0.0642** *	0.0107	0.0859** *	0.1001	0.0052	0.0256	0.0067	0.3633	
2018	0.0469**	0.1375	-0.0113	0.1069	0.0505** *	0.0355	0.0917** *	0.3124	0.0523**	0.0658	0.0291	1.2010	
2019	0.0739** *	0.2683	0.0197	0.0003	0.1461** *	0.1147	0.0798** *	0.0997	0.0302	0.0226	0.0599*	4.0719*	
2020	0.0876** *	0.1461	0.0602** *	0.0449	0.0726** *	0.0593	0.0166	0.7186	0.0089	0.0023	0.0899**	0.1754	
2021	0.1729** *	0.9773	0.0952** *	0.0580	0.1176** *	0.2781	0.1043** *	0.0032	0.0260	0.2494	0.1709** *	0.5898	

Nota: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATET ajustado por covariables, efectos de grupo y año de presentación del examen. Errores estándar ajustados por programa.