Programas Ondas y Jóvenes Investigadores e Innovadores

Informe 4: Resultados de la evaluación de impacto del programa Ondas y Jóvenes investigadores

Mayo de 2024

Contenido

Abreviaturas 4

Programa Ondas 5

1. Introducción 5

2. El programa Ondas 6

2.1. Descripción del programa 6

2.2. Contexto normativo del Programa Ondas 6

3. Características de los NNA vinculados a Ondas 8

3.1. Participación y cobertura del programa 9

3.2. Características demográficas de los NNA vinculados a ONDAS 11

3.3. NNA vinculados a Ondas certificados en procesos de fortalecimiento de capacidades en investigación y creación 12

3.4. Grupos de investigación de NNA vinculados a Ondas certificados 13

3.5. Deserción 16

3.6. Resultados en pruebas Saber 18

3.6.1. Resultados en pruebas Saber individuales 18

3.6.2. Resultados en pruebas Saber a nivel de colegio 20

3.7. Graduación de la educación superior 22

4. Estrategia para la evaluación de impacto del programa ONDAS 26

4.1. Grupos de tratamiento y control con información individual de los NNA 26

4.2. Grupos de tratamiento y control con información a nivel de establecimiento educativo 27

4.3. Fuentes de información y variables de resultado 27

4.4. Variables de resultado a nivel de NNA 28

4.5. Variables de resultado a nivel de establecimiento educativo 29

4.6. Técnicas de estimación 29

4.6.1. Emparejamiento 29

4.6.2. Diferencia en diferencias 31

4.6.3. Limitaciones 33

5. Resultados de la evaluación a nivel de NNA 34

5.1. Descripción de la muestra de evaluación 34

5.2. Resultados 34

6. Resultados de la evaluación a nivel de colegios 38

6.1. Descripción de la muestra de evaluación 38

6.2. Resultados 41

7. Conclusiones 44

Programa Jóvenes Investigadores e Innovadores 45

1. Introducción 45
2. Contexto de la implementación del programa JII 46
3. Caracterización de los JII 47
3.1. Convocatoria y selección 48
3.2. Características de los JII seleccionados 48
3.2.1. Lugar de nacimiento 48
3.2.2. Características demográficas 52
3.3. Área de conocimiento 54
3.4. Nivel educativo de los JII 56
3.5. Productos de nuevo conocimiento de los JII 59
3.6. Salarios 61
3.7. Vinculación a grupos de investigación 62
4. Metodología para la evaluación de impacto 63
4.1. Variables de resultado 63
Producción de nuevo conocimiento. 63
Mercado laboral. 63
4.2. Grupos de tratamiento y control 63
4.3. Técnica de estimación 65
4.4. Comparación entre el grupo de tratamiento y control 67
5. Resultados 69
5.1. Impactos preliminares del programa en productos de nuevo conocimiento 69
5.2. Impactos preliminares del programa en variables de mercado laboral 75
6. Análisis costo-beneficio del programa Jóvenes Investigadores e Innovadores 78
6.1. Costos 78
6.2. Beneficios 80
6.3. Valor presente neto del flujo de beneficios y costos 83
7. Conclusiones 84
Bibliografía 86
Anexos – Programa Ondas 88
Anexo 1. Fuentes de información y limitaciones de la información para el programa Ondas 88
A1.1. Base de datos nacional 88
A1.4. Cruces de información a nivel de NNA 92
A1.3. Datos a nivel de colegios 96
Anexo 2. Resultados adicionales de la evaluación de impacto del programa Ondas 99
Anexos – Programa Jóvenes Investigadores e Innovadores 103
Anexo 3. Compilación de bases de datos y limitaciones de la información 103

A3.1. Base de datos de JII y aspirantes 103
A3.2. Cruces de información con la base de datos de inscritos 103
Anexo 4. Resultados adicionales del Programa JII 106

Abreviaturas

Abreviatura	Significado
ACB	Análisis Costo-Beneficio
DID	Diferencia en Diferencias
GEIH	Gran Encuesta Integrada de Hogares
IBC	Ingreso Base de Cotización
ICFES	Instituto Colombiano para la Evaluación de la Educación
JII	Jóvenes Investigadores e Innovadores
MEN	Ministerio de Educación Nacional
MinCiencias	Ministerio de Ciencia, Tecnología e Innovación
NARP	Población Negra, Afrocolombiana, Raizal y Palenquera
NNA	Niños, Niñas y Adolescentes
PSM	Propensity Score Matching
SIMAT	Sistema de Matrículas Estudiantil
SINEB	Sistema de Información Nacional de Educación Básica y Media
SNCTeI	Sistema Nacional de Ciencia, Tecnología e Innovación
SNIES	Sistema Nacional de Información de la Educación Superior
STEM	Science, Technology, Engineering and Mathematics (Ciencia, tecnología, ingeniería y matemáticas)
VPN	Valor Presente Neto

Programa Ondas

1. Introducción

El programa Ondas inició su operación en 2001 con el objetivo de promover el interés de niños, niñas y adolescentes (NNA) por la investigación, así como para impulsar el desarrollo de actitudes y habilidades que les permitan participar activamente en la cultura de la ciencia, la tecnología y la innovación (MinCiencias, 2020). A lo largo de su historia, el Programa Ondas ha consolidado una propuesta pedagógica y metodológica que permite a los NNA preguntar, experimentar, verificar y explorar nuevas formas de comprender el mundo a través de la investigación y con la mediación de sus maestros, la interacción con profesionales expertos en diferentes áreas de conocimiento y el diálogo con su comunidad (MinCiencias, 2020).

En esencia, durante la implementación del programa Ondas, los NNA de establecimientos educativos de educación básica y media conforman grupos de investigación acompañados por un docente coinvestigador y desarrollan ideas de investigación a partir de preguntas que surjan de sus intereses o de su entorno inmediato. Este documento presenta los resultados de la evaluación de impacto del programa Ondas. Para ello, se compiló la información de los NNA que han estado vinculados al programa entre los años 2001 y 2023. Se emplearon técnicas de emparejamiento para estimar los efectos del programa en variables como deserción y repitencia. También se empleó información agregada de las pruebas Saber a nivel de establecimiento educativo para determinar los impactos en los resultados en pruebas estandarizadas, como una aproximación a los impactos del programa en la calidad de la educación. Así mismo, se incluyeron los resultados del levantamiento de información cualitativa que se realizó a partir de entrevistas y grupos focales a actores vinculados al programa, como docentes, estudiantes, padres de familia, entidades coordinadoras que operan el programa en los departamentos y funcionarios de instancias departamentales y municipales que tuvieran relación con el programa. Esta información cualitativa se empleó para tener una comprensión más profunda de los resultados cuantitativos de la evaluación y enriquecer su interpretación.

Los resultados de la evaluación muestran que el programa Ondas tiene impactos en reducción de la deserción escolar y de la repitencia de los NNA vinculados al programa Ondas. Así mismo, se encuentran resultados positivos en pruebas Saber de matemáticas de quinto grado en los establecimientos educativos que han participado del Programa Ondas. En la siguiente sección de este documento se presenta una descripción del programa y luego se hace una caracterización del programa y de los NNA vinculados a Ondas para tener un mayor contexto sobre la intervención. En la cuarta sección se presenta la estrategia para la evaluación de impacto, junto con la definición de los grupos de tratamiento y control y en las siguientes secciones se describen los resultados. Por último, se hace una discusión sobre los resultados.

2. El programa Ondas

2.1. Descripción del programa[1]

El Programa Ondas tiene el propósito de fomentar la vocación y el entusiasmo por la investigación en niños, niñas y adolescentes (NNA), así como promover la indagación científica y tecnológica desde educación básica y media en Colombia. MinCiencias coordina el Programa Ondas a nivel nacional y da los lineamientos metodológicos, pedagógicos y operativos del programa, al tiempo que realiza el seguimiento a su ejecución técnica y financiera. En algunos casos, MinCiencias también financia el programa, si bien la financiación también puede provenir del Sistema General de Regalías en los departamentos. El programa Ondas lo implementa en cada departamento una Entidad Coordinadora que son instituciones de educación superior, centros de investigación y desarrollo tecnológico, centros de innovación o similares con presencia en el territorio. En cada departamento también existe un Comité Departamental del Programa Ondas que está integrado por MinCiencias, aliados estratégicos locales—como secretarías de educación y planeación, cajas de compensación familiar, centros de investigación, entre otros—quienes aportan a la implementación del programa, convocatoria y selección de grupos de investigación, brindar acompañamiento y diseñar y promover estrategias para el fortalecimiento del Programa Ondas en el departamento.

El Programa Ondas se desarrolla en cuatro fases, junto con actividades trasversales de sistematización, evaluación y administración. En una primera fase se realiza el alistamiento, que incluye todas las actividades administrativas y la conformación de los equipos territoriales para la implementación del Programa Ondas. Estos equipos territoriales los selecciona la Entidad Coordinadora y se componen de personal pedagógico, administrativo y de investigación para la ejecución de las actividades propias del Programa Ondas; este equipo es capacitado en todo lo referente a los principios pedagógicos y metodológicos del programa Ondas, así como en temáticas relacionadas con la ruta de investigación que se propone para el desarrollo de los proyectos de investigación de los NNA. Posteriormente, la entidad territorial hace la invitación a las sedes educativas para hacer parte del programa y para que la comunidad educativa conforme grupos de investigación, los cuales pueden estar conformados por entre 20 y 30 estudiantes acompañados por un maestro coinvestigador.

En la segunda fase de la implementación del Programa Ondas se desarrollan los proyectos de investigación. Durante esta etapa se realiza un acompañamiento a los grupos de investigación a través de asesorías metodológicas y se realizan procesos de capacitación y fortalecimiento de capacidades a los maestros que acompañan el proceso. En la tercera fase se realiza la socialización y comunicación de los avances y resultados de los proyectos de investigación que desarrollan los NNA. Por último, en la cuarta fase se hace el seguimiento, cierro y sistematización de la implementación del programa por parte de las Entidades Coordinadoras.

2.2. Contexto normativo del Programa Ondas

El Programa Ondas tiene el objetivo de promover en niños, niñas y adolescentes el interés por la investigación y el desarrollo de actitudes y habilidades que les permitan insertarse activamente en una cultura de la Ciencia, la Tecnología y la Innovación (CTeI). Para este fin, se financian proyectos de investigación en los establecimientos de educación básica y media y se brinda asesoría a maestros y grupos de investigación, creándose así habilidades en los NNA y en sus entornos y fortaleciendo la investigación y la apropiación de la CTeI desde edades tempranas.

[1] Esta sección se basa en el documento Lineamientos para fomentar la vocación científica en niños, niñas y adolescentes de MinCiencias.

El Programa Ondas fue creado en 2001 y desde entonces la normatividad en CTeI ha destacado la relevancia del programa como un motor del desarrollo de vocaciones científicas desde la educación básica y media. En 2008 se formuló la Política Nacional de Fomento a la Investigación y la Innovación, titulada *Colombia construye y siembra futuro,* que buscaba crear condiciones para que el conocimiento sea un instrumento del desarrollo. En dicha política se destacó la importancia del programa como medida para la creación de vocaciones científicas en niños y niñas, así como la necesidad de consolidarlo, profundizarlo y ampliarlo (Colciencias, 2008). Un año después, la Ley 1286 de 2009 -que derogó parcialmente la Ley 29 de 1990- transformó a Colciencias (hoy Minciencias) en Departamento Administrativo y estableció unos lineamientos generales para fortalecer el SNCTeI. Entre otras disposiciones, se definió como uno de los objetivos del nuevo Departamento Administrativo el fortalecimiento de la capacidad científica del país y promover la formación de investigadores, haciendo especial énfasis en programas que impactaran sectores estratégicos para la transformación del país.

En ese mismo año se publicó el documento CONPES 3582 de 2009 con la Política Nacional de Ciencia, Tecnología e Innovación, el cual tiene entre sus estrategias principales el fortalecimiento del recurso humano para la investigación y la innovación. Una de las limitaciones identificadas en este documento es el insuficiente recurso humano para la investigación y la innovación, considerando como uno de sus factores asociados la falta de desarrollo de competencias científicas en el sistema educativo. Si bien, se reconoce el aporte del programa Ondas en la formación de vocaciones científicas y tecnológicas en NNA; en este documento también se solicitó al MEN y a Colciencias – hoy MinCiencias el incremento de la cobertura nacional del programa (Consejo Nacional de Política Económica y Social, 2009).

Un hito importante para el programa Ondas fue la creación del Fondo de Ciencia, Tecnología e Innovación (FCTeI) al cual se le asignó el 10% de los recursos generados por el Sistema General de Regalías (SGR), según Acto Legislativo 005 de 2011. La creación del FCTeI constituyó un hito para la financiación de proyectos para el desarrollo de capacidades regionales en CTeI y significó una significativa fuente de recursos para la implementación del programa Ondas en los departamentos. Entre 2012 y 2018 se financiaron 32 proyectos del Programa Ondas, por un total de $2.111 billones – equivalentes al 18,3% del monto total de recursos del SGR para CTeI (Vicepresidencia de la República, Ministerio de Ciencia, Tecnología e Innovación, 2020).

Ahora bien, no toda la normatividad asociada al programa Ondas se deriva del sector de CTeI. Por el lado del sistema educativo, en el Plan Decenal de Educación 2016-2026 estableció como uno de sus desafíos estratégicos fomentar la investigación que lleve a la generación de conocimiento en todos los niveles de la educación. Para este fin, resulta importante fortalecer los programas curriculares actuales dentro del país en términos de educación superior y garantizar ambientes propicios para impulsar una investigación de calidad en todos los niveles educativos; y el programa Ondas cumple un rol importante en la formación de vocaciones en investigación en la educación básica y media.

Posteriormente, en el Plan Nacional de Desarrollo 2018-2022 se destacó el bajo nivel de desarrollo de vocaciones científicas en la población en edades escolares, así como un bajo nivel de apropiación social de la CTeI. Por tanto, el Plan definió la necesidad de fomentar las capacidades, la mentalidad y la cultura para el sector y algunos objetivos como la articulación de iniciativas en instituciones de educación superior con otros niveles de formación (por ejemplo, básica y media) por medio de los programas Ondas y Jóvenes Investigadores e Innovadores.

En 2019 se creó el Ministerio de Ciencia, Tecnología e Innovación (MinCiencias) en busca del fortalecimiento del sector de CTeI para que sea un motor que impulse el desarrollo y el crecimiento

del país. Entre los objetivos del Ministerio, se destacan la formulación de la política en CTeI para el país o establecer las estrategias de transferencia y apropiación social de la CTeI. Dentro del Plan Estratégico Institucional y el Plan de Acción Institucional de MinCiencias se establece la importancia de fomentar una cultura alrededor de la CTeI para el fortalecimiento de las vocaciones científicas en los territorios y se indica la importancia del acercamiento de los NNA a la investigación y a la innovación por medio del programa Ondas.

En ese mismo año, el gobierno nacional convocó y conformó la Misión Internacional de Sabios de 2019, la cual ayudó a trazar una ruta para el avance de la CTeI en el país. En los documentos de la Misión se reconoce la importancia de los diferentes actores como el Estado, las instituciones educativas y las comunidades como responsables del desarrollo de un sistema nacional que permita el fortalecimiento de capacidades científicas en el país. Una de las propuestas de la Misión tiene que ver con la inserción social del conocimiento para el uso efectivo en los problemas del país y ahí se reconoce la necesidad de transformaciones curriculares en las instituciones educativas para insertar el concepto de ciencia desde la primera infancia, así como fomentar el relacionamiento entre investigadores y maestros y estudiantes de educación básica y media; para ello, resaltan la necesidad de consolidar y reforzar el modelo de la investigación como estrategia pedagógica del programa Ondas (Vicepresidencia de la República, Ministerio de Ciencia, Tecnología e Innovación, 2020).

Derivado de los lineamientos y estrategias definidas por la Misión Internacional de Sabios, se actualizó la Política Nacional de Ciencia, Tecnología e Innovación, a través del CONPES 4069 de 2021. En este documento se destacan los esfuerzos del programa Ondas para estimular la curiosidad científica de los NNA, pero se señala la falta de cobertura del programa en todos los departamentos del país, ya que solo se implementa en 9 de los 32 departamentos del país. La política también define una estrategia para incrementar las vocaciones científicas en la población infantil y juvenil, para lo cual designa al MinCiencias para que, con el apoyo del Ministerio de Educación Nacional, implementen una ruta estratégica para dinamizar las vocaciones de CTeI en el marco de los Proyectos Educativos Institucionales dentro de los colegios y escuelas (Consejo Nacional de Política Económica y Social , 2021).

3. Características de los NNA vinculados a Ondas

La principal fuente de información para caracterizar a los NNA que han estado vinculados a Ondas es la base de datos compilada a partir de los registros administrativos del Programa que MinCiencias entregó a Fedesarrollo en el marco de la evaluación. También se empleó la información de la base de datos de los NNA certificados en procesos de fortalecimiento de sus capacidades en investigación y creación, la cual tiene 34.180 registros entre 2019 y 2022 (los cuales hacen parte también de la base de datos nacional).[2] Esta base de datos registra a los NNA certificados, que han estado en el programa por un período mínimo de dos años, y cuenta con variables que detallan los proyectos de investigación y los grupos participantes.[3] Adicionalmente, se usó la información de matrícula para aproximadamente 51 mil NNA vinculados a Ondas entre 2001 y 2013 para caracterizarlos en temas de deserción escolar y repitencia, así como los datos de 927 NNA que estuvieron vinculados a Ondas entre 2001 y 2013 y que se graduaron de educación superior entre 2020 y 2022. En el Anexo 1 se detallan las fuentes de información.

[2] Esta base de datos cuenta con 8.500 niños en 2022 que todavía no aparecían como efectivamente certificados.
[3] Si bien muchas de las variables se completaron como “No disponible”.

3.1. Participación y cobertura del programa

La participación en el programa Ondas ha variado mucho a lo largo del período estudiado. En años como 2011, 2012 y 2022 se encontraron más de 40.000 NNA vinculados a Ondas, mientras que en años como 2017 se registraron menos de 100.[4] Si bien Ondas ha operado por más de dos décadas, su alcance es todavía pequeño y limitado frente a la población total que estudia en Colombia, que para 2022 equivalía a aproximadamente 9,7 millones de alumnos de educación preescolar, básica, media y ciclos lectivos (DANE, 2022).

Se destaca que el programa ha llegado a todos los departamentos del país en diferentes momentos durante sus más de dos décadas de operación (ver Tabla 1). El departamento de Bolívar concentra un alto porcentaje de NNA vinculados al programa entre 2001 y 2023, con el 18,9%, seguido por Sucre, con el 14,54%, y Arauca, con el 7,17%, lo que destaca una alta participación de la región Caribe en el programa. En contraste, departamentos como Vaupés y Vichada tienen una baja representación con 0,17% y 0,25%, respectivamente. Antioquia y Bogotá tienen una participación más bien moderada, con porcentajes del 2,43% y 1,75% respectivamente, lo que muestra que el programa tiende a enfocarse más en regiones menos avanzadas en términos de ciencia, tecnología e innovación.

Tabla 1. NNA vinculados a Ondas entre 2001 y 2023 según el departamento

Departamento	**NNA vinculados a Ondas**	**Porcentaje**
Amazonas	2.598	0,7
Antioquia	8.994	2,43
Arauca	26.602	7,17
Atlántico	9.597	2,59
Bogotá	6.488	1,75
Bolívar	70.133	18,91
Boyacá	3.597	0,97
Caldas	19.550	5,27
Caquetá	28.175	7,6
Casanare	1.900	0,51
Cauca	3.783	1,02
Cesar	6.561	1,77
Chocó	5.803	1,56
Córdoba	7.691	2,07
Cundinamarca	8.590	2,32
Guainía	1.579	0,43
Guaviare	3.124	0,84
Huila	20.998	5,66
La Guajira	1.042	0,28
Magdalena	1.931	0,52
Meta	3.553	1,0
Nariño	4.924	1,33
Norte de Santander	5.046	1,36
Putumayo	1.151	0,31
Quindío	4.252	1,15
Risaralda	5.275	1,4
San Andrés	2.541	0,69
Santander	26.550	7,16

[4] Ver más detalles en el Anexo 1.

Departamento	NNA vinculados a Ondas	Porcentaje
Sucre	53.938	14,54
Tolima	3.615	0,97
Valle del Cauca	19.743	5,32
Vaupés	648	0,17
Vichada	914	0,25

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas.

Ahora bien, el programa no opera todos los años en todos los departamentos. La información recogida revela que en 2021 se realizaron convocatorias en 10 departamentos: Antioquia, Arauca, Atlántico, Caldas, Caquetá, Cundinamarca, Guaviare, Huila, Sucre y Tolima de Sucre, Caldas, Arauca, Huila, Caquetá, Cundinamarca, Antioquia y Guaviare; mientras que en 2022 se realizaron en el doble departamentos (en 22), donde participaron, además, los departamentos de Bogotá, Bolívar, Boyacá, Cauca, Chocó, Meta, Nariño, Putumayo, Quindío, Risaralda, San Andrés y Valle del Cauca (ver Tabla 26 del Anexo). En línea con lo anterior, el peso de cada departamento en el número total de NNA vinculados al programa Ondas varía también cada año. Por ejemplo, Atlántico concentró el 94% de los NNA vinculados al programa Ondas en el país en el 2004, porcentaje que se redujo al 37,6% en 2005, al 34,2% en 2006, y a menos del 10% en años subsiguientes. Por su parte, Bolívar concentraba aproximadamente el 50% de los NNA vinculados al programa Ondas en 2013 y en 2018, pero tuvo años como 2010 y 2011 donde concentró entre el 34% y el 38% de los NNA vinculados al programa Ondas, o como 2019 donde tuvo el 9,6% del total de NNA vinculados al programa Ondas.

Según la información cualitativa recabada con entidades coordinadoras, la implementación de las convocatorias en los diferentes departamentos depende en gran medida de la disponibilidad de recursos y de la capacidad de las entidades coordinadoras para gestionarlos. En los últimos años, el Sistema General de Regalías se ha convertido en una fuente crucial de financiamiento, permitiendo a los departamentos ampliar la cobertura de Ondas y garantizar su continuidad.

"hay una diferencia frente a que si tú tienes recursos de regalías a no tener recursos de regalías, porque si tú tienes recursos de regalías, tú puedes estar en los 27 municipios que Caldas tiene [....], cuando hemos tenido los recursos de regalías, hemos garantizado estar en los 27 municipios, cuando ya son recursos o de la secretaría o los recursos nación que da el Ministerio ya nos toca restringirnos a 10-12 municipios porque los recursos no dan"

(Entrevista a coordinadora de entidad coordinadora, 2023)

Otro aspecto muy importante para destacar en cuanto a la cobertura del programa es que no solo ha llegado a todos los departamentos sin concentrarse en los que tienen mayor desarrollo en ciencia y tecnología, sino que ha conseguido llegar a zonas rurales y periféricas, donde existen rezagos educativos significativos y donde el desarrollo de actividades que fomenten la investigación cobra amplia relevancia.

"lo que uno percibe es como esa inclusión de la población, de las instituciones educativas rurales a los proyectos de vida, no solamente de los estudiantes, sino también de los docentes"

(Entrevista a funcionario de gobierno local, 2023)

"en el caso de las zonas rurales más que todo que es una parte en donde es difícil llegar, tienen difícil acceso y no cuentan con todos los recursos para esas zonas, la investigación se ha fortalecido,

porque los asesores hemos tratado de llegar hasta allá y enseñarles a los niños sobre investigación. Entonces, para la zona rural ha sido muy satisfactorio en ese aspecto".

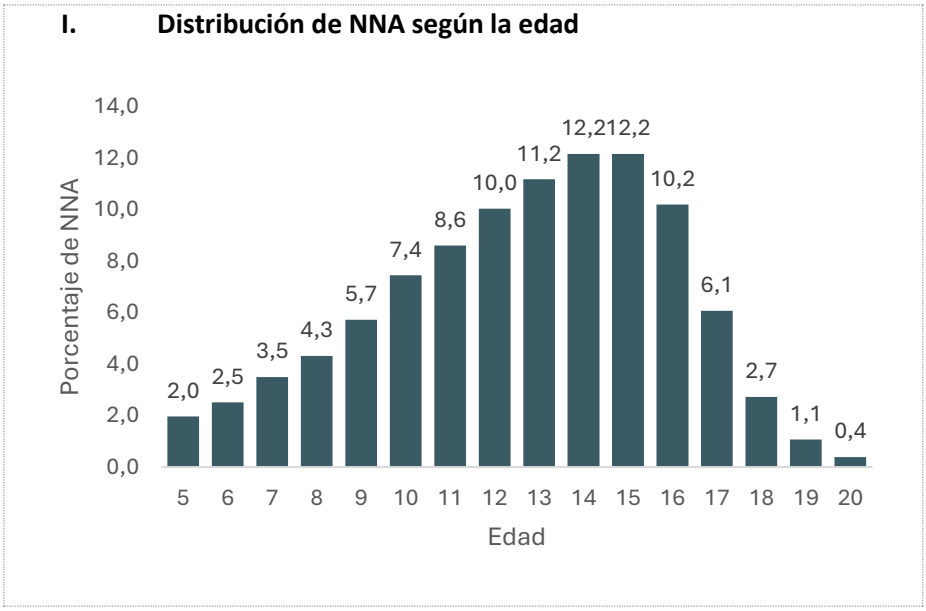
(Entrevista a asesora de línea de entidad coordinadora, 2023)

3.2. Características demográficas de los NNA vinculados a ONDAS

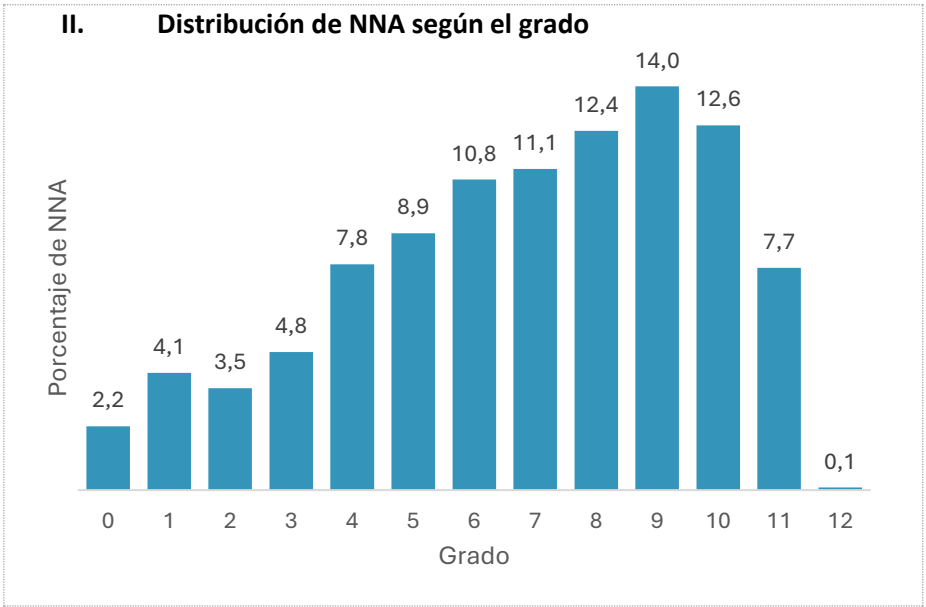
La edad promedio de los NNA vinculados al Programa Ondas es de 12,7 años y el 52,4% son niñas. El 55,7% de los NNA vinculados al Programa Ondas son adolescentes con edades entre 12 y 16 años (ver Figura 1, panel I). Se destaca que hay un porcentaje considerable de niños de 5 a 9 años, creciente del 2% al 5,7%.[5]

En cuanto al grado, el porcentaje de NNA que se vincula al Programa Ondas aumenta conforme aumenta el grado escolar, empezando en los niveles iniciales con un 2,2% en grados como transición, jardín o prejardín (categorizados como grado 0). Los grados escolares intermedios, específicamente del grado 6 al 9, muestran una concentración más alta de participación, lo cual es consistente con la mayor participación de NNA vinculados al Programa Ondas en ese rango etario. Esto pudiera reflejar una mayor integración y fomento de la curiosidad científica y la investigación en estas etapas educativas. Los porcentajes de participación más elevados se encuentran en los grados 9 y 10, con un 14,0% y 12,6% respectivamente.

Figura 1. Distribución de los NNA vinculados a Ondas según la edad o el grado



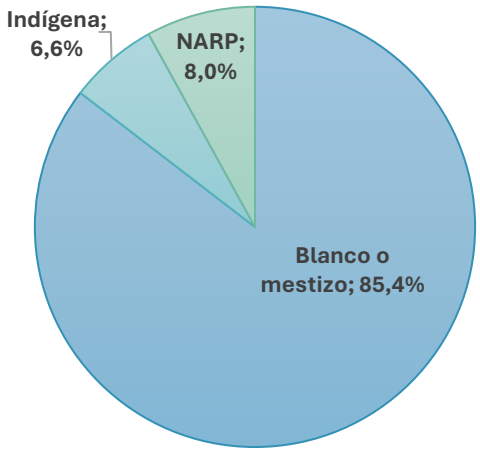
[5] Solo se toman edades entre los 5 y 20 años, hay edades por fuera de estos valores que son atípicas y que se deben probablemente a una mala digitación de la edad o la fecha de nacimiento.



Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas. Grado cero corresponde a caminadores, transición, jardín, prejardín.

Se cuenta con información disponible de reconocimiento étnico desde 2017. El 85,4% de los NNA vinculados al Programa Ondas se identifican como blancos o mestizos—un total de 82.247 NNA (ver Figura 2). La población indígena representa un porcentaje importante, del 6,6% en el programa (6.302 NNA) y el grupo NARP (Negros, Afrocolombianos, Raizales y Palenqueros) representa el 8,0% con 7.718 NNA.

Figura 2. Distribución de los NNA de Ondas según la etnia



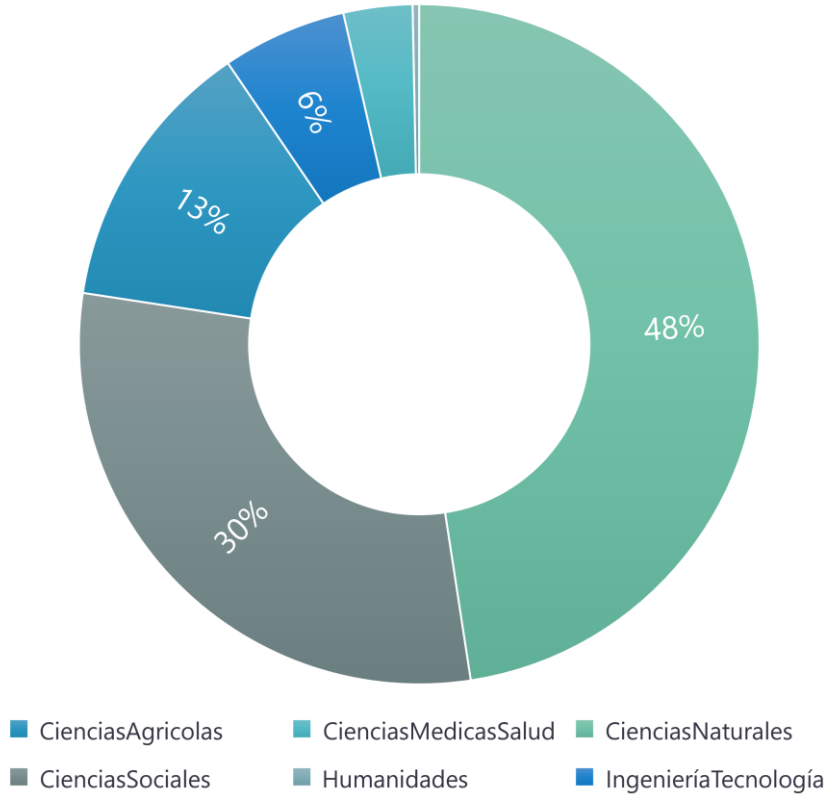
Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas.

3.3. NNA vinculados a Ondas certificados en procesos de fortalecimiento de capacidades en investigación y creación

Entre 2019 y 2021 se certificaron 25.689 NNA vinculados al Programa Ondas en procesos de fortalecimiento de capacidades en investigación y creación; el 76% de estos NNA se vinculó a Ondas a

través de un proyecto abierto.[6] El 48% de los NNA participó en un proyecto en el área de Ciencias Naturales, el 30% en uno de Ciencias Sociales y el 13% en uno de Ciencias Agrícolas. El 2021 fue el año en el que más NNA se certificaron en el programa, con casi 17.000 NNA. De ellos, el 69% participó en un proyecto de tipología abierto y casi la mitad participó en un proyecto alineados con la línea temática de Ciencias Naturales (ver Figura 3).

Figura 3. Distribución de los NNA certificados entre 2019 y 2021 según el área de conocimiento del proyecto en el que participaron



Fuente: elaboración propia con base en la base de datos de NNA certificados entre 2019 y 2021.

3.4. Grupos de investigación de NNA vinculados a Ondas certificados

Se cuenta con información de 926 grupos de investigación de NNA vinculados al Programa Ondas y certificados entre 2019 y 2021 (151 en 2019, 103 en 2020 y 672 en 2021). El número de estudiantes vinculados y certificados en el Programa Ondas por grupo de investigación fue de 31 en 2019 y de 46 en 2021 (ver Tabla 2). La cifra incrementó a 181 en 2020 por un dato atípico de un colegio y un grupo de investigación en el que se registran 1.299 niños. Se trata de un colegio público en Florencia-Caquetá cuya pregunta de investigación es "¿Como podemos contribuir desde el área técnica, a mejorar el bienestar de la población canina y felina presentes en la comunidad educativa?".

[6] Un total de 19.444 NNA de los 25.689 NNA vinculados al Programa Ondas y con información de certificación se vinculó a través de un proyecto abierto.

Tabla 2. Número de estudiantes vinculados y certificados en ONDAS por grupo de investigación (mediana)

Año	Número de estudiantes por grupo de investigación vinculado a ONDAS y certificado (mediana)
2019	31
2020	181
2021	46

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas y certificados.

De acuerdo con la información cualitativa, la conformación de los grupos de investigación depende en gran medida de la iniciativa de los docentes, quienes guían la dinámica del grupo al inicio. Luego, sobre la marcha los estudiantes mantienen el grupo en el tiempo y consolidan su dinámica, crean líneas de investigación e incluso semilleros a partir de sus intereses. Luego, los docentes juegan un rol de coinvestigadores, orientadores y guías, lo que posibilita la emergencia de experiencias de trabajo autónomo por parte de los estudiantes.

"inicialmente se trabaja intereses y motivaciones de los niños, que sean ellos mismos quienes decidan qué quieren investigar, qué quieren aprender. Se hace una lluvia de ideas a través de actividades de sensibilización. Por ejemplo, yo pongo centros de interés, en una mesa pongo por ejemplo animales, en otra pongo experimentos, en otra pongo maticas y trato que ellos vayan pasando por los diferentes centros de interés. A través de los videos también, entonces les pongo videos de diferentes temáticas. Y finalmente, pongo cinco temáticas que yo sé que a ellos les gusta, que les impacta y las sometemos a votación. Entonces se hace una votación democrática en donde ellos dicen cuál les gusta más"

(Entrevista a docente, 2023)

"ya desde la mitad del proceso hasta terminar, yo ya paso a un segundo plano. Entonces mi rol fue solamente orientar. Inclusive ni orientar, porque yo ya no daba las clases, sino que entre ellos se daba una dinámica de retroalimentación, y yo solamente era enrutar las conclusiones finales"

(Entrevista a docente, 2023)

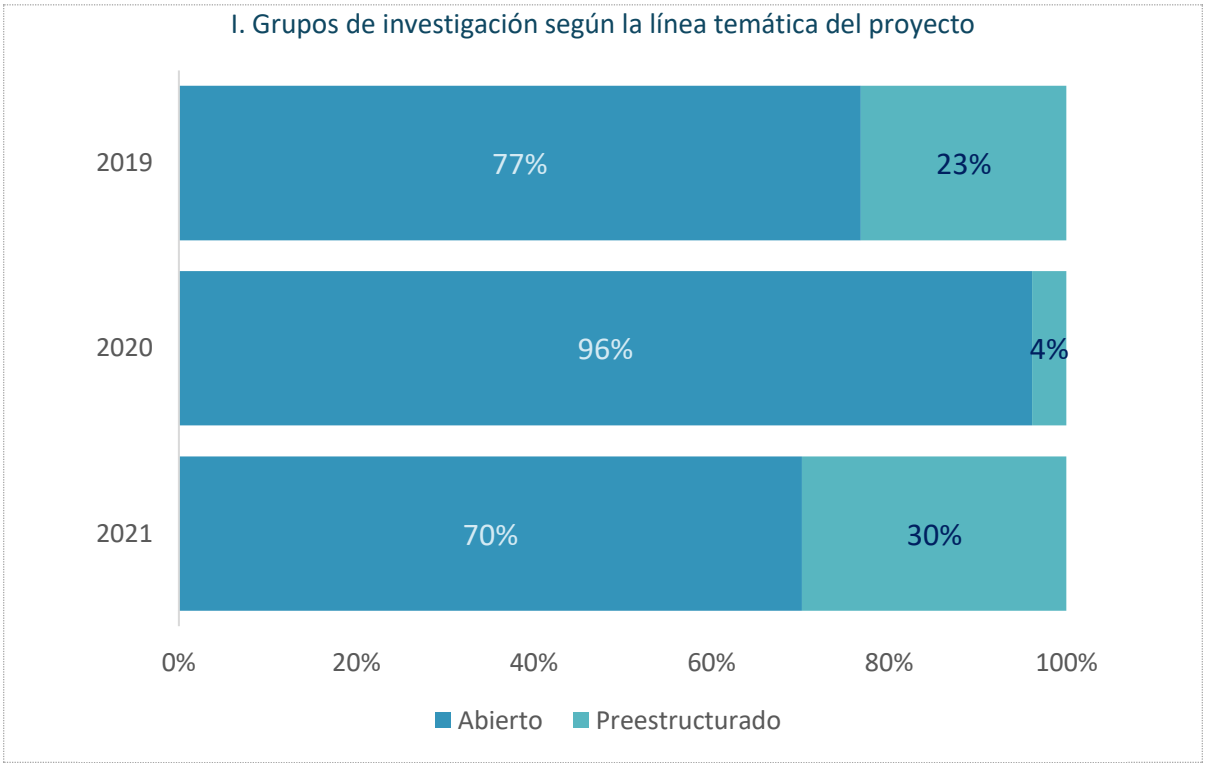
En cuanto a las áreas de conocimiento, la mayoría de los proyectos de investigación se realizaron en el área de ciencias naturales, seguido de las ciencias sociales (ver Figura 4, panel I). El número de proyectos en Ingeniería y Tecnología se incrementó en los tres años evaluados, de 10 en 2019, a 23 en 2020 y a 44 en 2021; si bien, el porcentaje de proyectos en esta área fue sustancialmente más alto en 2020 cuando representaron en 22% del total de proyectos si se compara con 2019 o 2021 cuando dicho porcentaje fue del 7% y 8%, respectivamente. En el trabajo cualitativo se identificó una gran diversidad de temas investigados, entre los que se encuentra la radiación producida por celulares, uso de plantas medicinales como alternativa a la medicina industrial, reutilización de papel, descontaminación de fuentes hídricas, microturbinas hidráulicas, reducción y reutilización de residuos, sistemas de riego automatizado, inclusión de población no oyente, o generación de conciencia ambiental a partir de trabajo comunitario, entre otros.

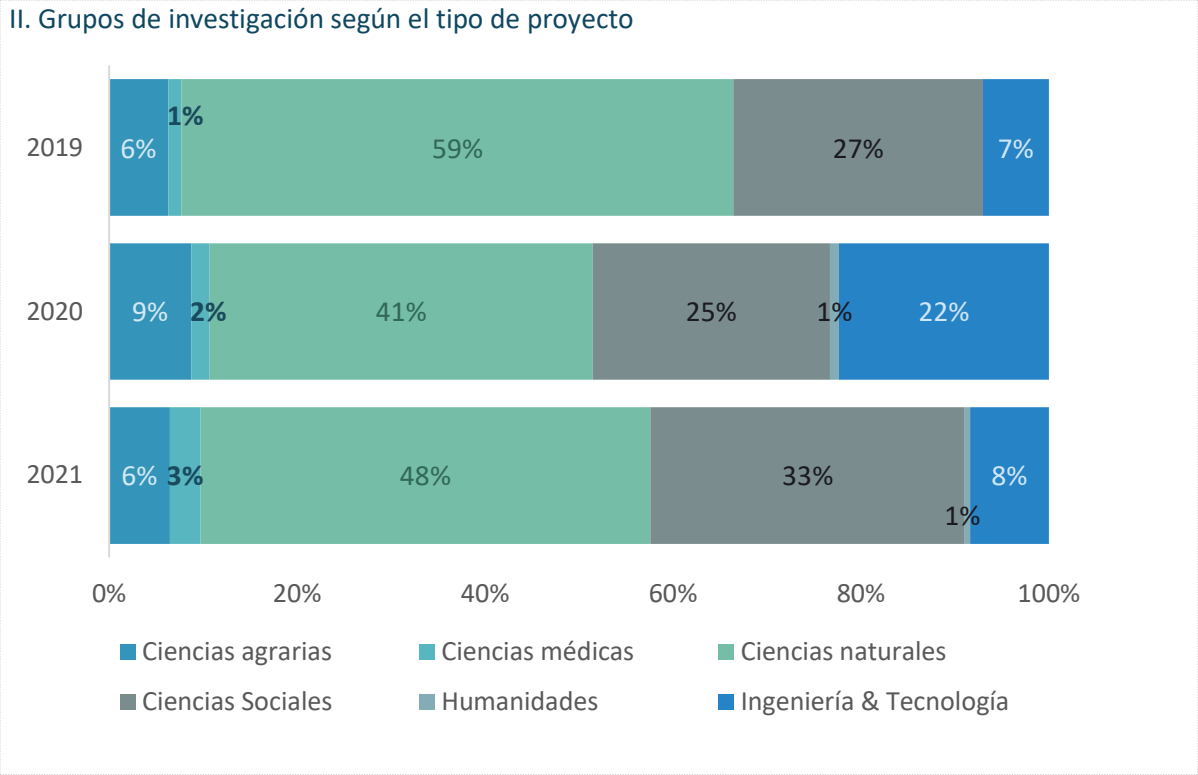
Los temas – y, por ende, las áreas—de los proyectos de investigación de los NNA suelen estar guiados por el docente que le dedica tiempo a las actividades del Programa Ondas y por su área de conocimiento. Sin embargo, se destaca que las preguntas puedan surgir de las experiencias y observaciones de los NNA del entorno en el que viven, alineando la experiencia en investigación con sus observaciones y experiencias diarias.

"La temática general la propongo yo, en este caso fue el agua, entonces nosotros tomamos tres ejes fundamentales, el aire, el suelo y el agua. Entonces dentro de ellos mismos, de los tres, seleccionan el agua. Entonces […] tengo una muchacha en especial que ella tenía una curiosidad, unas ganas de aprender desde hace muchos años el tema del agua, pero especialmente en la quebrada de donde sacan el agua para suministrar a la casa de ella, […] de ahí toman el agua. ¿Por qué? Porque ella vive en un sector […] de invasión. Y que a través del tiempo fruto de una negligencia administrativa, política y policial del Departamento, pues generó un desplazamiento, y todo ese proceso de desplazamiento, se focalizó acá en la ciudad capital, y se creó una invasión grandísima, […] se hizo el crecimiento vertical sobre la quebrada, […] y allá no les llega el servicio de agua potable, no tienen servicio de acueducto, no tienen servicio de alcantarillado"

(Entrevista a docente, 2023)

Figura 4. Grupos de investigación de Ondas según el tipo de proyecto y su línea temática





Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas y certificados.

Los grupos de investigación que participan en el Programa Ondas pueden presentar proyectos de dos tipos: abiertos y preestructurados. Los proyectos abiertos son propuestos y diseñados por los grupos de investigación formados por los NNA a partir de sus propios intereses y necesidades con la guía de un docente; entre tanto, los proyectos preestructurados ya tienen unas temáticas de interés nacional definidas por el Programa Ondas en colaboración con diferentes actores y entidades y cuentan con guías metodológicas establecidas (MinCiencias, 2020). La mayoría de los grupos de investigación participaron en el Programa Ondas a través de un proyecto abierto (ver Figura 4, panel I). En las entrevistas cualitativas se destacó la importancia de los proyectos preestructurados para los grupos de investigación que ingresan por primera vez al Programa Ondas y les sirve para ganar experiencia y ritmo de trabajo.

"generalmente usamos ese preestructurado para los grupos que son nuevos en Ondas. Entonces como el camino está ya escrito, solamente hay que seguirlo, entonces aprenden más fácilmente… o sea, no tienen que diseñar la pregunta, la pregunta ya está diseñada, entonces más bien lo que hacen es contextualizarla. No tienen que diseñar la metodología, ya existe la metodología, lo que tienen que hacer es aprenderla. Entonces es como menos traumático para la primera vez que el grupo se enfrente a la investigación, que lo haga a través de un proyecto pre estructurado, aprenden la dinámica, hacen el ejercicio"

(Entrevista a coordinadora de entidad coordinadora 2023)

3.5. Deserción

El 1,4% de los NNA vinculados al programa Ondas han desertado temporalmente de la educación básica y media, esto, entendido como que dejaron de aparecer en la base de datos de matrícula entre

2007 y 2015 y entre 2018 y 2022 y, entre esos años, no avanzaron de grado escolar.[7] Este porcentaje es ligeramente mayor en hombres que en mujeres (1,6 vs 1,2%); también es más alto en NNA indígenas (2,5%), NARP (1,4%) que para el resto de los NNA vinculados al Programa Ondas que no hacen parte de ningún grupo étnico (ver Tabla 3). Por su parte, el 10,5% de los NNA vinculados al Programa Ondas entre 2001 y 2013 repitieron al menos un grado escolar durante el tiempo en el que han estado matriculados en educación básica y media. Esta cifra también es más alta en hombres (12,6%) que en mujeres (8,5%) y en NNA indígenas y NARP (12,1% y 10,5%, respectivamente) que para el resto de los NNA.

Tabla 3. Deserción y repitencia de NNA vinculados a Ondas

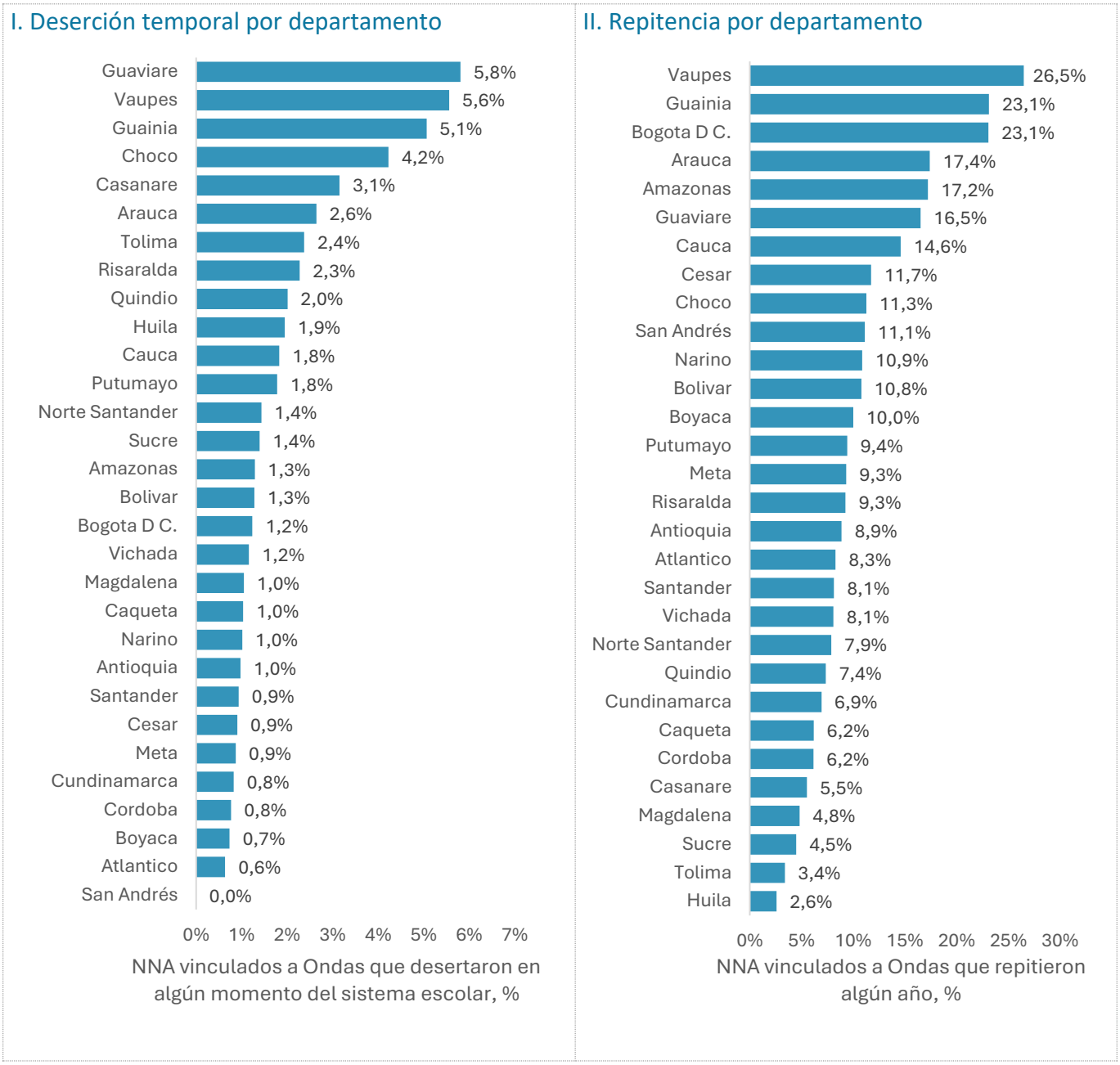
Variable	Porcentaje de NNA vinculados a Ondas	Sexo		Etnia		
		Mujeres	Hombres	Indígena	NARP	Ningún grupo étnico
Deserción temporal, %	1,4%	1,2%	1,6%	2,5%	1,4%	0,6%
Repitente, %	10,5%	8,5%	12,6%	12,1%	10,5%	9,1%

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas entre 2001 y 2013 y SIMAT (2007-2015;2018-2022).

Los NNA vinculados al Programa Ondas en departamentos como Guaviare, Vaupés y Guainía presentan las tasas más altas de deserción temporal, con porcentajes que varían entre 5,1% y 5,8% (ver Figura 5, panel I). Los NNA vinculados al Programa Ondas en departamentos como Chocó (4,2%), Casanare (3,1%), Arauca (2,6%) y Tolima (2,4%) presentan cifras de deserción temporal de un punto porcentual por encima del promedio nacional. Por su parte, en los departamentos de Vaupés, Guainía y Bogotá se observan las cifras más altas de repitencia para los NNA vinculados al Programa Ondas, de entre el 23,1% y el 26,5% (ver Figura 5, panel II).

[7] No se cuenta con una variable en SIMAT que clasifique a un joven como desertor, por lo que se establece esta aproximación. Por ejemplo, si un NNA aparece en octavo grado en SIMAT en 2018 y luego vuelve a aparecer en 2020 en noveno grado se asume como desertor temporal en la medida en que desapareció de la base de datos y no avanzó hasta décimo grado como hubiese debido en ausencia de deserción. La base de datos de NNA vinculados al Programa Ondas y con información de SIMAT cuenta con 51.014 observaciones, de ahí que el 1,4% de NNA vinculados al Programa Ondas que desertaron en algún momento de la educación básica y media corresponde a 717 observaciones.

Figura 5. Deserción y repitencia de NNA vinculados a Ondas por departamento



Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas entre 2001 y 2013 y SIMAT (2007-2015;2018-2022). El departamento corresponde al registrado cuando el NNA se vinculó al Programa Ondas.

3.6. Resultados en pruebas Saber

Los resultados obtenidos en las pruebas Saber constituyen un indicador inmediato del resultado del programa Ondas en educación; entendiéndolo como un resultado indirecto en la medida en que el Programa Ondas pueda influir positivamente en el rendimiento académico de los NNA al fomentar su curiosidad, motivación y capacidades como el pensamiento crítico y la resolución de problemas.

3.6.1. Resultados en pruebas Saber individuales

Los resultados de las pruebas individuales Saber 11 de los NNA vinculados al Programa Ondas entre 2001 y 2013 muestran que tienen puntajes promedio más altos que el promedio nacional. La *Tabla 4* presenta la desviación promedio de los puntajes de los NNA vinculados al programa Ondas con

relación al promedio nacional en el año que presentaron la prueba en cuatro áreas de conocimiento: ciencias naturales, lenguaje, matemáticas, ciencias sociales y para el agregado de las áreas evaluadas. Es importante destacar que los resultados de las pruebas Saber 11 no son comparables año tras año por lo que se empleó la diferencia entre el puntaje de cada NNA vinculado al Programa Ondas con respecto al valor promedio nacional en cada año para así eliminar las diferencias interanuales y proporcionar una medida consistente del desempeño de cada NNA vinculado al Programa Ondas a través de los años. De esta manera, una desviación positiva significa que el NNA vinculado al Programa Ondas tuvo un desempeño mejor que el promedio nacional, mientras que si la desviación es negativa indica que el NNA vinculado al Programa Ondas tuvo un mejor desempeño que el promedio nacional de los NNA que presentaron la prueba en ese año. En suma, las desviaciones de los puntajes con respecto al promedio nacional reflejan las diferencias en el rendimiento académico de los estudiantes del programa Ondas comparados con el estándar nacional, proporcionando una medida comparativa de los resultados en educación del programa para los NNA vinculados a él.[8]

Los NNA vinculados al Programa Ondas tienen mejor rendimiento en las pruebas Saber 11 que el promedio de estudiantes que presentan las prueba Saber en cada año, tal como lo muestran las desviaciones positivas para el promedio y para cada una de las áreas presentadas en la *Tabla 4*. En el área de lenguaje es donde se observan mejores puntajes de los NNA vinculados al Programa Ondas con respecto al resultado nacional, con puntajes promedio 2,02 puntos por encima, le sigue el área de ciencias (1,64) y el resultado agregado (1,61). Las desviaciones son más altas para los niños vinculados al Programa Ondas que para las niñas, sobre todo en el área de lenguaje, donde los hombres tienen una desviación promedio de 3,5 puntos y las mujeres una de 1,01 puntos. Las desviaciones en los puntajes de los NNA vinculados al Programa Ondas son más altas para los que participaron un solo año en el Programa Ondas que para los que participaron más de un año; así, el puntaje promedio agregado de los NNA que participaron un año en el Programa Ondas es 1,67 puntos más alto que el promedio nacional, mientras que para los que participaron dos veces es 0,41 puntos más altos y para los que participaron tres veces es más bajo (0,22 puntos por debajo). La diferencia entre los NNA vinculados al Programa Ondas de colegios públicos y privados es muy alta, ya que la desviación en el puntaje agregado es de 6,08 puntos para los que vienen de colegios privados y de 0,96 para los que vienen de colegios públicos, esto puede ser el reflejo de diferencias en factores económicos y acceso a recursos educativos.

De acuerdo con el departamento de residencia de los NNA vinculados al Programa Ondas, las mayores desviaciones en el puntaje agregado se observaron en Sucre (6,47 puntos), Santander (5,8 puntos), Nariño (4,53 puntos), Meta (4,51 puntos) y Norte de Santander (4,50) puntos). En contraste, hay departamentos donde los NNA vinculados al Programa Ondas tienen menores puntajes que el promedio nacional, como Chocó (8,1 puntos por debajo del promedio nacional para el puntaje agregado), Vaupés (5,6 puntos), Bolívar (1,89 puntos), Caldas (1,80 puntos). Estos resultados pueden deberse a diferencias en los recursos socioeconómicos y educativos de los estudiantes en cada departamento, o a diferencias en la implementación del Programa Ondas en cada departamento, en caso de que se validara que existen impactos del programa en los resultados de la prueba Saber. En ese sentido, la evaluación de impacto del programa Ondas realizada en 2017 arrojó impactos en general positivos en los resultados en pruebas Saber 11, particularmente altos en ciencias naturales,

[8] Se emplea información de NNA vinculados al Programa Ondas entre 2001 y 2013 y sus puntajes en pruebas Saber 11 entre 2000 y 2015; esta información se empleó para la evaluación de impacto del Programa Ondas realizada en 2017 (Fedesarrollo, 2017).

una vez se comparaban estos NNA vinculados al Programa Ondas con su grupo de control (Fedesarrollo, 2017).

Tabla 4. Diferencia promedio entre el puntaje en la prueba Saber 11 de los NNA vinculados al programa Ondas y el puntaje nacional en cada año

Desagregación	Ciencias	Matemáticas	Lenguaje	Ciencias sociales	Agregado
Total nacional	1,64	1,21	2,02	1,46	1,61
Mujer	0,94	1,50	1,01	1,32	1,12
Hombre	2,65	0,87	3,47	1,67	2,36
Participó una vez (un año) en Ondas	1,70	1,25	2,09	1,49	1,67
Dos veces	0,43	0,45	0,46	0,66	0,41
Tres veces	0,19	-0,79	-1,04	2,76	-0,22
Colegio privado	5,70	5,25	7,47	5,14	6,08
Colegio público	1,06	0,60	1,23	0,89	0,96
Amazonas	-1,01	-1,88	-2,51	-1,18	-1,68
Antioquia	3,06	4,60	3,77	3,98	3,61
Arauca	2,28	1,05	2,21	2,02	1,89
San Andrés y Providencia	-0,92	3,64	0,16	0,55	0,75
Atlántico	-0,02	-0,21	-0,47	0,60	-0,10
Bogotá	3,25	3,28	3,47	3,21	3,20
Bolívar	-2,13	-1,99	-1,66	-2,14	-1,89
Boyacá	2,11	1,00	2,38	2,97	1,81
Caldas	-2,31	0,93	-5,37	0,60	-1,80
Caquetá	-0,56	-0,98	-0,86	-0,25	-0,67
Casanare	3,02	3,25	3,27	4,49	3,08
Cauca	1,91	0,75	2,46	4,08	1,70
Cesar	0,44	-0,27	-0,05	0,20	0,22
Chocó	-9,84	-8,53	-5,93	-14,61	-8,10
Córdoba	-1,26	-1,67	-2,13	-1,03	-1,64
Cundinamarca	2,47	2,10	2,74	3,14	2,35
Guainía	-0,77	-1,12	-4,03	0,21	-1,75
Huila	2,18	1,48	3,17	2,39	2,13
La Guajira	-4,78	3,39	6,78	-7,71	-0,84
Magdalena	-0,60	-1,48	-2,13	-2,13	-1,15
Meta	4,14	5,26	5,15	6,41	4,51
Nariño	5,51	2,26	5,57	4,41	4,53
Norte de Santander	4,32	3,92	5,32	4,25	4,50
Putumayo	2,56	0,65	3,27	1,77	2,24
Quindío	0,59	1,67	0,35	2,69	0,77
Risaralda	1,66	1,54	0,78	1,29	1,31
Santander	5,67	5,14	6,94	4,85	5,80
Sucre	6,35	4,62	8,50	4,99	6,47
Tolima	3,68	4,14	3,89	4,99	3,68
Vaupés	-3,16	-5,65	-8,11	-6,14	-5,64
Vichada	0,46	0,18	1,84	1,96	0,73

Fuente: Elaboración propia con base en Fedesarrollo (2017).

3.6.2. Resultados en pruebas Saber a nivel de colegio

Los resultados de las pruebas Saber 3, 5, y 9 para los colegios con NNA vinculados a Ondas muestran que el porcentaje de estudiantes con puntajes satisfactorios o avanzados es, por lo general, comparable al rendimiento nacional en las áreas de matemáticas y lenguaje (ver Tabla 5). Por ejemplo, el 24,6% de los estudiantes en los colegios vinculados al programa Ondas tuvo puntaje de matemáticas

satisfactorio y el 23% tuvo un puntaje avanzado, estos porcentajes son muy similares al resultado nacional de 25% y 23%, respectivamente. Ahora bien, se destacan algunas diferencias en el resultado en lenguaje, sobre todo en quinto y en noveno grado, donde el porcentaje de estudiantes con nivel avanzado es más bajo para los colegios con NNA vinculados al Programa Ondas. Los colegios privados vinculados al Programa Ondas muestran un desempeño alto, con un porcentaje considerablemente más alto de estudiantes que alcanzan niveles satisfactorios y avanzados en ambas áreas, lo cual puede ser un reflejo del acceso a mejores recursos y metodologías de enseñanza. Según el departamento, los colegios vinculados al Programa Ondas en Antioquia, Caquetá, Huila y Valle del Cauca tienden a tener porcentajes más bajos de estudiantes con resultados satisfactorios o avanzados; contrario a lo que ocurre en Risaralda o Santander.

Tabla 5. Resultados en las pruebas Saber 3, 5, 9 de 2017 para los colegios vinculados a Ondas

	Porcentaje de estudiantes con puntaje Satisfactorio o Avanzado en la prueba											
	Matemáticas tercer grado		**Lenguaje tercer grado**		**Matemáticas quinto grado**		**Lenguaje quinto grado**		**Matemáticas noveno grado**		**Lenguaje noveno grado**	
	% Satisfactorio	**% Avanzado**	**% Satisfactorio**	**% Avanzado**	**% Satisfactorio**	**% Avanzado**	**% Satisfactorio**	**% Avanzado**	**% Satisfactorio**	**% Avanzado**	**% Satisfactorio**	**% Avanzado**
Resultado nacional	**25**	**23**	**27**	**19**	**16**	**12**	**29**	**14**	**20**	**6**	**41**	**7**
Total colegios vinculados a Ondas	24,6	23,0	27,7	17,7	15,8	10,9	29,3	11,8	18,4	4,5	37,9	4,4
Colegios privados vinculados a Ondas	28,4	50,0	33,6	47,0	22,9	32,4	36,1	38,2	32,2	23,6	56,9	22,9
Colegios oficiales vinculados a Ondas	24,4	21,6	27,4	16,2	15,5	9,8	29,0	10,4	17,9	3,7	37,1	3,6
Colegios vinculados a Ondas en Antioquia	23,2	16,5	25,9	13,5	12,7	6,9	28,0	8,4	17,4	3,6	43,2	4,0
Colegios vinculados a Ondas en Caquetá	25,0	17,8	23,0	12,0	14,7	8,5	28,3	4,2	15,8	3,7	24,0	1,2
Colegios vinculados a Ondas en Huila	26,3	21,3	28,2	15,8	16,8	8,1	30,7	8,9	19,9	2,5	37,9	3,3
Colegios vinculados a Ondas en Risaralda	26,6	23,1	31,5	16,7	16,2	9,4	31,2	12,2	18,9	3,7	39,7	4,3
Colegios vinculados a Ondas en Santander	27,5	34,0	31,2	26,0	22,9	21,5	36,7	21,1	32,0	10,6	49,3	8,7
Colegios vinculados a Ondas en Valle	22,9	20,0	26,6	16,3	13,1	8,4	27,4	10,0	14,1	3,3	36,7	3,3

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas agregados por establecimiento educativo, resultados de las pruebas Saber a nivel de establecimiento educativo (en el Anexo 1 se detallan las fuentes de información y los cruces de datos requeridos para llegar a estas cifras). Los resultados corresponden a 2017 y solo se tienen en cuenta los establecimientos educativos con NNA vinculados a Ondas entre 2013 y 2017. Los agregados nacionales, para todos los estudiantes se obtuvieron de (ICFES, 2018). Solo se presenta información para los departamentos con al menos 20 establecimientos educativos con NNA vinculados al programa Ondas entre 2013 y 2017 y con resultados para 2017.

Por su parte, los resultados de la prueba Saber 11 para los colegios con NNA vinculados al Programa Ondas dejan ver puntajes más altos en el área de lectura crítica (0,709), seguido de matemáticas (0.688), tal como se muestra en la Tabla 6. Los colegios privados vinculados al Programa Ondas tienen un mejor desempeño que los públicos, y las diferencias más altas se observan en el área de inglés y ciencias sociales y ciudadanas. Por lo general, los colegios con NNA vinculados al Programa Ondas en departamentos de Sucre, Cesar, Bolívar, Caquetá y Arauca tienden a tener los índices promedio más bajos (menos de 0,64 para los tres primeros departamentos en el puntaje total). Entre tanto, los colegios con NNA vinculados al Programa Ondas en departamentos de Cundinamarca y Santander tienen índices promedio muy altos en todas las áreas, con un índice promedio total por encima de 0,7.

Tabla 6. Resultados en las pruebas Saber 11 de 2020 y 2021 para los colegios vinculados a Ondas (Índice por colegio)

Desagregación	Matemáticas	Ciencias naturales	Sociales y ciudadanas	Lectura crítica	Inglés	Total
Total colegios	0,688	0,661	0,646	0,709	0,656	0,675
Colegios privados	0,759	0,740	0,733	0,776	0,758	0,753
Colegios oficiales	0,685	0,657	0,642	0,706	0,651	0,671
Antioquia	0,691	0,657	0,655	0,720	0,669	0,680
Arauca	0,672	0,647	0,608	0,670	0,598	0,645
Atlántico	0,664	0,640	0,610	0,684	0,641	0,649
Bolívar	0,646	0,621	0,610	0,682	0,624	0,639
Caldas	0,675	0,654	0,637	0,700	0,647	0,665
Caquetá	0,658	0,626	0,612	0,673	0,607	0,640
Cesar	0,651	0,632	0,598	0,671	0,621	0,637
Cundinamarca	0,718	0,689	0,676	0,729	0,681	0,701
Guaviare	0,682	0,643	0,617	0,686	0,615	0,654
Huila	0,686	0,665	0,636	0,698	0,635	0,668
Risaralda	0,687	0,664	0,650	0,710	0,656	0,676
Santander	0,736	0,702	0,678	0,735	0,677	0,710
Sucre	0,648	0,641	0,598	0,668	0,596	0,636
Tolima	0,689	0,663	0,648	0,709	0,660	0,676
Valle del Cauca	0,686	0,661	0,654	0,716	0,659	0,678

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas agregados por establecimiento educativo, resultados de las pruebas Saber a nivel de establecimiento educativo (en el Anexo 1 se detallan las fuentes de información y los cruces de datos requeridos para llegar a estas cifras). Los datos corresponden a 2020 y 2021 solo para los colegios con NNA vinculados a Ondas en los cinco años anteriores a la presentación de la prueba; los puntajes están ponderados por el número de alumnos evaluados. El índice por colegio va de cero a uno.

3.7. Graduación de la educación superior

Se identificaron 927 NNA vinculados a Ondas entre 2001 y 2013 que se graduaron de educación superior entre 2020 y 2022.[9] El número de graduados de educación superior aumenta anualmente, lo que es el reflejo de que los NNA que alguna vez participaron en el Programa Ondas alcanzan la edad para adelantar estudios de educación superior; específicamente, los graduados de educación superior pasaron de 150 en 2020 a 458 en 2022 (ver *Tabla 7*). Además, el número de mujeres que alguna vez estuvieron vinculadas al Programa Ondas y que se graduaron de educación superior en el período analizado es más alto que el número de hombres (560 vs. 367); estas cifras son consistentes con las cifras de deserción y repitencia por sexo vistas anteriormente donde se observaba menor repitencia y deserción por parte de las niñas, lo que a su vez puede traducirse en mayores tasas futuras de graduación de la educación media, acceso a la educación superior y graduación. Por último, la mayoría de los jóvenes que estuvieron vinculados al Programa Ondas y se graduaron de educación superior entre 2020 y 2022 no pertenecen a ningún grupo étnico.

Se destaca que casi la mitad de los NNA vinculados al Programa Ondas que se graduaron de educación superior entre 2020 y 2022 participaron de convocatorias del Programa Ondas en el departamento de Bolívar (ver Figura 6). Esto no es de extrañarse ya que un alto porcentaje de los NNA que se vincularon al Programa Ondas antes del 2014 lo hicieron en el departamento de Bolívar (la mitad del total de NNA en 2013, o más de un tercio en 2010 y 2012, tal como se muestra en la Tabla 26 del Anexo 1. Otros departamentos con alto número de jóvenes que estuvieron en el Programa Ondas y que se

[9] Según el resultado de cruce entre la base de datos de NNA vinculados a Ondas entre 2001 y 2013 y SNIES entre 2018-2022.

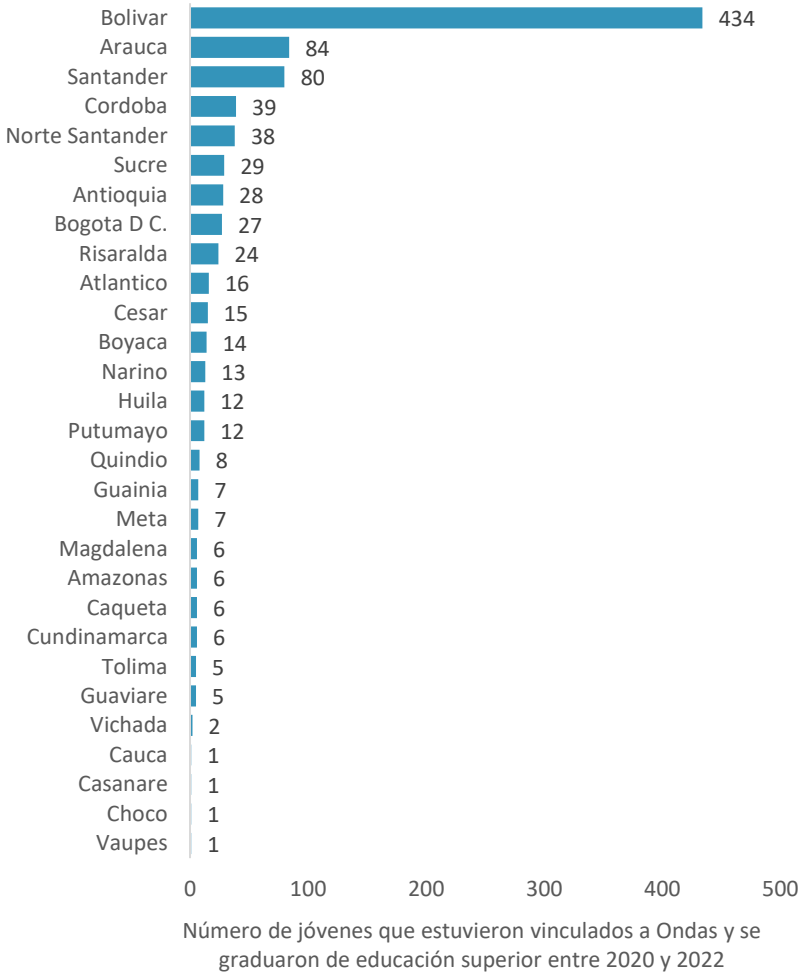
graduaron de educación superior son Arauca (84), Santander (80), Córdoba (39) y Norte de Santander (38).

Tabla 7. Jóvenes que estuvieron vinculados a Ondas y que se graduaron de educación superior entre 2020 y 2022

Desagregación		Número de jóvenes que estuvieron vinculados a Ondas y se graduaron de educación superior entre 2020 y 2022
Por año	2020	150
	2021	319
	2022	458
Género	Mujeres	560
	Hombres	367
Grupo étnico	Indígena	21
	NARP	18
	Ningún grupo étnico	888

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas entre 2001 y 2013 y SNIES (2020-2022).

Figura 6. Jóvenes que estuvieron vinculados a Ondas y que se graduaron de educación superior entre 2020 y 2022, por departamento

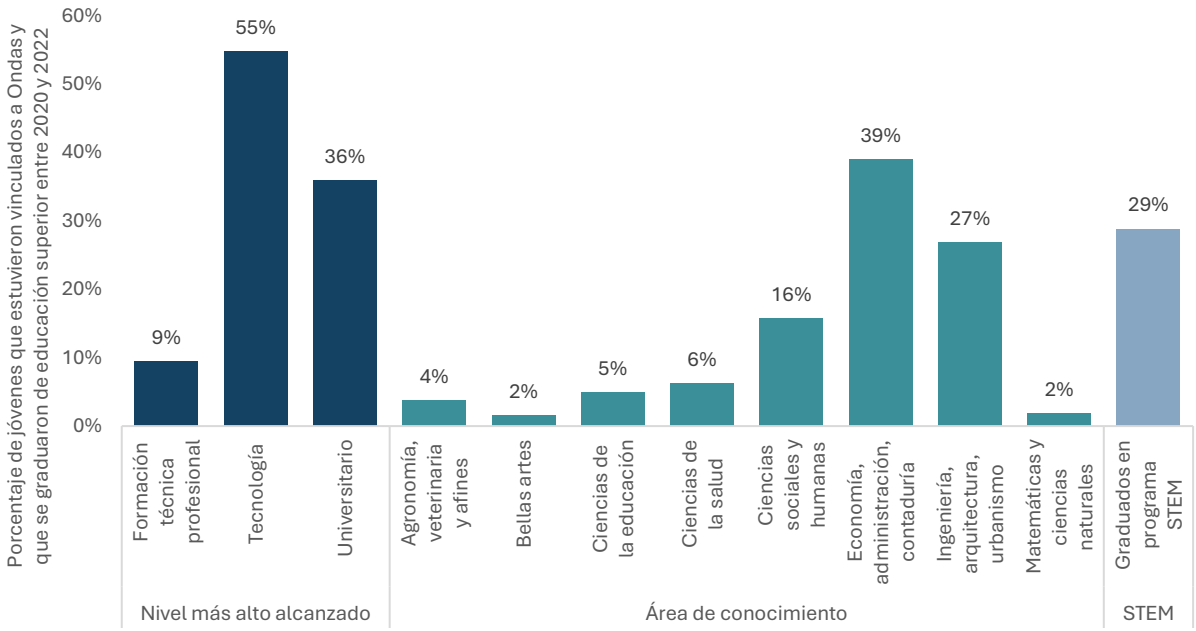


Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas entre 2001 y 2013 y SNIES (2020-2022). El departamento corresponde al registrado cuando el NNA se vinculó a Ondas.

El 11,5% de los NNA graduados entre 2020 y 2021 hizo una carrera técnica, el 56% una tecnología y el 36% un programa de formación universitaria. Ahora bien, cuando se toma en consideración solo el nivel más alto alcanzado, de los jóvenes que pertenecieron al Programa Ondas el 9% alcanzó formación técnica profesional, el 55% una tecnología y el 36% una carrera universitaria (ver *Figura 7*). Para este nivel más alto alcanzado, el 39% escogió una carrera en economía, administración y contaduría, el 27% una carrera en ingeniería, arquitectura y urbanismo, y el 16% en ciencias sociales y humanas.

Aproximadamente uno de cada tres jóvenes –el 29% que estuvieron vinculados al programa Ondas y que se graduaron entre 2020 y 2022 lo hicieron de un programa STEM, los cuales corresponden a las áreas de ingeniería, arquitectura, urbanismo y afines o en matemáticas y ciencias naturales. Esta cifra es considerablemente alta si se tiene en cuenta que en 2022 el 19,6% de los graduados de educación superior del país lo hicieron de una carrera de educación superior en ciencias naturales, matemáticas, estadística o ingeniería, industria y construcción.[10]

Figura 7. Jóvenes que estuvieron vinculados a Ondas y que se graduaron de educación superior entre 2020 y 2022



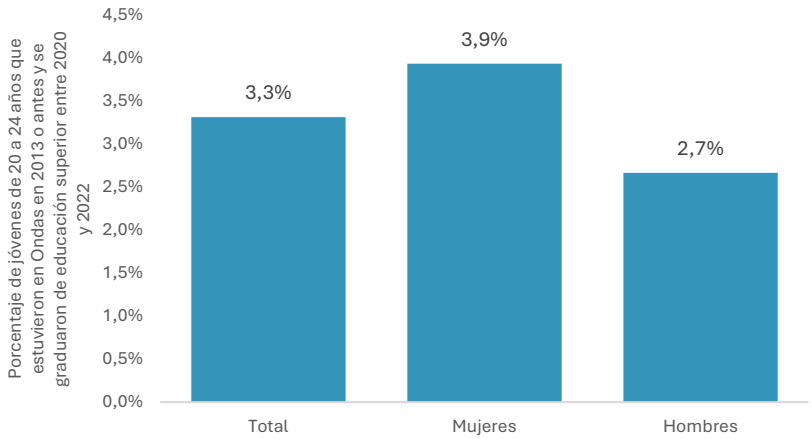
Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas entre 2001 y 2013 y SNIES (2020-2022).

Como se mencionó anteriormente, 927 NNA vinculados al Programa Ondas se graduaron de educación superior entre 2020 y 2022. Para obtener una medida de la proporción de jóvenes que han estado vinculados al Programa Ondas y que tienen edad de haberse graduado de educación superior, se calculó el número de graduados como porcentaje de jóvenes que estuvieron en el Programa Ondas y

[10] Para mayor información sobre los graduados en educación superior en 2022 consultar la siguiente página del Ministerio de Educación: https://hecaa.mineducacion.gov.co/consultaspublicas/content/poblacional/index.jsf

que tenían entre 20 y 24 años en 2022.[11] Así, el 3,3% de los jóvenes de 20-24 años que estuvieron vinculados al Programa Ondas se graduaron de educación superior entre 2020 y 2022. Además, este porcentaje es más alto en mujeres (3,3%) que en hombres (2,7%).

Figura 8. Porcentaje de jóvenes entre 20 y 24 años que se graduaron de educación superior entre 2020 y 2022



Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas entre 2001 y 2013 y SNIES (2020-2022).

Los entrevistados en el marco de la recolección de información cualitativa del estudio destacaron que la experiencia del Programa Ondas sirve como escenario de orientación vocacional, puesto que, al fomentar los intereses de los estudiantes, les permite confirmar si les gustaría seguir ese camino en la educación superior. El Programa Ondas también les sirve a los NNA para orientar su proyecto de vida, identificar sus fortalezas y debilidades y enfocarlo a la selección de la carrera universitaria, en algunos casos el Programa Ondas ha influido en el proyecto de vida y la decisión sobre qué estudiar.

"es importante para ellos porque ya están en una etapa de sus últimos años de colegio y es que a veces ellos no saben cómo orientarse […]. estos espacios son excelentes porque ahí ellos están mirando sus fortalezas, sus debilidades, para qué realmente quieren hacer o que les va a llamar la atención qué puede ser potencial para ellos para una carrera universitaria"

(Entrevista a padre de familia, 2023)

"Yo era de las personas que dicen: yo quiero ser futbolista, me gusta mucho el fútbol. Pero entré, hice todo el proceso, primero empecé por robótica, después por investigación, después un poco de ambas a ver qué era lo que me gustaba, si era lo que quería. […] Entonces a mí me apasionó esto, me llamó mucho la atención y ahora quiero estudiar Ingeniería Ambiental"

[11] Este rango etario parece arbitrario, pero los jóvenes Ondas graduados de educación superior entre 2020 y 2022 tenían edades entre 20 y 24 años (correspondiente a los percentiles 10 y 90) o entre 19 y 28 (si uno considera los percentiles 1 y 99). Ahora bien, mientras más amplio el rango etario, más se corre el riesgo de incluir como no graduados a jóvenes que pudieron haberse graduado antes.

(Grupo focal con participantes, 2023)

4. Estrategia para la evaluación de impacto del programa ONDAS

El Programa Ondas puede tener impactos directos sobre los NNA vinculados al programa al ofrecerles una propuesta pedagógica y metodológica que les permite fomentar la curiosidad y la exploración científica. Esto no solo tiene el potencial de mejorar su adaptación al entorno escolar, reduciendo las tasas de deserción y repitencia, sino que también puede elevar su rendimiento académico en pruebas estandarizadas SABER e incrementar su motivación para continuar su trayectoria educativa en la educación superior. Adicionalmente, el programa también puede contribuir al desarrollo profesional de los docentes, quienes a su vez pueden influir positivamente en el desempeño académico de sus estudiantes.

Así, la evaluación de impacto del programa permitirá tener una visión de los cambios reales y significativos derivados de la vinculación al Programa Ondas en la trayectoria educativa de los NNA que estuvieron vinculados en él. Con la evaluación de impacto se busca responder qué hubiera pasado con los NNA que se vincularon al Programa Ondas si no hubieran sido parte del programa; sin embargo, como no es posible tener este contrafactual, se debe identificar claramente un grupo de tratamiento, que en este caso estaría conformado por los NNA vinculados al Programa Ondas, y un grupo de control, el cual se conformaría con NNA que no estuvieron vinculados al programa pero que sean similares en características observables y no observables a los NNA vinculados al programa Ondas.

El principal reto de esta evaluación es poder establecer un grupo de control adecuado para poder tener resultados confiables y robustos. En un escenario ideal se contaría con el listado de los NNA que se inscriben en la convocatoria pero que no resultan seleccionados, lo que permitiría también eliminar el sesgo de autoselección derivado de la motivación intrínseca que tienen los NNA que deciden vincularse a los grupos de investigación que se vinculan al Programa Ondas. Al no contar con esta información, se proponen dos alternativas para la construcción de los grupos de tratamiento y control, cuya viabilidad recae en la disponibilidad de información.

4.1. Grupos de tratamiento y control con información individual de los NNA

Esta alternativa plantea tomar como grupo de tratamiento a los NNA vinculados al Programa Ondas y cuya información se encuentra disponible en los listados de las convocatorias; mientras que el grupo de control serían los niños que estaban en el mismo curso y grado que los NNA vinculados al Programa Ondas. Es decir, si un grupo de investigación de un establecimiento educativo en el que participan cinco NNA de tercer grado y dos de quinto grado se postula al programa y resulta seleccionado, estos harían parte del grupo de tratamiento; mientras que para el grupo de control se tomarían los niños de tercero y quinto grado en el mismo curso que los tratados. Esta aproximación no elimina el sesgo de autoselección derivado de las características que hacen a los NNA vincularse a un grupo de investigación, por lo que requiere amplia información de características de los NNA para tratar de controlar lo mejor posible por este sesgo.

Esta estrategia se empleó en Fedesarrollo (2017) para adelantar la evaluación de impacto de Ondas en NNA que estuvieron vinculados al Programa Ondas entre 2001 y 2013. Para los NNA que se vincularon al Programa Ondas entre 2014 y 2023 no se pudo acceder a la información de los compañeros de clase en el SIMAT. Por tanto, solo se cuenta con información de NNA vinculados al Programa Ondas y de sus controles para el período entre 2001 y 2013. Para estos NNA se cuenta con la información histórica de matrícula entre 2007-2015 derivado de la anterior evaluación realizada por

Fedesarrollo (2017) y se cuenta con la actualización de esta información para el período entre 2018 y 2022. En el Anexo 1 se presentan más detalles sobre las fuentes de información empleadas para la evaluación de impacto y los cruces de información llevados a cabo con registros administrativos.

4.2. Grupos de tratamiento y control con información a nivel de establecimiento educativo

Como alternativa a la evaluación de impacto empleando registros individuales de los NNA se propone una evaluación de impacto empleando información secundaria y el establecimiento educativo como unidad de análisis. Bajo este enfoque, el grupo de tratamiento corresponde a los establecimientos educativos que han implementado el programa Ondas entre 2001 y 2023. Por su parte, el grupo de control corresponde a los establecimientos educativos del país que no han implementado el programa Ondas y que son similares al grupo de tratamiento en variables observables.

La ventaja de este enfoque es que recae fuertemente en el uso de información secundaria disponible en línea: el Sistema de Información Nacional de Educación Básica y Media (SINEB) cuenta con la información de todos los establecimientos educativos y, la información de las pruebas Saber agregadas a nivel de establecimiento educativo, están disponibles en el sistema DataIcfes, con nombre y código de establecimiento educativo. No obstante, este enfoque no limita el sesgo de autoselección, que bien pudiera mitigarse mediante la inclusión de un amplio conjunto de covariables.

Bajo esta estrategia, la evaluación se centra en determinar los impactos del Programa Ondas en la calidad de la educación básica y media de los establecimientos educativos. En este caso, la calidad de la educación se entiende como el logro de las competencias básicas que se espera que los estudiantes adquieran en cada grado y área del conocimiento, y que en Colombia se mide a través de los resultados en las pruebas Saber. Es posible esperar impactos del Programa Ondas en la calidad de la educación en la medida en que puede fomentar el desarrollo de habilidades críticas y analíticas en los NNA cuando estos se involucran en actividades de investigación, por lo que pudiera también tener impactos en la forma en la que los niños abordan el aprendizaje y, por tanto, en sus resultados en pruebas estandarizadas. Ahora bien, al realizar la evaluación a nivel de establecimiento educativo, el supuesto subyacente es que hay efectos desborde del programa, donde el programa impacta no solo a los NNA vinculados a grupos de investigación, sino también a sus compañeros, quienes se ven favorecidos por el intercambio de conocimientos y experiencias y por la mayor formación y aprendizaje del docente que acompaña el proyecto de investigación.

4.3. Fuentes de información y variables de resultado

Para adelantar la evaluación de impacto del Programa Ondas en la calidad de los establecimientos educativos que implementan el programa se emplearon las siguientes fuentes de información:[12]

Base de datos del programa Ondas. Tiene la información de los NNA vinculados al Programa Ondas entre 2001 y 2023. Esta información se agregó a nivel de establecimiento educativo para tener el grupo de tratamiento en las estimaciones a nivel de establecimiento educativo.

Sistema de Matrícula Estudiantil (SIMAT). Cuenta con la información de matrícula en educación básica y media y permitió la identificación de un grupo de control conformado por los compañeros de clase

[12] En el Anexo 1 se presenta con mayor detalle las fuentes de información empleada, las limitaciones de la información y los cruces de información llevados a cabo para poder tener la información de matrícula en educación básica y media, así como de graduación de educación superior. Así mismo, se detalla cómo se depuró la información de beneficiarios vinculados al Programa Ondas entre 2001 y 2023 para agregar la información a nivel de colegios y cómo se obtuvo el grupo de control.

de los NNA vinculados al Programa Ondas entre 2001 y 2013, así como su continuidad año a año en la base de datos.[13]

Sistema Nacional de Información de la Educación Superior (SNIES). El Ministerio de Educación realizó el cruce de la información de los NNA vinculados al Programa Ondas entre 2001 y 2013 y sus compañeros de clase con los graduados de educación superior entre 2020 y 2022.

Sistema de Información Nacional de Educación Básica y Media – SINEB. Contiene el directorio de instituciones educativas oficiales y no oficiales de educación preescolar, básica y media. Se requería esta base de datos para completar la información faltante de códigos de establecimiento educativo que no estaba disponible en la base de datos del Programa Ondas.[14]

Base de datos de las pruebas Saber 3, 5, 9 a nivel de establecimiento para 2009 y 2012-2017. Cuenta con la información de los resultados de estas pruebas Saber 3, 5, 9 a nivel de establecimiento educativo. Se accedió a la información a través del sistema *DataIcfes*.

Base de datos de las pruebas Saber 11 a nivel de establecimiento para 2014-2021. Cuenta con la información de los resultados de las pruebas Saber 11 a nivel de establecimiento educativo, y características del establecimiento como departamento, municipio, naturaleza, grado y calendario. Se accedió a la información a través del sistema *DataIcfes*.[15]

4.4. Variables de resultado a nivel de NNA

El programa Ondas se configura como un espacio formativo adicional al del desarrollo de las actividades escolares, donde los NNA pueden aprender otras habilidades, cultivar su curiosidad, y aumentar la motivación con la educación. Es por eso por lo que se evalúan los impactos del Programa Ondas en variables como:

Deserción temporal. Dado que se cuenta con la información año a año de SIMAT, un NNA se clasifica como desertor temporal si deja de aparecer en algún momento en la matrícula y no avanza en grado académico durante ese tiempo, ni tampoco es repitente. Por ejemplo, si un NNA aparece en octavo grado en SIMAT en 2018 y luego vuelve a aparecer en 2020 en noveno grado se asume como desertor temporal en la medida en que desapareció de la base de datos y no avanzó hasta décimo grado como hubiese sido en ausencia de deserción.

Repitencia. Un NNA clasifica como repitente si en algún momento en su trayectoria educativa aparece como repitente en la base de datos de SIMAT.

Graduación de educación superior. Se cuenta con la información de graduados de la educación superior entre 2020 y 2022, pero no se tienen datos de los NNA que ya ingresaron a la educación superior. Para esta evaluación se propone evaluar la tasa de graduación de educación superior empleando a toda la población que tenía entre 20 y 24 años en 2022.[16]

---

[13] Se cuenta con los compañeros de clase de los NNA vinculados al Programa Ondas entre 2001 y 2013 de la evaluación realizada por Fedesarrollo en 2017, además se cuenta con la información año a año de matrícula tanto para los tratados como controles entre 2007 y 2015. En el marco de esta evaluación, el MEN realizó una actualización del cruce, complementando la información con la matrícula año a año entre 2018 y 2022.

[14] La información se descargó de:
https://sineb.mineducacion.gov.co/bcol/app?service=direct/0/Home/$DirectLink&sp=IDsede=52649-

[15] También está disponible la información de las pruebas Saber 11 a nivel de colegio para períodos anteriores; sin embargo, no siguen la misma estructura ni son comparables con los resultados de 2014 en adelante.

[16] Este rango etario parece arbitrario, pero los jóvenes Ondas graduados de educación superior entre 2020 y 2022 tenían edades entre 20 y 24 años (correspondiente a los percentiles 10 y 90).

4.5. Variables de resultado a nivel de establecimiento educativo

De acuerdo con la información cualitativa recabada, el programa Ondas permite que se abarquen problemáticas del contexto inmediato de los NNA y que sean ellos quienes planteen soluciones a través del ejercicio investigativo. Así, el desarrollo de los proyectos de investigación les permite plantear hipótesis, buscar información y analizarla, lo que fomenta el desarrollo del pensamiento crítico y la capacidad para discernir información. Estas habilidades son directamente aplicables en otros ámbitos de la vida de los NNA, particularmente en el ámbito educativo, teniendo el potencial de influir positivamente en su desempeño académico (medido por las pruebas estandarizadas Saber). El desarrollo de habilidades de pensamiento crítico puede llevar a los NNA a interpretar, analizar y evaluar mejor los textos—y, por tanto, mejores puntajes en lectura crítica y lenguaje; formular preguntas y diseñar experimentos que puedan mejorar su comprensión de las ciencias naturales; o abordar los problemas matemáticos de mejor manera, aplicando razonamiento lógico en ellos. Es por ello que se evalúan los resultados en las pruebas Saber en Lenguaje o lectura crítica, matemáticas y ciencias naturales. A continuación, se detallan las variables de resultado construidas:

Porcentaje de niños en el establecimiento con puntaje avanzado o satisfactorio en la prueba Saber. Para las pruebas Saber 3, 5, 9 el ICFES determina cuatro posibles niveles de desempeño: insuficiente, mínimo, satisfactorio y avanzado dependiendo de las competencias que alcanzan los estudiantes de cada grado en cada prueba. El nivel de desempeño "Avanzado" se refiere a un desempeño sobresaliente y "Satisfactorio" se refiere a un desempeño adecuado de acuerdo con las competencias exigidas en el área evaluada (ICFES, n.d.).

Índice para la clasificación de los planteles educativos por área. Es un índice entre cero y uno que evalúa el nivel de desempeño promedio de los niños del establecimiento educativo en las áreas de matemáticas, lectura crítica y ciencias naturales. (ICFES, n.d.)

4.6. Técnicas de estimación

4.6.1. Emparejamiento

Los métodos de emparejamiento se emplean usualmente para construir el grupo de comparación empleando características observables, bajo el supuesto de que las diferencias no observadas entre el grupo de tratamiento y control no se relacionan con las variables de resultados. Los métodos de emparejamiento emplean técnicas estadísticas para crear un grupo de control artificial a partir de observaciones que son similares a cada observación del grupo de tratamiento. Sin embargo, a medida que aumenta el número de características o de individuos, el emparejamiento se puede hacer inviable debido a la dificultad de encontrar coincidencias exactas. Por tanto, Rosenbaum y Rubin (1983) desarrollaron el método de "*Propensity Score Matching*" (PSM) o de emparejamiento por puntajes de propensión, el cual consiste en construir un grupo de control constituido por individuos que sean similares a los individuos tratados en características observables, pero que por alguna razón (ajena a ellos) no recibieron el tratamiento. Específicamente, con el método de PSM se busca emparejar cada individuo tratado con un individuo no tratado con base en sus características observables y, al ser similares, la asignación del tratamiento sea aleatoria. Sin embargo, esto requiere cumplir con el supuesto de que los individuos son idénticos en todas sus características observables.

Al ser individuos similares en todas sus características observables, la metodología permite asumir que la asignación del tratamiento entre cada pareja de individuos fue aleatoria. Claramente, para poder cumplir con el supuesto de que los individuos son idénticos en todas sus características observables, la metodología exige que se cuente con información muy detallada tanto de los individuos tratados como de los no tratados.

Siguiendo a Rosembaum y Rubin (1983), la metodología de emparejamiento basada en la probabilidad de participación en el programa (*propensity score*), se estima a partir de las características observadas de cada individuo. En este sentido, el primer paso consiste en la estimación de la probabilidad de ser tratado mediante un modelo *logit*:

$$P(x) = P(X_i) = \ f(\alpha + \beta X_i + \varepsilon_i)$$

donde $P(X)$ es la probabilidad de que el individuo haya sido intervenido, y X_i es un vector de características observables del individuo. Luego, se aplica un algoritmo para emparejar a los individuos del grupo de tratamiento y control con puntajes más parecidos, para lo cual se pueden emplear diferentes aproximaciones:

Vecino más cercano. Este algoritmo consiste en buscar para cada individuo de tratamiento un individuo de control con la probabilidad de participación más cercana. El algoritmo busca entonces minimizar la distancia entre la probabilidad predicha del individuo del grupo de tratamiento y la probabilidad predicha del individuo de control. El individuo de control para el individuo de tratamiento i está dado por:

$$C(i) = \{j \in D = 0 | \|P_i(X) - P_j(X)\| \}$$

Distancia máxima (radio): Este algoritmo es una refinación del algoritmo anterior, en el caso en que se estén emparejando individuos que no sean suficientemente similares. Metodológicamente consiste en imponer una distancia máxima (k) entre las probabilidades de tratamiento de los individuos de tratamiento y de control y emparejar únicamente cuando la distancia entre dos vecinos sea inferior o igual a este umbral. En este caso, el grupo de comparación está dado por:

$$C(i) = \{j \in D = 0 | \|P_i(X) - P_j(X)\| \leq k\}$$

Kernel: Éste es un método no paramétrico que empareja a los individuos del grupo de tratamiento con un promedio ponderado del grupo de control. En este promedio ponderado otorga más peso a las unidades con probabilidad de participación más cercanas dentro de un ancho de banda con respecto a la probabilidad predicha de participación del individuo de tratamiento. Este ponderador (w_{ij}^{KERNEL}) es una función de qué tan lejos está el individuo de tratamiento i de cada individuo del grupo de control j.

$$w_{ij}^{KERNEL} = \frac{G\left(\frac{P_j(X) - P_i(X)}{h}\right)}{\sum_{k \in D=0}^{K} G\left(\frac{P_k(X) - P_i(X)}{h}\right)}$$

Regresión Lineal Local: este método es una variación paramétrica del método anterior, construyendo una regresión lineal de la variable objetivo del grupo de control en el vecindario de cada observación del grupo de tratamiento sobre un intercepto. El algoritmo RLL utiliza el mismo método de estimación anterior, pero se asume que la variable de resultado cambia linealmente con la probabilidad de participación.

Una vez emparejados los individuos del grupo de tratamiento (con alguno de los algoritmos descritos), se obtiene un grupo de control con el cual es posible hacer las comparaciones de promedios de

variables de resultado. Así, el estimador de interés para la evaluación de impacto de la intervención, es de nuevo el *ATT:*

$$\tau^{ATT} = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1]$$

En la anterior ecuación, $E[Y_i(0)|D_i = 1]$ es el resultado contrafactual, es decir, el valor esperado de la variable de resultado en el grupo de tratamiento en ausencia del tratamiento. Dado que, como se ha mencionado, este valor es imposible de observar en la vida real, nos aproximamos mediante el grupo de control resultante del proceso de emparejamiento descrito anteriormente.

$$\tau^{ATT} = E[Y_i|D_i = 1] - E\left[Y_{c(i)}|D_{c(i)} = 0\right] \qquad (17)$$

Esta metodología se empleó para estimar los impactos del programa Ondas en deserción escolar y graduación de educación superior de los NNA. Además, se empleó el emparejamiento como una etapa previa en las estimaciones a nivel de establecimiento educativo, en combinación con la metodología de diferencias en diferencias que se plantea a continuación.

4.6.2. Diferencia en diferencias

Se empleó la metodología de diferencia en diferencias para estimar los impactos del programa Ondas en la calidad educativa, medidas empleando los resultados en pruebas Saber a nivel de colegio. Una vez se estableció un grupo de control con una distribución de covariables similar al grupo de tratamiento (dentro del soporte común), se estimaron los impactos del programa Ondas en los resultados de las pruebas Saber de los establecimientos educativos utilizando el estimador de diferencia en diferencias (DID). Ahora bien, aprovechando que se cuenta con información de las pruebas Saber para varios años y que los establecimientos educativos se vinculan al programa Ondas en diferentes momentos del tiempo, se usó la generalización del modelo de DID desarrollada por Callaway y Sant'Anna (2021) para múltiples tratamientos en múltiples momentos de tiempo. Este estimador facilita el estudio de los efectos del programa sobre cohortes específicas y su evolución temporal.

Callaway y Sant'Anna (2020) desarrollaron una generalización del modelo de diferencias en diferencias que permite analizar el efecto del programa en el tiempo y para diferentes cohortes (en este caso, cohortes de participación en el Programa Ondas). Los autores parten de la representación del modelo de diferencias clásico, donde el efecto promedio sobre los tratados (ATT por sus siglas en inglés) viene dado por:

$$ATT = E\left[Y_{T=1;t=1} - Y_{T=0;t=1}|P_{T=1} = 1\right]$$

La generalización del estimador se basa en la idea de que cada observación pertenece a una cohorte *h* en cada momento *t* del tiempo, por lo que el estimador ATT puede representarse como:

$$ATT(h,t) = E\left[Y_{T=h;t=t} - Y_{T=0;t=t}|P_{T=h} = 1\right]$$

Los autores no imponen restricciones sobre los grupos de tratamiento o los períodos de tiempo, por lo que pueden ser varios. Este estimador se conoce como el efecto promedio del tratamiento en el grupo y tiempo (group-time average treatment effect), y mide el impacto de la intervención en los participantes del programa a lo largo del tiempo. Una ventaja del método es que permite examinar cómo es el impacto en cohortes específicas y cómo evoluciona en el tiempo, además de que permite tener un grupo de control dinámico conformado no solo por los individuos que no participan del programa, sino también por los futuros tratados del programa en cada momento del tiempo.

Callaway y Sant'Anna (2020) demuestran que para obtener estimadores insesgados y consistentes utilizando esta metodología, es necesario imponer restricciones a la anticipación del tratamiento por parte de los tratados en el tiempo, lo cual puede influir en su comportamiento previo al tratamiento y en el resultado de las estimaciones. En caso de que esto ocurra, se debe recurrir a funciones de ponderación para corregir las estimaciones. Para ellos los autores desarrollaron tres métodos de estimación: regresión de resultados (OR), ponderación por probabilidad inversa (IPW) y estimadores doblemente robustos (DR). OR modela los resultados del grupo de control, IPW modela la probabilidad condicional de que una unidad sea asignada al grupo de tratamiento y DR combina ambos enfoques.

Para ilustrar la estimación mediante el método de diferencia en diferencias con múltiples períodos y tratamientos en el tiempo, partimos del caso en el cual contamos con información de tipo panel en la cual:

$$\{(Y_{i,1}, Y_{i,2}, \dots, Y_{i,\tau}, D_{i,1}, D_{i,2}, \dots, D_{i,\tau}, X_i)\}_{i=1}^{n}$$

Donde la cual $D_{i,t}$ es una variable dummy que toma el valor de uno si el joven *i* participa en JII en el periodo *t,* en otro caso, toma el valor de cero. Adicionalmente, se define una variable dummy $G_{i,g}$ que toma el valor de uno para el joven *i* si el periodo *t* es su primer período de participación en el programa. Esto es, la variable $G_{i,g}$ es equivalente a una variable que podemos determinar *inicio del tratamiento.* Por último, se define una variable C la cual toma el valor de uno para los Establecimientos educativos que nunca implementan el programa.

Los autores asumen que una vez un individuo es tratado, se asume como tratado para siempre. Esto es:

$$D_{i,t} = 1 \Rightarrow D_{i,t+1}, para\ t = 1,2, \dots \tau$$

Bajo estos supuestos, el parámetro de interés puede escribirse como:

$$ATT(g,t) = E[Y_t(g) - Y_t(0) | G_g = 1], para\ t \geq g$$

el cual representa el efecto promedio sobre la población tratada por primera vez en el momento *g,* dentro del conjunto de periodos *t.*

En este punto, y dado que se cuenta con múltiples periodos *t,* y con múltiples grupos de unidades tratadas en diferentes momentos *g.* El ATT para la aproximación para la estimación de los impactos con un grupo de control con jóvenes que nunca participaron en el programa sería:

$$ATT_{unc}^{ny}(g,t) = E[Y_t - Y_{g-1} | G_g = 1] - E[Y_t - Y_{g-1} | C = 1]$$

Mientras que para el caso en el que se utiliza a los futuros tratados sería, el *ATT* puede escribirse como:

$$ATT_{unc}^{ny}(g,t) = E[Y_t - Y_{g-1} | G_g = 1] - E[Y_t - Y_{g-1} | D_t = 0, G_g = 0]$$

Nótese que, en ambos casos, la estimación del impacto implica realizar tantas estimaciones de diferencia en diferencia como periodos *g* estén disponibles. Es decir, con esta aproximación del modelo de diferencia en diferencias propuesta por Callaway y Sant'Anna (2020) es posible reconstruir la trayectoria del impacto del programa en cada periodo de tiempo. A diferencia del modelo de diferencia en diferencias tradicional, en el cual se comparan los individuos tratados y los controles únicamente en dos periodos de tiempo. Esto no solo produce estimaciones más robustas, sino que permite identificar los posibles efectos heterogéneos o la acumulación de los impactos.

Ahora bien, se puede obtener el impacto promedio del programa mediante la agregación de las múltiples estimaciones de impacto mediante la metodología propuesta por Callaway y Sant'Anna (2020). Para ello se puede tomar los promedios ponderados de todos los impactos estimados para todos los individuos en todos los momentos del tiempo. Esto es:

$$\theta_W^{simple} = \frac{1}{K}\sum_{g=2}^{\tau}\sum_{t=2}^{\tau} 1\{g \leq t\} ATT(g,t) P(G = g | C \neq 1)$$

Por último, el método de diferencia en diferencias descrito requiere, al igual que en su versión tradicional, del cumplimiento del supuesto de tendencias paralelas, el cual permite establecer que la evolución de las variables de resultado entre los grupos de tratamiento y control previo a la intervención, tenían la misma tendencia, y por tanto cualquier cambio en la tendencia es atribuible al programa. En esta versión del modelo de diferencia en diferencias, surgen dos aproximaciones al supuesto de tendencias paralelas, de acuerdo con la metodología de conformación del grupo de control establecida. Si se realizan estimaciones con base en el grupo de control con individuos que nunca participaron en el programa, el supuesto de tendencias paralelas vendrá dado por:

$$E\left[Y_t(0) - Y_{t-1}(0) | X, G_g = 1\right] = E[Y_t(0) - Y_{t-1}(0) | X, C = 1]$$

para todo $t \in \{2, \dots, \tau\}, g \in \varrho \; para \; todo \; t \geq g$

De otro lado, si las estimaciones se basan en el grupo de control conformado por los futuros controles, el supuesto de tendencias paralelas vendrá dado por:

$$E\left[Y_t(0) - Y_{t-1}(0) | X, G_g = 1\right] = E\left[Y_t(0) - Y_{t-1}(0) | X, D_s = 0, G_g = 0\right]$$

para todo $(s,t) \in \{2, \dots, \tau\} \times \{2, \dots, \tau\}, \; g \in \varrho \; para \; todo \; t \geq g, s \geq t$ se cumple el supuesto de tendencias paralelas. La implementación del modelo de DID para múltiples períodos y múltiples tratamientos requiere del cumplimiento del supuesto de tendencias paralelas, asegurando que cualquier desviación en las tendencias de los resultados entre los grupos de tratamiento y control se debe a que el establecimiento desarrolló el programa Ondas. Este supuesto es fundamental para obtener estimaciones insesgadas y consistentes del impacto del programa.

4.6.3. Limitaciones

La estrategia de evaluación presenta algunas limitaciones que deben ser consideradas. En primer lugar, sería ideal contar con un grupo de NNA que se inscribieron a las convocatorias del programa pero que no fueron seleccionados, ya que se podría controlar mejor el sesgo de auto selección. En esa misma dirección, hay características intrínsecas, como la motivación por participar en el programa, que pueden no controlarse completamente con las características disponibles. Otra limitación para la evaluación con información a nivel de NNA es que, pese a las gestiones realizadas con el MEN, solo se pudo contar con el grupo de control para los NNA vinculados al Programa Ondas entre 2001 y 2013. En cuanto a la evaluación a nivel de establecimiento educativo, la principal limitación recae en el supuesto de que hay efectos de desborde y que los efectos del programa no se limitan a sus beneficiarios sino a los compañeros de clase. Por último, otras limitaciones identificadas en las fuentes de información se presentan con mayor detalle en el Anexo 1.

5. Resultados de la evaluación a nivel de NNA

5.1. Descripción de la muestra de evaluación

En la Tabla 27 del Anexo 1 se presentan las características generales de la muestra de evaluación para el año en el cual se hizo el emparejamiento en SIMAT, que es el mismo año en el que fueron beneficiarios o el primer año en el que aparecían en SIMAT en caso de que el año del programa Ondas fuera anterior. La edad promedio de los NNA en la muestra de evaluación es de 16,4 años en 2016, el 47% es hombre, el 76% reside en zona urbana, el 94% no pertenece a ningún grupo étnico, el 13% estudia en jornada completa, y el 23% fue beneficiario en el departamento de Bolívar.

5.2. Resultados

Las pruebas de diferencias de medias de las variables de resultado de la Tabla 8 revelan que los NNA vinculados al programa Ondas tienen menor tasa de deserción escolar que los controles (1,4% vs 1,96%), así como menor tasa de repitencia (10,5% vs. 13,1%). En cuanto a la graduación de educación superior, se observa que el porcentaje de jóvenes entre 20 y 24 años que se graduaron entre 2020 y 2022 es significativamente más alto para los controles que para los tratados (3,5% vs. 3,3%, respectivamente).

Tabla 8. Prueba de diferencia de medias de las variables de resultado en deserción, repitencia y graduación de la educación superior

Variable	Controles	Tratados	Diferencia	Error Est.
Proporción de NNA que desertan temporalmente	0,01966	0,01405	0,01***	0,00
Proporción de NNA que repiten algún grado escolar	0,1306	0,1053	0,03***	0,00
Proporción de jóvenes de 20 a 24 años graduados de educación superior (20-22)	0,0354	0,0331	0,00*	0,00

Fuente: Elaboración propia con información de la base de datos de Ondas y el cruce con SIMAT y SNIES

Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente

Ahora, se emplea la metodología de PSM para emparejar los grupos de tratamiento y control y validar si estas diferencias de medias se mantienen una vez se garantiza que los grupos de tratamiento y control son similares en las características observables empleadas. La Tabla 9 cuantifica las diferencias en las características observadas entre los grupos de tratamiento y control, empleando la diferencia estandarizada para cada una de estas variables; esto, tanto para la muestra sin emparejar y ya emparejada. Las diferencias estandarizadas se reducen considerablemente después de hacer el emparejamiento y fluctúan entre –0,03 y 0,05 desviaciones estándar. De hecho, en la muestra sin emparejar solo hay una covariable con diferencia estandarizada mayor a 0,25, que es el valor crítico definido en la literatura para diferencias muy altas (Imbens y Rubin, 2015). Se destaca también que en la muestra emparejada se logra hacer llegar las razones de varianza a valores cercanos a uno (promedio 1,06 para todas las covariables).

Tabla 9. Diferencias estandarizadas entre las características observadas de tratados y controles antes y después del emparejamiento

Variable	Media estandarizada		Razón de varianzas	
	Crudo	Emparejado	Crudo	Emparejado
Antioquia	-0,20	0,00	0,44	0,99
Atlántico	-0,09	-0,01	0,62	0,96
Bogotá	-0,22	-0,01	0,39	0,95
Bolívar	0,47	-0,01	1,57	1,00

Variable	Media estandarizada		Razón de varianzas	
	Crudo	Emparejado	Crudo	Emparejado
Boyacá	-0,09	0,00	0,48	1,00
Caldas	-0,05	0,01	0,42	1,25
Caquetá	-0,02	0,00	0,83	0,98
Cauca	-0,09	0,00	0,38	0,98
Cesar	-0,10	0,00	0,49	0,98
Córdoba	-0,10	0,00	0,67	1,00
Cundinamarca	-0,07	0,00	0,65	0,99
Chocó	-0,05	0,00	0,40	0,95
Huila	-0,09	-0,01	0,47	0,93
La Guajira	-0,06	0,00	0,27	1,00
Magdalena	-0,04	0,01	0,77	1,04
Meta	-0,10	0,00	0,45	1,04
Nariño	-0,03	0,01	0,79	1,05
Norte de Santander	-0,05	0,00	0,77	1,00
Quindío	-0,08	0,00	0,43	1,05
Risaralda	-0,06	0,00	0,64	1,02
Santander	0,08	-0,01	1,22	0,98
Sucre	0,04	0,01	1,24	1,07
Tolima	-0,06	0,01	0,57	1,16
Valle del Cauca	-0,12	0,00	0,26	0,96
Arauca	0,18	0,00	1,97	1,00
Casanare	-0,11	0,00	0,23	0,95
Putumayo	-0,04	0,00	0,65	1,01
San Andrés	-0,03	0,01	0,54	1,21
Amazonas	-0,01	0,02	0,94	1,15
Guainía	-0,01	0,00	0,86	1,05
Guaviare	-0,05	0,01	0,53	1,26
Vaupés	0,01	0,01	1,18	1,11
Participó en el Programa Ondas en 2006	-0,05	0,00	0,32	1,10
Participó en el Programa Ondas en 2009	-0,24	0,01	0,38	1,03
Participó en el Programa Ondas en 2010	-0,04	0,00	0,93	1,00
Participó en el Programa Ondas en 2011	-0,07	-0,01	0,86	0,99
Participó en el Programa Ondas en 2012	-0,03	0,00	0,96	1,00
Edad en 2016	0,10	0,00	0,85	0,98
Variable =1 si el NNA es hombre	0,06	0,01	1,00	1,00
Estrato	-0,06	0,01	1,01	1,05
NNA tiene capacidades excepcionales	-0,01	0,01	0,74	1,52
Variable=1 si el estudiante presenta algún tipo de discapacidad	0,01	0,05	1,05	1,60
NNA víctima del conflicto	0,00	0,03	0,99	1,15
NNA reside en zona urbana	0,02	-0,02	0,98	1,02
NNA no pertenece a un grupo étnico	0,02	-0,02	0,94	1,08
NNA indígena	-0,03	0,01	0,86	1,07
NNA es NARP	0,01	0,02	1,09	1,12
Grado	0,11	0,00	0,90	1,01
Ciclo adultos	-0,12	0,01	0,42	1,05
Variable =1 si el estudiante está en un colegio que ofrece metodología de educación tradicional	0,09	-0,03	0,77	1,11
Variable =1 si el estudiante está en un colegio que ofrece programa para jóvenes extra edad	-0,10	0,00	0,44	1,04
Variable =1 si el estudiante está en un colegio que ofrece metodología escuela nueva	-0,04	0,02	0,76	1,15
Variable =1 si el estudiante está en un colegio que ofrece metodología postprimaria	0,02	0,02	1,15	1,16
Variable =1 si el estudiante está en un colegio que ofrece metodología de educación	-0,01	0,00	0,95	1,03
NNA estudia en jornada completa o única	-0,04	0,03	0,92	1,07

Variable	Media estandarizada		Razón de varianzas	
	Crudo	Emparejado	Crudo	Emparejado
NNA estudia en jornada de la mañana	0,04	-0,01	0,98	1,01
NNA estudia en jornada de la tarde	0,02	-0,01	1,02	0,98

Fuente: Elaboración propia con información de la base de datos del Programa Ondas y el cruce con SIMAT y SNIES

Se encuentran impactos positivos y significativos en menor deserción y en repitencia, más no en graduación de la educación superior. Específicamente, el programa genera una reducción de 0,18 puntos porcentuales en el porcentaje de NNA que desertan en algún momento de la educación básica y media. Así mismo, el programa genera reducciones de 1,76 puntos porcentuales en el porcentaje de NNA que han repetido algún año (ver Tabla 10). No se encontraron efectos en la variable de graduación de la educación superior, si bien hay que tener en cuenta las limitaciones para la construcción de la variable, tales como la población de referencia.

Tabla 10. Impactos estimados del Programa Ondas en deserción, repitencia y graduación de la educación superior

Variable	ATT	Error estándar	p-valor
Proporción de NNA que desertan temporalmente	-0,0018**	0,0008	0,021
Proporción de NNA que repiten algún grado escolar	-0,0176***	0,0019	0,000
Proporción de jóvenes de 20 a 24 años graduados de educación superior (20-22)	0,0006	0,0016	0,709

Fuente: Elaboración propia con información de la base de datos de Ondas y el cruce con SIMAT y SNIES

Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente

Los hallazgos cualitativos refuerzan que el programa Ondas puede influir positivamente en la motivación y en la curiosidad de los estudiantes por el aprendizaje. Esto pudiera traducirse en los efectos que se observaron en menores tasas de deserción y repitencia en la educación básica y media. El programa Ondas representa un espacio formativo diferente a la escuela tradicional, donde la motivación no depende de una calificación o de la obligación de entregar un trabajo, sino del entusiasmo y el interés de los NNA por aprender investigando. Además, el programa Ondas genera otras dinámicas adicionales a las escolares para sus beneficiarios, en la medida en que les permite visitar universidades u otros espacios académicos diferentes de sus colegios o de su contexto inmediato, o la oportunidad de participar en los encuentros.

"llevar a Ondas al currículum y a la evaluación del aprendizaje cambia completamente. […] ¿por qué? Porque a nivel cognitivo, incluso psicológico el niño no está sometido a una evaluación, a un seguimiento, a toda esta rigurosidad que le pone el sistema educativo, sino está involucrado es una actividad didáctica optativa que así como se metió al grupo de básquet o se metió al coro del colegio, pues se metió al semillero. Entonces ahí como que no hay esa presión, nadie le está pidiendo que al final debe mostrar una competencia específica a desarrollar, entonces cambia completamente"

(Entrevista a coordinador de entidad coordinadora, 2023)

En la información cualitativa se destacó que el programa genera cambios en los proyectos de vida de los NNA vinculados a este; y es que les permite explorar opciones de vida diferentes a las que habían visto en su familia y en su contexto cercano. Esto es muy notable, sobre todo, en zonas rurales donde hay un mayor campo de oportunidad para generar esos cambios y transformaciones de vida. Estos cambios en los proyectos de vida pueden influir entonces en la motivación de los NNA vinculados al programa Ondas por terminar su educación media y desertar menos, e incluso considerar la educación superior como un propósito alcanzable.

"estos programas despiertan a que el estudiante cambie su proyecto de vida, a que no se conforme con un bachillerato, a que no se conforme a ganar un día a día, sino que se pueda proyectar como un profesional universitario y que pueda mejorar la calidad de vida de su entorno familiar"

(Entrevista a docente, 2023)

"mis compañeros de colegio y mis amigos teníamos muy pocas aspiraciones, nuestra aspiración era graduarnos e irnos del colegio a trabajar. Y llegan estas profes con esas ideas locas y nos meten en la cabeza que se pueden hacer cosas increíbles. Desde cuando estábamos dentro del colegio lo veíamos como algo chévere porque era esta actividad social y tenía esa relación con ciencia y nos parecía interesante"

(Entrevista a beneficiario graduado, 2023)

"en la parte rural siempre se les preguntaba qué querían ser, entonces las niñas decían, yo voy a quedarme trabajando, cocinando, labreros para ayudar a mi papá o a mi mamá en la siembra. Hoy en día ya se ven a los chicos y chicas que ya piensan en ser doctores, en ser científicos, en estudiar, en aprender"

(Entrevista a asesora de línea de entidad coordinadora, 2023)

Aunque no se observaron impactos en la graduación de educación superior, el programa puede influenciar las aspiraciones de los NNA e incentivarlos a alcanzar este nivel educativo, así sea difícil por las barreras estructurales para acceder a ella. Dejando de lado los problemas de accesibilidad que son ajenos al programa. Por otro lado, el programa Ondas tiene también el potencial de ayudar a los NNA que acceden a educación superior a tener mejores resultados en su formación, en la medida en que se generan habilidades que les pueden servir para triunfar en el pregrado, tales como el pensamiento crítico, la capacidad de abordar problemas complejos, el trabajo en equipo, las habilidades comunicativas y la generación de confianza.

"hay habilidades que los chicos pueden desarrollar a temprana edad. Entonces, ¿Qué pasa en las universidades? Lo digo porque yo soy profesora de matemáticas. Los chicos cuando llegan a la universidad, tienen muchísimas falencias. [..] Cuando llegan a la universidad, el problema ya está armado, ya aquí no lo podemos solucionar, entonces por eso es importante la intervención a temprana edad"

(Entrevista a coordinadora de entidad coordinadora, 2023)

"Esa es la mayor ventaja de Ondas. Y los muchachos de bachillerato que han estado en proyectos Ondas, van a tener una facilidad mucho más que los que no han participado, o sea van a tener unas bases significativas para afrontar la universidad y enfrentarse a los procesos de investigación universitaria. O sea, ya llevan un terreno ganado"

(Entrevista a docente, 2023)

"Yo recuerdo que en las primeras materias del primer semestre la mayoría de cosas que yo estaba viendo en las clases por ejemplo biología o química, ya las había visto en el Club o en la misma especialidad"

(Entrevista a beneficiario graduado, 2023)

6. Resultados de la evaluación a nivel de colegios

6.1. Descripción de la muestra de evaluación

La unidad de análisis para esta evaluación es el establecimiento educativo, por lo que se agregó la información de la base de datos de NNA vinculados al programa Ondas a dicho nivel. Se consolidó una base de datos de 9.285 establecimientos educativos para todo el período de tiempo (ver *Tabla 11*). Es importante aclarar que estos establecimientos educativos pudieron haber desarrollado el programa en varias ocasiones. El año 2022 reporta el mayor número de establecimientos educativos que desarrollan el Programa Ondas, con un total de 1.092 (11,3%), seguido del año 2009 con 1.051 establecimientos educativos (11,8%). Por el contrario, entre 2001-2008 se cuenta con información de menos de 200 establecimientos educativos que desarrollaron el programa, así como en 2017 que hubo 10 y en 2023 que hubo 76.

Tabla 11. Número de establecimientos educativos que desarrollan el programa Ondas por año

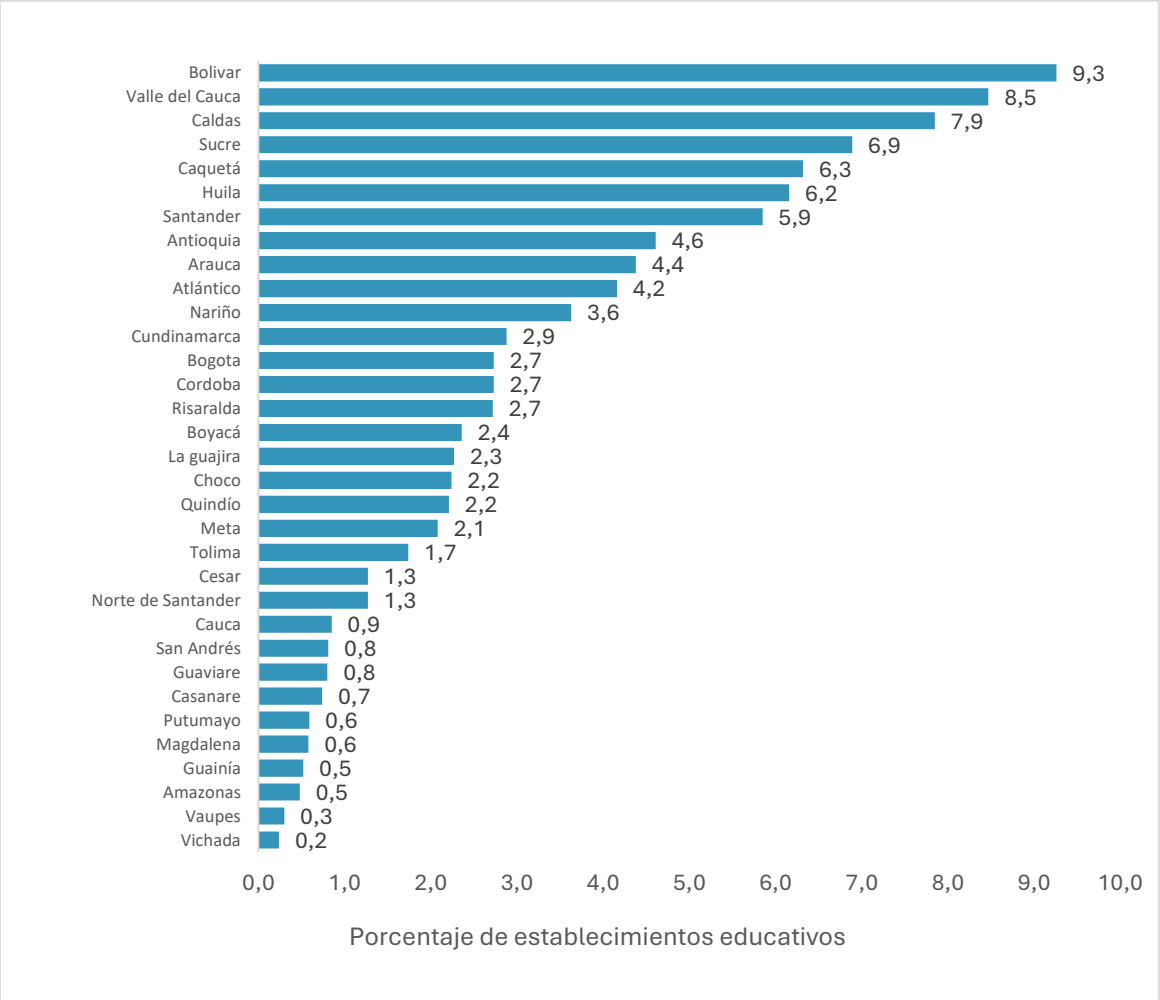
Año	Número	Porcentaje
Sin información	153	1,65
2001-2005	88	0,95
2002	56	0,6
2003	129	1,39
2004	135	1,45
2004-2005	73	0,79
2005	143	1,54
2005-2006	20	0,22
2006	26	0,28
2007	72	0,78
2008	20	0,22
2009	1.051	11,32
2010	551	5,93
2011	737	7,94
2012	931	10,03
2013	605	6,52
2014	209	2,25
2015	605	6,52
2016	378	4,07
2017	10	0,11
2018	155	1,67
2019	510	5,49
2020	858	9,24
2021	602	6,48
2022	1.092	11,76
2023	76	0,82
Total	**9.285**	**100**

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados al programa Ondas.

Los establecimientos educativos que desarrollan el Programa Ondas se concentran principalmente en los departamentos de Bolívar (9,3%), Valle del Cauca (8,5%), Caldas (7,9%), y Sucre (6,9%); si bien, tienen presencia en todos los departamentos del país (ver Tabla 1). De otro lado, se destaca que los establecimientos educativos que desarrollan el Programa Ondas tienen un promedio de 39,3 NNA vinculados al programa (17,9 hombres y 9,3 mujeres). Ahora bien, existe mucha variabilidad en el promedio de NNA vinculados al Programa Ondas por año en el establecimiento educativo, con años

como 2003 donde hubo 5,5 NNA por establecimiento educativo, y otros como 2013 con 82,6 NNA en promedio.

Figura 9. Porcentaje establecimientos educativos vinculados al programa Ondas



Fuente: Elaboración propia con base en las bases de datos de NNA vinculados al Programa Ondas.

Ahora bien, tal como se menciona en el Anexo 1, el 92% de los NNA vinculados al programa Ondas no tenía el código del establecimiento. Esta variable es fundamental para poder cruzar la información del Programa Ondas con las bases de datos de las pruebas Saber. Por tanto, se hizo un cruce con la base de datos de SINEB para poder obtener el código del establecimiento y eliminar duplicidades en establecimientos educativos y unificar registros que, pese a presentar ligeras variaciones en cómo se escribe el nombre entre diferentes años, corresponden al mismo establecimiento.[17]

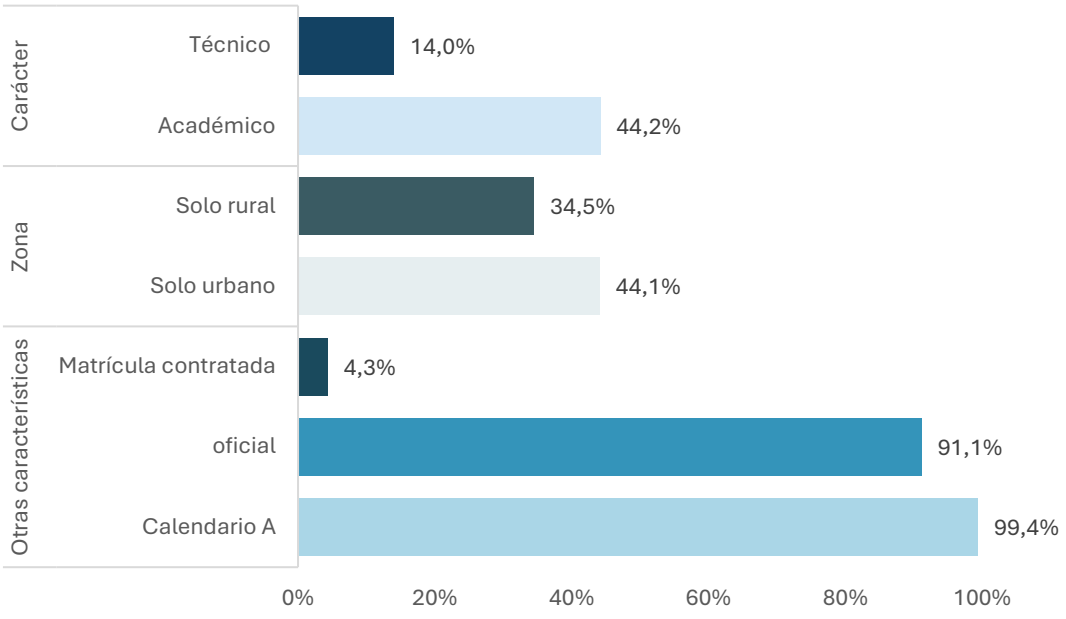
Como resultado de este cruce, se encontraron 5.985 establecimientos educativos en la base de datos de SINEB, equivalente al 64% del total de establecimientos educativos de la base de datos nacional del Programa Ondas. Estos establecimientos educativos corresponden a 3.160 observaciones únicas que componen la muestra de evaluación, ya que varios de estos establecimientos educativos desarrollaron el programa varios años. De hecho, el 51,4% de los establecimientos educativos tratados de la muestra de evaluación desarrollaron el programa Ondas solo una vez, el 25,5% participó dos veces, el 12,5% participó tres veces, el 6% participó cuatro veces y el resto de los establecimientos

[17] En la práctica, se puede hacer directamente este cruce con la base de datos de Saber que también tiene nombre y código DANE de establecimiento, pero se recuperaron más códigos empleando el SINEB.

desarrollaron el programa Ondas entre cinco y ocho ocasiones. En total, estos establecimientos educativos abarcan a 280.892 NNA vinculados al programa Ondas lo que corresponde al 75,7% del total, por lo que la muestra de evaluación recoge un porcentaje muy importante de NNA vinculados al programa.

Así, la muestra de 3.160 establecimientos educativos que conforma el grupo de tratamiento corresponde a instituciones del sector oficial (el 91,1%) y de calendario A (el 99,4%). La mayoría de los establecimientos educativos son de carácter académico (el 44,2%) y el 14% son de carácter técnico. El 44,1% de los establecimientos educativos son solamente urbanos y el 34,5% son únicamente rurales; el resto de los establecimientos se clasifican tanto urbanos como rurales por las sedes con las que cuentan (ver Figura 10).

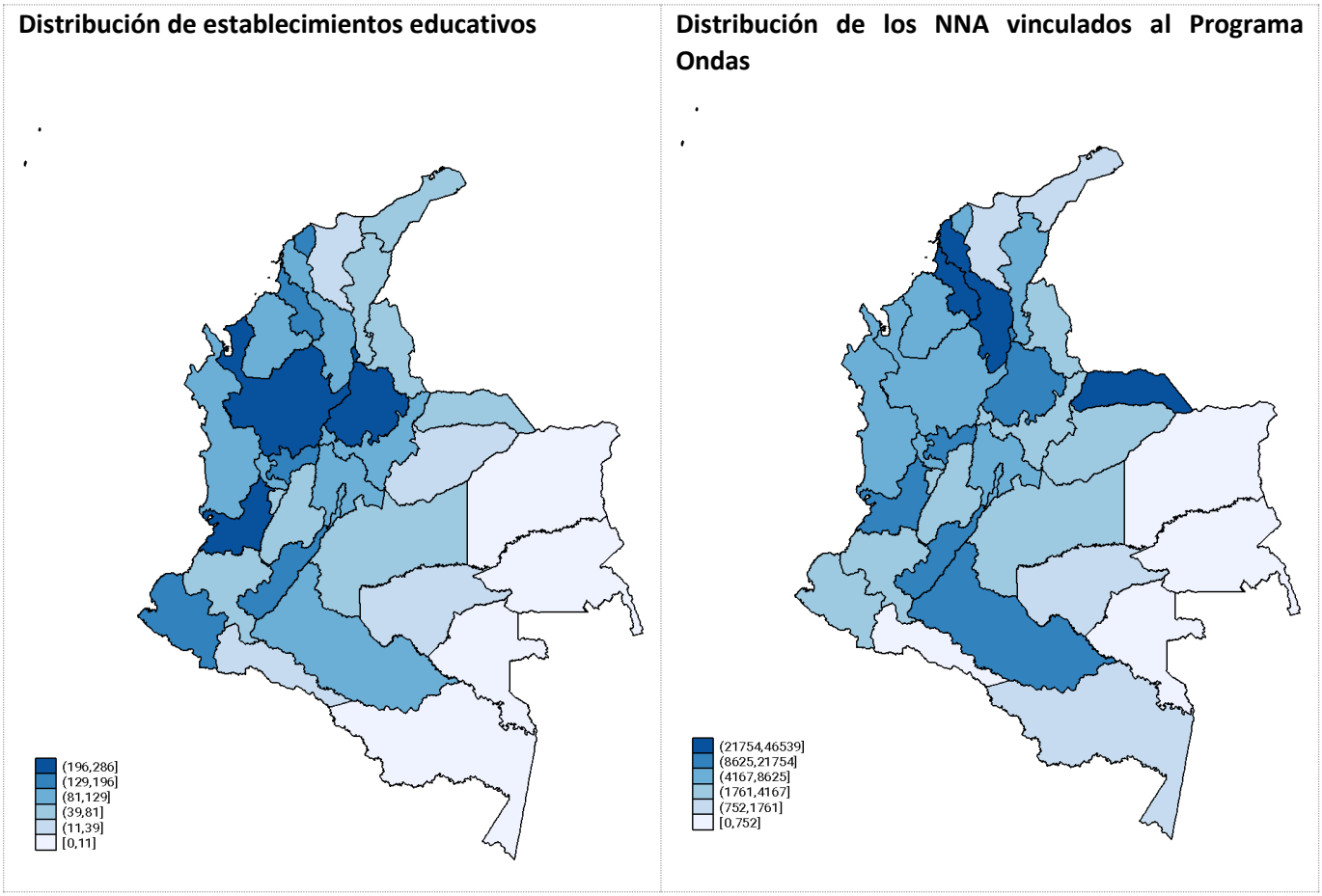
Figura 10. Características de los establecimientos educativos que desarrollan el Programa Ondas y que quedaron en la muestra de evaluación



Fuente: Elaboración propia con información de la muestra de evaluación.

Valle del Cauca es el departamento con mayor número de establecimientos educativos en la muestra de tratados para la evaluación, seguido de Santander y Antioquia, todos ellos con más de 200 establecimientos educativos que alguna vez desarrollaron el programa Ondas (ver Figura 10, panel A). Sin embargo, al observar el número de estudiantes en los establecimientos educativos de la muestra de evaluación, los departamentos con mayor número de NNA vinculados al Programa Ondas son Sucre, Bolívar, Arauca, Santander y Caquetá, con más de 20.000 NNA vinculados al programa Ondas cada uno durante todo el período de evaluación (ver Figura 11, panel B).

Figura 11. Distribución de los establecimientos educativos tratados en la muestra de evaluación por departamento y número de NNA que hacen parte de ellos



Fuente: Elaboración propia con información de la muestra de evaluación.

Esta base de datos de establecimientos educativos se cruzó con las bases de datos de Saber para determinar los grupos de tratamiento y control finales según la variable de resultado. En el Anexo 1 se detallan todos los procedimientos implementados.

6.2. Resultados

Los resultados preliminares de las estimaciones de impactos en calidad educativa de los establecimientos educativos que desarrollan el programa Ondas son positivos para los resultados de las pruebas Saber de matemáticas de quinto grado. El porcentaje de estudiantes con puntaje satisfactorio o avanzado en la prueba de matemáticas es 4,15 puntos porcentuales más alto para los establecimientos educativos con NNA vinculados al Programa Ondas que para el grupo de control (ver Tabla 12). Los resultados se observan sobre todo en el segundo y tercer año posteriores a la vinculación de los establecimientos educativos al programa, con impactos de 9 y 3,4 puntos porcentuales, respectivamente (ver Figura 30 del Anexo). Cuando se toma como grupo de tratamiento solo a los establecimientos educativos con NNA que estaban en cuarto y quinto en el año en que se vincularon al programa Ondas, que son quienes efectivamente presentan la prueba, los resultados son algo más de 2 puntos porcentuales más altos y se observan desde el mismo año de vinculación al programa y hasta el tercer año luego de haber estado vinculado. En promedio, los establecimientos educativos que desarrollan el programa Ondas y que tienen niños en cuarto y quinto grado

incrementan en 6,55 puntos porcentuales el porcentaje de niños con puntaje satisfactorio o avanzado en las pruebas Saber 5 de matemáticas.

No se encontraron impactos en los resultados agregados de las pruebas Saber de Lenguaje, ni para quinto ni para noveno grado, incluso cuando se restringe la muestra de establecimientos educativos a aquellos que contaban con NNA vinculados al programa Ondas cercanos a la presentación de la prueba. Tampoco se encontraron impactos en el resultado agregado de la prueba Saber de matemáticas de noveno grado.

Tabla 12.Impacto en pruebas Saber 5 y 9 agregadas a nivel de establecimiento educativo

Porcentaje de niños con puntaje satisfactorio o avanzado en el establecimiento educativo	**Lenguaje- ATT**	**Matemáticas ATT**
Quinto grado	1,40	**4,15*** **
Quinto grado – sólo se incluyen establecimientos educativos con NNA vinculados al Programa Ondas en 4º y 5º	1,40	**6,55*** **
Noveno grado	0,76	0,79
Noveno grado - solo se incluyen establecimientos educativos con NNA vinculados al Programa Ondas entre 6º y 9º	1,02	1,05

Fuente: Elaboración propia con información de las bases de datos del Programa Ondas y Saber 3, 5,9. Se emplean controles de naturaleza del establecimiento educativo, zona, carácter, si es matrícula contratada, calendario A y departamento. *** significativo al 1%.

Ahora bien, cuando se analizan los resultados agregados en pruebas Saber 11 no se encuentran impactos significativos para el índice total del establecimiento educativo ni para índice por área (ver Tabla 13 y Figura 31). Incluso, cuando se utilizan como tratados sólo los establecimientos educativos que tuvieron NNA en décimo y undécimo cuando estaban vinculados al Programa Ondas (ver Figura 32).

Tabla 13. Impacto en pruebas 11 agregadas a nivel de establecimiento educativo

Prueba	**ATT**	**Establecimientos educativos con NNA vinculados al Programa Ondas en 10 y 11 - ATT**
Matemáticas	0,00	0,00
Ciencias naturales	-0,01	-0,01
Lectura	0,00	0,00

Fuente: Elaboración propia con información de las bases de datos del Programa Ondas y Saber 11. Se emplean controles de naturaleza del establecimiento educativo, zona, carácter, si es matrícula contratada, calendario A y departamento. ** Significativo al 5%.

Los resultados en calidad de la educación básica y media, medida a través de los puntajes en las pruebas Saber de los establecimientos educativos, no son contundentes. Sin embargo, es importante tener en cuenta que la esencia del aporte del Programa Ondas al sistema escolar tiene que ver con la formación de capacidades investigativas y vocaciones científicas en los NNA. De la información

cualitativa recabada se destaca que el programa Ondas tiene incidencia en la generación de capacidades de pensamiento crítico y resolución de problemas de los NNA vinculados a él. La generación de la idea de investigación y el desarrollo de la misma hace que los NNA desarrollen habilidades para discernir información, comparar puntos de vista, sopesar opiniones y argumentar con base en la evidencia con la que cuentan. Más aún, el programa permite que los NNA se apropien de conceptos que vienen del ámbito escolar y que los asocien con su investigación, lo que les permite aplicar los conocimientos adquiridos de una forma práctica y contextualizada, facilitando su comprensión. De ahí que el programa Ondas permite enriquecer la experiencia educativa de los NNA.

"las variables de interacción que ofrece el programa de Ondas y la metodología da unas capacidades en el desarrollo cognitivo, en el desarrollo académico de los muchachos que intervienen en el programa, mejorando como una consecuencia casi que a la par en la percepción de su rol en la sociedad, de su rol en el mundo, de ser más críticos y más sensibles a los problemas del contexto de nuestra comunidad, de nuestra sociedad, del departamento"

(Entrevista a funcionario de gobierno local, 2023)

"Cuando yo veo un grupo que trabaja Nacho Derecho de primaria y ese grupo tenía unos niños que tenían que salir de la casa a las 5 de la mañana… cuando llegaban al colegio ya habían repartido el desayuno. Y a través de Nacho Derecho ellos entendieron que tenían el derecho a la alimentación, pidieron la cita en el Concejo e hicieron que cambiaran la hora que servían el desayuno para ellos tener el derecho a la alimentación. Esos niños ya quedan transformados en su pensamiento y en su razonamiento de cómo son las vías por las cuales yo puedo actuar para que mi calidad de vida pueda ser mejor"

(Entrevista a coordinadora de entidad coordinadora, 2023)

Si bien se encontraron impactos positivos en las pruebas Saber de matemáticas de quinto grado, la falta de efectos en otras áreas y niveles es solo el reflejo de que el programa puede ser pequeño en una institución educativa como para generar impactos agregados, así como de la ausencia de efectos derrame sobre los NNA que no se vinculan al Programa Ondas. Aunque mejorar la calidad de la educación no es un objetivo directo del Programa Ondas, el programa sí tiene el potencial de lograrlo a través de la generación de capacidades para la investigación, formulación de preguntas, pensamiento crítico, análisis de información, resolución de problemas, entre otras. Ahora bien, de la información cualitativa surgieron algunas barreras a la articulación entre el MEN y el programa que pueden limitar los impactos del Programa Ondas en la calidad de la educación. Uno de los problemas reportados es que los docentes participan en el programa como parte de una actividad extracurricular, que no implica reducción de horas de las actividades docentes o cambios en la asignación de trabajo para destinar tiempo a las sesiones de asesoría y a las reuniones propias del grupo de investigación. Los docentes también reportaron la sensación de que el trabajo que realizan en el Programa Ondas no es reconocido. Por lo que ambos problemas pudieran estar desincentivando la participación de los docentes y dificultando el desarrollo del programa como un espacio de investigación formativa.

"yo soy de los que muchas veces sábados y domingos estoy con los muchachos, y lo que hago es que me llevo la familia mía para el acompañamiento"

(Entrevista con docente, 2023)

7. Conclusiones

El Programa Ondas opera desde 2001 y ha llegado a todos los departamentos del país. Sin embargo, su alcance es limitado ya que su cobertura no llega al 1% de los estudiantes matriculados en educación preescolar, básica, media y en ciclos lectivos en cada año. Sin embargo, es notable que llegue a regiones menos prósperas y con menor nivel de desarrollo en términos de ciencia y tecnología.

En cuanto a los impactos del programa, se encontró que participar en el programa Ondas genera reducciones en deserción escolar y repitencia. Estos impactos pueden ser el resultado de que los NNA estén más motivados e interesados en su proceso educativo, como consecuencia directa del programa. Los hallazgos cualitativos muestran que el programa Ondas fomenta la motivación intrínseca en los NNA de un modo diferente a lo que experimentan en su establecimiento educativo, puesto que cobra importancia el interés y el entusiasmo por el desarrollo de los proyectos, empoderando a los NNA de su proceso de aprender investigando. Más aún, en el marco del programa Ondas se generan espacios que permiten a los NNA explotar su curiosidad, indagar y explorar los problemas de su entorno y fomentar su creatividad.

Respecto a los resultados en calidad de la educación, es importante tener en cuenta que no es un objetivo per se del programa, pero puede incidir indirectamente. Específicamente, el programa facilita el desarrollo de habilidades como el pensamiento crítico, la capacidad para la resolución de problemas, la formulación de preguntas, el diseño de experimentos o el análisis de información. Estos, a su vez, pueden influir en mejores resultados en pruebas estandarizadas de educación. Al analizar los impactos del programa Ondas en la calidad de la educación se encontraron impactos positivos en el resultado agregado a nivel de establecimiento educativo de las pruebas Saber de 5to grado en matemáticas, pero no en lenguaje o en ciencias naturales. Tampoco se encontraron impactos en los resultados en las pruebas Saber en noveno grado o en 11.

Los impactos del programa Ondas van más allá de las variables que se pueden medir con registros administrativos. En las entrevistas y grupos focales se destacó la generación de habilidades para la investigación y habilidades blandas como un resultado clave del programa, esto con el potencial de afectar no solo el desempeño académico actual de los NNA sino en el futuro en la educación superior. Un aspecto aún más significativo es el efecto en los proyectos de vida de los NNA, ampliando sus posibilidades y permitiéndoles explorar opciones de vida diferentes a las que su contexto les delimita.

Por último, hay otros efectos intangibles del Programa Ondas relacionados con la democratización del acceso al conocimiento. El Programa Ondas permite ampliar las capacidades científicas de los NNA y fomentar una ciudadanía crítica y consciente del valor del conocimiento científico, quienes además pueden utilizarlo sin importar el rol futuro que desempeñen en la sociedad.

"van a tener una incidencia en la sociedad como ciudadanos, entonces nosotros decimos, si estos niños o niñas [así no tomen] el camino de la ciencia se van a convertir en líderes, empresarios, incluso empleados o gobernantes que van a entender el uso y a valorar la importancia de la ciencia y el conocimiento científico para la toma de decisiones. Incluso yo te lo digo en temas rurales, el porcentaje de estudiantes de zonas rurales que logran ingresar a la universidad es muy bajo, pero el porcentaje de estudiantes rurales de Ondas que aprenden a investigar, transfieren automáticamente sus conocimientos a la producción agrícola por ejemplo"

(Entrevista a coordinador de entidad coordinadora, 2023)

Programa Jóvenes Investigadores e Innovadores

1. Introducción

El programa Jóvenes Investigadores e Innovadores (JII) es una iniciativa estratégica del país para apoyar el fortalecimiento de las capacidades de ciencia, tecnología e innovación, especialmente en la población joven. El programa tiene el propósito de acercar a los jóvenes con la investigación y la innovación y promover su integración en los grupos de investigación reconocidos dentro del Sistema Nacional de Ciencia, Tecnología e Innovación. Al mismo tiempo, apoya a los grupos de investigación a través de la incorporación de talento joven, facilitando su formación en investigación y participación en dinámicas y redes de conocimiento especializadas.

Este documento presenta los resultados de la evaluación de impacto del programa. Para ello, primero se hace una caracterización de los JII, la cual arroja que la mayoría de ellos nació en uno de los tres principales centros urbanos: Bogotá, Medellín y Cali; con cerca de un quinto de población que nació en el primero. Sin embargo, se destaca que hay amplia dispersión en los lugares de nacimiento de los jóvenes investigadores e innovadores y que algunos provienen de municipios principalmente rurales y dispersos (el 3,6%). Más de la mitad de los jóvenes investigadores son mujeres y tienen una edad promedio de alrededor de 24 años. Por área de conocimiento, casi un tercio de los jóvenes investigadores está vinculado a grupos de ingeniería y tecnología, en tanto que un porcentaje menor se concentra en áreas de Ciencias agrícolas (6,8%) y humanidades (5%).

Se empleó el método de diferencias (DID) con múltiples períodos y múltiples tratamientos para estimar los impactos en la producción de nuevo conocimiento y los resultados del mercado laboral de los jóvenes. Dada la orientación del programa hacia la generación de capacidades vocacionales e investigativas y a la inmersión en entornos de investigación, se espera que active la generación de nuevo conocimiento, evidenciada en la producción de artículos científicos y libros. De otro lado, dada la experiencia práctica que adquieren los participantes y la exposición a redes profesionales, es razonable esperar que incremente su empleabilidad, su estabilidad laboral y los ingresos que reciben. También se incluyeron algunos resultados del levantamiento de información cualitativa que se realizó a partir de entrevistas a directores de los grupos de investigación, a personas que hubieran estado vinculadas al programa, y a funcionarios de centros o institutos de investigación, unidades de I+D+i, centros de innovación y productividad, centros de desarrollo tecnológico y empresas altamente innovadoras.

Los resultados preliminares muestran impactos positivos de JII en la publicación de artículos, especialmente artículos A1 y A2; aunque al desagregar la población según el género, los resultados solo se mantienen para los hombres. También se encontraron impactos positivos en salarios –medidos por el Ingreso Base de Cotización (IBC) y en el tiempo que cotizan los jóvenes a seguridad social en cada año.

Este documento cuenta con cinco secciones además de esta introducción. En la siguiente sección se presentan unas características generales de los JII para contextualizar el programa que se va a evaluar. En la tercera sección se describe con detalle la metodología para la evaluación de impacto y en la cuarta se presentan los resultados preliminares. La quinta sección explica el análisis costo-beneficio del programa y, en la última parte del documento, se presentan las principales conclusiones.

2. Contexto de la implementación del programa JII

El programa JII fue creado en 1995 con el objetivo de facilitar la inserción de jóvenes profesionales a grupos de investigación y centros de investigación y desarrollo tecnológico, mediante el establecimiento de vínculos que permitan promover la apropiación de los métodos y los conocimientos del quehacer científico y tecnológico, orientados por la metodología de "aprender haciendo". A partir de esta metodología, los jóvenes desarrollan sus capacidades para la ciencia, la tecnología y la innovación a través de su vinculación a grupos de investigación mediante el desarrollo de una beca – pasantía. El programa JII focaliza a profesionales recién egresados o aquéllos que solo tienen pendiente la ceremonia de grado, que se han caracterizado por su buen desempeño académico, y que quieren iniciar una carrera en investigación trabajando en grupos o unidades de investigación con el acompañamiento de un tutor.

El Programa JII busca entonces formar investigadores de excelencia con el fin de incrementar la capacidad nacional para competir con los más altos estándares internacionales, prioritariamente en aquellas áreas señaladas como estratégicas para el país. El programa JII también puede aportar en el relevo generacional de los investigadores en el país.

A lo largo de la historia del programa JII, la normatividad asociada a la política de Ciencia, Tecnología e Innovación ha reconocido la importancia del programa como un medio para el acercamiento de jóvenes recién egresados a la innovación tecnológica. Así, en 2008 se formuló la Política Nacional de Fomento a la Investigación y la Innovación, titulada *Colombia construye y siembra futuro,* la cual buscaba crear condiciones para que el conocimiento fuese un instrumento del desarrollo económico del país. Este documento definía metas con respecto a la aceleración del crecimiento económico y a la disminución de la inequidad por medio de la investigación y la innovación en los principales problemas sociales del país, explorando alternativas a corto, mediano y largo plazo. Además, se estableció que Colciencias –hoy MinCiencias desarrollaría el Programa Nacional de Formación de Investigadores, orientado a fortalecer el desarrollo de las áreas prioritarias definidas en la política y en los planes y programas de CTeI; uno de los ejes centrales de la formación de investigadores estaba constituido por el apoyo a la formación de JII (Colciencias, 2008).

Posteriormente, la Ley 1286 de 2009 -que derogó parcialmente la Ley 29 de 1990- transformó a Colciencias (hoy MinCiencias) en Departamento Administrativo y fortaleció el SNCTeI. Allí, se estableció como uno de los objetivos generales de Colciencias fortalecer la capacidad científica del país y promover la formación de investigadores, haciendo especial énfasis en programas que impactaran sectores estratégicos para la transformación del país. En ese contexto, el programa JII continuó siendo un eje fundamental en la formación de investigadores.

En 2009, el Documento CONPES 3582 *Política Nacional de Ciencia, Tecnología e Innovación*, declaró de importancia estratégica para el país el proyecto de inversión Capacitación de Recursos Humanos para la Investigación para implementar la estrategia de apoyo a la formación para la I+D+i. Este documento realizó un diagnóstico del SNCTeI e identificó seis grandes limitantes: i) bajos niveles de innovación de las empresas; ii) débil institucionalidad del sistema; iii) escasez de recurso humano para realizar investigación e innovación; iv) ausencia de focalización de la política en áreas estratégicas; v) baja apropiación social del conocimiento y disparidades regionales en capacidades científicas y tecnológicas; lo que en conjunto genera una vi) baja capacidad para generar y usar conocimiento. De ahí que una de las estrategias fundamentales en este CONPES fuera la formación de JII como un medio esencial para dar un salto en la formación de investigadores.

Más recientemente, el Plan Nacional de Desarrollo 2018-2022 ratificó la necesidad de contar con formación y vinculación de capital humano altamente calificado como un medio para incrementar la producción de nuevo conocimiento y el desarrollo tecnológico, lo que redundaría en incrementos en productividad y crecimiento económico. Específicamente, el Pacto por la Ciencia, la Tecnología y la Innovación se definió como una de las metas duplicar el número de participantes en el programa JII.

Luego, la creación del Ministerio de Ciencia, Tecnología e Innovación en 2019 buscaría una mayor promoción y fortalecimiento del sector de CTeI y, por ende, de los programas que venían funcionando – como el programa JII, como elementos esenciales para el fortalecimiento de la competitividad del país. Es importante destacar que el Plan Estratégico Institucional y el Plan de Acción Institucional de MinCiencias establecen la importancia de fomentar una cultura alrededor de la ciencia y la tecnología para el fortalecimiento de las vocaciones científicas en los territorios. Aunque estos documentos no sean leyes o normas en sentido estricto, también orientan las acciones del Ministerio en estos aspectos y define actividades específicas para el logro de las metas definidas en otros documentos normativos. En este sentido, las versiones más recientes de estos planes indican la importancia del acercamiento de los jóvenes a la investigación a través del programa JII.

Ahora bien, en 2019 el gobierno nacional convocó y conformó la Misión Internacional de Sabios de 2019, cuyos documentos finales permitieron obtener un diagnóstico del estado del sector de ciencia, tecnología e innovación. Allí se reconoce la importancia de los diferentes actores como el Estado, las instituciones educativas y las comunidades como responsables del desarrollo de un sistema nacional que permita el fortalecimiento de capacidades científicas en el país. En ese sentido, la Misión identificó que una de las prioridades debe ser la incorporación de jóvenes investigadores e innovadores para la formación de talento humano en todas las áreas del conocimiento.

Para acoger los lineamientos de la Misión de Sabios, se actualizó la *Política Nacional de Ciencia, Tecnología e Innovación 2022 – 2031* a través del documento CONPES 4069 de 2021. De acuerdo con este documento, el objetivo de la política es incrementar la contribución del sector de CTeI al desarrollo económico, social y ambiental del país, considerando un enfoque diferencia, territorial y participativo. Por ello, se plantean acciones dirigidas a consolidar los sistemas nacionales y regionales de CTeI a través de la dinamización de la producción y transferencia de conocimiento y del fortalecimiento de los procesos de investigación y creación. Este CONPES también destaca el esfuerzo realizado en la última década para el fortalecimiento del talento humano en CTeI a través de programas como JII; sin embargo, reconoce también que no se logra financiar toda la demanda de jóvenes que quieren participar en el programa a nivel nacional.

3. Caracterización de los JII

La principal fuente de información es la base de datos de JII en la cual se compilaron 19.086 registros de jóvenes inscritos, seleccionados y no seleccionados para la beca-pasantía entre 2004 y 2021.[18] Esta información se cruzó con otros sistemas de información del Gobierno Nacional para poder las variables de resultado de la evaluación, particularmente con el CvLac de MinCiencias y la Planilla Integrada de Liquidación de Aportes (PILA) del Ministerio de Salud; de este modo, se pudo obtener

[18] En adelante, cuando se habla de *inscritos a JII* se hace referencia a los jóvenes que aplicaron a las convocatorias, independientemente de si fueron seleccionados o no para la beca-pasantía. Cuando se habla de *inscritos no seleccionados* se hace referencia a los que se inscribieron a las convocatorias pero que no fueron seleccionados.

información de producción e ingreso base de cotización para los JII y los inscritos al programa que no fueron seleccionados. En el Anexo 2 se ofrecen más detalles de la información.

3.1. Convocatoria y selección

Se cuenta con información para 19.086 inscritos a JII, de los cuales 5.292 corresponden a aspirantes no seleccionados y 13.784 a seleccionados para la beca-pasantía (ver Tabla 14). Se destaca que 2012 tuvo el mayor número de inscritos con un total de 2.213,[19] de los cuales 1.031 fueron seleccionados para participar y 1.182 no. Sin embargo, los años 2011 y 2021 resaltan por tener el mayor número de jóvenes que participaron en JII, alcanzando 1.735 y 1.643 respectivamente. Para 2015 solo se logró compilar información de 21 inscritos, pero también se cuenta con una inscripción baja en las convocatorias de 2004 a 2008. Es importante mencionar que falta información sobre los inscritos no seleccionados para el programa entre 2013 y 2017.

Tabla 14. Información compilada de inscritos por cohorte

Año de la convocatoria	Aspirantes no seleccionados	Participantes en JII	Total
2004	394	38	432
2005	252	32	284
2006	337	202	539
2007	172	296	468
2008	268	272	540
2009	123	1.107	1.230
2010	86	1.380	1.466
2011	63	1.735	1.798
2012	1.182	1.031	2.213
2013	0	796	796
2014	0	847	847
2015	0	21	21
2016	0	605	605
2017	0	1.498	1.498
2018	166	791	957
2019	1.275	644	1.919
2020	565	856	1.421
2021	409	1.643	2.052
Total	**5.292**	**13.794**	**19.086**

Fuente: Elaboración propia con base en los datos de las convocatorias del programa JII

En 2021 se inscribieron 2.052 jóvenes y se seleccionó al 80% de ellos para participar. La mediana del monto reportado en la base de datos de las convocatorias es de $2.500.000 pesos en ese año por mes.[20]

3.2. Características de los JII seleccionados

3.2.1. Lugar de nacimiento

La mayoría de los jóvenes que participan en JII nacieron en Colombia, con un 0,6% que reportan haber nacido en otros países, principalmente Estados Unidos, Venezuela y México. Ahora bien, al analizar la

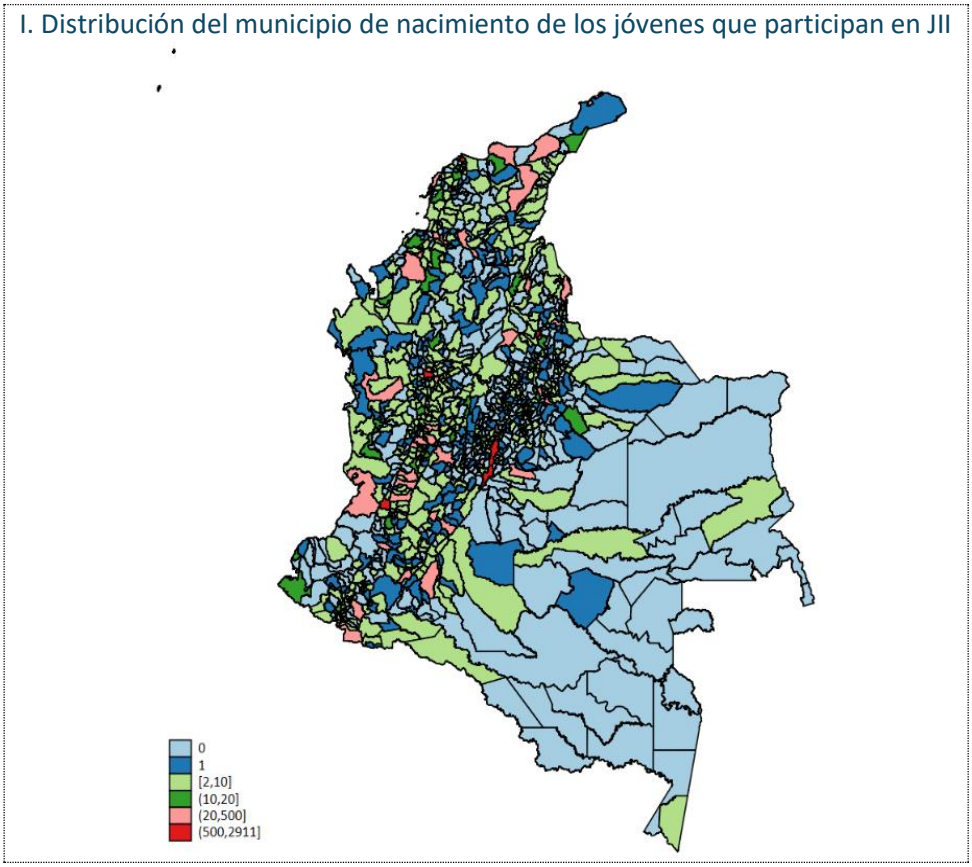
[19] En adelante, cuando se habla de *inscritos a JII* se hace referencia a los jóvenes que aplicaron a las convocatorias, independientemente de si fueron seleccionados o no para la beca-pasantía. Cuando se habla de *inscritos no seleccionados* se hace referencia a los que se inscribieron a las convocatorias pero que no fueron seleccionados.

[20] Solo 725 de los 1643 beneficiarios tienen reporte de monto.

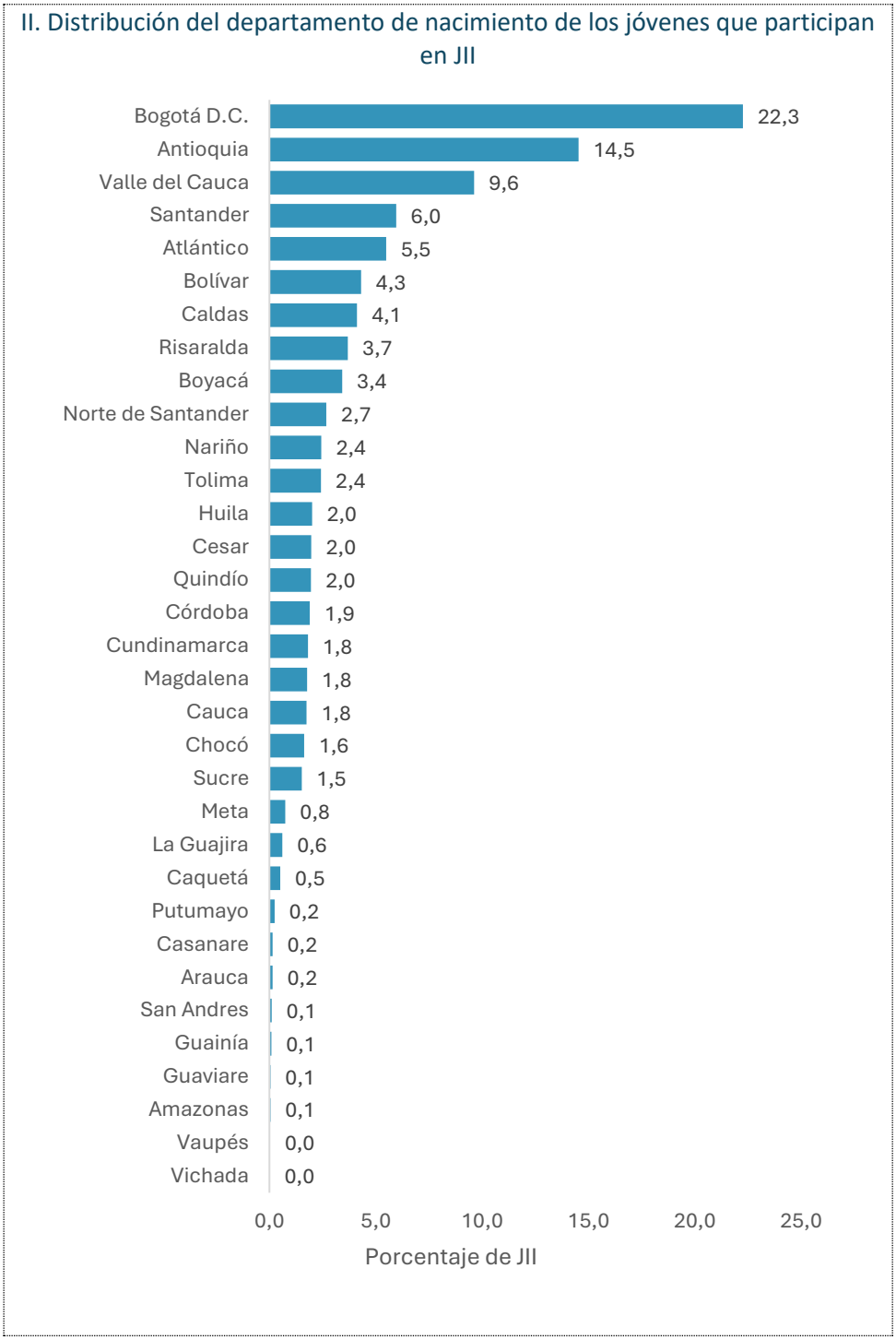
distribución del municipio o el departamento de nacimiento de los jóvenes se encuentra cierta diversidad en el origen de los jóvenes que participan en JII, abarcando departamentos como Amazonas, Vaupés, Vichada, Guainía y Guaviare con menos del 0,1% y otros como Valle del Cauca y Antioquia, en los que nacieron el 9,6% y el 14,5% de los participantes en JII (ver *Figura 12*).

Más de la quinta parte de los jóvenes investigadores e innovadores nacieron en Bogotá, un 11% en Medellín, y un 7% en Cali, que son los lugares donde se concentra la oferta educativa y de grupos de investigación; siendo esto un reflejo de la centralización de las oportunidades educativas y de investigación.[21] Esto contrasta con el grupo de municipios en los que no nació ningún joven investigador y que corresponden al 43% del total de los municipios del país. En el 20% de los municipios nació al menos un joven investigador y, en el 2% de los municipios nacieron al menos 100 jóvenes investigadores e innovadores (este último porcentaje corresponde a 20 municipios).

Figura 12. Lugar de nacimiento de los jóvenes que participan en JII



[21] Para recabar esta información se combinó el reporte de municipio de nacimiento en la base de datos de las convocatorias con el reportado en CvLac.



Fuente: Elaboración propia con información de las convocatorias y CvLac.

La base de datos de las convocatorias y de CvLac solo cuentan con información del municipio y departamento de nacimiento, pero no del departamento de residencia actual de los jóvenes investigadores e innovadores. Para tener un indicio de dónde se ubican actualmente los jóvenes investigadores e innovadores se tomó el departamento donde se ubica el o los grupos de investigación

en los que están o han estado vinculados, lo que permitió validar que la mayoría de ellos se encuentra en Bogotá, Antioquia y Valle del Cauca.[22]

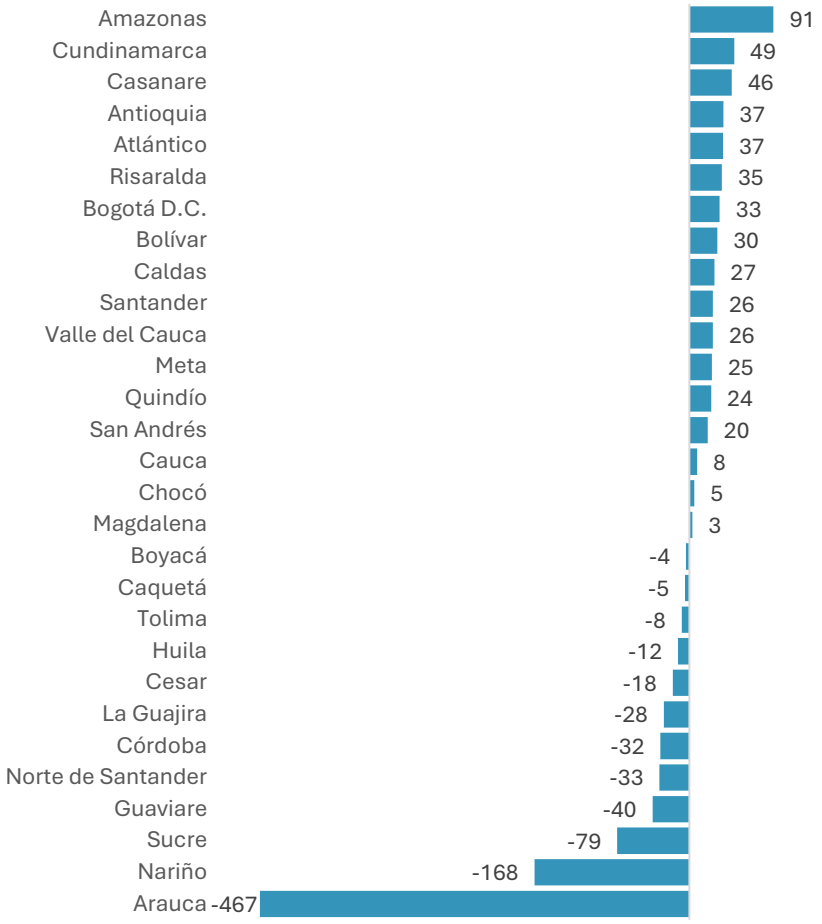
Se empleó el departamento de localización de grupos a los que se vinculan los jóvenes investigadores e innovadores para calcular una tasa de atracción neta de Jóvenes Investigadores e Innovadores por departamento. Esta tasa puede definirse como la diferencia entre el número de jóvenes investigadores e innovadores vinculados a grupos del departamento y el número de jóvenes investigadores e innovadores que nacieron en el departamento, dividido por el número de jóvenes investigadores e innovadores vinculados a grupos del departamento. Si la tasa es positiva, quiere decir que el departamento atrae a más jóvenes investigadores e innovadores de los que nacieron allí; mientras que si la tasa es negativa será un indicio de que hay pocas oportunidades para la atracción de jóvenes investigadores e innovadores.[23]

Los resultados presentan patrones interesantes, si bien hay algunos datos muy atípicos dado el poco número de observaciones (ver Figura 13). Amazonas representa el caso más interesante de atracción de jóvenes investigadores e innovadores; mientras que solo 3 jóvenes nacieron en ese departamento, 35 están o han estado vinculados a algún grupo de investigación de ese departamento. Le siguen Cundinamarca y Casanare con tasas del 49% y el 46%, respectivamente. En contraste, departamentos como Arauca y Nariño tienen tasas altamente negativas. En el caso de Arauca, 17 jóvenes investigadores e innovadores nacieron allí, mientras que solo 3 están vinculados a algún grupo de ese departamento. Por su parte, en Nariño nacieron 330 jóvenes investigadores e innovadores, pero solo 123 están vinculados en algún grupo de investigación de ese departamento.

[22] Una limitación de esta forma de recabar la información de departamento actual es que hay jóvenes en varios grupos que, a su vez, están en varios departamentos. Además, es importante tener en cuenta que no todos los jóvenes tienen información de grupo.

[23] El 79,8% de los jóvenes investigadores tenía información de grupo, por lo que este análisis se restringe a los jóvenes con información de departamento de nacimiento y del grupo. Si un joven investigador participa en varios grupos de investigación de varios departamentos, se cuenta en ambos departamentos. De ahí que la diferencia entre los que participan en grupos del departamento y los que nacieron ahí se pondere por este primer número.

Figura 13. Tasa neta de atracción de jóvenes investigadores e innovadores por departamento

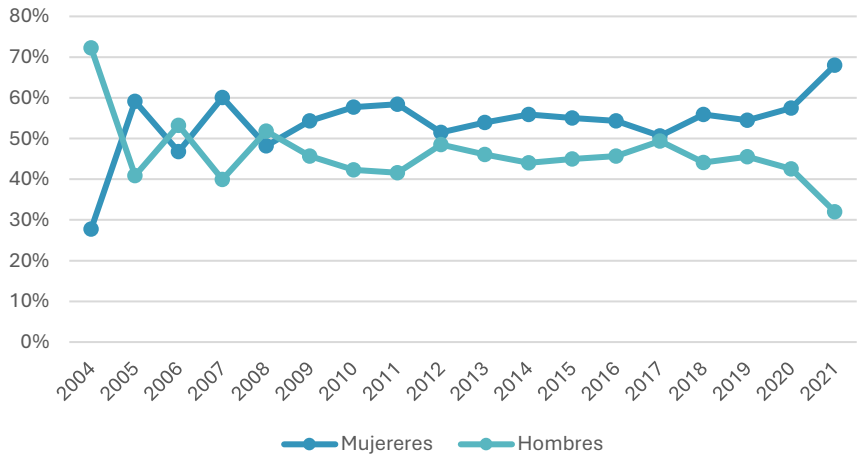


Fuente: Elaboración propia con información de las convocatorias y CvLac. Solo se consideran jóvenes investigadores e innovadores con información de grupo de investigación.

3.2.2. Características demográficas

El 56% de los jóvenes investigadores e innovadores son mujeres, pero este valor ha ido fluctuando a lo largo de las convocatorias, con una tendencia a incrementarse el porcentaje de mujeres que hacen parte del programa en cada cohorte (ver *Figura 14*). El porcentaje de mujeres que participan en el programa fluctuó mucho en las primeras convocatorias, entre el 28% y el 60%, y llegó a niveles del 68% para 2021. La participación de las mujeres en el total de jóvenes investigadores e innovadores está entre el 60% y el 68% para las que nacieron en Caquetá, Boyacá, La Guajira, Bogotá, Santander y Meta; en tanto que su participación está entre el 45% y el 50% en departamentos como Risaralda, Cesar, Caldas y Quindío.

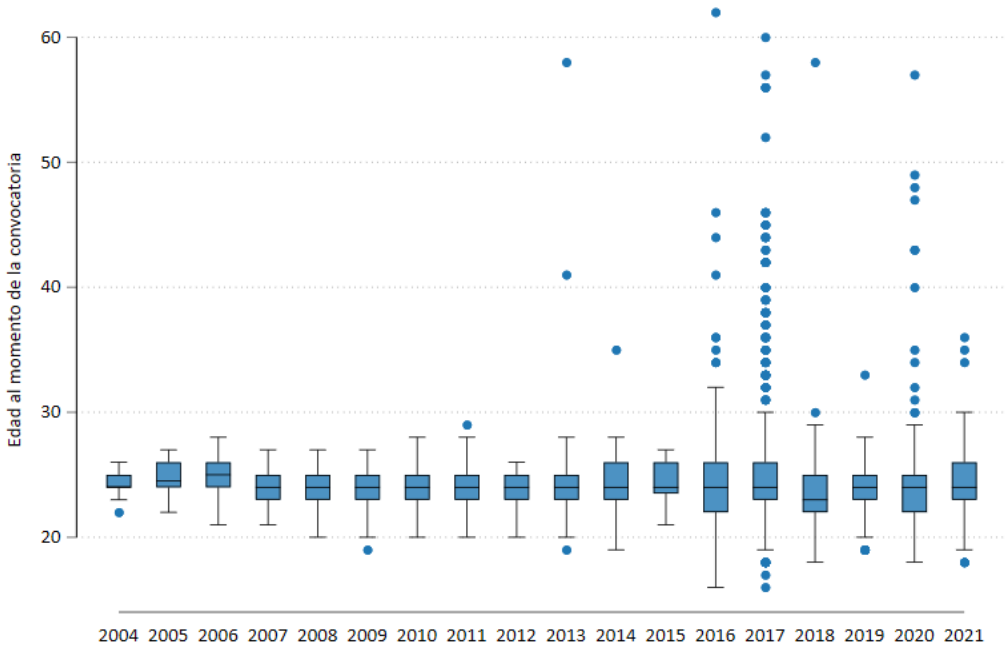
Figura 14. Porcentaje de jóvenes investigadores e innovadores que son hombres o mujeres por año de la convocatoria



Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac.

La edad promedio de los participantes está alrededor de los 24 años y, tal como se muestra en la Figura 15, la mayoría de las edades está entre 20 y 28 años, que es el límite de edad de la convocatoria. Se destaca que hay algunas edades atípicas, lo que pudiera deberse a inconsistencia en el reporte de la fecha de nacimiento para hacer este cálculo. Los valores extremos y muy atípicos encontrados en la edad se presentaron, sobre todo, en las convocatorias de 2016, 2017, 2020 y 2021.

Figura 15. Distribución de la edad al momento de la convocatoria

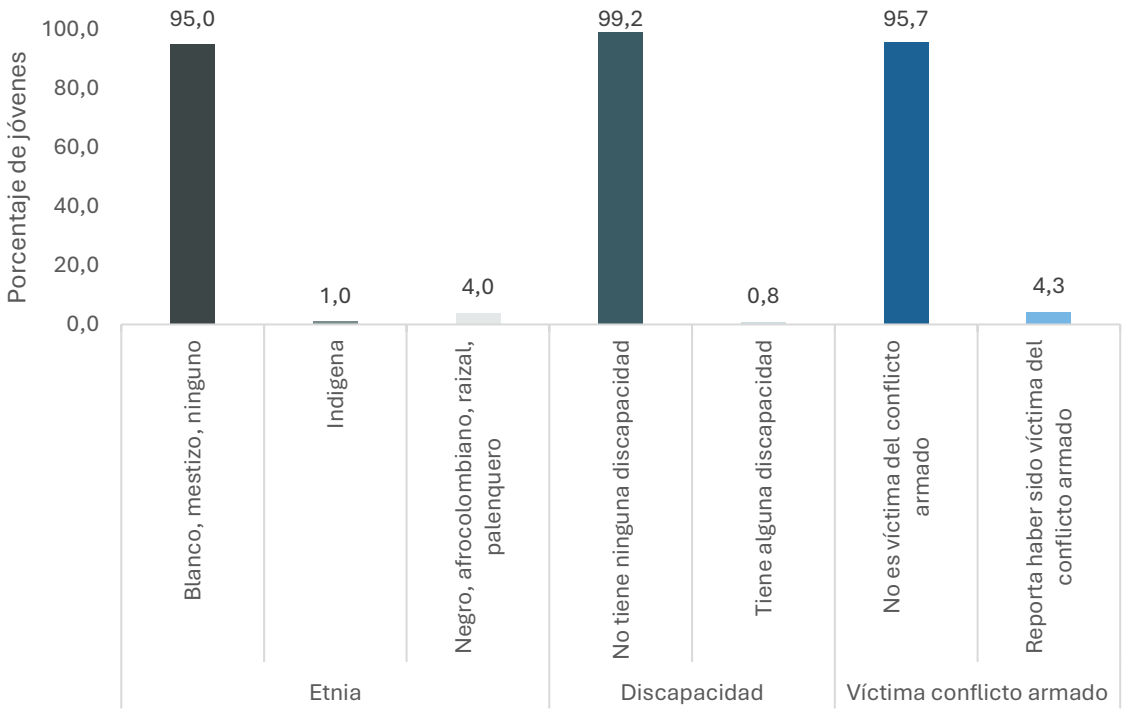


Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac.

La pertenencia étnica de los JII muestra que el 95,0% de los participantes se identifica como blanco, mestizo o sin afiliación étnica específica, mientras que un 4,0% se reconoce como Negro, afrocolombiano, raizal o palenquero, y un 1,0% como indígena. Este panorama etnocultural refleja la

composición predominante en el país, aunque también señala una participación minoritaria de grupos étnicos. En relación a las personas con discapacidad, la mayoría de los jóvenes, el 99,2%, mencionan no tener ninguna discapacidad, en comparación con un 0,8% que indica tener alguna. La mayoría de los que tienen alguna discapacidad reportan que es una discapacidad visual. Además, un 4,3% de los jóvenes reportan haber sido víctimas del conflicto armado en Colombia, lo que muestra también que el programa cubre poblaciones vulnerables. Ahora bien, es importante tener en cuenta que estas tres variables no se reportaron todos los años ni están completas en los años en los que se reportó en la base de datos de CvLac, por lo que tienen un subreporte.[24]

Figura 16. Pertenencia étnica, condición de discapacidad y condición de víctima del conflicto armado de los jóvenes investigadores e innovadores



Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac. Los datos de pertenencia étnicay condición de víctima del conflicto armado cuentan con 5.769 observaciones y los de personas con discapacidad con 5.765.

3.3. Área de conocimiento

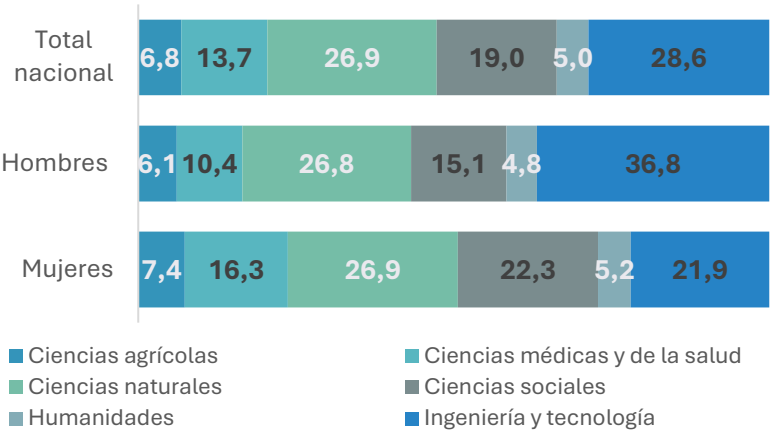
Para identificar el área de conocimiento del joven investigador se tomó la información del grupo de investigación en el que está vinculado.[25] A nivel nacional, el área de Ingeniería y Tecnología es la que tiene mayor participación entre los jóvenes investigadores e innovadores con un 28,6%. Sin embargo, se observa una brecha de género en esta área: mientras que la ingeniería representa un 36,8% de los jóvenes investigadores e innovadores hombres, en el caso de las mujeres es solo el 21,9%. Sigue en

[24] Solo se cuenta con información para 5.769 JII para esta variable.

[25] Un total de 9.251 jóvenes investigadores e innovadores tienen información de área de conocimiento. Para obtener el área de conocimiento se cruzó la información de grupo y se tomó el área de conocimiento del grupo. Para los jóvenes que han participado o están participando en varios grupos de investigación se tomó priorizó un solo grupo siguiendo estos criterios: (i) mayor tiempo de vinculación, (ii) mayor dedicación horaria, (iii) antigüedad del grupo (solo para unas pocas observaciones para las que los anteriores criterios no servían para discriminar).

importancia el área de Ciencias Naturales, donde hay un 26,9% de jóvenes investigadores e innovadores. Los jóvenes investigadores e innovadores participan en menor proporción en grupos de investigación vinculados a las áreas de Humanidades (5%) y Ciencias Agrícolas (6,8%). Se destaca que el área de las Ciencias Sociales representa el 22,3% de las jóvenes investigadoras mujeres, pero solo el 15,1% de los hombres.

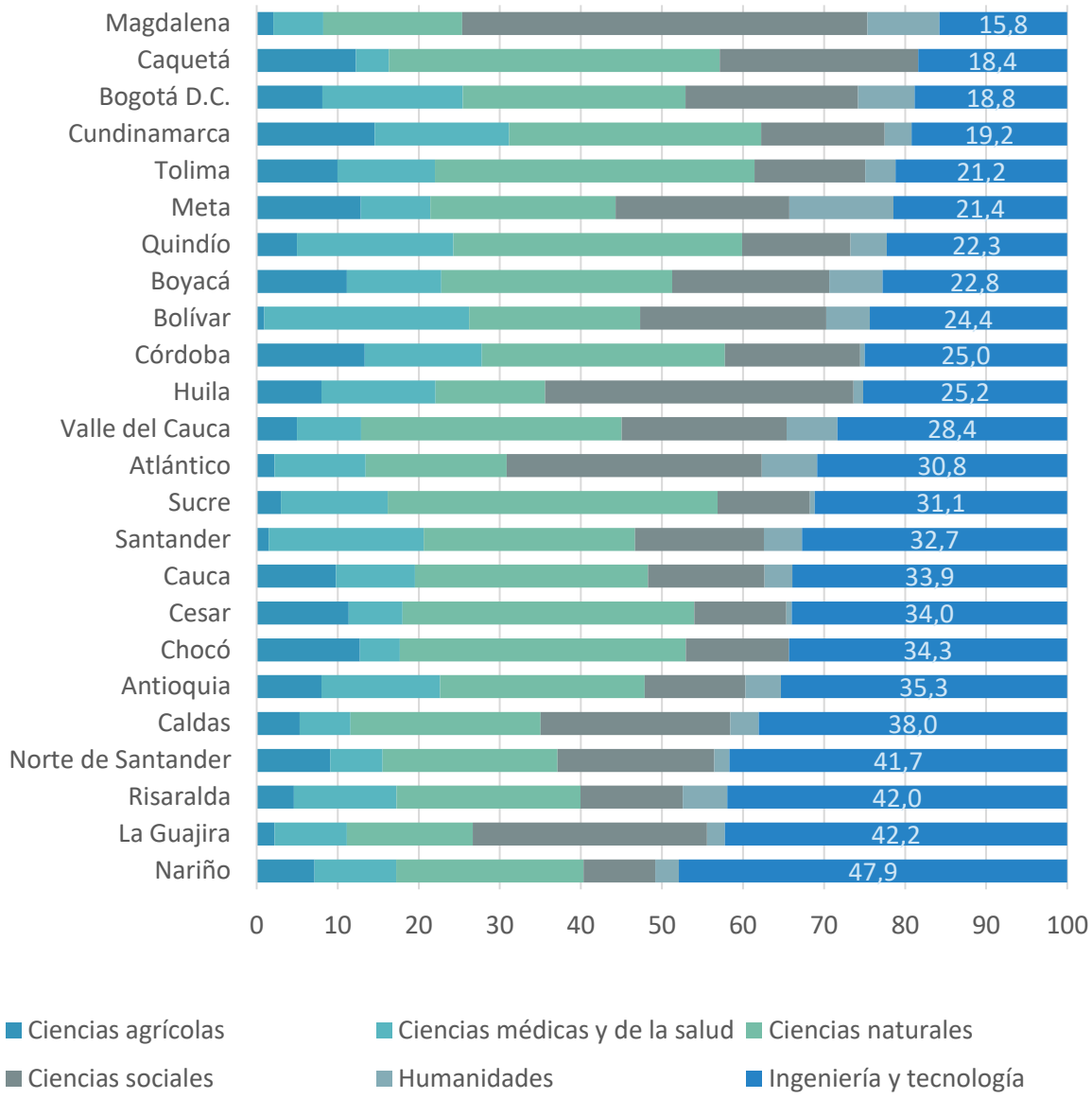
Figura 17. Distribución de jóvenes investigadores e innovadores según área de conocimiento



Fuente: Elaboración propia con información de los jóvenes investigadores que cruzaron con CvLac y Grupos de investigación reconocidos.

La distribución de las áreas de conocimiento de los jóvenes investigadores e innovadores por departamento muestra variaciones considerables (ver Figura 18). Ingeniería y tecnología, que es el área con mayor participación a nivel nacional, representa el 48% del total de jóvenes investigadores e innovadores en Nariño comparado con el 15% en Magdalena. El área de Ciencias Agrícolas representa el 15% del total de jóvenes investigadores e innovadores en Cundinamarca, pero menos del 1% en Bolívar. El área de Ciencias Médicas y de la Salud es predominante entre los jóvenes investigadores e innovadores de Bolívar, donde uno de cada cuatro jóvenes investigadores e innovadores está vinculado a esa área. Por su parte, el área de Ciencias Naturales tiene una alta participación en departamentos como Sucre y Caquetá, donde representa el 41% del total de los jóvenes. Los porcentajes más altos de jóvenes en grupos asociados a las Ciencias Sociales se encuentran en los departamentos de Atlántico (31%), Huila (38%), y Magdalena (50%). Por último, el área de Humanidades tiene una participación de 5% a nivel nacional entre los jóvenes investigadores e innovadores, pero esta cifra es considerablemente más alta en Bogotá (7%), Magdalena (9%) y Meta (13%).

Figura 18. Distribución de jóvenes investigadores e innovadores según el área de conocimiento del grupo de investigación



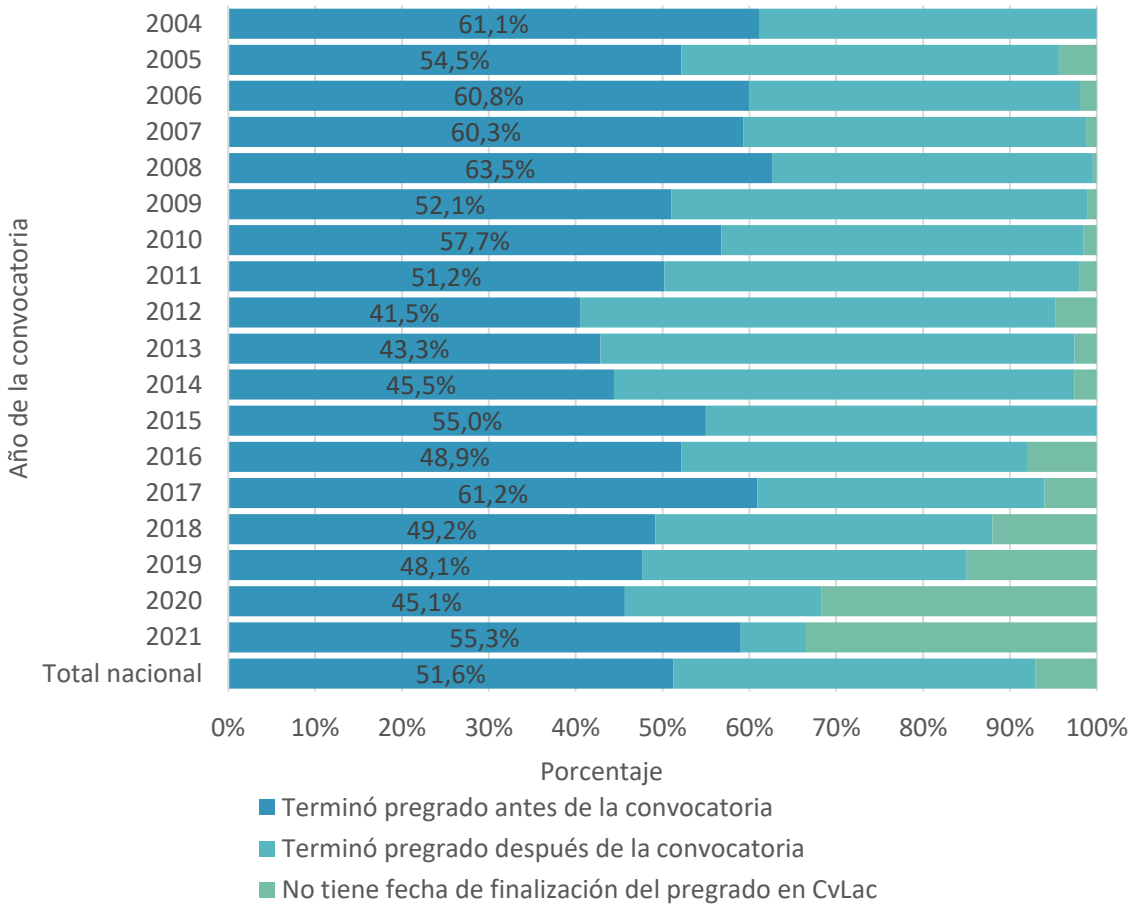
Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac y Grupos de investigación reconocidos. Solo se conservan los departamentos con más de 30 observaciones para contar con datos agregados más robustos y confiables y reducir la influencia de variaciones aleatorias en departamentos con muy pocas observaciones.

3.4. Nivel educativo de los JII

Con relación a la formación académica de los JII, se encontró que el 51,6% de ellos terminó su pregrado antes de la convocatoria en la que participaron.[26] Dicho porcentaje ha fluctuado entre el 41,5% para la convocatoria de 2012 y el 63,5% para la del 2008 (ver Figura 19). El porcentaje de jóvenes que no ha finalizado el pregrado es muy alto y creciente en las últimas cinco convocatorias, lo cual es un comportamiento esperable en la medida en que no han finalizado sus estudios; lo que sí muestra este hecho, es que hay un porcentaje inferior al 4% de jóvenes investigadores e innovadores de las primeras cohortes que no termina estudios de pregrado o que nunca reporta su terminación.

[26] El 91,1% de los jóvenes reporta información de pregrado en CvLac, pero no todos reportan la fecha de finalización, que bien puede ser porque no lo han terminado o porque no reportaron su finalización.

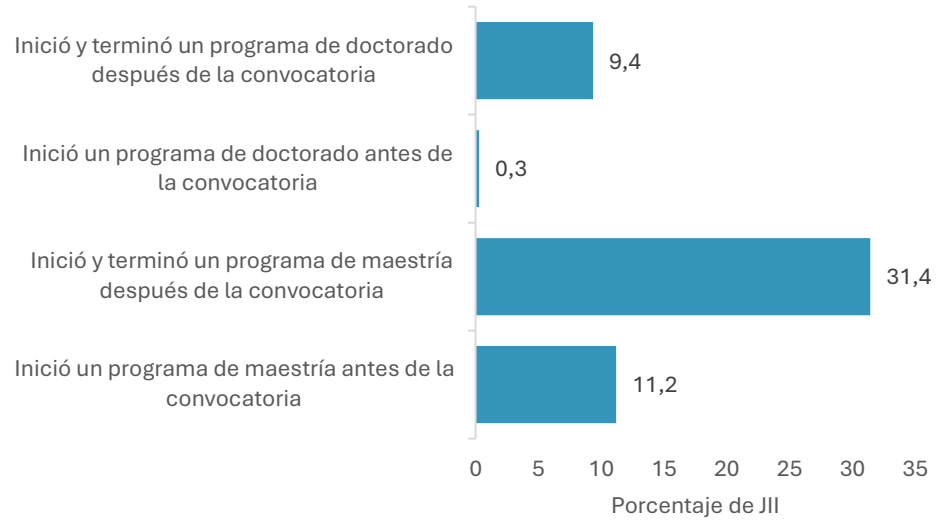
Figura 19. Porcentaje de JII según el estado de finalización del pregrado en el momento de la convocatoria



Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac.

En cuanto a la realización de estudios de posgrado, los datos revelan que el 11,2% de los jóvenes investigadores e innovadores ya había iniciado un programa de maestría antes de la convocatoria, y un 31,4% inició y terminó un programa de maestría después de la convocatoria. En cuanto a los doctorados, hay un porcentaje mucho menor que había iniciado (0,3%) un programa después de la convocatoria o que lo inició y lo terminó (9,4%) después de la convocatoria (ver Figura 20).

Figura 20. Porcentaje de jóvenes investigadores e innovadores que iniciaron un programa de maestría o doctorado según el momento de la convocatoria



Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac.

En suma, uno de cada tres JII adelanta estudios de maestría después de terminar la pasantía y uno de cada diez realiza estudios de doctorado. En las entrevistas con actores del programa se destacó que la mayoría de los JII inician sus estudios de maestría al poco tiempo de finalizar la pasantía, o incluso durante ese tiempo. También se señaló que el programa facilita el acceso a estudios de posgrado, en la medida en que la vinculación al grupo de investigación facilita la búsqueda de apoyos financieros o becas para la continuación de los estudios de posgrado.

"todos los muchachos que yo conozco de nuestro grupo generalmente hicieron su maestría como mínimo, como mínimo"

(entrevista a director de grupo de investigación, 2023)

"nosotros la experiencia que hemos tenido con el programa es que casi todos los Jóvenes Investigadores posterior al año de Jóvenes Investigadores o inclusive durante en el año de joven investigador empiezan sus estudios posgraduales, empiezan al menos la maestría o el doctorado con el programa. Yo hice la maestría en el primer año de joven investigador y seguí pues posteriormente con una asistencia de docencia"

(entrevista a directora de grupo de investigación, 2023)

También se destacó que el programa contribuye con el desarrollo de capacidades para realizar programas de doctorado, ya sea en Colombia o en el extranjero; además, contribuye a facilitar el acceso a este nivel de formación por la experiencia y herramientas que adquieren durante la pasantía.

"yo entré al Programa, pero no pensé que iba a ir a los eventos, proyectos y todo lo que hice como interactuar con otros investigadores. De hecho; eso me lleva a querer seguir estudiando, a seguir un posgrado que de pronto en el primer momento no tenía pensado. También fue un punto de inicio para decir quiero seguir investigando en lo que me gusta"

(grupo focal con participantes, 2023)

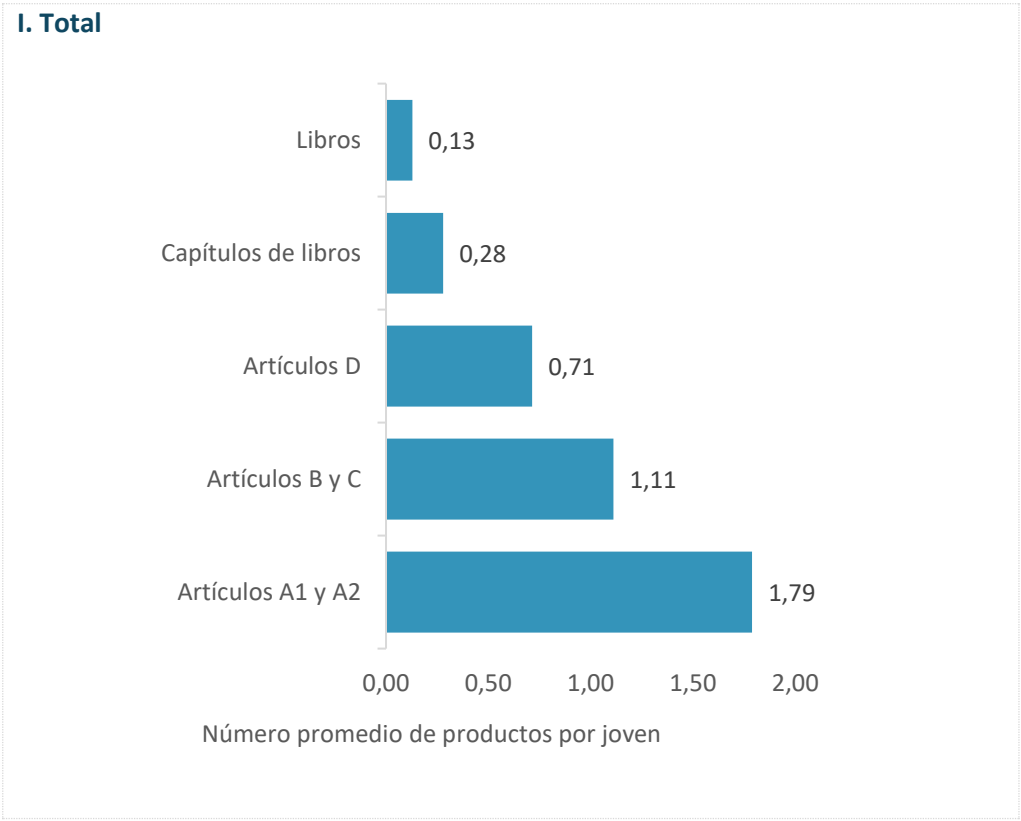
"Por ejemplo, yo tengo en este momento una estudiante que está ahorita en Polonia, está terminando su doctorado en bioquímica. ¿Cómo se logró eso?, se logró gracias a que ella tuvo su experiencia, su primera experiencia como joven investigador. Eso nos permitió tenerla un buen tiempo y hacer que tuviera una hoja de vida, porque el gran problema cuando un chico sale a fuera es que usted tiene que competir y si usted no tiene las herramientas para competir así sea bueno, no llega, no consigue la meta. Entonces el Programa Jóvenes Investigadores para nosotros nos benefició en el sentido que podíamos darle la formación adecuada y darle a los muchachos un estándar competitivo para cuando siguieran al siguiente paso que era el doctorado o quisieran aplicar a becas en el caso de otro país tuvieran al menos uno o dos artículos publicados en revistas buenas"

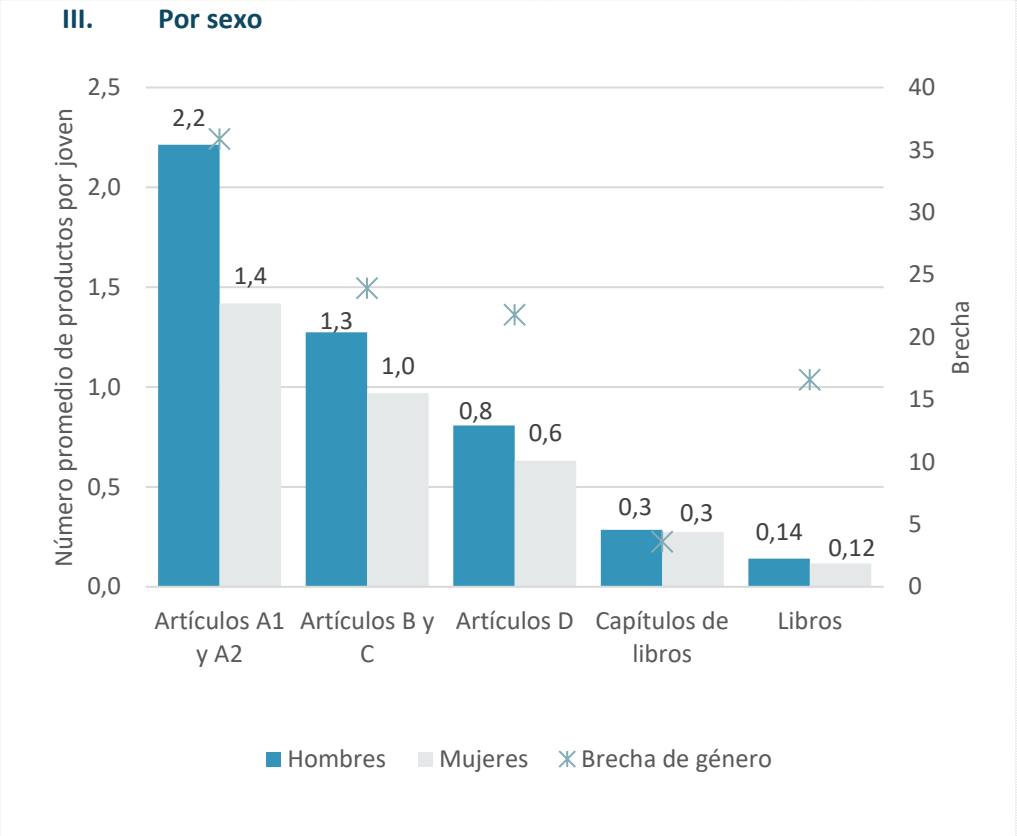
(entrevista a director de grupo de investigación, 2023)

3.5. Productos de nuevo conocimiento de los JII

Los jóvenes investigadores e innovadores publican un promedio de 1,79 artículos A1 y A2 después de participar en la convocatoria, siendo el tipo de artículo que más publican en su trayectoria profesional después de participar en JII (Figura 21, panel I). Le siguen los artículos en categorías B y C, con un promedio de 1,1 productos, y los artículos en categoría D con un promedio de 0,7. Además, publican un promedio de 0,28 capítulos de libros y 0,13 libros. Ahora bien, se observa una brecha de género alta en los productos de nuevo conocimiento considerados (Figura 21, panel II). Los hombres publican 35,9% más de artículos clasificados como A1 y A2 que las mujeres, puesto que registran un promedio de 2,2 artículos A1 y A2 publicados después de participar en el programa, comparado con 1,4 artículos de las mujeres clasificados en esa categoría. La brecha de género es también alta en productos clasificados como B y C (24%) y artículos D (22%), pero muy baja en el caso de los capítulos de libros (3,6%).

Figura 21. Número promedio de productos publicados por los jóvenes investigadores e innovadores después de la convocatoria



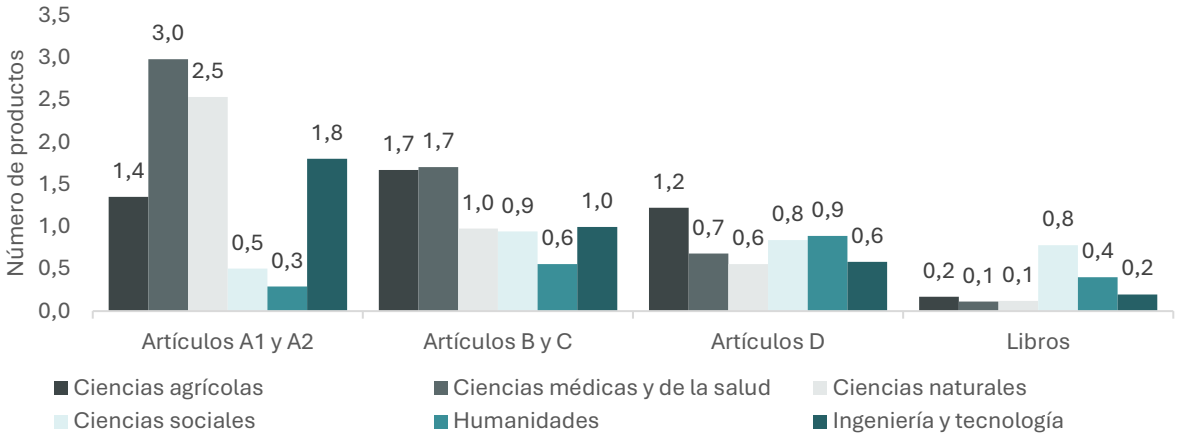


Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac y su componente de productos. La brecha de género se calcula como la diferencia entre el promedio de la sumatoria del total de productos de los hombres y las mujeres, dividido por los productos de los hombres.

En cuanto a la brecha de género en la producción académica, se observa que, aunque las mujeres participan activamente en todas las categorías, los hombres presentan mayor frecuencia en las publicaciones de mayor impacto, con una diferencia del 42.4% en artículos A1 y A2. Las mujeres tienen una participación más alta en las Ciencias Médicas y de la Salud y Ciencias Sociales, comparativamente con su participación en Ingeniería y Tecnología.

Por área de conocimiento, los jóvenes vinculados a grupos en ciencias médicas y de la salud publican más artículos A1 y A2, con un promedio de 3 artículos publicados después de haber participado en JII; le siguen las áreas de ciencias naturales (2,5) e ingeniería y tecnología (1,8). En cuanto a los artículos clasificados en las categorías B, C y D, se destaca un alto predominio de las ciencias agrícolas y las ciencias médicas y de la salud, puesto que los jóvenes en estas áreas publican en promedio 1,7 artículos después de haber participado en el programa. Por su parte, las ciencias sociales y las humanidades tienen un rol más destacado en la publicación de libros, con un promedio de 0,4 y 0,8 respectivamente, sugiriendo una preferencia por este formato para la diseminación del conocimiento en estas áreas.

Figura 22. Número promedio de productos publicados por los jóvenes investigadores e innovadores después de la convocatoria



Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con CvLac y su componente de productos.

3.6. Salarios

El IBC promedio de los jóvenes investigadores e innovadores es de $2.071.734, el cual corresponde al valor promedio que se emplea como base para aportar a pensiones.[27] Sin embargo, cuando se desglosa por género, se observa una disparidad: las mujeres tienen un IBC promedio de $1,919,902, mientras que los hombres uno de $2,255,909; lo que lleva a una brecha del 14,9% a favor de los hombres (ver Tabla 15). Los jóvenes investigadores e innovadores cotizan un promedio de 9,5 meses en el año, casi el 79% del total año, y no se observan diferencias considerables entre hombres y mujeres, si bien estas últimas tienen un valor ligeramente más bajo.

Tabla 15. Ingreso Base de Cotización y tiempo cotizado en el año de los jóvenes investigadores e innovadores

Desagregación	IBC para el pago de pensiones	Promedio de meses cotizados a pensión en el año
Total nacional	$2.071.734	9,5
Mujeres	$1.919.902	9,4
Hombres	$2.255.909	9,5

Fuente: Elaboración propia con información de los jóvenes investigadores e innovadores que cruzaron con PILA. Los valores del IBC están deflactados a pesos de 2018.

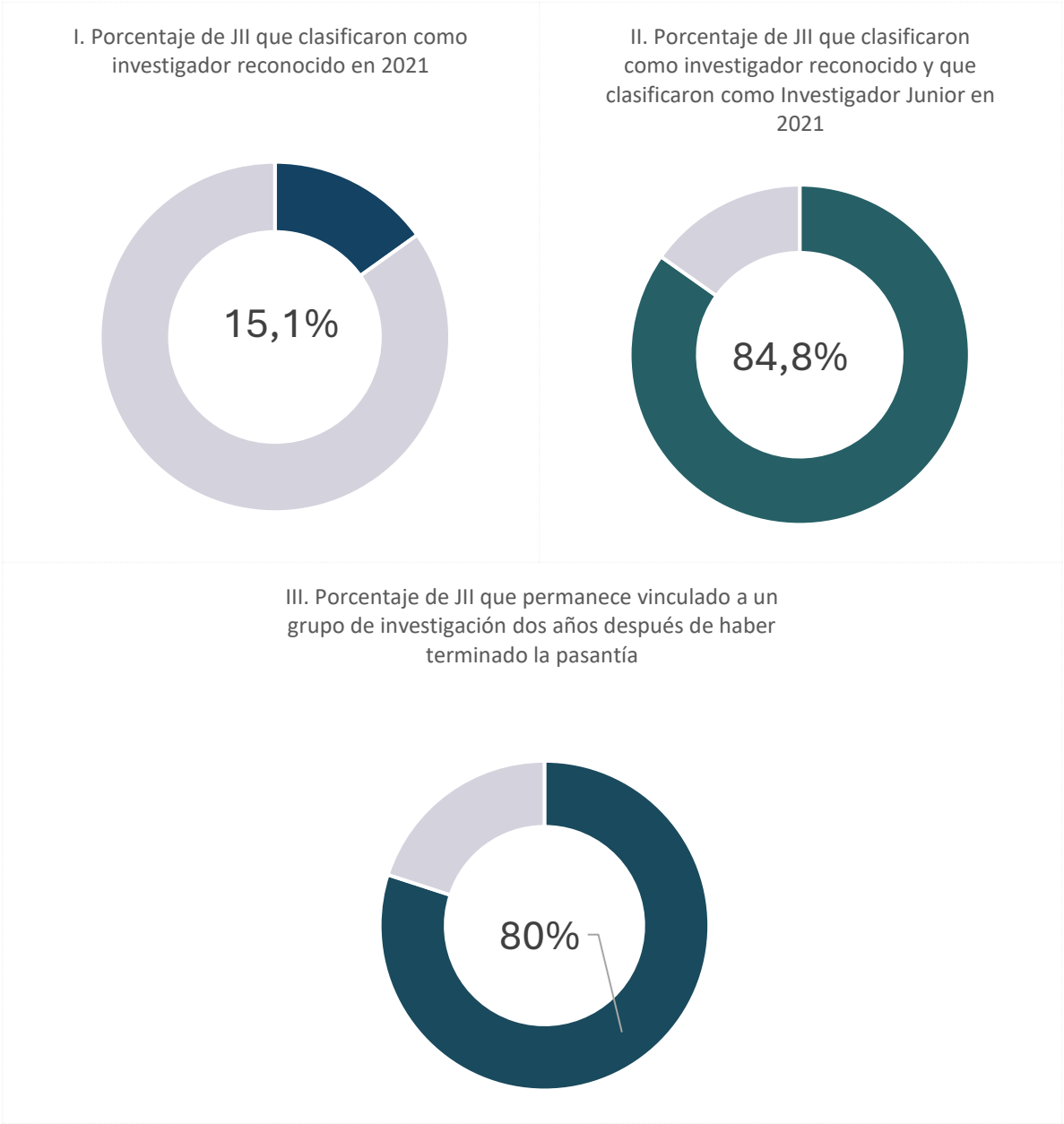
El programa tiene un efecto positivo inmediato positivo en las condiciones económicas de los participantes, sobre todo en el caso de los que son de universidades públicas o de regiones menos desarrolladas, tal como lo mencionaron en las entrevistas con actores involucrados con el programa. El monto entregado mensualmente a los JII representa una fuente importante para su sostenimiento y, en algunos casos, les facilita la financiación de sus programas de maestría.

[27] El IBC corresponde al salario de los trabajadores dependientes o a la proporción del ingreso laboral que los independientes deben utilizar como base para aplicar el porcentaje respectivo de la cotización a pensiones. Se asemeja mucho al ingreso laboral promedio de las personas con educación superior (pregrado o posgrado) en el país ($2.087.320, en pesos de 2018 según la GEIH de 2022). Solo se presenta el IBC para la cotización a pensiones por simplicidad, pero el IBC promedio para salud es también muy similar (la diferencia es de $15.975 y puede estar relacionada con diferencias en el mes del aporte).

3.7. Vinculación a grupos de investigación

El 15,1% de los JII con información de vinculación en CVLAC clasificó como investigador reconocido en la última convocatoria de medición en 2021, de ellos, el 80% fue reconocido como investigador Junior. Además, el 84,8% de los JII permanece vinculado a un grupo de investigación dos años después de haber terminado la pasantía. Esto demuestra una tendencia positiva en la inserción de los JII en el campo de la investigación y la innovación.

Figura 23. JII que son reconocidos como investigador junior o que permanecen vinculados al programa dos años después de finalizada la pasantía



En las entrevistas con actores vinculados al programa se destacó que la experiencia en JII permite iniciar el camino para vincularse como docentes universitarios, lo que facilita también que permanezcan vinculados en grupos de investigación.

"toda esta experiencia fue la que me empezó como a alinear con el tema académico. Yo no tenía experiencia docente. Entonces pues como en su momento apenas si estaba como viendo primeros pinitos, asistiendo algunas clases y participando en algunas actividades, y de cierta manera es lo que me tiene hoy aquí en la Universidad ya como docente tiempo completo"

(grupo focal con participantes, 2023)

"hemos tenido 56 reportes de Jóvenes Investigadores, […] el impacto ha sido muy importante porque algunos de ellos después se han ido a hacer doctorado y en este momento están vinculados por la institución. O sea de hecho nuestro director general es una persona que fue joven investigador de Minciencias en su momento y el director del Programa de Variedades que es un grupo de investigación está en A1 también fue joven investigador de MinCiencias"

(entrevista a funcionario de centro/instituto de investigación, 2023)

4. Metodología para la evaluación de impacto

4.1. Variables de resultado

El programa JII tiene como objetivo acercar a los jóvenes a la investigación y a la innovación a través de una beca-pasantía que les permite vincularse a grupos de investigación. En ese sentido, se analizan dos dimensiones en el marco de esta evaluación de impacto: la producción de nuevo conocimiento y los resultados en el mercado laboral.

Producción de nuevo conocimiento. El programa JII facilita que sus participantes tengan acceso a mentores y grupos de investigación, les permite trabajar en proyectos de investigación e innovación y tiene como requisito la generación de un producto derivado del proyecto en el que participaron. Por tanto, y aunque la publicación del producto resultado de la beca-pasantía no es un requisito, se espera que el programa impacte positivamente en la generación de productos de nuevo conocimiento. Así, como variables de resultado en esta dimensión se utilizó el número de artículos—y sus clasificaciones A1, A2, B, C y D— y libros que producen los jóvenes.

Mercado laboral. El programa JII ofrece a sus participantes una forma de entrada al mercado laboral formal, proporcionando experiencia y ampliando sus redes profesionales. Como resultado, se esperaría una mayor empleabilidad e ingresos de los JII. Específicamente para esta evaluación se empleó el Ingreso Base de Cotización (IBC) para el Sistema General de Seguridad Social en Salud en pensiones que es una buena aproximación del ingreso laboral,[28] y el número de meses cotizados a pensión en el año y que sirve como proxy de estabilidad en el mercado laboral, ya que refleja la continuidad y permanencia de una persona en un puesto de trabajo.

4.2. Grupos de tratamiento y control

La evaluación de impacto trata de responder qué hubiera pasado con los participantes del programa en ausencia del mismo. Esto implicaría tener que observar los resultados en producción científica y en el mercado laboral de un joven que recibió la beca-pasantía, y del mismo joven sin ella, de modo que se pudiera obtener el resultado con y sin intervención. Sin embargo, en la práctica solo se puede observar al joven con la beca-pasantía o sin ella. Por tanto, para adelantar la evaluación se deben identificar dos grupos de individuos diferentes: (i) un "grupo de tratamiento" conformado por todos los jóvenes que participaron en JII; y (ii) un "grupo de control" conformado por individuos que no participaron del programa, pero que sean similares (iguales) en características observables y no

28

observables a los jóvenes que participaron en JII, de tal forma que puedan asemejarse lo mejor posible al escenario contrafactual ideal (el mismo individuo con y sin participación en el programa).

En ese sentido, el principal reto de la evaluación de impacto es encontrar el grupo de control que cumpla con las características de ser "idéntico" al grupo de tratamiento. Esto puede ser muy desafiante sobre todo en programas como JII en los que los individuos se autoseleccionan para participar en el programa. Una alternativa es emplear como grupo de control los individuos que se hayan postulado para participar del programa pero que no han sido seleccionados siempre que exista la información disponible. En el caso de JII se cuenta con la información de los jóvenes que aplicaron a cada una de las convocatorias entre 2004 y 2021 (si bien no se cuenta con aspirantes que no se seleccionaron para las cohortes de 2013 a 2017). El uso de estos aspirantes que no se seleccionaron como grupo de control permite controlar el sesgo de autoselección ya que se comparan jóvenes que se postularon al programa.

Ahora bien, la selección final de los jóvenes que van a participar de la beca-pasantía puede estar influenciada por múltiples factores, como el incumplimiento de los requisitos por parte de los postulantes o las instituciones, la elección del postulante de no participar después de la postulación o la limitación de recursos que impide aceptar a todos los candidatos calificados. Por lo que uno esperaría poder contar con información detallada sobre los motivos por los cuales los inscritos no resultaron seleccionados y poderlos incorporar en las regresiones para controlar por sesgos adicionales. Sin embargo, para JII no se cuenta con esta información sino solamente con el listado de inscritos (tanto seleccionados como no seleccionados), por lo que se propuso la conformación de dos grupos de control en función de la información disponible y que se detalló en la sección anterior:

Futuros tratados. Como se verá más adelante, una de las técnicas de estimación propuestas es la metodología de diferencias en diferencias con múltiples períodos y múltiples tratamientos en el tiempo, la cual permite utilizar a los futuros tratados en cada momento del tiempo como grupo de control.[29] Por ejemplo, para la cohorte de jóvenes investigadores e innovadores que fueron seleccionados en 2012 se emplea como grupo de control a los futuros tratados del programa, que son quienes fueron seleccionados en las convocatorias de 2013 en adelante; siguiendo con el ejemplo, para la cohorte de jóvenes investigadores e innovadores seleccionados en 2019, el grupo de control serían los jóvenes que se seleccionaron en 2020 y 2021 y que, en 2019, no participaban todavía del programa. Así, se conforma un grupo de control dinámico a partir de los futuros tratados, lo que permite eliminar el sesgo de selección puesto que también se postulan al programa, solo que en un momento diferente del tiempo, pero se origina el sesgo de por qué un individuo se presentó en un momento del tiempo y no en otro. Como se verá más adelante, Callaway y Sant'Anna (2020) muestran que, bajo el cumplimiento del supuesto de tendencias paralelas, este sesgo se controla con la estimación del modelo de diferencias en diferencias con múltiples períodos y múltiples tratamientos en el tiempo. El grupo de control conformado por los futuros tratados (o los no tratados aún) se empleó para las estimaciones de los impactos de JII en variables de mercado laboral, puesto que solo se cuenta con información para el grupo de tratamiento, si bien se cuenta con la heterogeneidad dada por la diferencia en el año de la convocatoria.

***Futuros tratados y jóvenes que se inscribieron en las convocatorias del JII pero que no resultaron seleccionados*.** Emplear a los jóvenes que se inscribieron a las convocatorias de JII como grupo de control permite eliminar el sesgo de selección inherente a los programas en los que los individuos deben postularse para participar. En este caso, se utilizan los 4.496 jóvenes que se inscribieron y no fueron seleccionados para la beca-pasantía y que cuentan con información en CvLac, junto con los

[29] Se les denomina *"aún no tratados"* o *"not yet treated"* en inglés.

futuros tratados (cuya información esté disponible también en CvLac). Por ejemplo, en el año 2019 el grupo de control corresponde al grupo de control original más los jóvenes que participaron en JII en 2020 y 2021 que no han iniciado la pasantía. Este grupo de control conformado por los jóvenes que no resultaron seleccionados en las convocatorias y los futuros tratados se empleó para las estimaciones del impacto de JII en productos de nuevo conocimiento.

4.3. Técnica de estimación

Se empleó el modelo de diferencia en diferencias (DID) ya que es la metodología más idónea y robusta para estimar los impactos de JII.[30] Esta es una estrategia robusta en ausencia de diseños aleatorizados o de puntajes que determinan la elegibilidad al programa – y que son esenciales para la aplicación de técnicas de regresión discontinua, y tiene muchas más ventajas frente a métodos de emparejamiento. De manera general, el estimador de DID compara el resultado del grupo de tratamiento antes y después de la intervención (primera diferencia), con la diferencia en los resultados obtenidos por el grupo de control (segunda diferencia). Ahora bien, dado que la evaluación considera a los jóvenes que participaron en JII de más de una década, de 2004 a 2021, sería inconsistente comparar los tratados de los últimos períodos con los iniciales; por lo que se empleó el estimador de diferencias en diferencias con múltiples tratamientos en múltiples momentos de tiempo; este estimador permite entender los efectos de la intervención a lo largo del tiempo (en múltiples períodos) y para múltiples grupos de tratamientos (algunas unidades tratadas pueden serlo una vez, o varias).

Callaway y Sant'Anna (2020) desarrollaron una generalización del modelo de diferencias en diferencias que permite analizar el efecto del programa en el tiempo y para diferentes cohortes (en este caso, cohortes de participación en JII). Los autores parten de la representación del modelo de diferencias clásico, donde el efecto promedio sobre los tratados (ATT por sus siglas en inglés) viene dado por:

$$ATT = E\left[Y_{T=1;t=1} - Y_{T=0;t=1} | P_{T=1} = 1\right]$$

La generalización del estimador se basa en la idea de que cada observación pertenece a una cohorte *h* en cada momento *t* del tiempo, por lo que el estimador ATT puede representarse como:

$$ATT(h,t) = E\left[Y_{T=h;t=t} - Y_{T=0;t=t} | P_{T=h} = 1\right]$$

Los autores no imponen restricciones sobre los grupos de tratamiento o los períodos de tiempo, por lo que pueden ser varios. Este estimador se conoce como el efecto promedio del tratamiento en el grupo y tiempo (group-time average treatment effect), y mide el impacto de la intervención en los participantes del programa a lo largo del tiempo. Una ventaja del método es que permite examinar cómo es el impacto en cohortes específicas y cómo evoluciona en el tiempo, además de que permite tener un grupo de control dinámico conformado no solo por los individuos que no participan del programa, sino también por los futuros tratados del programa en cada momento del tiempo.

Callaway y Sant'Anna (2020) demuestran que para obtener estimadores insesgados y consistentes utilizando esta metodología, es necesario imponer restricciones a la anticipación del tratamiento por parte de los tratados en el tiempo, lo cual puede influir en su comportamiento previo al tratamiento y en el resultado de las estimaciones. En caso de que esto ocurra, se debe recurrir a funciones de ponderación para corregir las estimaciones. Para ellos los autores desarrollaron tres métodos de estimación: regresión de resultados (OR), ponderación por probabilidad inversa (IPW) y estimadores

[30] No se cuenta con un diseño aleatorizado de la intervención; tampoco se cuenta con un puntaje y un punto de corte que determine la elegibilidad al programa y que es crucial para la aplicación de técnicas de Regresión Discontinua; y el estimador de DID es mucho más robusto en comparación con métodos de emparejamiento en la medida en que permite controlar por factores no observables constantes en el tiempo.

doblemente robustos (DR). OR modela los resultados del grupo de control, IPW modela la probabilidad condicional de que una unidad sea asignada al grupo de tratamiento y DR combina ambos enfoques.

Para ilustrar la estimación mediante el método de diferencia en diferencias con múltiples períodos y tratamientos en el tiempo, partimos del caso en el cual contamos con información de tipo panel en la cual:

$$\{(Y_{i,1}, Y_{i,2}, \dots, Y_{i,\tau}, D_{i,1}, D_{i,2}, \dots, D_{i,\tau}, X_i)\}_{i=1}^{n} \qquad (3)$$

Donde la cual $D_{i,t}$ es una variable dummy que toma el valor de uno si el joven *i* participa en JII en el periodo *t,* en otro caso, toma el valor de cero. Adicionalmente, se define una variable dummy $G_{i,g}$ que toma el valor de uno para el joven *i* si el periodo *t* es su primer período de participación en el programa. Esto es, la variable $G_{i,g}$ es equivalente a una variable que podemos determinar como *inicio del tratamiento.* Por último, se define una variable C la cual toma el valor de uno para los jóvenes que nunca son seleccionados para la beca-pasantía.

Los autores asumen que una vez un individuo es tratado, se asume como tratado para siempre. Esto es:

$$D_{i,t} = 1 \Rightarrow D_{i,t+1}, para\ t = 1,2, \dots \tau \qquad (4)$$

Bajo estos supuestos, el parámetro de interés puede escribirse como:

$$ATT(g,t) = E\left[Y_t(g) - Y_t(0) | G_g = 1\right], para\ t \geq g \qquad (5)$$

el cual representa el efecto promedio sobre la población tratada por primera vez en el momento *g,* dentro del conjunto de periodos *t.*

En este punto, y dado que se cuenta con múltiples periodos *t,* y con múltiples grupos de unidades tratadas en diferentes momentos *g.* El ATT para la aproximación para la estimación de los impactos con un grupo de control con jóvenes que nunca participaron en el programa sería:

$$ATT_{unc}^{ny}(g,t) = E\left[Y_t - Y_{g-1} | G_g = 1\right] - E\left[Y_t - Y_{g-1} | C = 1\right] \qquad (6)$$

Mientras que para el caso en el que se utiliza a los futuros tratados sería, el *ATT* puede escribirse como:

$$ATT_{unc}^{ny}(g,t) = E\left[Y_t - Y_{g-1} | G_g = 1\right] - E\left[Y_t - Y_{g-1} | D_t = 0, G_g = 0\right] \qquad (7)$$

Nótese que, en ambos casos, la estimación del impacto implica realizar tantas estimaciones de diferencia en diferencia como periodos *g* estén disponibles. Es decir, con esta aproximación del modelo de diferencia en diferencias propuesta por Callaway y Sant'Anna (2020) es posible reconstruir la trayectoria del impacto del programa en cada periodo de tiempo. A diferencia del modelo de diferencia en diferencias tradicional, en el cual se comparan los individuos tratados y los controles únicamente en dos periodos de tiempo. Esto no solo produce estimaciones más robustas, sino que permite identificar los posibles efectos heterogéneos o la acumulación de los impactos.

Ahora bien, se puede obtener el impacto promedio del programa mediante la agregación de las múltiples estimaciones de impacto mediante la metodología propuesta por Callaway y Sant'Anna (2020). Para ello se puede tomar los promedios ponderados de todos los impactos estimados para todos los individuos en todos los momentos del tiempo. Esto es:

$$\theta_W^{simple} = \frac{1}{K} \sum_{g=2}^{\tau} \sum_{t=2}^{\tau} 1\{g \leq t\} ATT(g,t) P(G = g | C \neq 1) \qquad (8)$$

Por último, el método de diferencia en diferencias descrito requiere, al igual que en su versión tradicional, del cumplimiento del supuesto de tendencias paralelas, el cual permite establecer que la evolución de las variables de resultado entre los grupos de tratamiento y control previo a la intervención, tenían la misma tendencia, y por tanto cualquier cambio en la tendencia es atribuible al programa. En esta versión del modelo de diferencia en diferencias, surgen dos aproximaciones al supuesto de tendencias paralelas, de acuerdo con la metodología de conformación del grupo de control establecida. Si se realizan estimaciones con base en el grupo de control con individuos que nunca participaron en el programa, el supuesto de tendencias paralelas vendrá dado por:

$$E\left[Y_t(0) - Y_{t-1}(0)|X, G_g = 1\right] = E[Y_t(0) - Y_{t-1}(0)|X, C = 1] \qquad (10)$$

para todo $t \in \{2, \dots, \tau\}, g \in \varrho \; para \; todo \; t \geq g$

De otro lado, si las estimaciones se basan en el grupo de control conformado por los futuros controles, el supuesto de tendencias paralelas vendrá dado por:

$$E\left[Y_t(0) - Y_{t-1}(0)|X, G_g = 1\right] = E\left[Y_t(0) - Y_{t-1}(0)|X, D_s = 0, G_g = 0\right] \qquad (11)$$

para todo $(s,t) \in \{2, \dots, \tau\} \times \{2, \dots, \tau\}, \; g \in \varrho \; para \; todo \; t \geq g, s \geq t$ se cumple el supuesto de tendencias paralelas descrito anteriormente, la estimación mediante el método de diferencia en diferencias arrojará estimadores robustos.

4.4. Comparación entre el grupo de tratamiento y control

Con la información final de la base de datos de inscritos cruzada con la base de datos de CvLac se sacaron algunas estadísticas descriptivas de los grupos de tratamiento y control y la diferencia entre ellos. La edad promedio de los tratados es 24,1 y el 56% son mujeres, estas dos cifras son significativamente más altas para los tratados (ver Tabla 16). También se observa que hay una mayor proporción de tratados que nacieron en Antioquia, Chocó, Magdalena, Córdoba y Nariño; mientras que hay un mayor porcentaje de controles que tratados en Arauca, Bolívar, Boyacá, Casanare, Guainía, La Guajira, Quindío y Vichada. Estas variables de edad, sexo y departamento de nacimiento se incluyeron como controles en las estimaciones de impactos.

Tabla 16. Tabla de diferencias de medias: edad, género y porcentaje de jóvenes que nacieron en cada departamento

Variable	Controles	Tratados	Diferencia	Error Est.
Edad en el año de convocatoria	23,82	24,15	-0,33***	0,04
Mujer	53,70	56,00	-2,29***	0,87
Amazonas	0,02	0,02	-0,00	0,03
Antioquia	12,63	14,71	-2,09***	0,61
Arauca	0,31	0,14	0,17**	0,08
Atlántico	5,91	5,41	0,50	0,40
Bogotá D.C.	21,25	22,15	-0,89	0,73
Bolívar	6,70	4,43	2,27***	0,38
Boyacá	4,67	3,58	1,10***	0,34
Caldas	3,86	4,08	-0,22	0,34
Caquetá	0,40	0,50	-0,09	0,12
Casanare	0,40	0,17	0,24***	0,08
Cauca	1,80	1,80	0,00	0,23
Cesar	1,86	1,92	-0,06	0,24
Choco	0,74	1,14	-0,40**	0,18
Córdoba	1,48	1,92	-0,44*	0,23
Cundinamarca	1,64	1,67	-0,03	0,22
Guainía	0,16	0,02	0,13***	0,04

Variable	Controles	Tratados	Diferencia	Error Est.
Guaviare	0,00	0,05	-0,05	0,03
Huila	1,98	1,98	-0,00	0,24
La Guajira	0,90	0,60	0,30**	0,14
Magdalena	1,08	1,81	-0,73***	0,22
Meta	0,99	0,80	0,19	0,16
Nariño	2,02	2,47	-0,45*	0,27
Norte De Santander	2,85	2,80	0,05	0,29
Putumayo	0,22	0,20	0,03	0,08
Quindío	2,58	2,01	0,57**	0,26
Risaralda	3,57	3,78	-0,20	0,33
San Andrés	0,00	0,03	-0,03	0,03
Santander	5,73	6,03	-0,30	0,41
Sucre	1,57	1,61	-0,04	0,22
Tolima	2,54	2,42	0,12	0,27
Valle Del Cauca	10,04	9,75	0,29	0,52
Vichada	0,07	0,00	0,07***	0,02

Fuente: Elaboración propia con información de las convocatorias de JII y CvLac.

Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente

En cuanto a la generación de productos de nuevo conocimiento por parte del grupo de tratamiento y control, se destacan que no hay diferencias significativas entre ambos grupos en el número promedio de productos que tenían antes de inscribirse en la convocatoria de JII, salvo para artículos con clasificación D para los que se observa que son significativamente más altos para los tratados (ver Tabla 17). Después de la intervención, la sumatoria de productos es mucho más alta tanto para tratados como controles, siendo los artículos A1 y A2 el tipo de producto que más publican los jóvenes que se inscribieron a JII; sin embargo, el promedio del total de artículos es significativamente más alto para los controles que para los tratados.

Tabla 17. Diferencia de medias en la sumatoria de productos de nuevo conocimiento entre tratados y controles antes y después de inscribirse a la convocatoria de JII

	Antes				Después			
Variable	Controles	Tratados	Diferencia	Error Est.	Controles	Tratados	Diferencia	Error Est.
Artículos A1 y A2	0,05	0,05	-0,01	0,01	2,12	1,79	0,33*	0,19
Artículos B y C	0,06	0,07	-0,01	0,01	1,25	1,11	0,14*	0,08
Artículos D	0,07	0,10	-0,03***	0,01	0,94	0,71	0,22***	0,05
Libros	0,01	0,01	-0,00	0,00	0,19	0,13	0,06***	0,02

Fuente: Elaboración propia con información de las convocatorias de JII y CvLac. Solo se emplea la submuestra de jóvenes que reportan productos publicados en CvLac.
Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente

Ahora bien, una vez se restringe la sumatoria de productos de nuevo conocimiento entre tratados y controles a los que se publicaron cinco años antes de la intervención y cinco años después de la intervención, se mantiene la diferencia estadísticamente significativa en el número de artículos D antes de la intervención (ver Tabla 18). No obstante, al comparar los artículos A1 y A2 y B y C se encuentran valores significativamente más altos para los tratados que para los controles. Esto demuestra entonces la dificultad de agregar los resultados para individuos que participaron en la intervención en diferentes momentos y sustenta la necesidad de utilizar el estimador de diferencias en múltiples períodos de tiempo y para múltiples grupos.

Tabla 18. Diferencia de medias en la sumatoria de productos de nuevo conocimiento entre tratados y controles 5 años antes y 5 años después de inscribirse a la convocatoria de JII

	Antes				Después			
Variable	**Controles**	**Tratados**	**Diferencia**	**Error Est.**	**Controles**	**Tratados**	**Diferencia**	**Error Est.**
Artículos A1 y A2	0,19	0,20	-0,01	0,03	0,61	0,79	-0,18***	0,06
Artículos B y C	0,22	0,27	-0,05	0,03	0,49	0,65	-0,16***	0,05
Artículos D	0,24	0,32	-0,08***	0,03	0,56	0,58	-0,02	0,05
Libros	0,02	0,02	0,00	0,01	0,05	0,06	-0,02	0,01

Fuente: Elaboración propia con información de las convocatorias de JII y CvLac. Solo se emplea la submuestra de jóvenes que reportan productos publicados en CvLac.
Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente

5. Resultados

5.1. Impactos preliminares del programa en productos de nuevo conocimiento

Los resultados indican que el programa genera incrementos en la publicación de artículos al cuarto año posterior a la convocatoria; a partir de este momento se generan impactos positivos y significativos de alrededor de 0,5 artículos que llegan a ser de 0,95 al octavo año, tal como se muestra en el panel I de la *Figura 24*. También se observa que los impactos en los tres primeros años antes del tratamiento son estadísticamente diferentes de cero, salvo para t=-3 cuyo valor es estadísticamente no significativo al 10% (z=2,03), por lo que se valida el supuesto de tendencias paralelas. En promedio, un joven investigador e innovador publica 0,3 más artículos que aquellos jóvenes que se inscribieron al programa y no fueron seleccionados (ver Tabla 19). Ahora bien, los impactos en publicaciones solo se dan para los hombres, quienes publican 0,56 artículos adicionales, en promedio, luego de participar en el programa. Los impactos se empiezan a observar en el tercer año posterior a la intervención (alrededor de 0,9) y se incrementan hasta llegar a 1,4 al octavo año. Para las mujeres el impacto promedio no es significativo, aunque al observar los impactos por año se observan algunos impactos significativos alrededor de los años 4 y 5 posteriores a la convocatoria.

Al desagregar los impactos por tipo de artículo publicado se encuentra que los impactos se concentran en los artículos A1 y A2 que son los más especializados. El impacto promedio observado es de 0,26 artículos A1 y A2; no obstante, el impacto es creciente en el tiempo posterior a la convocatoria, pasando de 0,2 en el cuarto año a 0,7 en el octavo año. Ahora bien, el impacto del programa en Artículos A1 y A2 solo se observa en los hombres, quienes incrementan el número de publicaciones de este tipo en 0,34 en el segundo año posterior a la convocatoria, hasta llegar a 1,2 artículos en el octavo año. En promedio, un hombre en el grupo de tratamiento publica 0,41 artículos A1 y A2 más que inscritos que no fueron seleccionados para el programa. Para las mujeres el impacto promedio no es significativo y solo se observan algunos efectos alrededor del octavo año posterior a la convocatoria. No se encuentran efectos en otro tipo de artículos (B y C) y tampoco en libros, salvo en algunos períodos posteriores al tratamiento.

Tabla 19. Impacto agregado del programa JII en productos de nuevo conocimiento

Variable de resultado	Total	Hombres	Mujeres
Total artículos	0,30***	0,56***	0,15
Artículos A1 y A2	0,26***	0,41***	0,06
Artículos B y C	0,06	0,13	0,06
Artículos D	- 0,02	0,02	0,03
Libros	0,01	0,02	0,01

Fuente: Elaboración propia con información de las convocatorias de JII y CvLac. Se emplearon variables de control como edad, sexo, departamento de nacimiento y si tiene estudios de maestría o doctorado en cada año y se empleó el estimador DRIPW que utiliza como ponderador la probabilidad inversa estabilizada. Solo se emplea la submuestra de jóvenes que reportan productos publicados en CvLac.
Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente

Es importante destacar que los impactos del programa en publicaciones se dan justo en aquellos productos de mayor calidad, en los artículos A1 y A2, y que no ocurren inmediatamente, sino que toman varios años en materializarse. Estos impactos son el reflejo de varias fuerzas/aspectos que confluyen durante la beca-pasantía. En primer lugar, durante el desarrollo de la pasantía los JII reciben mentoría y acompañamiento por parte de los investigadores que acompañan el proceso de los jóvenes. En el levantamiento de información cualitativa se destacó la importancia del proceso formativo durante la ejecución de la pasantía, ya que no solo desarrollan actividades de investigación, sino que cuentan con el acompañamiento y la guía de los docentes involucrados. Un resultado de este proceso formativo es el fortalecimiento de las habilidades técnicas en investigación y el fortalecimiento de capacidades para la investigación de alta calidad.

"para mí el centro, más que terminen produciendo un artículo -que en muchos casos se les exige-, para mí el centro es la formación. Creo que es el artículo que publica, pero el centro es que se formó para poder elaborar un artículo, etcétera, etcétera. Entonces me parece que diría que está el valor agregado es en la formación, más que en los productos y su impacto"

(entrevista a directora de grupo de investigación, 2023)

"claramente el Programa de Jóvenes Investigadores sí les permite a ellos adquirir unas habilidades en investigación porque se enfrentan a unas competencias que ellos no han adquirido claramente durante su proceso de formación en pregrado. Porque cuando inician en Jóvenes Investigadores, pues ya es un contexto diferente donde están enmarcados en un proyecto de investigación, deben responder por unas actividades relacionadas con el proyecto de investigación y eso hace que ellos desarrollen esas competencias en investigación"

(entrevista a director de grupo de investigación, 2023)

Por tanto, el programa, como un espacio inherentemente formador en el que se generan y cualifican capacidades académicas tiene la capacidad de generar el desarrollo de aptitudes y habilidades investigativas en los participantes, lo que se ha traducido en un mayor número de publicaciones de alta calidad. Algunos actores incluso consideran el desarrollo de la pasantía como un escenario introductorio a las dinámicas de un doctorado.

"Por las mismas características del programa, normalmente las personas que acceden a ese programa o ya tienen contacto con el grupo y ya tienen como un feeling hacia lo que está haciendo el grupo de investigación. Entonces, es un poco como un programa de doctorado en el primer nivel"

(entrevista a directora de grupo de investigación, 2023)

Adicionalmente, los JII tienen la posibilidad de insertarse en las dinámicas de los grupos de investigación, lo que les ayuda a confirmar o no el gusto y las aptitudes por la carrera investigativa; y,

dependiendo de las dinámicas particulares de cada grupo, los JII pueden incluso tener más publicaciones al ser coautores de los productos derivados de otros proyectos de investigación en el grupo.

"¿qué nos ha pasado a nosotros y cuál es nuestra experiencia? Todos los jóvenes investigadores que han pasado por la Fundación terminan con publicaciones, la publicación principal derivada del estudio financiada por el Ministerio pero nosotros los invitamos a que se vinculen a más actividades dentro del grupo de investigación por decir a otros proyectos de investigación y terminan saliendo con 2 o 3 publicaciones en journals internacionales, entonces no solamente ellos adquieren una experiencia dentro del grupo de investigación y es una experiencia certificada, sino que adicionalmente terminan con productos de investigación, artículos científicos, ponencias etcétera, etcétera, etcétera, entonces pues para un joven investigador eso es realmente muy bueno"

(entrevista a funcionario de centro/instituto de investigación, 2023)

Vale la pena mencionar que los impactos en publicaciones no son únicamente para los participantes en el programa, sino también para los grupos que se vinculan a él, ya que pueden fortalecer su capacidad de producción de artículos, libros, entre otros. En particular, para las instituciones dedicadas a investigación que participan en el programa estiman altamente la experiencia en el programa porque les permite adelantar un entrenamiento específico al tiempo que ahorran tiempo y recursos en la incorporación de nuevo talento humano.

"Jóvenes Investigadores nos brindaba la capacidad de tener esos muchachos acá al 100% o sea una persona brindándole las condiciones económicas al 100%, lo que pueda hacer usted con ellos es lo que existe, patentes, artículos de alta calidad"

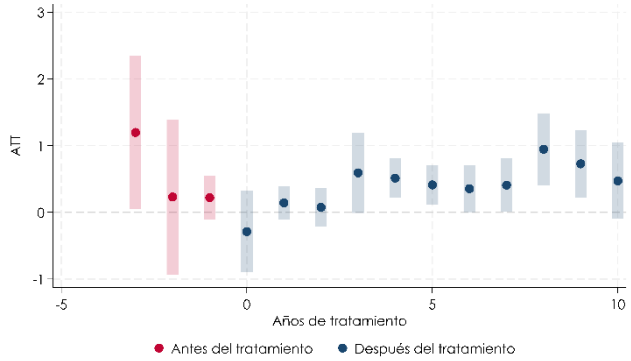
(entrevista a director de grupo de investigación, 2023)

"La producción realmente con Jóvenes Investigadores […] garantiza que haya un nivel de producción mínimo y que esas habilidades del joven investigador se vean reflejadas precisamente en producto, no solamente publicaciones, lo digo no solamente por mi grupo, sino por otros grupos de ingeniería que han tenido también Jóvenes Investigadores donde terminan realizando prototipos, software, o sea otro tipo de productos que no son simplemente publicaciones"

(entrevista a directora de grupo de investigación, 2023)

Figura 24. Impactos del programa JII en productos de nuevo conocimiento en cada año

I. Impactos de JII en el número de artículos publicados por año



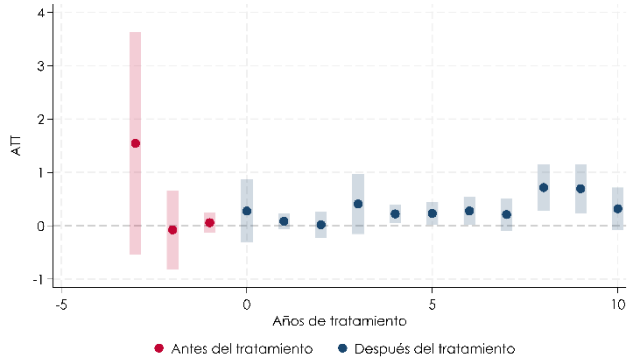
II. Impactos de JII en el número de artículos publicados por año – Hombres



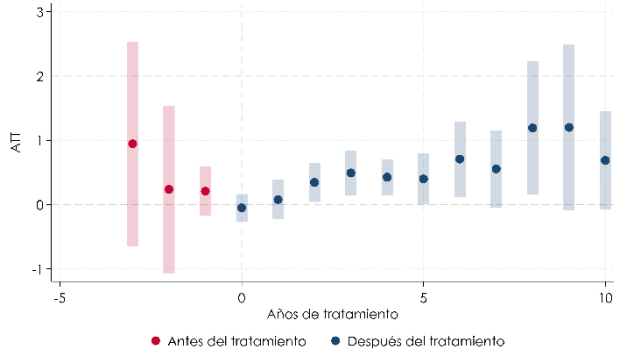
III. Impactos de JII en el número de artículos publicados por año - Mujeres



IV. Impactos de JII en el número de artículos A1 y A2 publicados por año



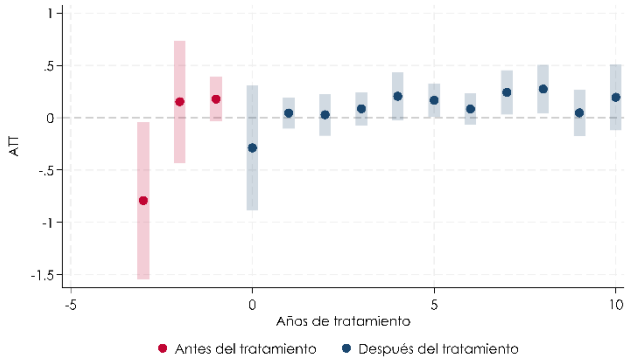
V. Impactos de JII en el número de artículos A1 y A2 publicados por año – Hombres



VI. Impactos de JII en el número de artículos A1 y A2 publicados por año – Mujeres



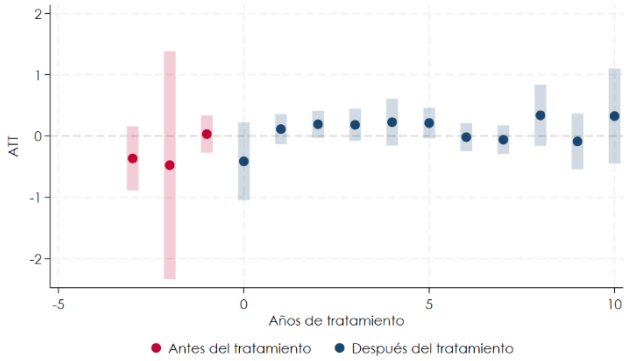
VII. Impactos de JII en el número de artículos B y C publicados por año



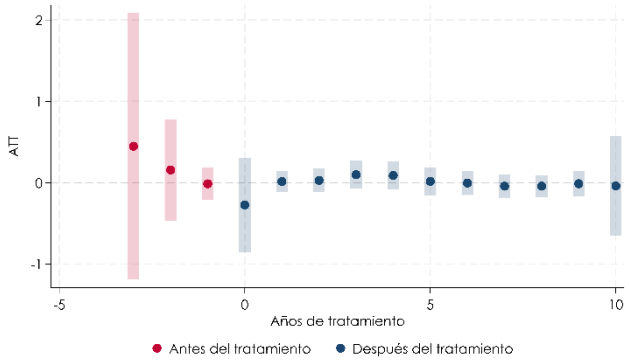
VIII. Impactos de JII en el número de artículos B y C publicados por año - Hombres



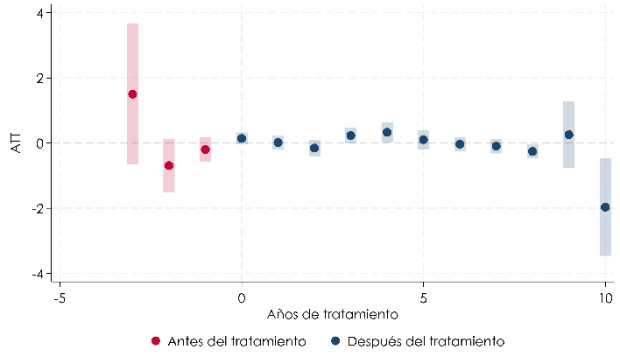
IX. Impactos de JII en el número de artículos B y C publicados por año – Mujeres



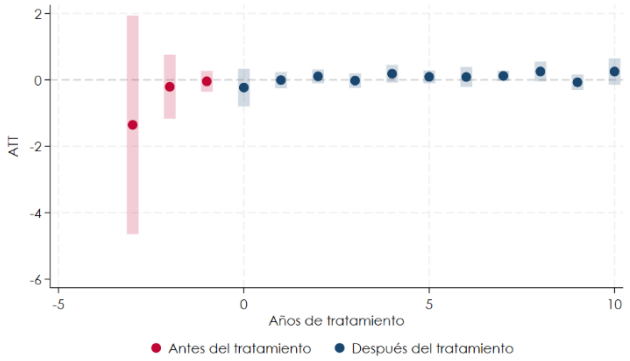
X. Impactos de JII en el número de artículos D publicados por año



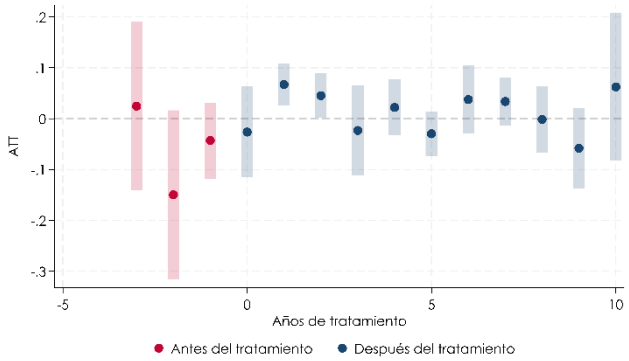
XI. Impactos de JII en el número de artículos D publicados por año– Hombres



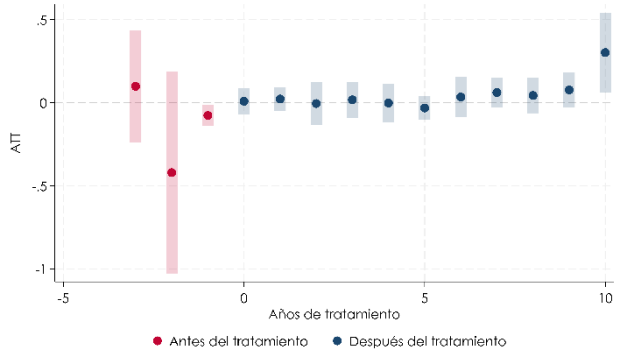
XII. Impactos de JII en el número de artículos D publicados por año – Mujeres



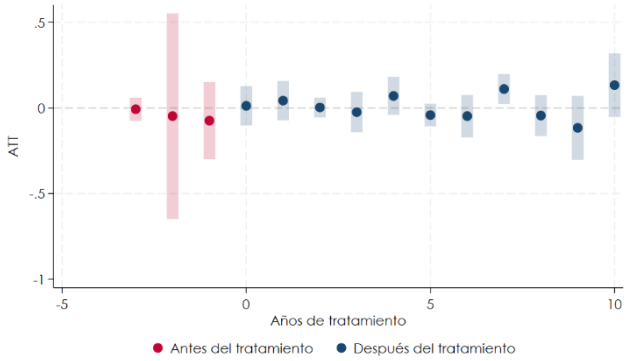
XII. Impactos de JII en el número de libros publicados por año



XIV. Impactos de JII en el número de libros publicados por año – Hombres



XV. Impactos de JII en el número de libros publicados por año– Mujeres



Fuente: Elaboración propia con información de las convocatorias de JII y CvLac. Se emplearon variables de control como edad, sexo, departamento de nacimiento y si tiene estudios de maestría o doctorado en cada año. Solo se emplea la submuestra de jóvenes que reportan productos publicados en CvLac. Los intervalos de confianza corresponden al 95%.

5.2. Impactos preliminares del programa en variables de mercado laboral

Los resultados de las estimaciones empleando a los futuros tratados como grupo de control arrojan resultados positivos del programa en variables del mercado laboral. El impacto promedio del programa en el IBC empleado para la cotización a pensión es positivo y tiende a crecer en los años posteriores al tratamiento (Figura 25, ver panel A). En promedio, este impacto es de $213.574 (en pesos de 2018); esto representa alrededor del 27% del salario mínimo de 2018. Además, los tratados cotizan 1,4 meses más al año después de haber participado en el programa. La Figura 25 muestra que el impacto en meses cotizados es de alrededor de 1-1,5 meses en los dos años posteriores al tratamiento, pero se dispara en el tercer año. Nótese que todos los coeficientes estimados antes del tratamiento, los cuales se muestran en color rojo en las gráficas de la Figura 25 son no significativos y los intervalos de confianza se cruzan con el cero, validando el supuesto de tendencias paralelas.

Cuando se desagrega la información entre hombres y mujeres, se encuentra que el efecto es ligeramente mayor para hombres cuando se emplea el IBC como variable de resultado, pero la diferencia es más grande cuando se considera el número de meses cotizados. El impacto en aportes a pensiones para los hombres es positivo y creciente en los años posteriores al tratamiento, siendo, en promedio $259.679. En el caso de las mujeres, el impacto es también creciente en los años posteriores al tratamiento, arrojando un valor promedio de $256.691. En cuanto a los meses cotizados, se encuentra que los hombres del grupo de tratamiento aportan 1,5 meses más luego de la beca-pasantía mientras que las mujeres realizan aportes a pensión por 1,38 meses más que el grupo de control.[31]

Tabla 20. Impactos del programa JII en IBC y cotización a pensión

Variable de resultado		**ATT**
Total	IBC pensión	213.574,49***
	Meses cotizados a pensión	1,4***
Hombres	IBC pensión	259.679,48***
	Meses cotizados a pensión	1,5***
Mujeres	IBC pensión	256.691,64***
	Meses cotizados a pensión	1,3***

Fuente: Elaboración propia con información de las convocatorias de JII y PILA. Los valores de IBC están deflactados a pesos de 2018. Se emplearon variables de control como edad, sexo, departamento y municipio y se empleó el estimador OR para los resultados totales y de mujeres y DRIPW para los hombres en la medida que ponderaba mejor sus resultados pretratamiento.
Nota: *, **, *** Corresponde a los niveles de significancia del 10%, 5% y 1% respectivamente.

Uno de los beneficios inmediatos del programa es el monto mensual que se entrega a los JII, por lo que no es de extrañarse los impactos en esta dimensión. Ahora bien, los resultados muestran que en caso de no tener la beca-pasantía sino cualquier otra alternativa laboral, los jóvenes tendrían ingresos más bajos y menor estabilidad en el mercado laboral. La beca-pasantía representa una valiosa oportunidad para que los jóvenes se inserten en el mercado laboral y nutran su hoja de vida, demostrando experiencia investigativa e incluso, dependiendo de la institución, experiencia laboral. Esto es una gran ventaja considerando los desafíos que enfrentan los jóvenes recién graduados para obtener un primer empleo.[32] Además, en muchos casos, los ingresos de la beca-pasantía permiten a los jóvenes financiar o complementar los gastos de los estudios de maestría, los cuales pueden adelantar al tiempo que desarrollan sus actividades en el grupo de investigación.

[31] También se intentaron estimaciones con el IBC de cotización al SGSS en Salud, pero no siempre se pudieron validar los supuestos del modelo, por lo que los resultados no son concluyentes.

[32] Según el Observatorio Laboral para la Educación (OLE), el 69,3% de los graduados de educación superior en el sector oficial eran cotizantes en 2021, mientras que el 79,1% de los graduados de educación superior en el sector privado en 2020 era cotizantes en 2021. Ver: https://ole.mineducacion.gov.co/portal/ (consultado el 29 de mayo de 2024).

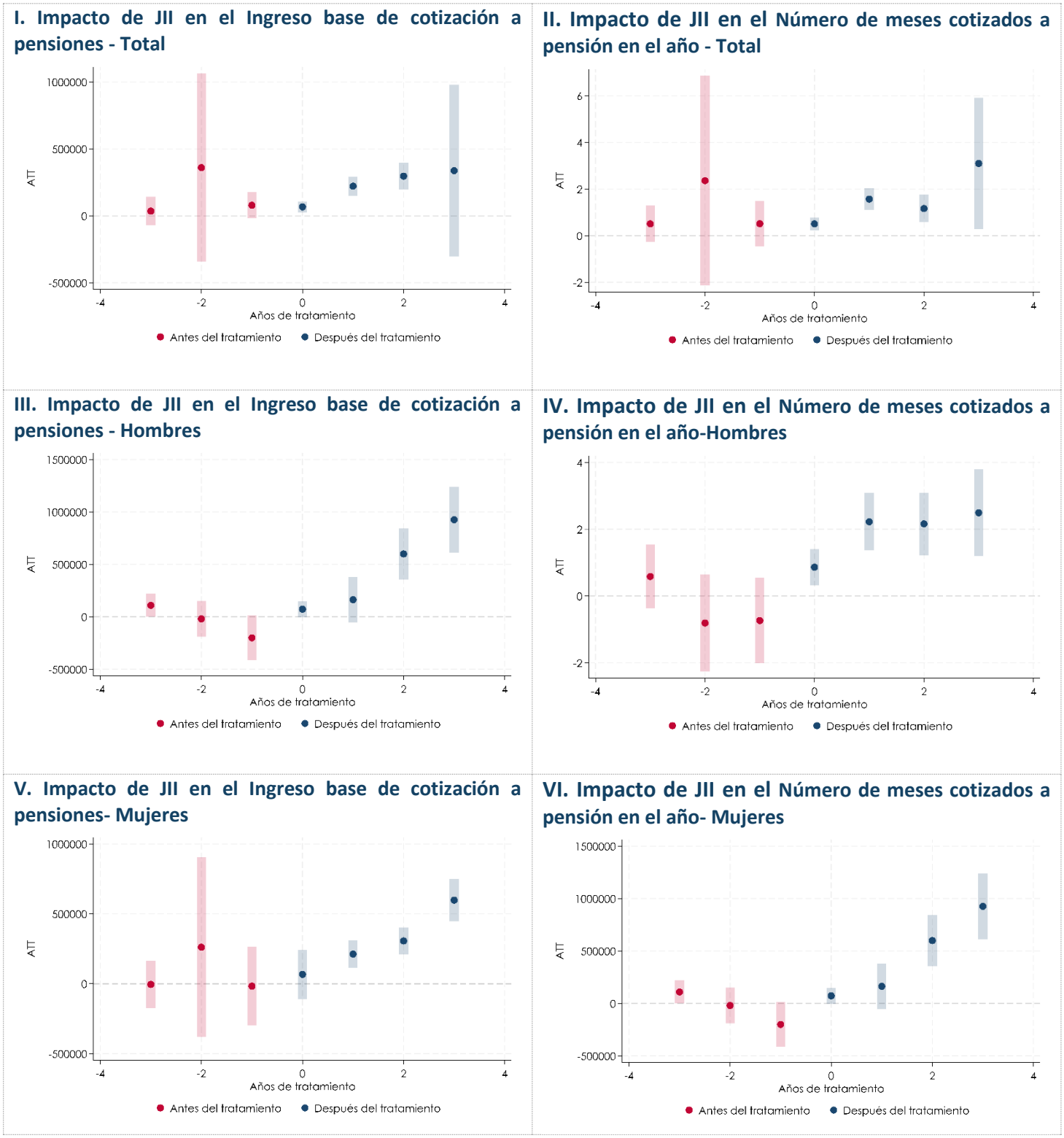
"le permite a la persona hacer algo que es importantísimo que es ir generando un currículo porque en la mayoría de los casos uno se encuentra siempre con la misma sensación digamos como de ansiedad en algunos perfiles de personas que están buscando ubicarse y que siempre piden experiencia. Entonces un joven investigador pues ya lleva un acervo, se pueden generar algunos artículos, que eso también enriquece el currículo. Así que a mí me parece que la figura es una figura bastante robusta e interesante para todas las áreas de generación de conocimiento y adicionalmente los diferentes formatos, para los profesores de una universidad pues claramente, pero para un centro como el nuestro, que es un centro de investigación, pues obviamente también nos permite acceder a profesionales jóvenes que pueden robustecerse"

(entrevista a funcionaria de centro/instituto de investigación, 2023)

"el poder tener al menos uno de los dos años pagado como joven investigador, eso es un alivio muy grande para la matrícula. Yo pude tener varios de estos estudiantes que han pasado por el grupo como jóvenes investigadores que hacían simultáneamente maestría y que con esto se podían financiar"

(entrevista a director de grupo de investigación, 2023)

Figura 25. Impactos del programa JII en IBC y cotización a pensión



Fuente: Elaboración propia con información de las convocatorias de JII y PILA. Los valores de IBC están deflactados a pesos de 2018. Se emplearon variables de control como edad, sexo, departamento y municipio y se empleó el estimador OR para los resultados totales y de mujeres y DRIPW para los hombres en la medida que ponderaba mejor sus resultados pretratamiento. Los intervalos de confianza corresponden al 95%.

6. Análisis costo-beneficio del programa Jóvenes Investigadores e Innovadores

El Análisis Costo-Beneficio (ACB) es una herramienta que permite comparar sistemáticamente los beneficios generados por una intervención con los costos en que se incurrieron para llevarla a cabo. Para ello, es necesario cuantificar y monetizar detalladamente los costos asociados al programa, entre los cuales se incluyen los costos privados asumidos por los participantes y las entidades vinculadas, los costos de implementación del programa y los costos de oportunidad incurridos por los diferentes actores, tanto en términos de tiempo como de otras alternativas al programa que no se aprovecharon. Por el lado de los beneficios, se debe cuantificar y monetizar tanto la retribución directa que reciben los participantes, como los impactos de la intervención, incluyendo todo beneficio para la sociedad en general. A continuación, se detallan los costos y los beneficios cuantificados para JII, así como la comparación entre ellos.

6.1. Costos

El monto destinado a cada JII como parte de la beca-pasantía puede variar según la convocatoria y el joven que se postule. Por ejemplo, la Convocatoria 874 de 2020 "para el fortalecimiento de proyectos en ejecución de CTeI en ciencias de la salud con talento joven e impacto regional" vinculó a jóvenes profesionales y de pregrado; la diferencia entre ambos es que los primeros ya están graduados, mientras que los segundos deben tener mínimo el 60% del pregrado completo. En este caso, el valor de la beca pasantía es de $2.500.000 por 12 meses para los jóvenes profesionales, para un total de $30.000.000; mientras que para los de pregrado es de $1.000.000 por 12 meses, para un total de $12.000.000.[33] Por su parte, para la Convocatoria 907 de 2021 destinada a jóvenes investigadores e innovadores en el marco de la reactivación económica se vincularon jóvenes profesionales con beca-pasantía de $2.000.000 y jóvenes de pregrado con beca-pasantía de $1.000.000, ambos por 12 meses.

En 2009 el monto promedio destinado a los JII era de casi $1.500.000, mientras que en 2021 era de $2.623.612 (o de $2.500.000 para la mediana). Al analizar los montos de la beca-pasantía en precios constantes, se observa que el monto ha permanecido relativamente constante por encima de los $2.000.000. Así, el costo directo promedio mensual por JII ha sido de $2.229.583 (cuando se usa la mediana para evitar la influencia de valores extremos). Por tanto, el costo directo anual del programa por beneficiario es de $26.755.002.

Tabla 21. Monto destinado a los JII

	Precios corrientes		**Precios diciembre 2018**	
Año de la convocatoria	**Monto promedio**	**Mediana**	**Monto promedio**	**Mediana**
2009	1.490.700	1.490.700	2.093.680	2.093.680
2010	1.545.000	1.545.000	2.103.472	2.103.472
2011	1.606.800	1.606.800	2.108.938	2.108.938
2012	1.700.100	1.700.100	2.178.219	2.178.219
2013	1.768.500	1.768.500	2.222.851	2.222.851
2014	1.848.000	1.848.000	2.240.815	2.240.815
2015	1.933.000	1.933.000	2.195.344	2.195.344
2016	2.077.349	2.068.362	2.231.070	2.221.418

[33] https://minciencias.gov.co/convocatorias/vocaciones-cientificas-ctei/convocatoria-para-el-fortalecimiento-proyectos-en

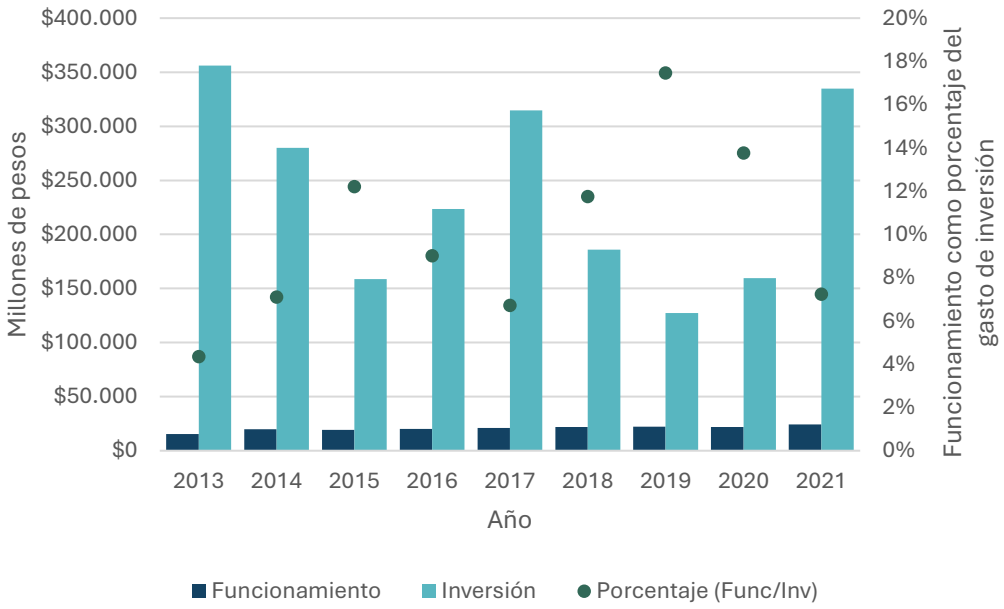
2017	2.213.151	2.213.151	2.283.482	2.283.482
2018	2.032.455	2.343.726	2.032.455	2.343.726
2019	2.509.304	2.484.348	2.417.441	2.393.399
2020	2.218.255	2.484.348	2.103.010	2.355.279
2021	2.623.612	2.500.000	2.354.916	2.243.964
Promedio			**2.197.361**	**2.229.583**

Fuente: Elaboración propia con información de las convocatorias de JII.

El monto total destinado a cada JII no tiene en cuenta los costos adicionales en los que incurre MinCiencias para la implementación del programa. Estos costos adicionales abarcan salarios del personal u otros gastos en que se deba incurrir para llevar a cabo todo el proceso de convocatoria, selección y seguimiento de los JII. La Figura 26 presenta el gasto total de funcionamiento e inversión de MinCiencias en los últimos nueve años. Los gastos de funcionamiento incluyen partidas como salarios, contribuciones de nómina, adquisición de bienes y servicios, impuestos y otras tasas; mientras que los gastos de inversión incluyen lo que gastó el Ministerio en programas. La información de ejecuciones presupuestales no permite desagregar estos gastos entre los diferentes programas de MinCiencias, ni desligar el gasto de inversión en JII de otros programas de desarrollo de vocaciones científicas y capacidades para la investigación en NNA a nivel nacional.

Por tanto, se propone utilizar el porcentaje de gasto de funcionamiento dentro del gasto de inversión como una aproximación de los gastos de funcionamiento en los que se incurre para llevar a cabo el programa. Es decir, el gasto en funcionamiento como proporción del gasto de inversión ha fluctuado entre el 4% y el 17% entre 2013 y 2021, con un promedio de 10%. Así, se propone adicionar un valor del 10% del monto total de la beca al costo total para dar cuenta de los gastos de operación anuales. De esta manera, el monto de la beca es de $26.755.002 por año, por lo que el costo anual de operación equivaldría a $2.675.500 (el 10%), para un costo total del programa por beneficiario de $29.430.502.

Figura 26. *Gasto de inversión y de funcionamiento MinCiencias (pagos)*



Fuente: Ejecuciones presupuestales de MinCiencias

Es importante mencionar que el costo total de aproximadamente $29 millones subestima el costo total del programa, en la medida en que no incluye los costos en los que incurren las instituciones que acogen a los jóvenes, tales como el tiempo de los profesores y/o investigadores para hacer la mentoría y el acompañamiento, costos de entrenamiento, materiales, productividad perdida, entre otros.[34]

6.2. Beneficios

Un primer paso para la metodología de costo-beneficio es monetizar los efectos de la intervención y estimar el flujo de beneficios esperados en el tiempo para poderlo comparar con los costos.

Beneficios en el ingreso laboral. Como se mencionó en secciones anteriores, un joven investigador e innovador tiene un IBC promedio de $2.071.734 por mes, y cotiza 9,5 meses en el año (ver Tabla 15). Esto se traduce en ingresos anuales de $19.681.473 para los JII. Los impactos estimados del programa incluyen un aumento de $213.574 en el IBC y de 1,4 meses en el número de meses cotizados en un año; lo que implica que, sin la intervención, los jóvenes tendrían un IBC promedio de $1.858.160 y cotizaría un promedio de 8,1 meses en el año.[35] Así, en ausencia de la intervención, un joven que se inscribió al programa pero que no fue seleccionado tendría un IBC promedio de $15.051.092. Por lo tanto, el beneficio monetario derivado de los impactos en IBC y en meses de cotización son de $4.630.381 en un año, resultado de la diferencia entre el ingreso laboral anual de los JII y los jóvenes que no fueron seleccionados.[36]

Beneficios por publicaciones. Otro impacto de JII es una mayor producción de publicaciones académicas, específicamente 0,26 artículos A1 y A2 adicionales. Siguiendo a (Núñez, et al., 2014), las publicaciones pueden monetizarse utilizando el sistema de puntos salariales para profesores de universidades estatales definido en el Decreto 1279 de 2002. Este decreto asigna bonificaciones de 15 puntos salariales por artículos de tipo A1 y 12 puntos salariales por artículos de tipo A2 (ver artículo 10). Por su parte, el Decreto 208 de 2024 fijó el valor del punto salarial en $20.895 (ver artículo 2). Así, un joven investigador publica, en promedio, 0,26 artículos A1 y A2 más que los que no fueron seleccionados en JII y, asumiendo un enfoque conservador, de 12 puntos salariales por artículo A2, el beneficio en puntos salariales sería de 3,12 (0,26*12). Estos puntos salariales se traducirían en un beneficio monetizado anual de $619.328.

Ahora bien, un joven investigador publica 0,26 artículos A1 y A2 más que los jóvenes que se inscribieron y no fueron seleccionados en JII. Asumiendo un enfoque conservador, de incrementos de 12 puntos salariales por artículo A2, el aumento total para los JII sería de 12*0,26 puntos salariales, equivalente a 3.12 puntos salariales, lo que, valorado en $20.895 arroja un beneficio monetizado derivado de publicaciones equivalente a $65.192 pesos por mes. Dado que, los JII cotizan en promedio 9,5 meses en el año, el beneficio por publicaciones en un año sería de $619.324.

Flujo futuro de beneficios. Un aspecto crucial es determinar por cuánto tiempo duran los impactos en ingresos laborales de los JII. Las ganancias derivadas de la mayor experiencia laboral, salarios de los beneficiarios y del proceso formativo en investigación pudieran durar toda la vida laboral de los JII.

---

[34] Estimar estos costos requeriría información más detallada del número de horas que dedica el mentor al JII, la escala salarial del mentor, los materiales entregados por la institución que acoge a los JII, entre otros factores para los cuales no se cuenta con la información.

[35] Esta cifra viene de restar al IBC promedio de los JII de 2.071.734 el impacto estimado de la intervención ($213.574) para obtener el IBC que tendría un joven que no participó en el programa ($1.858.160). De la misma forma se procede para obtener el número de meses que cotizaría un joven promedio que no participó en el programa: 9,5-1,4=8,1 meses.

[36] Dicho de otra forma, los beneficios en un año equivalen a $213.574 en los 9,5 meses que cotiza un JII más los 1,4 meses de salario que los controles no reciben por no estar en el programa.

Sin embargo, hay autores que destacan que la acumulación de capital humano derivado de procesos formativos, como el que ocurre con las actividades de investigación con JII, puede perderse si se está inactivo o si no se usan las habilidades que se han adquirido (Blomberg, 1989). Más aún, Almeida y Carneiro (2005) argumentan que las habilidades adquiridas en programas de capacitación se deprecian puesto que el conocimiento se vuelve obsoleto con el tiempo, por lo que emplean una tasa de depreciación de habilidades del 17%. [37]

Teniendo estas consideraciones, para JII se asume un escenario en el que los beneficios laborales en IBC se asumen durante toda la trayectoria laboral, desde los 24 años – que es la edad promedio de los JII—hasta los 60 años de edad –que es la edad promedio de jubilación (promedio entre la edad de hombres, de 62 años, y mujeres, de 57 años). También se asume que los ingresos laborales aumentan a una tasa promedio del 5,5%, basada en el incremento promedio del salario mínimo en los 10 años previos a la pandemia—entre 2011 y 2020 (ver Tabla 22). No se toma para este promedio los valores del aumento del salario mínimo entre 2021 y 2024 debido a las distorsiones provocadas por la pandemia y a la elevada inflación subsecuente, la cual se esperaría que retorne a los niveles prepandemia en el corto plazo.

Tabla 22. Aumento del salario mínimo e inflación

Año	Aumento del salario mínimo	Inflación
2010	3,6%	3,2%
2011	4,0%	3,7%
2012	5,8%	2,4%
2013	4,0%	1,9%
2014	4,5%	3,7%
2015	4,6%	6,8%
2016	7,0%	5,8%
2017	7,0%	4,1%
2018	5,9%	3,2%
2019	6,0%	3,8%
2020	6,0%	1,6%
2021	3,5%	5,6%
2022	10,7%	13,1%
2023	16,0%	9,3%
2024	12,1%	

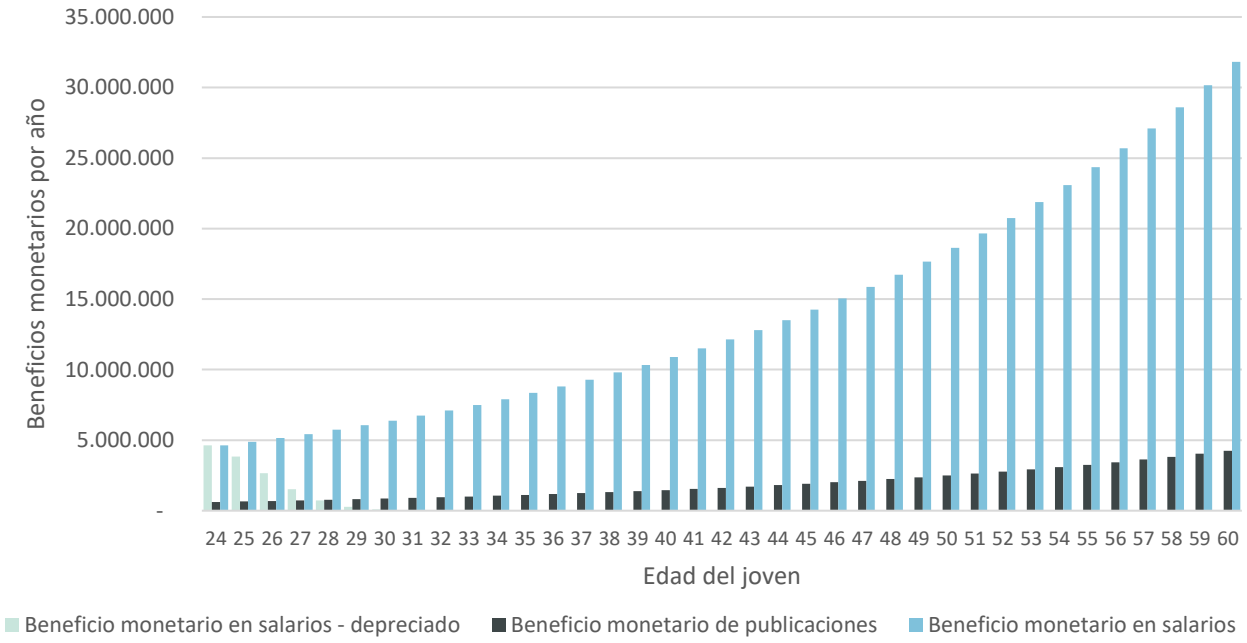
Fuente: Banco de la República.

Adicional a los supuestos anteriores, se asume que los beneficios monetarios del salario se deprecian a una tasa del 17% puesto que las habilidades adquiridas pueden perderse a lo largo del tiempo, si bien, también se considera un escenario sin esta depreciación como parte del análisis de sensibilidad. Con los supuestos anteriores, en la Figura 27 se pueden observar el flujo proyectado de beneficios monetarios derivados de la intervención a lo largo de la vida laboral del JII (entre los 24 y 60 años). Nótese que al asumir que los beneficios monetarios derivados de aumentos del salario se deprecian, estos beneficios solo se observarían en los primeros 13 años de la intervención, lo cual es un escenario

[37] Este estudio estima la tasa interna de retorno de inversiones de *"on-the-job training"* que son programas de formación en los que los trabajadores aprenden y desarrollan habilidades en su lugar de trabajo mientras realizan sus actividades diarias en la firma en que trabajan. Los autores afirman que las habilidades se pueden depreciar porque los empleados abandonan la firma—y en el caso de JII puede pasar porque los jóvenes abandonan los grupos de investigación—o por la rotación de empleados.

conservador para la estimación de los beneficios (si se compara con el caso en el que dicho beneficio se mantiene a lo largo de la vida laboral). Es importante aclarar que no se deprecia el beneficio monetario de publicaciones puesto que las publicaciones de los JII no dejan de existir y pueden seguir contribuyendo al prestigio y a las oportunidades profesionales de los JII.

Figura 27. Proyección del flujo de beneficios futuro



Fuente: elaboración propia.

Consideraciones adicionales. Es importante tener en cuenta que la estimación de los beneficios es conservadora en la medida en que se toma el IBC como proxy de los ingresos laborales. El IBC es igual al salario en el caso de los asalariados, pero es solo el 40% de los honorarios de los trabajadores independientes. De acuerdo con la GEIH 2022, el 26,8% de la población con educación superior completa y con edades entre los 24 y 60 años es trabajador independiente, por lo que para al menos un cuarto de los JII se puede estar subestimando el ingreso laboral total y, por tanto, los beneficios totales del programa. Adicionalmente se asume que solo se dan 3,12 puntos salariales por las 0,26 publicaciones adicionales en toda la historia laboral.

Beneficios sociales. Es razonable esperar beneficios derivados de la vinculación de los JII a los grupos de investigación que van más allá de los beneficios monetarios que pueden recibir los JII y que no son fáciles de monetizar. En primer lugar, el programa JII influye en el fortalecimiento de los grupos, centros e institutos de investigación que acogen a los participantes del programa, en la medida en que los JII contribuyen a la estabilidad de los grupos de investigación y a la producción de conocimiento, tal como se mencionó en las entrevistas cualitativas:

"Hay un valor agregado para el grupo de investigación, porque da fortalecimiento y categorización y sostenibilidad de los grupos de investigación ¿en qué sentido? da un valor agregado en que ellos siguen generando conocimiento y productos al grupo de investigación para su categorización. […] ¿qué le favorece al grupo de investigación categorizado? Generar más proyectos, generar más

alianzas nacionales e internacionales y seguir fortaleciendo la investigación, la innovación y el desarrollo tecnológico"

(entrevista a directora de grupo de investigación, 2023)

El aporte del programa JII al fortalecimiento de los grupos de investigación y a la generación de productos de nuevo conocimiento puede también tener efectos económicos en el largo plazo en la medida en que puede contribuir al aumento de la competitividad, la innovación y el desarrollo tecnológico, con sus consecuentes efectos en crecimiento económico.

Otro de los beneficios sociales destacados por los informantes durante las entrevistas para la recolección de información cualitativa es el aporte del Programa JII a la inclusión social en el país, en la medida en que permite a los jóvenes que participan tener un año de financiación de actividades de investigación, realizar intercambios y aprovechar oportunidades en el exterior que de otro modo no hubieran podido acceder. Este sin duda es un aporte valioso a la construcción de una sociedad más equitativa y con más oportunidades.

"es dramático, eso sí que produce cambios de movilidad social, pues estuve yo en Purdue y tuve la oportunidad de ver un muchacho que se fue de acá, él era hijo de una vendedora ambulante y vivían en un cuarto con la mamá y otro hermanito y la mamá lo pudo mantener hasta que el pelao salió y él se fue a Purdue con el Programa de Nexo Global y empezó a hacer allá una rotación. Inmediatamente lo vincularon al doctorado y pasó de la Nacional a Purdue al doctorado y allá se quedó, imagínese lo que es eso para una familia en Colombia"

(entrevista a director de grupo de investigación, 2023)

6.3. Valor presente neto del flujo de beneficios y costos

Una vez identificados y monetizados los costos y beneficios del programa, el siguiente paso consiste en cuantificar los beneficios netos de costos a lo largo del tiempo, esto es, comparar los flujos de beneficios contra los flujos de costos. Para hacer esta comparación se calculan tres indicadores:

Valor Presente Neto (VPN): es la sumatoria de todos los beneficios B_t en el tiempo descontados a una tasa social de descuento s, menos la sumatoria de todos los costos C_t en el tiempo descontados a la misma tasa social de descuento. Cuando esta resta es positiva, quiere decir que el programa es rentable para la sociedad.

$$VPN = \sum_{t=0}^{T} \frac{B_t}{(1+s)^t} - \sum_{t=0}^{T} \frac{C_t}{(1+s)^t}$$

Razón entre beneficios y costos: es la razón entre el valor presente de todos los beneficios y el valor presente de todos los costos, ambos descontados a una tasa social de descuento s. El programa es beneficioso para la sociedad cuando esta razón es mayor que uno.

$$RatioBC = \frac{\sum_{t=0}^{T} \frac{B_t}{(1+s)^t}}{\sum_{t=0}^{T} \frac{C_t}{(1+s)^t}}$$

Tasa Interna de Retorno (TIR): esta tasa da cuenta de la rentabilidad que debe tener el programa para generar beneficios a la sociedad y es igual a la tasa que iguala los beneficios y costos de la intervención.

$$\sum_{t=0}^{T} \frac{B_t}{(1+i)^t} - \sum_{t=0}^{T} \frac{C_t}{(1+i)^t} = 0$$

Un parámetro fundamental para el cálculo de estos indicadores y para poder comparar el flujo de beneficios y costos es la tasa social de descuento. Esta tasa mide la tasa a la cual la sociedad estaría dispuesta a intercambiar consumo presente por consumo futuro. López (2008) estima que la tasa de descuento para Latinoamérica está alrededor del 3-4% en escenarios de bajo crecimiento futuro, o de 5-7% en escenarios de alto crecimiento; además, este autor plantea que la tasa de descuento adecuada también depende del horizonte del proyecto, siendo 4,4% una tasa adecuada para proyectos con horizontes de 25 años y 4,2% para proyectos con horizontes de 50 años. Dado que en este estudio se están asumiendo beneficios a lo largo de la vida laboral de los jóvenes (aproximadamente 40 años), se emplea esta última tasa del 4,2% como tasa de descuento.

La Tabla 23 presenta los resultados del ACB empleando una tasa de descuento del 4,2% y para dos escenarios para la cuantificación del flujo de beneficios privados de los JII: sin depreciación de habilidades y con depreciación de habilidades (del 17%). Además, se presentan cuatro escenarios adicionales para diferentes tasas de descuento. En términos generales, el programa presenta resultados positivos para la sociedad y sus beneficios superan ampliamente los costos sin importar el escenario escogido; sin embargo, los resultados fluctúan mucho según el supuesto que se asuma. Cuando se asume que el beneficio en salario y tiempo de cotización se mantiene a lo largo de la vida laboral de los JII, la TIR es del 11% y la ratio B-C es de 8,32 – empleando una tasa de descuento del 4,2%; mientras que, si se asume que las habilidades adquiridas se deprecian, y que, por tanto, los beneficios en salario no se mantienen a lo largo de la vida laboral, la TIR se reduce a 1,5% y la ratio B-C a 1,42.

Tabla 23. Resultado del ACB

	Tasa de descuento: 4,2%			**Tasa de descuento: 3%**			**Tasa de descuento: 5%**		
	VPN	**Ratio B-C**	**TIR**	**VPN**	**Ratio B-C**	**TIR**	**VPN**	**Ratio B-C**	**TIR**
Sin depreciación de habilidades	215.510.328	8,32	11,0%	279.567.852	10,50	11,6%	182.421.393	7,20	10,6%
Con depreciación de habilidades	12.497.162	1,42	1.5%	20.251.624	1,69	2,1%	8.466.539	1,29	1,1%

En ese sentido, la tasa de retorno a la inversión hecha en JII fluctúa entre el 1,5% y el 11%. Almeida y Carneiro (2004) estiman los retornos a las inversiones en programas de capacitación en el trabajo que implican entrenamiento de manera práctica de trabajadores, incluyendo acompañamiento y mentoría, encontrando que la mediana de la tasa de retorno es del 10%. Los autores plantean que la mayoría de los programas de este estilo no logran medir propiamente los costos o la depreciación de las habilidades adquiridas. Así, para el caso de JII, la tasa de retorno pudiera estar más cerca del 1,5% que del 11% en la medida en que se subestiman los costos asociados a la intervención al no considerar aquellos incurridos por la institución y los mentores.

7. Conclusiones

El programa JII tiene beneficios positivos para sus participantes en términos académicos, así como para las entidades y grupos de investigación que los vinculan y para la sociedad en general. Durante la beca-pasantía los JII reciben acompañamiento y mentoría y se insertan en las dinámicas propias de los grupos de investigación, lo que les permite adquirir experiencia, mejorar sus habilidades

científicas, ser incluidos como coautores en productos de nuevo conocimiento, y prepararse mejor para la educación posgradual y para la vinculación en grupos de investigación en general. Uno de los aspectos positivos destacados del programa es que funciona como preparación para programas de posgrado, e incluso facilita la realización de ellos, ya sea porque el monto financiado ayuda a cubrir los costos de maestrías, porque se les facilita la búsqueda de apoyos y becas, o porque les abre las puertas a estudios de doctorado. No es de extrañarse entonces que uno de cada tres JII realiza estudios de maestría después de terminar la beca-pasantía y cerca de uno de cada 10 realiza estudios de doctorado después de finalizado el programa.

La estimación de los impactos arrojó resultados positivos en producción académica, específicamente en artículos A1 y A2 que son de alta calidad. Estos impactos se derivan no solo del entrenamiento práctico que reciben los jóvenes en el grupo de investigación, sino de las mismas dinámicas de los grupos que les permite participar no solo del proyecto de investigación que presentaron para la beca-pasantía, sino de otros proyectos del grupo, permitiéndoles ser coautores de otras publicaciones. En ese sentido, el programa es un gana-gana tanto para JII como para los grupos de investigación, los cuales pueden fortalecer su capacidad de producción.

También se encontraron impactos positivos en variables de mercado laboral, demostrados en mayores salarios (medidos por el IBC) y en mayor tiempo cotizado por año, por lo que el programa facilita la inserción de los jóvenes en el mercado de trabajo y permite disminuir las barreras que estos enfrentan para obtener un primer trabajo después de graduarse. Estos ingresos adicionales son aún más importantes en los casos de jóvenes de más bajos recursos, de zonas más rurales o de universidades públicas, ya que contribuyen con su sostenimiento económico y, en muchos casos, con la financiación de estudios de posgrado. La participación en el programa también se relaciona con la vinculación posterior en grupos de investigación, ya que la mayoría de los jóvenes permanece en un grupo de investigación dos años después de haber terminado la pasantía. Muchos de los JII hacen una carrera académica y terminan como docentes-investigadores.

Los beneficios del programa superan los costos en que se incurre para proveerlo. La tasa interna de retorno del programa varía entre 1,5% y 11% dependiendo de los supuestos empleados en el cálculo del costo-beneficio. Los beneficios del programa no se limitan a sus participantes, sino que también llegan a los grupos de investigación y a la sociedad en general mediante la mayor producción de conocimiento. Como lo expresaron en las entrevistas con actores involucrados, el programa representa una experiencia que le cambia la vida a los jóvenes y que permite acceder a mejores oportunidades en el ámbito investigativo.

> *"es como una experiencia que les cambia la vida, […] Es un asunto muy como un parteaguas para algunos, no solamente porque tienen la posibilidad que no tienen muchas personas en el país de poder desarrollar un trabajo de investigación que les interese -contar con la financiación que eso es muy escaso en el país-, sino que además eso luego les posibilita para su hoja de vida, evidenciamos como eso les abre y da opciones para salir del país, para estudiar en otras universidades a nivel internacional"*
>
> *(entrevista a directora de grupo de investigación, 2023)*

Bibliografía

Colciencias. (2008). *Colombia construye y siembra futuro. Política Nacional de Fomento a la Investigación y a la Innovación.* Bogotá.

Consejo Nacional de Política Económica y Social . (2021). *Política Nacional de Ciencia, Tecnología e Innovación 2022-2031.* Bogotá.

Consejo Nacional de Política Económica y Social. (2009). *Documento CONPES 3582: Política Nacional de Ciencia, Tecnología e Innovación.* Bogotá.

Almeida, R., & Carneiro, P. (2004). *On-the-Job trining: Estimating Costs and Returns using Firm Level Data.*

Almeida, R., & Carneiro, P. (2005). *The internal rate of return to on-the-job training.*

Blomberg, R. (1989). Cost-Benefit Analyis of employee training: a literature review. *Adult Education Qarterly, 39*(2), 89-98.

Callaway, B., & Sant'Anna, P. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics, 225*(2), 200-230.

DANE. (2022). *Educación Formal (EDUC) 2022 - Boletín técnico.*

Fundación para la Educación Superior y el Desarrollo. (2017). *Evaluación de impacto del programa Ondas de Colciencias.* Obtenido de https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/3615/Repor_Enero_2017_Nu%c3%b1ez_INF_FIN.pdf?sequence=3&isAllowed=y

ICFES. (2018). *Resultados Nacionales Saber 3, 5, y 9 2012-2017.*

ICFES. (n.d.). *Documentación de la prueba Saber 3, 5, 9.*

ICFES. (n.d.). *Documentación del examen Saber 11.*

Imbens, G., & Rubin, D. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences. .* Cambridge: Cambridge University Press.

Lopez, H. (2008). The social discount rate: estimates for nine Latin American countries. *Policy Research Working Paper*.

Minciencias. (2020). *Lineamientos para fomentar la vocación científica en niños, niñas y adolescentes.*

Núñez, J., Castro, F., Estupiñán, F., Gordillo, A., Medina, C., Castañeda, C., . . . Zárate, T. (2014). *Evaluación de impacto y análisis costo-beneficio de los programas de formación de capital intelectual de Colciencias: Jóvenes Investigadores y Becas de Doctorados.* Bogotá: Fedesarrollo.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Vicepresidencia de la República, Ministerio de Ciencia, Tecnología e Innovación. (2020). *Colombia hacia las fronteras del conocimiento : síntesis de las recomendaciones (Colección Misión Internacional de Sabios, 2019).* Bogotá.

Vicepresidencia de la República, Ministerio de Ciencia, Tecnología e Innovación. (2020). *Colombia: hacia una socied el conocimiento (reflexiones y propuestas). Volumen I. (Colección Misión de Sabios 2019).* Bogotá.*ad d*

Anexos – Programa Ondas

Anexo 1. Fuentes de información y limitaciones de la información para el programa Ondas

A1.1. Base de datos nacional

Se compiló una base de datos de más de 370 mil NNA vinculados al programa Ondas entre 2001 y 2023. La Tabla 24 muestra el número de NNA vinculados al programa para cada año, donde se destacan años de alta participación como 2022 (cerca de 49 mil) y 2012 (con aproximadamente 56 mil).

Tabla 24. Base de datos compilada de NNA vinculados a Ondas

Año	Número	Porcentaje
Sin información	1.708	0,46
2001-2005	288	0,08
2002	493	0,13
2003	785	0,21
2004	1.236	0,33
2004-2005	400	0,11
2005	1.195	0,32
2005-2006	27	0,01
2006	729	0,2
2007	1.313	0,35
2008	23	0,01
2009	20.856	5,62
2010	22.115	5,96
2011	43.192	11,65
2012	46.507	12,54
2013	52.717	14,21
2014	5.721	1,54
2015	24.827	6,69
2016	14.810	3,99
2017	91	0,02
2018	4.923	1,33
2019	13.487	3,64
2020	35.735	9,64
2021	25.871	6,98
2022	49.194	13,26
2023	2.643	0,71
Total	**370.886**	**100**

Fuente: Elaboración propia con base en las bases de datos de NNA vinculados a Ondas.

Esta base de datos compilada cuenta con variables de identificación como el tipo, número de documento y nombres; así como fecha de nacimiento, edad, grado escolar, género, y datos de la institución educativa (código y nombre). Sin embargo, la calidad y completitud de la información es variable, lo que puede afectar la calidad de la evaluación. Especialmente en los registros previos a 2017, se observa una falta de estructuración en los datos proporcionados por los departamentos, con inconsistencias en el orden y la agrupación de los nombres y apellidos, complicando el proceso de identificación y seguimiento de los NNA vinculados a Ondas. Esta heterogeneidad en el registro de datos puede llevar a dificultades en los cruces de información.

En la compilación de la base de datos de los NNA vinculados al programa se identificaron varios vacíos de información críticos, lo que representa un desafío significativo para la integridad y el uso de los datos (ver Tabla 25). Por ejemplo, 1.708 registros no cuentan con la información del año en que estuvieron vinculados al programa y otros 715 solo ofrecen un rango de años. Además, más del 63% de los registros no contienen información completa sobre el tipo o número de documento de identidad. La edad es otro dato faltante en el 31% de los casos, y un 1% no tiene registrado el apellido. Estos vacíos se suman a la preocupación de que muchos documentos parecen ser inválidos, y un porcentaje menor de niños tiene edades registradas inusualmente altas, superiores a 20 años.

Ahora bien, un desafío encontrado es que el 1,6% de los niños no cuenta con información del nombre del establecimiento educativo y un porcentaje significativamente mayor no cuenta con el reporte del código de establecimiento. Específicamente, el 92% de los niños no tiene registro de código de establecimiento, lo que, como se verá más adelante, representa un reto importante para la agregación de la información a nivel de establecimiento educativo y para su posterior búsqueda en bases de datos de colegios; por ejemplo, varios niños pueden tener un colegio con el mismo nombre, pero algunas observaciones sí registran el código y otras no.

Tabla 25. Datos sobre la completitud de la información compilada de Ondas

Año	Sin tipo de documento		Sin número de documento		Sin primer nombre		Sin primer apellido		Sin edad		Sin nombre de establecimiento ed.		Sin información del código de colegio	
	Frec.	%	Frec.	%	Frec.	%	Frec.	%	Frec.	%	Frec.	%	Frec.	%
2001-2005	288	100%	288	100%	2	1%	2	1%	273	95%	84	29%	288	100%
2002	493	100%	493	100%	-	0%	1	0%	493	100%	-	0%	493	100%
2003	785	100%	785	100%	-	0%	-	0%	406	52%	26	3%	785	100%
2004	1.236	100%	1.232	100%	-	0%	7	1%	908	73%	14	1%	1.236	100%
2004-2005	400	100%	400	100%	-	0%	-	0%	201	50%	-	0%	400	100%
2005	1.195	100%	1.195	100%	-	0%	2	0%	279	23%	4	0%	1.195	100%
2005-2006	27	100%	27	100%	-	0%	-	0%	24	89%	-	0%	27	100%
2006	729	100%	729	100%	1	0%	2	0%	225	31%	8	1%	729	100%
2007	1.313	100%	1.313	100%	-	0%	-	0%	1.313	100%	36	3%	1.313	100%
2008	20	87%	15	65%	-	0%	-	0%	23	100%	1	4%	23	100%
2009	19.170	92%	18.538	89%	5	0%	36	0%	5.485	26%	335	2%	20.856	100%
2010	21.985	99%	21.769	98%	1	0%	19	0%	5.498	25%	1.454	7%	22.115	100%
2011	43.192	100%	43.192	100%	2	0%	50	0%	23.500	54%	589	1%	43.192	100%
2012	45.724	98%	44.837	96%	-	0%	33	0%	23.107	50%	452	1%	46.507	100%
2013	52.717	100%	52.717	100%	-	0%	18	0%	22.743	43%	2.736	5%	52.717	100%
2014	5.721	100%	5.721	100%	-	0%	-	0%	889	16%	-	0%	5.721	100%
2015	24.827	100%	24.827	100%	-	0%	19	0%	3.646	15%	-	0%	24.827	100%
2016	14.720	99%	13.213	89%	-	0%	16	0%	1.316	9%	-	0%	14.810	100%
2017	91	100%	-	0%	-	0%	-	0%	-	0%	-	0%	91	100%
2018	134	3%	-	0%	-	0%	1	0%	4.320	88%	-	0%	4.828	98%
2019	581	4%	-	0%	-	0%	29	0%	8.284	61%	15	0%	10.345	77%
2020	63	0%	-	0%	-	0%	5	0%	1.353	4%	-	0%	24.790	69%
2021	733	3%	-	0%	-	0%	7	0%	702	3%	-	0%	17.187	66%
2022	24	0%	1	0%	-	0%	2.316	5%	7.783	16%	-	0%	43.654	89%
2023	2	0%	-	0%	-	0%	-	0%	26	1%	-	0%	2.642	100%
Total	**236.170**	**64%**	**231.292**	**63%**	**11**	**0%**	**2.563**	**1%**	**112.797**	**31%**	**5.754**	**2%**	**340.771**	**92%**

Tabla 26. Distribución de los NNA vinculados a Ondas por departamento y año

Departamento	2001 - 2005	2002	2003	2004	2004 - 2005	2005	2005 - 2006	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Amazonas	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	0,0	1,4	2,3	1,4	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,0	0,0	0,0
Antioquia	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	16,6	0,0	0,0	3,0	2,6	0,0	0,0	0,0	100	2,6	13,2	0,1	2,6	0,1	0,0
Arauca	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	9,0	2,1	0,0	3,3	15,2	0,0	0,0	0,0	0,0	0,0	0,0	12,9	21,1	9,4	0,0
Atlántico	0,0	94,1	37,6	34,2	0,0	2,9	0,0	0,0	0,0	0,0	4,5	0,1	1,2	7,5	0,0	0,0	0,0	0,0	0,0	7,5	5,8	0,4	0,1	3,7	9,1
Bogotá	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	5,5	8,6	2,1	0,0	0,0	0,0	0,0	0,0	4,7	0,4	0,4	0,0	0,2	0,0
Bolívar	0,0	0,0	52,6	0,0	97,8	0,0	100	0,0	0,0	0,0	5,3	38,2	34,3	15,5	53,5	0,0	6,5	0,0	0,0	50,5	9,6	2,2	0,0	3,8	0,0
Boyacá	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,5	6,0	1,4	0,0	0,0	0,0	0,0	0,0	0,0	0,9	0,0	0,0	0,0	2,2	0,0
Caldas	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,7	1,6	0,0	0,0	0,0	0,0	0,0	20,6	9,8	19,5	22,0	1,6	89,1
Caquetá	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	6,3	0,0	8,7	0,0	0,0	3,8	18,1	29,3	0,0	0,0	0,1	35,0	5,8	0,1	0,0
Casanare	0,0	0,0	0,0	15,5	0,0	0,0	0,0	0,0	0,0	0,0	8,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Cauca	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,4	6,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,4	1,8
Cesar	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	8,8	11,2	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	4,9	4,3	0,0	0,0	0,0
Choco	0,0	0,0	0,0	0,0	0,0	75,1	0,0	0,0	0,0	100	0,2	0,0	0,0	0,0	0,2	0,0	0,0	2,8	0,0	0,0	0,0	0,0	0,0	8,8	0,0
Córdoba	15,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	4,2	0,0	0,0	5,7	7,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Cundinamarca	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	7,8	0,0	21,6	0,0	0,0	0,0	8,5	0,0	5,5	3,4	0,9	0,0
Guainía	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,2	0,0	0,0	2,8	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0
Guaviare	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,9	1,6	0,0	0,0	0,0	0,0	0,0	0,5	1,2	2,8	1,3	1,0	0,0
Huila	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,0	2,9	0,0	0,0	34,6	0,0	0,0	1,1	16,1	8,3	18,9	0,4	0,0
La guajira	0,0	5,9	9,8	5,7	0,0	8,8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,5	0,0
Magdalena	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,6	2,5	1,2	1,4	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Meta	0,0	0,0	0,0	31,9	0,0	0,0	0,0	0,0	0,0	0,0	7,6	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	2,2	0,0
Nariño	71,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	100	0,0	4,8	3,0	2,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	1,7	0,0
Norte Santander	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,8	1,4	1,5	0,7	6,8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Putumayo	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,8	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,3	0,0
Quindío	13,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	4,7	2,8	0,7	1,6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,2	0,0
Risaralda	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0	0,0	3,3	0,0	0,0	8,1	0,0	0,3	0,0	5,5	0,0	0,6	0,0
San Andrés	0,0	0,0	0,0	12,6	0,0	0,0	0,0	0,0	0,0	0,0	1,4	1,6	1,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,6	0,0
Santander	0,0	0,0	0,0	0,0	0,0	13,3	0,0	100	0,0	0,0	3,0	14,9	0,0	18,3	0,0	0,0	19,7	31,0	0,0	1,0	6,2	0,5	0,0	5,5	0,0
Sucre	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,2	0,0	33,3	17,2	6,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	24,6	43,7	0,0
Tolima	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,9	4,8	0,0	0,0	1,9	0,0	0,0	0,0	0,0	0,0	0,0	2,4	0,2	0,1	0,0
Valle del cauca	0,0	0,0	0,0	0,0	2,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	74,6	21,2	28,8	0,0	1,8	32,3	0,0	0,0	3,1	0,0
Vaupés	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Vichada	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	1,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

A1.4. Cruces de información a nivel de NNA

La información de la base de datos de Ondas se cruzó con otros sistemas de información para poder adelantar la evaluación. En particular, se realizó un cruce con el SIMAT para poder tener información sobre matrícula de los NNA vinculados a Ondas y de sus compañeros de clase que servirían como grupo de control, así como con SNIES para tener información de educación superior de los NNA vinculados a Ondas.

Sistema de Matrículas Estudiantil (SIMAT). En el marco de la evaluación del programa Ondas adelantada por Fedesarrollo (2017) se realizó un primer cruce de la base de datos de NNA vinculados a Ondas entre 2001 y 2013 con el SIMAT para tener la matrícula anual de los NNA vinculados a Ondas y para recabar la información de sus compañeros de clase, de modo tal que se pudiera tener un grupo de NNA que no se vincularon al programa, pero que presumiblemente tuvieran las mismas características de los que sí estuvieron en él. Esto permitió tener la información de matrícula de dichos niños entre 2007 y 2015. En el marco de esta evaluación, el Ministerio de Educación Nacional realizó un nuevo cruce de la información con el cual se pudo actualizar la información de matrícula de dichos NNA entre 2018 y 2022.[38] Luego de depuraciones, se cuenta con información de matrícula para 51.014 NNA vinculados a Ondas y para 238.297 compañeros de clase (grupo de control). Con esta información se construyeron variables de deserción, repitencia y rezago escolar.

Sistema Nacional de Información de la Educación Superior (SNIES). Sistema Nacional de Información de la Educación Superior (SNIES). El Ministerio de Educación realizó el cruce de la información de los NNA vinculados a Ondas entre 2001 y 2013 y sus compañeros de clase con los graduados de educación superior entre 2020 y 2022. Se encontraron 927 NNA vinculados a Ondas graduados en este período y 4.413 NNA del grupo de control.

La Figura 28 resume los cruces de información que se realizaron con otros sistemas de información del Ministerio de Educación, así como la muestra final de evaluación empleando variables construidas a partir del SIMAT y el SNIES. Nótese que para la caracterización se podían emplear registros de los más de 370 mil estudiantes siempre que contaran con la información de edad, grado, género, entre otras variables. Además, la Tabla 27 presenta las características de la muestra de estudio, combinando NNA tratados con controles.

[38] Para los beneficiarios entre 2014 y 2023 no se pudo obtener la información para los compañeros de clase de los NNA que se vincularon a Ondas entre esos años.

Figura 28. Cruces de información realizados con la información de NNA vinculados a Ondas

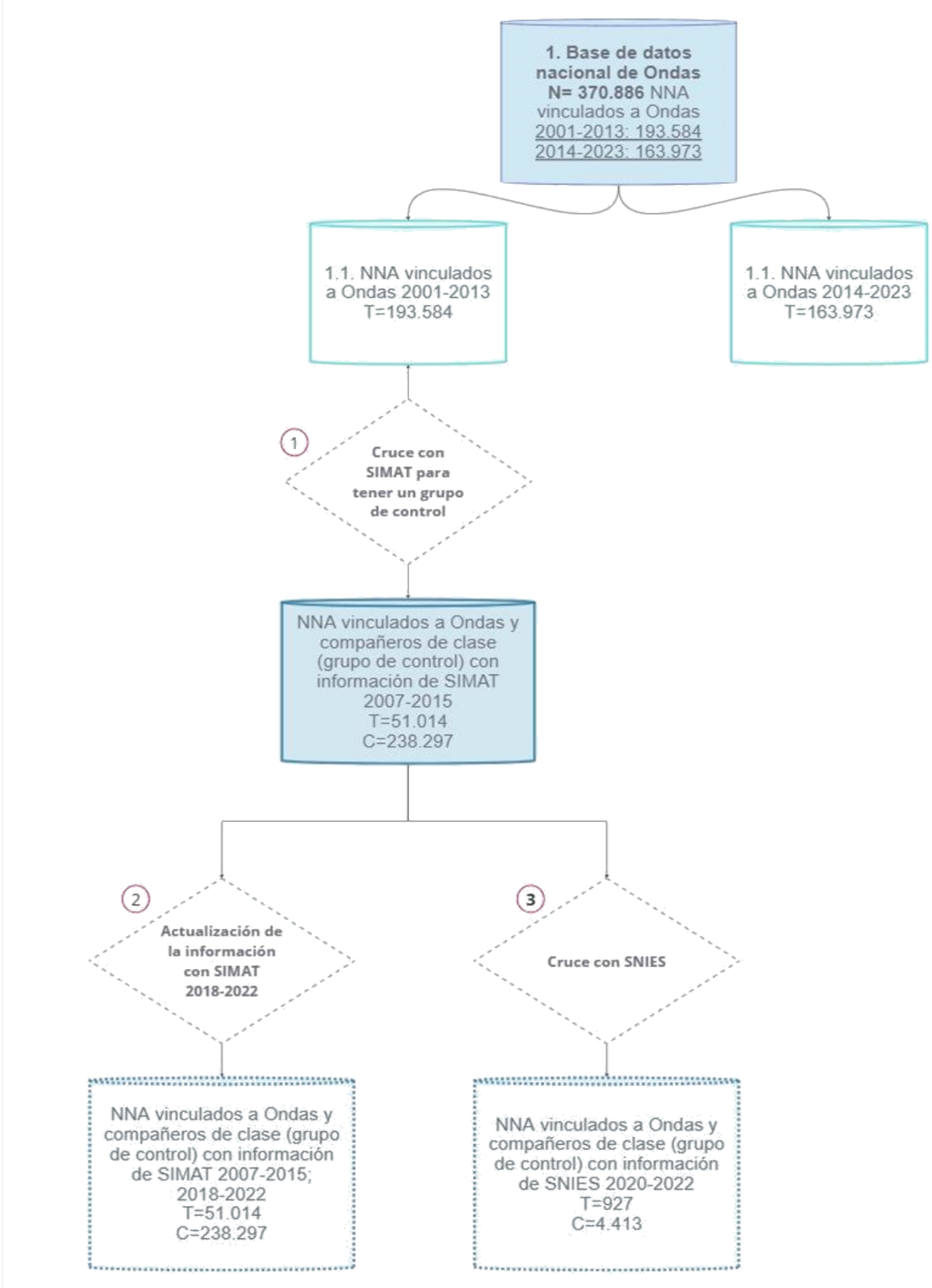


Tabla 27. Muestra de NNA para la evaluación de impacto

Variable	N	Promedio/ Proporción	Desviación estándar	Mínimo	Máximo
Edad en 2016	289189	16,4	3,78	3	31
Variable =1 si el NNA es hombre	289311	0,47	0,50	0	1
Estrato	288743	1,50	0,85	0	6
NNA tiene capacidades excepcionales	289311	0,00096	0,031	0	1
Variable=1 si el estudiante presenta algún tipo de discapacidad	289311	0,014	0,12	0	1
victima	289308	0,049	0,22	0	1
NNA reside en zona urbana	289244	0,76	0,43	0	1
NNA no pertenece a un grupo étnico	289308	0,94	0,24	0	1
NNA indígena	289308	0,039	0,19	0	1
NNA es NARP	289308	0,022	0,15	0	1
Grado	289311	6,68	3,09	0	13
Ciclo adultos	289310	0,024	0,15	0	1
Variable =1 si el estudiante está en un colegio que ofrece metodología de educación tradicional	288051	0,91	0,28	0	1
Variable =1 si el estudiante está en un colegio que ofrece programa para jóvenes extra edad	288051	0,020	0,14	0	1
Variable =1 si el estudiante está en un colegio que ofrece metodología escuela nueva	288051	0,027	0,16	0	1
Variable =1 si el estudiante está en un colegio que ofrece metodología postprimaria	288051	0,014	0,12	0	1
Variable =1 si el estudiante está en un colegio que ofrece metodología de educación	288051	0,015	0,12	0	1
NNA estudia en jornada completa o única	289311	0,13	0,33	0	1
NNA estudia en jornada de la mañana	289311	0,61	0,49	0	1
NNA estudia en jornada de la tarde	289311	0,24	0,43	0	1
NNA estudia en jornada nocturna o fin de semana	289311	0,027	0,16	0	1
Antioquia	289311	0,072	0,26	0	1
Atlántico	289311	0,039	0,19	0	1
Bogotá	289311	0,067	0,25	0	1
Bolívar	289311	0,23	0,42	0	1
Boyacá	289311	0,018	0,13	0	1
Caldas	289311	0,0044	0,066	0	1
Caquetá	289311	0,018	0,13	0	1
Cauca	289311	0,013	0,12	0	1
Cesar	289311	0,026	0,16	0	1
Córdoba	289311	0,058	0,23	0	1
Cundinamarca	289311	0,026	0,16	0	1
Chocó	289311	0,0039	0,063	0	1
Huila	289311	0,020	0,14	0	1
La Guajira	289311	0,0033	0,058	0	1
Magdalena	289311	0,026	0,16	0	1
Meta	289311	0,021	0,14	0	1
Nariño	289311	0,019	0,14	0	1
Norte de Santander	289311	0,042	0,20	0	1
Quindío	289311	0,011	0,11	0	1
Risaralda	289311	0,019	0,14	0	1
Santander	289311	0,098	0,30	0	1
Sucre	289311	0,035	0,18	0	1
Tolima	289311	0,012	0,11	0	1
Valle del Cauca	289311	0,013	0,11	0	1
Arauca	289311	0,049	0,22	0	1
Casanare	289311	0,011	0,10	0	1
Putumayo	289311	0,011	0,10	0	1
San Andrés	289311	0,0027	0,052	0	1
Amazonas	289311	0,016	0,13	0	1
Guainía	289311	0,0083	0,091	0	1
Guaviare	289311	0,0068	0,082	0	1
Vaupés	289311	0,0038	0,062	0	1
Vichada	289311	0,0029	0,054	0	1
Beneficiario de Ondas en 2006	289006	0,0040	0,063	0	1

Beneficiario de Ondas en 2009	289006	0,083	0,28	0	1
Beneficiario de Ondas en 2010	289006	0,14	0,34	0	1
Beneficiario de Ondas en 2011	289006	0,15	0,35	0	1
Beneficiario de Ondas en 2012	289006	0,25	0,43	0	1
Beneficiario de Ondas en 2013	289006	0,38	0,49	0	1

A1.3. Datos a nivel de colegios

Fuentes de información

Las fuentes de información empleadas para poder definir el grupo de tratamiento y control son:

Base de datos de Ondas. Se agregó la información de los NNA vinculados a Ondas entre 2001 y 2023 a nivel del establecimiento educativo reportado. Para tener esta información es fundamental tener el nombre del establecimiento, el código DANE del establecimiento, el departamento y el municipio para poder agregar el número de estudiantes a este nivel.

Sistema de Información Nacional de Educación Básica y Media – SINEB. De esta base de datos se emplearon las variables de código DANE del establecimiento, nombre del establecimiento, departamento, municipio; así como características de las instituciones como jornadas, carácter, especialidad, modelos educativos.

Base de datos de las pruebas Saber 3, 5, 9 a nivel de establecimiento para 2009 y 2012-2017. De esta base de datos se emplearon las variables de código DANE del establecimiento, nombre del establecimiento, y porcentaje de estudiantes con puntaje satisfactorio o avanzado en cada una de las áreas y niveles disponibles (lenguaje, matemáticas, ciencias naturales; y tercero, quinto y noveno).

Base de datos de las pruebas Saber 11 a nivel de establecimiento para 2014-2021. De esta base de datos se emplearon variables como el código DANE del establecimiento, características como el departamento, municipio, naturaleza, calendario, y el índice del resultado del colegio en cada una de las áreas.

Cruces de información

La Figura 29 ilustra todo el proceso, de nueve etapas, para cruzar la información de ONDAS y para poder definir los grupos de tratamiento y control para los resultados evaluados a nivel de colegio. En la primera etapa se hizo un cruce de la base de datos de Ondas de establecimientos educativos con SINEB, puesto que los establecimientos educativos de Ondas no tenían un código de establecimiento asociado y se pueden presentar muchos homónimos (establecimientos con el mismo nombre pero que no corresponden a la misma institución). Esto permitió identificar para cada establecimiento educativo de la base de ONDAS, su código de establecimiento y algunas características adicionales como carácter (oficial o privado), zona, especialidad, tipo de matrícula, entre otras.

El cruce de la base de datos de ONDAS con SINEB se realizó siguiendo estos pasos:

1. Se hizo un primer cruce empleando los códigos de los establecimientos disponibles. De ahí se reserva una base de datos con los que efectivamente cruzan y otra con los que no, la cual se emplea para un siguiente cruce en el paso dos.
2. Se hizo un cruce exacto por nombre completo del establecimiento y municipio para los que no cruzaron con código. De ahí se reserva una base de datos con los que efectivamente cruzan y otra con los que no, la cual se emplea para un siguiente cruce en el paso tres.
3. Se hizo un cruce dentro para de cada municipio identificando nombres parecidos. Este paso se realizó varias veces intentando variaciones del nombre, por ejemplo, eliminando palabras como “COLEGIO” o “INSTITUCIÓN EDUCATIVA”. En cualquier caso, se calculaba la probabilidad de que el nombre fuera el mismo y solo se dejaron establecimientos con muy alta probabilidad de ser los mismos (de más del 95% en el caso más laxo).
4. Luego, se combinan las bases de datos que se cruzaron en los pasos anteriores y se valida que un mismo establecimiento no haya cruzado con dos códigos diferentes, en cuyo caso se

eliminaría porque se trata de un homónimo. Como resultado se identifican 3.160 establecimientos tratados que conforman la muestra de evaluación.

Retomando la Figura 29, luego de definir la muestra de evaluación se realizaron dos procesos paralelos de cruce de información (etapas 2 y 6). Por un lado, se hizo un cruce de los datos con los resultados agregados a nivel de establecimiento educativo de las pruebas Saber 3, 5 y 9, y por otro lado, con los resultados de las pruebas Saber 11. El cruce de la muestra de evaluación con la base de datos de Saber 3, 5, 9 (etapa 2) arrojó como resultado una submuestra de 3.064 establecimientos educativos con información de las pruebas. Luego, la información de la base de datos de Saber 3, 5, 9 se volvió a cruzar, pero esta vez con la base de datos de establecimientos educativos de Ondas que no tenían código de establecimiento; este cruce se hizo con nombre de establecimiento con el objetivo de descartar del grupo de control aquellos establecimientos educativos que pudieran tener un parecido con los tratados (etapa 3); como resultado, se obtuvo un grupo de control de 14.077 establecimientos educativos. El siguiente procedimiento (etapa 4), fue implementar un Propensity Score Matching entre tratados y controles para definir el soporte común; por último, se seleccionaron solo los establecimientos educativos oficiales para realizar las estimaciones de impacto (etapa 5).

Para las etapas 6 a la 9, ilustradas en la Figura 29, se siguió una lógica similar a la empleada con las pruebas Saber 3, 5 y 9, pero con la base de datos de las pruebas Saber 11. En la etapa 6, se realizó un cruce entre la muestra de evaluación y la base de datos Saber 11, resultando en una submuestra de 1.112 establecimientos educativos tratados con puntajes de Saber. Posteriormente, se efectuó un nuevo cruce de los datos de Saber 11 con la base de datos de Ondas, específicamente con aquellos establecimientos sin código de establecimiento, empleando el nombre de la institución como filtro para excluir del grupo de control potencial cualquier establecimiento educativo con nombre parecido con los tratados (etapa 7). Esto dio como resultado un grupo de control preliminar compuesto por 6.593 establecimientos educativos. El proceso continuó con la implementación de un método de emparejamiento para dejar solo los establecimientos dentro del soporte común (etapa 8); para este emparejamiento se utilizaron variables de tipo de establecimiento, zona urbana, carácter (técnico, académico), si están en un resguardo, si tienen matricula contratada, si es calendario A, departamento. Finalmente, se seleccionaron solo los establecimientos oficiales para la evaluación.

Figura 29. Cruces de datos y procesos para construir el grupo de control con información secundaria

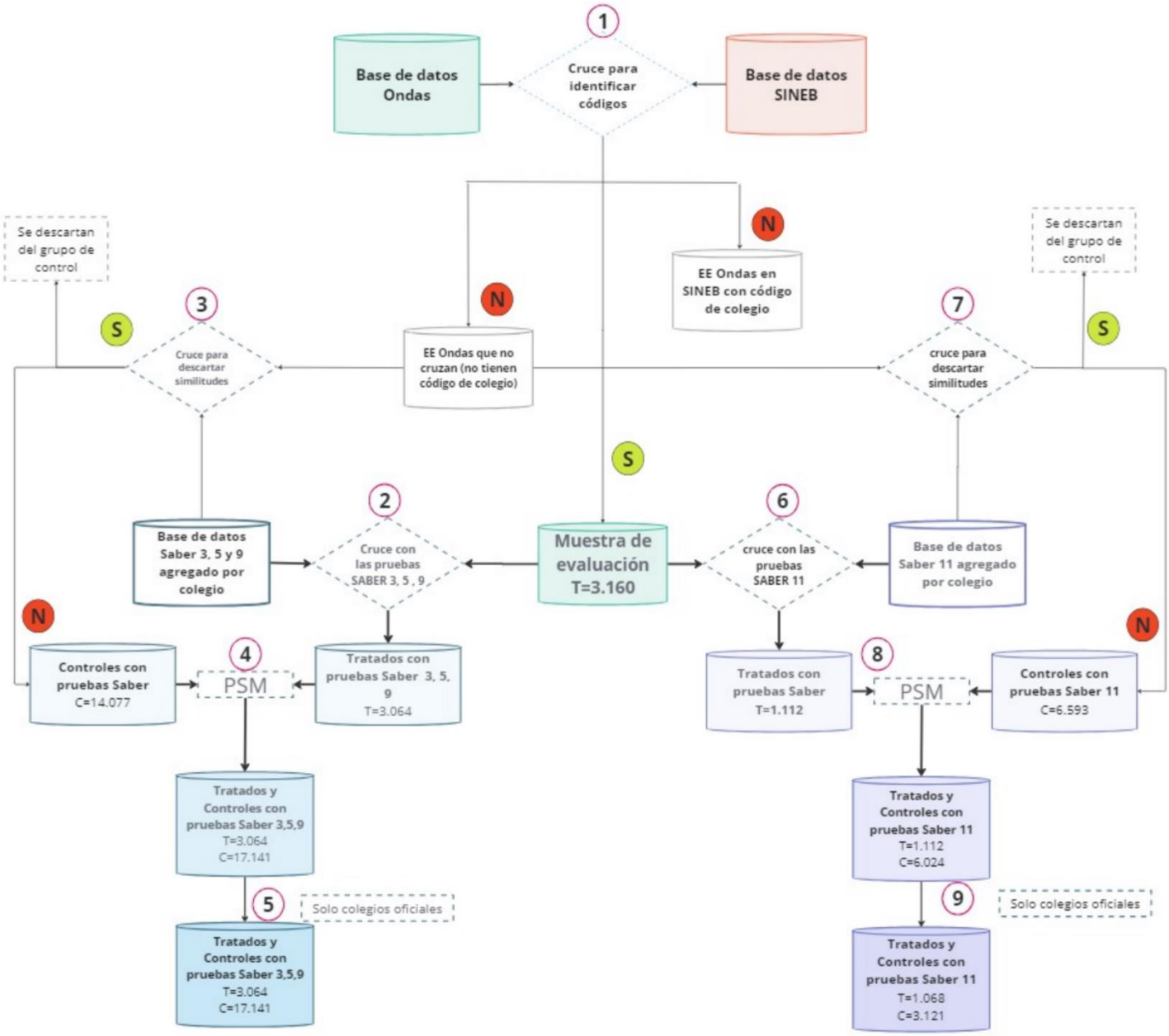
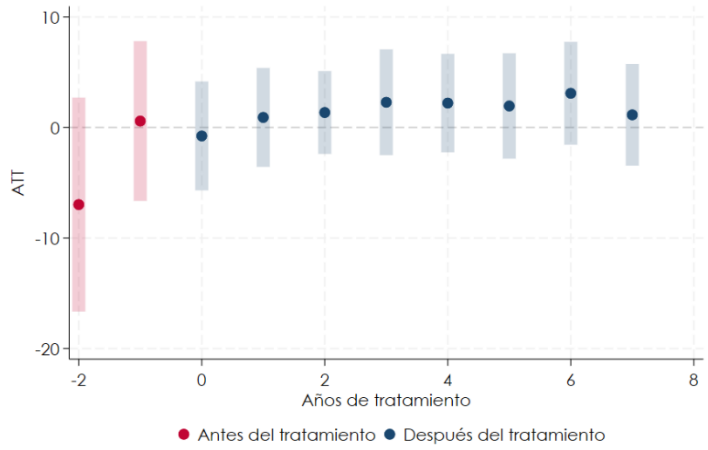
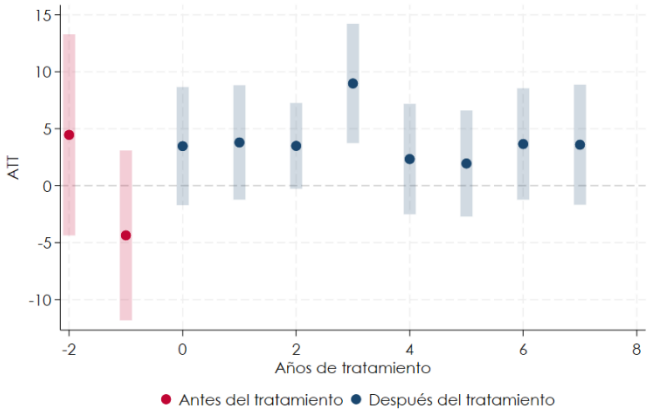


Figura 30. Impacto en pruebas Saber 5 y 9 agregadas a nivel de establecimiento educativo

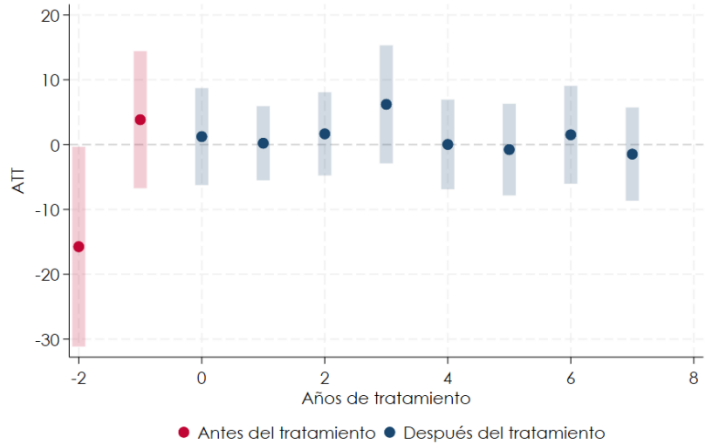
Lenguaje Grado 5º



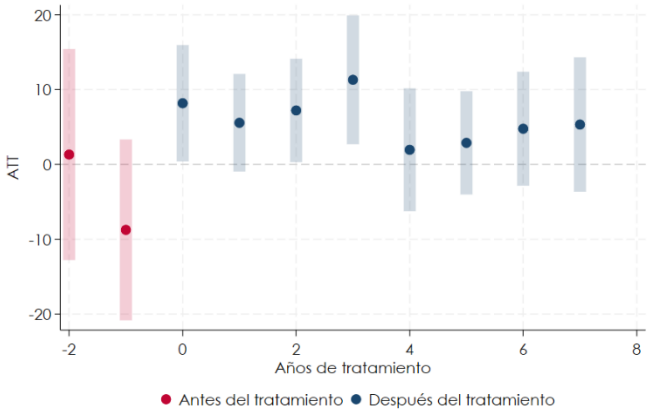
Matemáticas Grado 5º



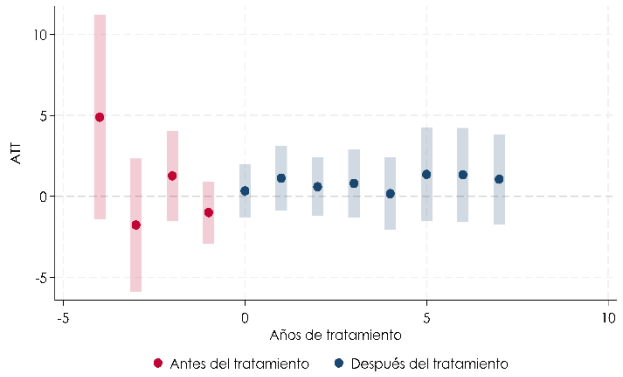
Lenguaje Grado 5º - Solo establecimientos educativos con NNA vinculados a Ondas en grados 4º y 5º



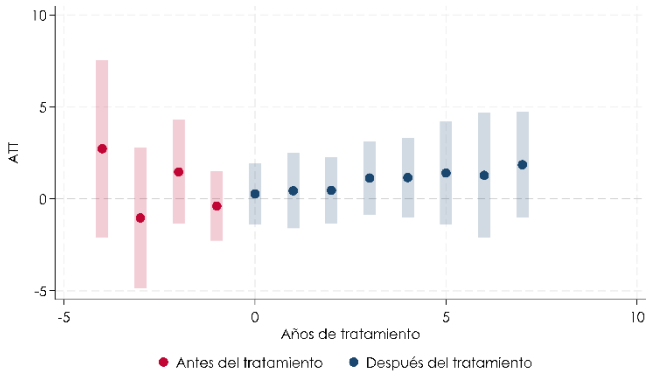
Matemáticas Grado 5º - Solo establecimientos educativos con NNA vinculados a Ondas en grados 4º y 5º



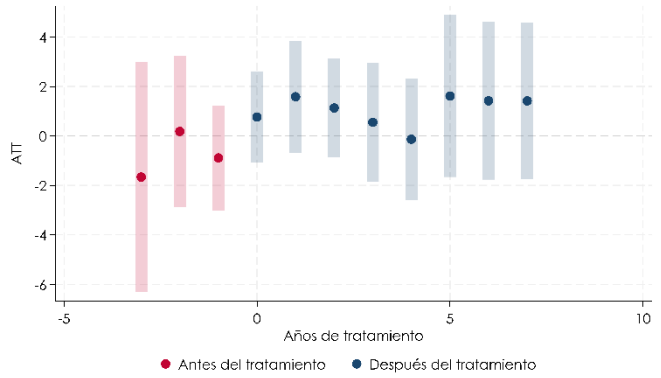
Lenguaje Grado 9º



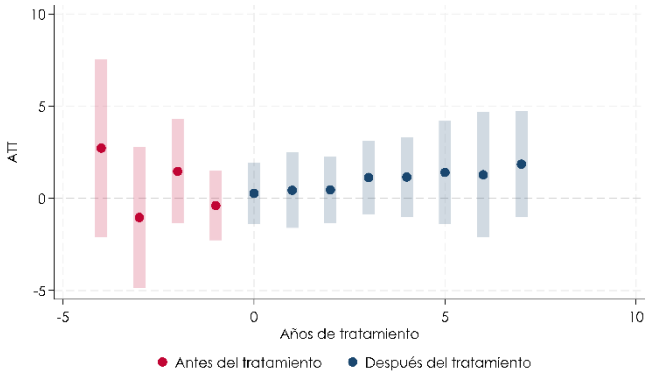
Matemáticas Grado 9º



Lenguaje Grado 9º - Solo establecimientos educativos con NNA vinculados a Ondas en grados entre 6º y 9º

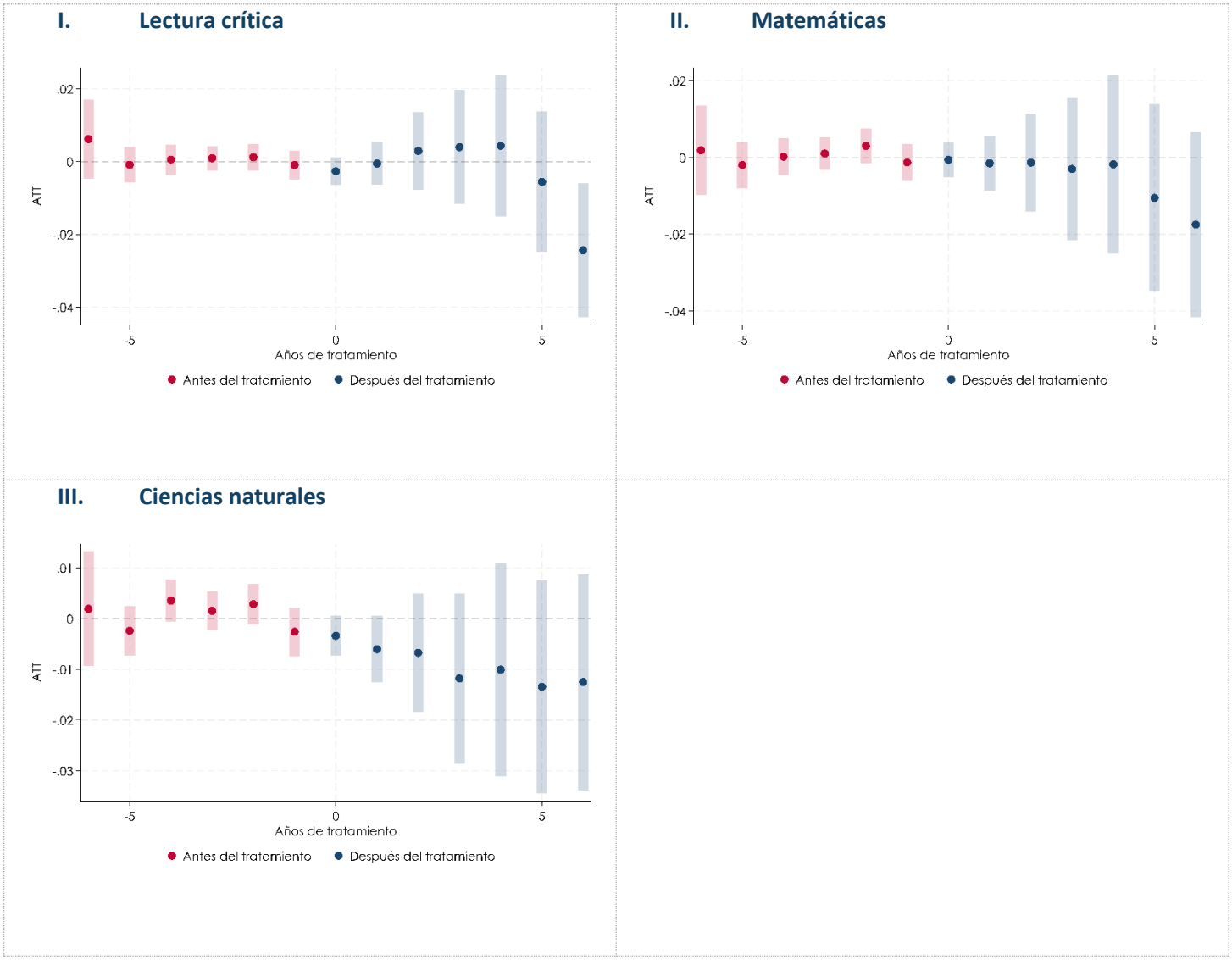


Matemáticas Grado 9º - Solo establecimientos educativos con NNA vinculados a Ondas en grados entre 6º y 9º



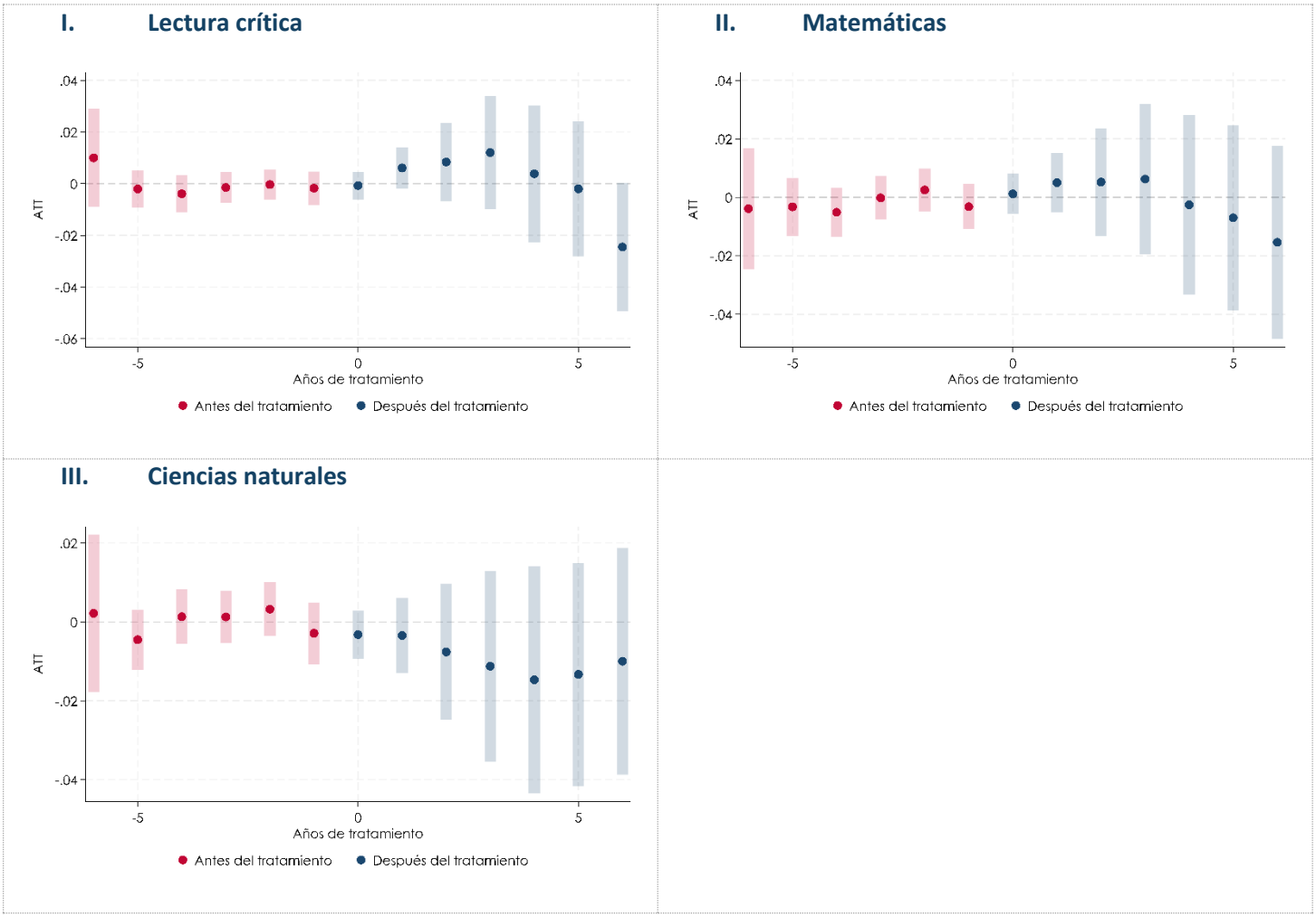
Fuente: Elaboración propia con información de las bases de datos de Ondas y Saber 3, 5,9. Se emplean controles de naturaleza del establecimiento educativo, zona, carácter, si es matrícula contratada, calendario A y departamento. Intervalos de confianza al 95%.

Figura 31. Impacto en pruebas 11 agregadas a nivel de establecimiento educativo



Fuente: Elaboración propia con información de las bases de datos de Ondas y Saber 11. Se emplean controles de naturaleza del establecimiento educativo, zona, carácter, si es matrícula contratada, calendario A y departamento. Intervalos de confianza al 95%.

Figura 32. Impacto en pruebas Saber 11 solo para establecimientos educativos con NNA vinculados a Ondas en grados 10 y 11



Fuente: Elaboración propia con información de las bases de datos de Ondas y Saber 11. Se emplean controles de naturaleza del establecimiento educativo, zona, carácter, si es matrícula contratada, calendario A y departamento. Intervalos de confianza al 95%.

Anexos – Programa Jóvenes Investigadores e Innovadores

Anexo 3. Compilación de bases de datos y limitaciones de la información

A3.1. Base de datos de JII y aspirantes

Se compiló la información para un total de 19.086 jóvenes que aplicaron a las convocatorias de JII entre 2004 y 2021. Sin embargo, esta base de datos cuenta con muchos datos faltantes que dificultaron los cruces de información que se realizaron. El grupo de jóvenes inscritos que no fueron seleccionados para participar en el programa solo cuenta con información del número de documento (ver *Tabla 28*). En tanto que los jóvenes que participaron en JII tienden a tener información más completa, por lo menos para el tipo de documento, departamento, género. Se destaca también que la información de grupo étnico, discapacidad, condición de víctima todavía no se completa bien en las convocatorias, si bien ha mejorado en los últimos dos años. Ahora bien, dado que esta base de datos se cruzó con CvLac empleando el número de documento, muchas de estas variables se pueden solventar o completar para los inscritos a JII que se encontraron en dicho sistema de información.

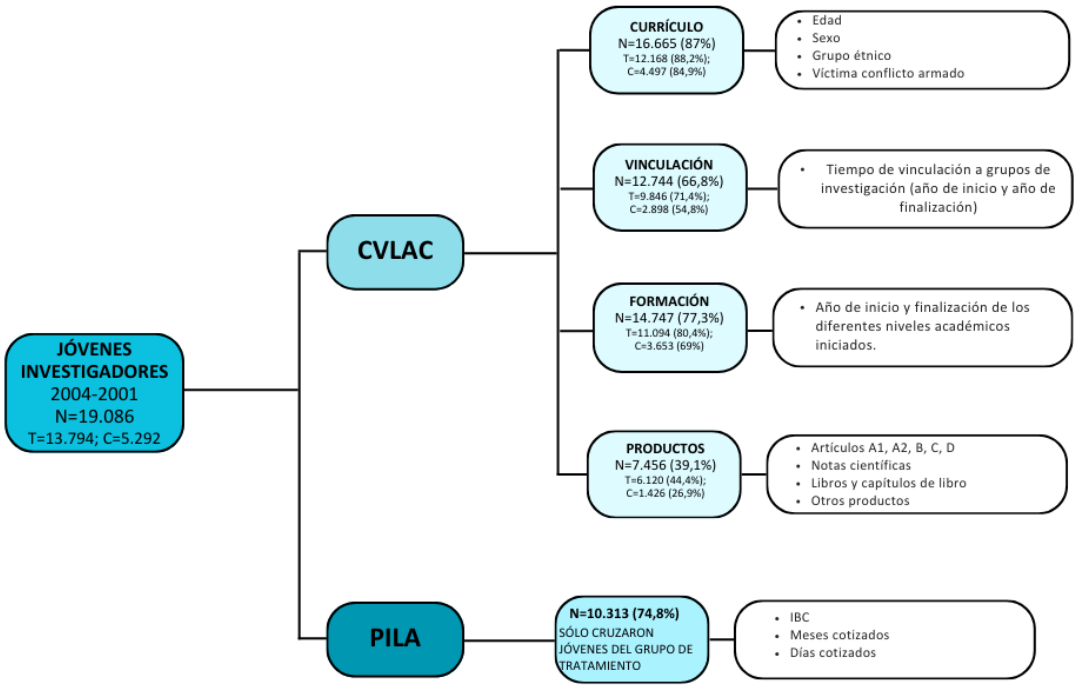
Tabla 28. Cifras sobre la completitud de la información

Variable	Inscritos a JII		Inscritos no seleccionados		Jóvenes que participaron en JII	
	Sum	Mean	Sum	Mean	Sum	Mean
Sin tipo de documento	8535	45%	5292	100%	3243	24%
Sin número de documento	0	0%	0	0%	0	0%
Sin nombre	8535	45%	5292	100%	3243	24%
Sin departamento	10254	54%	5292	100%	4962	36%
Sin municipio	10700	56%	5292	100%	5408	39%
Sin edad	12842	67%	5292	100%	7550	55%
Sin género	8362	44%	5126	97%	3236	23%
Sin información de grupo étnico	16810	88%	5292	100%	11518	84%
Sin información de discapacidad	16810	88%	5292	100%	11518	84%
Sin información de condición de víctima	16860	88%	5292	100%	11568	84%
Sin título	8967	47%	5292	100%	3675	27%

A3.2. Cruces de información con la base de datos de inscritos

La base de datos de inscritos a JII se cruzó con otros sistemas de información del Gobierno Nacional para poder construir las variables de resultado de la evaluación. En particular, se cruzó la información con el sistema CvLac de MinCiencias para obtener la producción académica de los jóvenes inscritos, así como con la Planilla Integrada de Liquidación de Aportes (PILA) del Ministerio de Salud para obtener las variables relacionadas con cotizaciones y estabilidad en el mercado laboral. En las siguientes secciones se presentan los resultados de estos cruces de información.

Figura 33. Cruces de información de la base de datos de inscritos a JII con CvLac y PILA



Fuente: Elaboración propia con base en los resultados del cruce entre la base de datos de inscritos, CvLac y PILA.

Cruce de información con CvLac

MinCiencias adelantó el cruce de la información de los inscritos a JII con la base de datos de CvLac. Tal como se muestra en la Figura 33 el formulario CvLac cuenta con la información de currículos de los investigadores, vinculación, formación, y productos. Los resultados del cruce de información indican que el 88% de los jóvenes que participaron en JII y el 85% de los inscritos no seleccionados contaban con información de currículo. Adicionalmente, del total de jóvenes el 66,8% tenía información de vinculación, el 77,3% tenía información de formación y el 39,1% tenía información de productos en CvLac.

En total se encontró al 87% de los inscritos a JII en la base de datos de CvLac, si bien este porcentaje varía entre las diferentes cohortes. Para las cohortes entre 2004 y 2009 el porcentaje de éxito en el cruce varía entre el 50% y el 70%, pero para años posteriores llega a ser casi del 100% (ver Anexo 3). Con esta información se conformó la base de datos para la caracterización de los jóvenes que participaron en JII y para la evaluación de impacto. Con la base de datos de currículo se obtuvo la información del lugar de nacimiento de los inscritos a JII, el sexo, grupo étnico, discapacidad, estado de víctima del conflicto armado. Luego, esta información se cruzó con la de vinculación, que permitió obtener el número de años en los que ha estado vinculado un joven a algún grupo de investigación. Esta base de datos de vinculación también permitió contar con la identificación del grupo de investigación, que luego se combinó con la base de datos pública de Grupos de investigación y que permitió contar con datos del área de conocimiento del o de los Grupos de investigación a los que está vinculado el joven inscrito a JII, así como el lugar de ubicación de los grupos.[39] Luego, la base de datos de formación detalla los estudios del joven, la fecha de inicio y la de fin.

[39] Se tomó la base de datos de Grupos de Investigación Reconocidos y Clasificados 2021 que cuenta con la información de 30.051 grupos de investigación, disponible en: https://minciencias.gov.co/ciudadano/datosabiertos

Tabla 29. Resultado del cruce entre la base de datos de inscritos a JII y CVLAC

Año de la convocatoria	Inscritos no seleccionados		Jóvenes que participaron en JII		Total	
	Número	% del total	Número	% del total	Número	% del total
2004	230	58%	18	47%	248	57%
2005	130	52%	22	69%	152	54%
2006	183	54%	154	76%	337	63%
2007	93	54%	243	82%	336	72%
2008	166	62%	220	81%	386	71%
2009	114	93%	1,084	98%	1,198	97%
2010	85	99%	1,352	98%	1,437	98%
2011	61	97%	1,722	99%	1,783	99%
2012	1,177	100%	1,026	100%	2,203	100%
2013	0		764	96%	764	96%
2014	0		808	95%	808	95%
2015	0		20	95%	20	95%
2016	0		392	65%	392	65%
2017	0		1,333	89%	1,333	89%
2018	165	99%	762	96%	927	97%
2019	1122	88%	569	88%	1691	88%
2020	562	99%	682	80%	1244	88%
2021	409	100%	997	61%	1406	69%
Total	**4,497**	**85%**	**12,168**	**88%**	**16,665**	**87%**

Cruce de información con la PILA

El resultado del cruce de la base de datos de inscritos a JII con la PILA arrojó un resultado de 10.313 jóvenes en ambas bases. Sin embargo, todos los jóvenes que cruzaron hacen parte de los que participaron en el programa y no hay ninguno de los que se inscribió pero que no quedó seleccionado. Con el sistema de información de la PILA se puede obtener la información del Ingreso Base de Cotización (IBC) de las personas que cotizan al Sistema General de Seguridad Social tanto a salud como a pensión y se puede calcular el número de meses que cotiza un joven en cada año.

El Ministerio de Salud solo cruzó la información de los inscritos a JII que contaban con la llave de información número y tipo de documento completo, pero solo el 56% de las observaciones contaba con este dato y todas correspondían a los Jóvenes que participaron. Para superar este percance, se envió una nueva base de datos al MinCiencias en noviembre de 2023 para tener un nuevo cruce. Esta nueva base completaba la información de tipo de documento de los jóvenes inscritos que se encontraron en la base de datos de CvLac, por lo que tenía el potencial de mejorar el cruce de información. De cualquier modo, se adelantaron las estimaciones para las variables construidas con la PILA empleando a los futuros participantes de la beca-pasantía, siguiendo la metodología descrita en el documento.

Anexo 4. Resultados adicionales del Programa JII

Tabla 30. Número de JII en un grupo de investigación de cada departamento, que nacieron en el departamento y tasa neta de atracción de JII

Departamento	JII en un grupo de investigación del departamento	JII que nacieron en el departamento	Tasa neta de atracción de JII
Amazonas	35	3	91
Cundinamarca	341	175	49
Casanare	26	14	46
Antioquia	2571	1621	37
Atlántico	882	560	37
Risaralda	633	409	35
Bogotá D.C.	3366	2256	33
Bolívar	726	506	30
Caldas	697	507	27
Santander	832	619	26
Valle del Cauca	1469	1094	26
Meta	126	95	25
Quindío	329	251	24
San Andrés	5	4	20
Cauca	241	221	8
Chocó	133	126	5
Magdalena	207	200	3
Boyacá	457	474	-4
Caquetá	65	68	-5
Tolima	287	311	-8
Huila	186	209	-12
Cesar	188	222	-18
La Guajira	47	60	-28
Córdoba	187	246	-32
Norte de Santander	239	317	-33
Guaviare	5	7	-40
Sucre	131	234	-79
Nariño	123	330	-168
Arauca	3	17	-467