

¿Quién Recibe la Ayuda Financiera del Programa *Ser Pilo Paga*? Un Análisis Multinivel

Juan Fernando Álvarez Correa¹

DIRECCIÓN
Juan Byron Correa²

TRABAJO DE GRADO PARA OPTAR AL TÍTULO DE
MAESTRÍA EN ECONOMÍA APLICADA
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS
UNIVERSIDAD DEL VALLE
CALI - COLOMBIA

Resumen

Existe evidencia empírica sobre los efectos positivos de la ayuda financiera estudiantil, en especial los relacionados con el acceso a la educación superior en las últimas tres décadas. Sin embargo, la literatura académica entorno a la distribución de este tipo de ayudas es escasa para países en desarrollo. Utilizando datos transversales de los estudiantes evaluados, en Colombia, en las pruebas SABER 11 entre 2014 y 2017, se estima un modelo logístico de interceptos aleatorios a fin analizar algunos factores determinantes de la propensión en favor de ser beneficiario de un programa de créditos-beca para el acceso a la educación superior que combina mecanismos de focalización basados en el nivel de pobreza (“necesidad”) y el puntaje en las pruebas de referencia (“mérito”). Este estudio encuentra que, el gobierno colombiano, fue más receptivo a la necesidad de los estudiantes elegibles al momento de otorgar la ayuda financiera.

Palabras clave: educación superior, ayuda financiera, desigualdad, Ser Pilo Paga.

Abstract

There is empirical evidence on the positive effects of student financial aid, especially those related to access to higher education in the last three decades. However, academic literature regarding the distribution of this type of aid is scarce for developing countries. Using cross-sectional data of the students evaluated, in Colombia, in the SABER 11 tests between 2014 and 2017, a logistic model of random intercepts is estimated in order to analyze some factors determining the propensity in favor of being a beneficiary of a credit program. scholarship for access to higher education that combines targeting mechanisms in the level of poverty (“need”) and score in the benchmark tests (“merit”). This study finds that the Colombian government was more receptive to the need of eligible students at the time of granting financial aid.

Keywords: postsecondary enrollment, financial aid, inequality, Ser Pilo Paga.

¹Estudiante Maestría en Economía Aplicada. E-mail: juan.alvarez.corea@correounivalle.edu.co

²Director trabajo de grado. Profesor Departamento de Economía, Universidad del Valle. E-mail: juan.correa@correounivalle.edu.co

1. Introducción

Tradicionalmente, el objetivo de la política de ayuda financiera para la educación ha sido romper el vínculo entre los antecedentes socioeconómicos familiares y las oportunidades universitarias (McPherson and Schapiro, 2006). En las últimas décadas, los formuladores de políticas de educación superior han estado profundamente interesados en identificar estrategias para mejorar el acceso a la educación superior para las poblaciones tradicionalmente subrepresentadas y de bajos ingresos (Toutkoushian and Shafiq, 2010). La expansión de estos programas, ha sido crucial para aumentar el acceso a la educación postsecundaria en las últimas tres décadas (Ziderman, 2003, Ferreyra, 2017); si bien, la ayuda financiera no ha eliminado por completo las restricciones relacionadas con los ingresos para los estudiantes que desean asistir a la universidad, ha permitido que muchos estudiantes, principalmente los pertenecientes a familias de bajo nivel socioeconómico (NSE), continúen su proceso de educación formal que de otra manera no lo hubieran logrado.

Esta escasa representación de estudiantes de bajo NSE en instituciones de educación superior (IES) también se hace visible en los países en desarrollo. Según el Banco Mundial, el costo prohibitivo de la educación terciaria es uno de los principales impulsores de las importantes desigualdades con respecto a su acceso en Colombia (OCDE, 2016). En dólares constantes de 2017, el costo promedio anual de la matrícula en una universidad pública en Colombia fue de casi \$900 dólares estadounidenses, mientras que para las universidades privadas es de aproximadamente \$7.300; esta última cifra supera el 40 % del ingreso nacional por adulto (Bernal and Penney, 2019). Por su parte, la asistencia a IES muestra un patrón claro por quintil de ingresos. La composición de los estudiantes es de aproximadamente el 8,6 % en el primer quintil, el 9,5 % en el segundo, pero más del 44 % para los del quinto (OCDE, 2016).

Respecto a IES con acreditación de calidad, la brecha es aún mayor. De acuerdo con Londono-Velez et al. (2017), en Colombia, para el año 2014, menos del 1,5 % de los estudiantes de bajo NSE se matricularon en instituciones educativas de calidad, inmediatamente después de graduarse de secundaria, en comparación con el 44,2 % de sus homólogos con alto NSE. Así mismo, cerca del 60 % de los puntajes más altos en la prueba SABER 11 de 2012 fueron obtenidos por estudiantes del estrato 1, 2 o 3, pero dos años después de haber presentado la prueba, cerca del 23 % de estos jóvenes no habían logrado acceder a una IES por falta de recursos económico (Álvarez et al., 2017).

Respecto a lo anterior, el gobierno ha desempeñado un papel importante en la provisión de fondos para hacer que la universidad sea más asequible para los estudiantes y, en el proceso, mejorar el acceso a la universidad para estudiantes de entornos poco representados y de bajos ingresos. Estos mecanismos de apoyo estatal pueden resumirse en: Las asignaciones directas a las universidades (instituciones públicas); y las asignaciones directas a los estudiantes en forma de “becas” basadas en la necesidad o en el mérito.

Teniendo en cuenta que la prioridad del gobierno nacional es brindar asistencia financiera para el acceso a la educación superior de calidad, en el año 2014 se creó el programa Ser Pilo Paga en el cual confluyen mecanismos de focalización de necesidad y mérito. Tomando en consideración la premisa anterior, esta investigación formula la siguiente pregunta: ¿Cuál es la prioridad del gobierno

colombiano al otorgar ayuda financiera para el acceso a la educación superior?. Para responder este interrogante, se analizó información oficial de los estudiantes elegibles para el programa Ser Pilo Paga entre 2014 y 2017, a fin de indagar algunos atributos (individuales, académicos e institucionales) que se correlacionan con la decisión del gobierno colombiano de otorgar recursos financieros para el ingreso a la educación superior, dando como resultado el siguiente objetivo de investigación: “*identificar y analizar los determinantes de la propensión en favor a ser beneficiario del programa Ser Pilo Paga entre 2014 y 2017*”.

A fin de lograr el objetivo definido anteriormente, el presente documento se compone de siete secciones incluida esta introducción. En la sección 2 se realiza una breve descripción del programa Ser Pilo Paga y el contexto del nivel socioeconómico en Colombia. La sección 3 presenta referencias literarias de estudios previos asociados al problema de investigación. En la sección 4 se expone el marco teórico bajo el cual se enmarca la investigación y la estrategia empírica empleada. La sección 5 presenta los datos y la estrategia de identificación para el umbral de necesidad. La sección 6 analiza los resultados econométricos. Finalmente, en la sección 7 se discuten los resultados y alcances de la investigación.

2. Antecedentes Institucionales

La educación superior cumple un papel estratégico en el proyecto de desarrollo económico, social y político en el que está comprometido el país, siendo la formación superior vital para alcanzar el éxito y la prosperidad de una nación (Bottia et al., 2012, Arango-Thomas et al., 2015). Pese al aumento sustancial en las tasas de matrícula universitaria en Colombia durante la última década, todavía hay grandes disparidades en el acceso por nivel socioeconómico (OCDE, 2016), debido principalmente a que la mayoría de las universidades son privadas y presentan altos costos de matrícula (OECD and Bank, 2013), así como la falta de un sistema de ayuda financiera bien desarrollado (De Wit et al., 2005).

En los últimos años, Colombia ha realizado esfuerzos a través de diversas políticas públicas para mejorar la calidad de la educación básica y secundaria³ (DNP, 2015). Recientemente, la política ha estado orientada a combatir la desigualdad en el acceso a la educación superior, desarrollando programas de gran escala dirigidos a la población vulnerable del país⁴. En este contexto, la iniciativa Ser Pilo Paga (SPP) buscaba expandir el acceso a la educación universitaria a estudiantes de bajo NSE.

2.1. El Programa *Ser Pilo Paga*

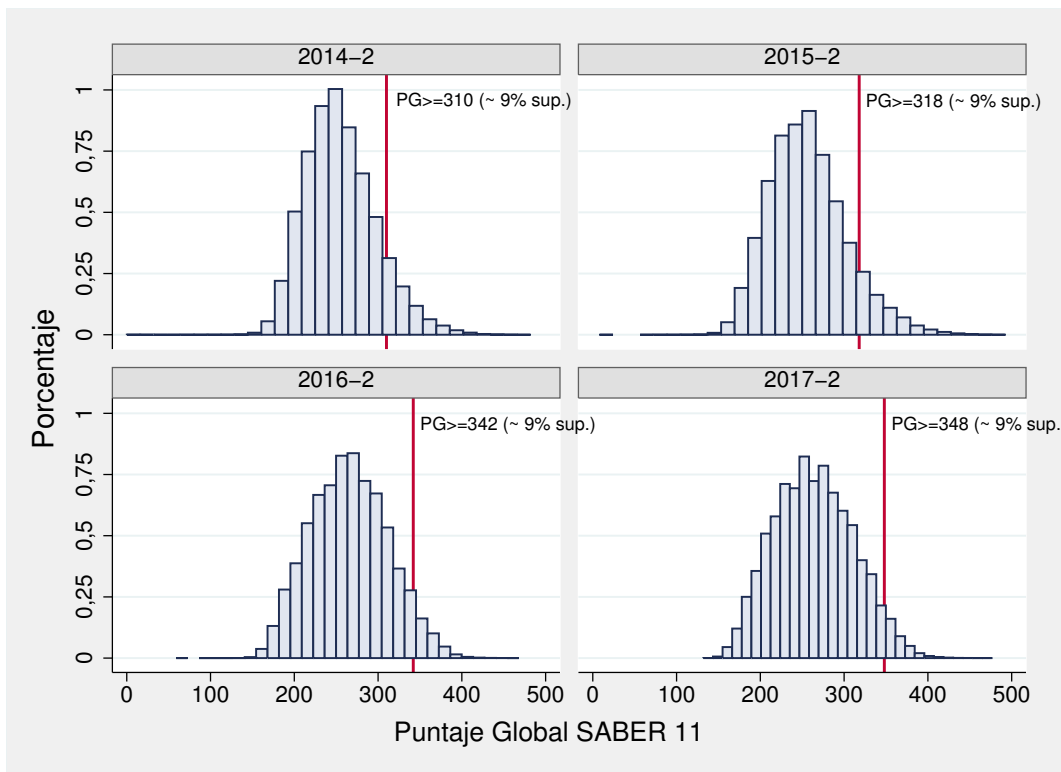
En un esfuerzo por contribuir al objetivo de convertir a Colombia en el país más educado de América Latina para el año 2025, el 1 de octubre de 2014 el Ministerio de Educación Nacional (MEN) lanzó *Ser Pilo Paga*, un programa nacional de ayuda financiera para personas de bajos ingresos

³Un ejemplo de estas iniciativas incluye el programa PACES, que proporcionó a más de 125.000 estudiantes, vales que cubrían más de la mitad del costo de la escuela secundaria privada entre 1991 y 1997.

⁴En 2002, con el apoyo de un préstamo del Banco Mundial, el gobierno diseñó y financió el programa Acceso con Calidad a la Educación Superior (ACCES), para aumentar el acceso a la educación superior a estudiantes de ingresos bajos y medios, a través de una línea de créditos de largo plazo. Actualmente, el programa se ha transformado en el crédito Tú Eliges.

basado en el mérito. Con esto, el gobierno de Juan Manuel Santos se comprometía a beneficiar con crédito-beca a 10.000 estudiantes ("Pilos", de ahora en adelante) por año, hasta completar los 40.000 en el 2018. Los préstamos, que pueden ser condonados después de graduarse, cubren la totalidad de la matrícula en cualquiera de las 44 universidades con acreditación de alta calidad en el país y en la especialidad de su elección. El costo académico de cada Pilo es transferido por el gobierno nacional a las universitarias donde fueron admitidos. Adicionalmente, los Pilos reciben directamente un estipendio de sostenimiento que oscila entre uno y cuatro veces el salario mínimo mensual por semestre, dependiendo de la necesidad de migrar o no a un área metropolitana diferente para cursar sus estudios de educación superior (Álvarez et al., 2017, Londono-Velez et al., 2017).

Figura 1: Criterio académico de elegibilidad SPP periodos 2014 - 2017



Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014-2017.

La elegibilidad para SPP se basa en tres criterios: [1] **Necesidad:** Los estudiantes deben provenir de un hogar con un puntaje por debajo del umbral en el índice socioeconómico SISBÉN⁵. La línea de corte varía según la ubicación geográfica de los hogares: 57,21 en las 14 áreas metropolitanas principales⁶; 56,32 en las demás áreas urbanas; y 40,75 en áreas rurales. [2] **Mérito** Los estudiantes deben obtener un puntaje global superior al percentil 91 en las evaluaciones nacionales de egreso

⁵El Sisbén es el Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales que, a través de un puntaje calculado por el Departamento Nacional de Planeación (DNP), clasifica a la población de acuerdo con sus condiciones socioeconómicas, permitiendo identificar aquellos en situación de pobreza y vulnerabilidad y así focalizar la inversión social.

⁶Las 14 ciudades principales son: Bogotá, Medellín, Cali, Barranquilla, Cartagena, Cúcuta, Bucaramanga, Ibagué, Pereira, Villavicencio, Pasto, Manizales y Santa Marta.

de la educación media *SABER 11*. Para el año 2014, esto significó obtener una puntuación de al menos 310 (sobre una escala de 500). Los umbrales de *SABER 11* se situarían en 318/500, 342/500 y 348/500 para los periodos 2015, 2016 y 2017, respectivamente (ver Figura 1). [3] Los solicitantes deben tener una oferta de admisión en alguno de los programas ofertados por una universidad con acreditación de alta calidad en el país.

2.2. El Nivel Socioeconómico

Una de las limitaciones que presenta investigación, es la imposibilidad de identificar el criterio de necesidad para los estudiantes evaluados en *SABER 11* en el último periodo del programa. Aunque la estrategia de identificación utilizada se expone con detalle en la subsección 5.1, presento en este apartado el marco teórico que respalda dicho proceso.

El gobierno colombiano utiliza dos medidas de nivel socioeconómico en la administración de sus políticas públicas: *Estrato* y *Sisbén*.

El *Estrato*, es una una medida del nivel socioeconómico que se usa principalmente para determinar el cobro diferencial de los servicios públicos (energía, agua y el saneamiento básico), aunque algunas empresas privadas lo utilizan para establecer precios para varios servicios (como televisión por cable y acceso a internet, entre otros). La representación actual del sistema de estratos provienen de la Ley 142 de 1994, en la cual el gobierno clasifica a las personas de acuerdo con la ubicación geográfica de su domicilio principal, en seis categorías diferentes: dos clases bajas (1 “baja-baja” y 2 “baja”), tres clases medias (3 “medio-bajo”, 4 “medio” y 5 “medio-alto”) y una clase Alta (6 “Alta”). El sistema está diseñado para que los estratos 1, 2 y 3 reciban subsidios de los estratos 5 y 6; los estratos 4 no reciben ni pagan subsidios. El cálculo del estrato se deriva utilizando una fórmula que incluye, tanto las características físicas de las viviendas como del entorno donde esta se encuentra ubicada, sin considerar el nivel de ingresos de quienes residen en el área. Los estratos pueden variar enormemente incluso a distancias pequeñas (por ejemplo, es posible que un área de estratos 1 bordee un área de estratos 5), aunque esto es poco común.

La segunda medida del nivel socioeconómico de una persona es su puntaje de *Sisbén*. Esta medida es utilizada por el gobierno para seleccionar los beneficiarios de programas sociales, como salud y educación. Está determinada por una fórmula clasificada que utiliza variables a nivel de hogar como un indicador del bienestar económico, destacándose las condiciones de la vivienda, el acceso a servicios públicos, la posesión de bienes durables, la asignación de capital humano, entre otras. Las medidas de ingreso familiar y estrato se excluyen en este cálculo; sin embargo, las personas que viven en barrios de estratos 4 o superiores no son elegibles para recibir un puntaje de *Sisbén*. La puntuación de *Sisbén* varía en una escala de 0 a 100, con el cual se divide la población en seis niveles (de inferior a mejor condición socioeconómica). El puntaje *Sisbén* para un hogar no caduca hasta que se publique una nueva versión del *Sisbén*, a pesar de existir cambios en algunas circunstancias, por ejemplo: un hogar podría mudarse a un vecindario de estratos 6 y mantener su mismo puntaje *Sisbén* a pesar de no ser elegible debido a su nuevo estrato.

Si bien, estos instrumentos corresponden a enfoques distintos (la estratificación socioeconómica para focalizar servicios domiciliarios y el *Sisbén* dirigido a servicios personalizados) ambos enfoques

concuerdan en priorizar la política social en aquellos individuos representados en los niveles más bajos de la clasificación. De acuerdo con [Bottia et al. \(2012\)](#) existe una alta correlación (0,8) entre el Sisbén III y la clasificación por estrato para el total de datos nacionales, haciendo que el error de exclusión sea prácticamente nulo.

3. Literatura Relacionada

La literatura ha estudiado ampliamente los efectos de la ayuda financiera para estudiantes ([Chen and DesJardins, 2010](#), [Agasisti and Murtinu, 2016](#), [Scott-Clayton and Zafar, 2019](#)). Recientemente, gran parte de la literatura sobre ayuda financiera estudiantil se ha interesado en los efectos diferenciales de la ayuda basada en la necesidad y en el mérito a nivel estatal e institucional. Estos estudios han analizado a través de métodos experimentales y cuasi-experimentales los efectos de este tipo de ayudas sobre la inscripción escolar ([Boatman and Long, 2016](#)), la persistencia en la graduación ([Goldrick-Rab et al., 2016](#)) y el logro académico ([Angrist et al., 2009](#)). Una contribución amplia y reciente de la evolución de la ayuda estudiantil estadounidense en las últimas cuatro décadas es documentada por [Dynarski and Scott-Clayton \(2013\)](#), quienes analizan la efectividad de los programas de créditos y becas estudiantiles sobre la inscripción, la persistencia y la finalización universitaria. Respecto a los efectos de la ayuda financiera sobre el rendimiento académico, las conclusiones aún se encuentran en discusión⁷. Sin embargo, [Agasisti and Murtinu \(2016\)](#) y [Barrera-Osorio and Filmer \(2016\)](#) concuerdan que los programas de ayudas basados en el mérito son más efectivos para elevar el rendimiento académico que aquellos basados en la necesidad pura.

Otro de los temas de relevancia dentro de la literatura de la ayuda financiera para la educación se basa en describir la distribución de la ayuda entre los estudiantes ([Heller, 2001](#), [McPherson and Schapiro, 2006](#), [Yang, 2010](#), [Kim, 2012](#), [Duffy and Goldberg, 2014](#)). Estos estudios generalmente investigan el grado en que las características asociadas a los estudiantes y a las instituciones educativas se correlacionan con cantidades mayores o menores de ayuda financiera. Diversos estudios de este tipo han examinado características individuales como el género, la raza/etnia, la edad, el nivel de ingresos y los puntajes de pruebas estandarizadas; al igual que atributos institucionales como la naturaleza de la institución educativa, la elegibilidad de los alumnos y las características del cuerpo estudiantil. La mayoría de los autores están interesados en la medida en que las cantidades de ayuda institucional siguen un patrón común.

Uno de los primeros estudios en esta área proviene de [Barnes and Neufeld \(1980\)](#). Utilizando datos de la Encuesta Nacional Longitudinal de 1972 (NLS-72, por sus siglas en inglés), los autores ajustan un modelo Tobit para estimar el efecto de las características de los estudiantes, incluidos los ingresos y la capacidad académica, en la concesión de ayudas financieras. Los investigadores encuentran que las variables relacionadas con la “necesidad” solo están débilmente correlacionadas tanto con la probabilidad de que un estudiante reciba una oferta de ayuda como con el tamaño de la ayuda una vez que es ofrecida. Estos resultados contrastan con los evidenciados por [Grubb and](#)

⁷Dicha evidencia abarca desde estudios que no encuentran efectos ([Angrist and Lavy, 2002](#)), efectos negativos ([Cornwell et al., 2003](#)), efectos positivos en el rendimiento de los estudiantes de alta capacidad, pero impacto negativo en el rendimiento de los estudiantes de baja capacidad ([Leuven et al., 2010](#), [Levitt et al., 2016](#)); a las investigaciones que encuentra efectos positivos y grandes en el rendimiento académico y la finalización ([Bettinger, 2004](#), [Angrist et al., 2009](#), [Barrera-Osorio et al., 2011](#), [Scott-Clayton, 2011](#), [Fack and Grenet, 2015](#)).

Tuma (1991) y Kane and Spizman (1994), quienes a partir de un modelo de probabilidad lineal y uno de datos censurados respectivamente, encuentran una disminución en la probabilidad de recibir ayudas financieras para estudiantes universitarios conforme aumentaban tanto el nivel educativo como de ingreso de los padres.

Con estudios posteriores, el vínculo entre la necesidad, la capacidad académica y las ofertas de ayuda financiera institucional se ha fortalecido. Utilizando datos de la Encuesta Nacional de Ayuda Estudiantil Postsecundaria (NAPSAS, por sus siglas en inglés) entre 1989 y 1990, Boschung et al. (1998), estiman un modelo Tobit, prediciendo la ayuda institucional por necesidad, etnia, género y características académicas. Los resultados sugieren que tanto la capacidad académica como la necesidad condujeron a mayores cantidades de ayuda financiera. También encontraron que los estudiantes asiáticos y afroamericanos recibieron mayores cantidades de ayuda institucional que los estudiantes blancos, mientras que los estudiantes hispanos recibieron menores cantidades de ayuda. En relación con el género, se encontró que los estudiantes masculinos tenían 2% más probabilidades de recibir becas, siendo este diferencial de \$187,4 dólares en promedio.

En otro estudio que utiliza datos de la NAPSAS, Heller (2001) estima el impacto de las características socioeconómicas de los estudiantes, como la raza, el género, el nivel de educación de la madre y los ingresos del hogar, en la probabilidad de recibir ayuda financiera. Usando un modelo logit con la recepción de ayuda como la variable dependiente, encuentra que la raza es un predictor importante del recibo de ayudas basada en la necesidad, particularmente para estudiantes afroamericanos. También encuentra que el GPA de la universidad está positivamente correlacionado con la probabilidad de recibir ayuda basada en la necesidad.

Singell Jr (2002) estima la capacidad de respuesta de la ayuda financiera estudiantil a las características de los estudiantes en la Universidad de Oregón entre 1995 y 1998. Las estimaciones de un modelo Tobit recursivo revelan que los puntajes de GPA y SAT de la escuela secundaria se asocian positivamente con la ayuda institucional, mientras que las medidas de la riqueza del hogar tales como los ingresos del área local y los montos de las hipotecas están negativamente asociados con la ayuda institucional. Así mismo, los resultados empíricos evidencian un alto grado de discreción en los paquetes de ayudas ofrecidos, donde ciertos tipos de estudiantes (por ejemplo, aquellos con especializaciones específicas) recibieron cantidades promedio de ayuda más altas que otros.

En una serie de libros sobre ayuda financiera para estudiantes, McPherson y Schapiro describen cómo la competencia institucional por el prestigio y la financiación ha impulsado cambios profundos en el sistema de ayuda financiera para estudiantes. En un capítulo McPherson and Schapiro (2006), documentan los cambios en la capacidad de respuesta de la ayuda institucional a las características de los estudiantes entre 1992-1993 y 1999-2000. Utilizando datos de la NAPSAS, los autores encuentran que el principio de otorgar ayuda financiera estrictamente en relación con la capacidad de pago se está convirtiendo en un factor cada vez menos importante en la distribución de ayuda en los colegios y universidades privadas de Estados Unidos. En esta línea, Doyle (2010) examina los cambios en la capacidad de respuesta institucional a las necesidades de los estudiantes y los puntajes

de las pruebas durante el período 1992–2003. Su análisis reveló que en tres de los cuatro sectores⁸ considerados en el estudio, la concesión de ayudas ha estado menos receptiva a las necesidades de los estudiantes y más receptiva a las características académicas con el tiempo. La única excepción a esta tendencia se presentó en instituciones públicas de doctorado de cuatro años, las cuales fueron consistentemente sensibles a las necesidades de los estudiantes en su provisión de ayuda institucional.

Estudios basados en muestras para Estados Unidos también han mostrado como las diferencias institucionales son determinantes críticos de la concesión de la ayuda financiera para la educación. Con relación a lo anterior, [Boschung et al. \(1998\)](#) encontraron que los estudiantes de instituciones privadas obtenían mayores montos ayuda en comparación con los estudiantes universitarios públicos. Del mismo modo, [Grubb and Tuma \(1991\)](#) evidenciaron como los estudiantes de universidades privadas tenían una probabilidad significativamente mayor de obtener ayudas financieras que sus contrapartes públicas.

Un estudio sobre la distribución de la ayuda financiera estatal para la educación postsecundaria en China merece una mención específica debido a su similitud con el enfoque adoptado en este documento. [Yang \(2010\)](#), utilizando datos de la Encuesta de Desarrollo de Estudiantes Universitarios (CSDS, por sus siglas en inglés) para el año 2008 ajustados a un Modelo Lineal Jerárquico de dos niveles, examinó la distribución del acceso a la ayuda financiera entre diferentes estudiantes e instituciones. Sus hallazgos sugieren una disposición del gobierno Chino por asignar la ayuda pública primando criterios de mérito más que de necesidad. Así, la investigación encontró que asistir a una institución selectiva con pares de alta capacidad, obtener puntajes altos en los exámenes de ingreso a la universidad y escoger carreras relacionadas con la ciencia se correlacionó positivamente la probabilidad de recibir ayuda. Contrario a lo esperado por el autor, los estudiantes de bajo NSE y hogares más pobres, no recibieron mayor cantidad de ayuda financiera.

Recientemente en Colombia, la investigación en torno a la ayuda financiera para la educación superior ha ganado protagonismo después del lanzamiento del programa Ser Pilo Paga en el año 2014. En este contexto, los investigadores se han interesado casi de manera exclusiva por intentar dilucidar los efectos causales del programa a partir de diseños quasi-experimentales de evaluación de impacto a corto plazo. Dentro de este grupo de investigaciones, la hipótesis subyacente sugiere que la introducción del programa SPP generó un efecto emocional que insentivó a los estudiantes a esforzarse y prepararse para obtener mejores resultados en la prueba estandarizada SABER 11 ([Arango, 2017](#), [Laajaj et al., 2018](#), [Basto Aguirre et al., 2019](#), [Bernal and Penney, 2019](#)).

Así mismo, la literatura reconoce otros efectos del programa SPP sobre la población beneficiada. [Londono-Velez et al. \(2017\)](#), a partir de un Diseño de Regresión Discontinua (RDD, por sus siglas en inglés), encuentra que el programa SPP aumentó la inscripción universitaria inmediata entre los estudiantes elegibles en más del 80 %, eliminó la brecha de matrícula socioeconómica entre los mejores estudiantes y promovió en un 46 % la diversidad de clases en las mejores universidades privadas del país. También, los autores sugieren que el programa parece haber elevado el rendimiento

⁸En este estudio, los cuatro sectores son: Instituciones doctorales públicas de cuatro años, instituciones no doctorales públicas de cuatro años, instituciones doctorales privadas de cuatro años e instituciones privadas no doctorales de cuatro años.

general de la prueba SABER 11 entre cohortes recientes.

Por su parte, [Álvarez et al. \(2017\)](#) indaga sobre los efectos causales del programa sobre variables de acceso a la educación superior y desempeño académico. Dentro de sus principales hallazgos se destacan: Un aumento en la probabilidad de acceso a educación superior de calidad de al menos 41 puntos porcentuales en las dos primeras cohortes del programa, una disminución del ausentismo intersemestral de 7,1 y 6 puntos porcentuales en las cohortes 1 y 2 respectivamente, además de mejoras en los índices de calidad de vida de los beneficiarios. Estos resultados se encuentran en línea con los hallazgos evidenciados por [Melguizo et al. \(2016\)](#), quien concluye que el programa ha sido efectivo en términos de aumentar el número potencial de estudiantes de bajos ingresos que se hubieran matriculado en la universidad, así como lograr disminuir el número de estudiantes que abandonaron. Así mismo, los efectos del programa de ayudas en la inscripción, exhiben la mayor magnitud y tienen un impacto heterogéneo claramente positivo en los solicitantes más pobres. Los resultados son menos convincentes para la deserción, pero en la dirección esperada ([Melguizo et al., 2016](#)).

Finalmente, dentro de la revisión bibliográfica realizada, solo un documento propone abordar los efectos del programa SPP sin la utilización de diseños experimentales de evaluación de impacto. En tal sentido, [Medina et al. \(2018\)](#) a partir de un análisis intensivo de los cambios en las distribuciones estadísticas de los puntajes de la prueba SABER 11 durante los periodos 2014-2017 para distintos NSE, proporciona evidencia de mejoras en el rendimiento de los estudiantes, no solo de los beneficiarios del programa, sino también de toda la población estudiantil. Por lo tanto, los autores concluyen que el programa abrió oportunidades similares para todos los estudiantes, especialmente para los más pobres.

Como se pudo notar, la literatura sobre el impacto de la ayuda financiera para la educación en Colombia ha sido suficientemente discutida mediante el diseño de métodos cuasi-experimentales, los cuales intentan recuperar el efecto causal de los datos observados derivados del programa Ser Pilo Paga. Sin embargo, continua sin ser explorada en torno a los aspectos relevantes que den cuenta de la manera en la ayuda financiera gubernamental se distribuye en los distintos estudiantes. En tal sentido, este documento pretende contribuir a llenar el vacío que existe en la literatura, al examinar como responden la ayuda financiera del programa SPP a las características de los estudiantes y de las instituciones educativas.

4. Teoría y Especificación del Modelo

En esta sección, se presenta una teoría básica de la respuesta gubernamental a las características de los estudiantes en relación con la ayuda financiera para el acceso a la educación superior. Las proposiciones en esta sección tienen como sustento el modelo de ayuda institucional de [Epple et al. \(2003\)](#) ajustado al caso Estatal.

El gobierno central utiliza la ayuda financiera para alentar a los estudiantes a asistir a una institución universitaria que de otro modo no podrían asistir. Para otorgar esta ayuda, el gobierno responde a una variedad de características de los estudiantes, algunas de las cuales son potenciadas

por las instituciones educativas (colegios) a las que pertenecen. En este estudio, las características de los estudiantes son principalmente de dos tipos: Necesidad y capacidad académica demostrada. En el caso de la necesidad, el Estado utiliza los fondos de ayuda financiera para proporcionar suficiente dinero para que el estudiante pueda asistir (reducir al brecha educativa superior). En el caso de la capacidad académica, el Estado utiliza sus recursos de ayuda financiera para incentivar a aquellos estudiantes sobresalientes cumplir estándares más altos.

Por lo tanto, se puede considerar que el gobierno tienen un nivel latente de demanda por estudiantes determinados, y^* , expresado a través del gasto en la ayuda financiera para la educación. La capacidad de respuesta gubernamental a la necesidad y el mérito académico se pueden resumir en los coeficientes β de la siguiente ecuación:

$$Y_i^* = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i \quad (1)$$

Donde Y^* es la demanda gubernamental por estudiantes, X_{1i} representa una medida del nivel de ingresos del hogar, X_{2i} es un indicador de la capacidad académica demostrada del estudiante (como el puntaje de las pruebas estandarizadas), β_{ki} son los coeficientes de las medidas de necesidad y capacidad académica respectivamente, y ϵ_i es el término de error aleatorio.

Suponemos además, que el Estado no ofrecerá recursos a todos estudiantes determinados que soliciten ayuda. Algunos gobiernos simplemente no pueden darse el lujo de hacerlo por cuestiones de capacidad presupuestal, otros simplemente tienen objetivos específicos que desean lograr y asignarán los recursos en consecuencia. De acuerdo con lo anterior, los gobiernos que tienen un alto nivel de demanda por estudiantes de bajos ingresos (o más bien gobiernos con una fuerte filosofía basada en la necesidad) tendrán un coeficiente para β_1 negativo. Por su parte, aquellos con un alto nivel de demanda por estudiantes con capacidad académica demostrada tendrán un coeficiente β_2 positivo.

Por otro lado, el nivel observado de demanda gubernamental puede verse a partir de su gasto en ayuda financiera, pero solo cuando el monto de la ayuda ofrecida es mayor que 0. Esto da lugar a observaciones *censuradas*, ya que una parte considerable de los estudiantes que califican no recibirán ayuda gubernamental alguna. De acuerdo con el objetivo de esta investigación, el cual se centra en analizar la distribución de la ayuda gubernamental para la educación superior sin considerar las diferencias en la cantidad de ayuda entre los receptores, un enfoque apropiado consiste en modelar el recibo de ayuda como un simple resultado binario.

Otro aspecto relevante a considerar dentro del diseño metodológico es la relación jerárquica entre las unidades de análisis. Dada la estructura anidada de los datos (estudiantes agrupados en colegios), los métodos estadísticas convencionales de un solo nivel no resultan apropiadas para determinar la asociación entre distintos factores y la variable objeto de estudio. Si bien la ayuda financiera es independientes entre los distintos estudiantes de diferentes colegios, esta no lo es para aquellos que pertenecen a un mismo colegios y comparten las mismas influencias específicas del grupo. En ese sentido, existe una heterogeneidad no observada a nivel de grupo que conduce a la dependencia entre las respuestas para las unidades en el mismo grupo después del condicionamiento de las covariables (Steele, 2009, Rabe-Hesketh and Skrondal, 2012). Esta heterogeneidad no observada se modela eficientemente al incluir efectos aleatorios en un modelo de regresión múltiple

(Hox et al., 2017).

En este documento, debido a la estructura jerárquica de los datos y la naturaleza discreta de la variable dependiente, se estima un modelo de regresión lineal jerárquico de dos niveles, siguiendo la notación y línea argumental de Rabe-Hesketh and Skrondal (2012), cuyo desarrollo estadístico del modelo jerárquico bietápico es el siguiente:

Denotaremos con $j = 1, \dots, J$ las unidades del segundo nivel (*los colegios*), $i = 1, \dots, n_j$ las unidades del primer nivel (*los estudiantes*⁹) y X_{ij} un conjunto de covariables del nivel 1. Para una respuesta binaria de Y_{ij} , $E(Y_{ij}|X_{ij}, u_j) = P_{ij} = \Pr(Y_{ij} = 1)$, el modelo logístico de interceptos aleatorios se puede escribir como:

$$\text{logit} [\Pr(Y_{ij} = 1|X_{ij}, u_j)] = \beta_0 + X'_{ij}\beta + u_j \quad (2)$$

Donde u_j son los efectos aleatorios del segundo nivel, los cuales suponemos independientes e idénticamente distribuidos con media 0 y varianza σ_u^2 [$u_j \sim N(0, \sigma_u^2)$].

En la literatura de modelos multinivel, la ecuación (2) o *modelo combinado*, puede escribirse alternativamente como:

$$\text{logit} [\Pr(Y_{ij} = 1|X_{ij}, u_j)] = \beta_{0j} + X'_{ij}\beta \quad (3)$$

donde

$$\beta_{0j} = \beta_0 + u_j \quad (4)$$

A fin de intentar explicar la variabilidad entre colegios, se incluye en la ecuación (4) un conjunto de covariables del nivel 2, denotadas por Z_j .

$$\beta_{0j} = \beta_0 + Z'_j\gamma + u_j \quad (5)$$

Al sustituir la ecuación (5) en ecuación (2) obtenemos la forma reducida:

$$\text{logit} [\Pr(Y_{ij} = 1|X_{ij}, Z_j, u_j)] = \beta_0 + X'_{ij}\beta + Z'_j\gamma + u_j \quad (6)$$

Usando el concepto de variable latente, la ecuación anterior puede escribirse equivalentemente como:

$$Y_{ij}^* = \beta_0 + X'_{ij}\beta + Z'_j\gamma + \underbrace{u_j + \varepsilon_{ij}}_{\xi_{ij}} \quad (7)$$

Donde $u_j \sim N(0, \sigma_u^2)$ y los residuos del nivel 1, ε_{ij} , tienen distribución logística estándar. La respuesta binaria Y_{ij} está determinada por las respuestas continuas latentes a través del modelo de umbral:

$$Y_{ij} = \begin{cases} 1 & \text{si } Y_{ij}^* > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Debido a que los interceptos aleatorios se comparten entre los estudiantes del mismo colegio, inducen dependencia entre las respuestas dicotómicas para los estudiantes dentro de la misma

⁹Dado que en la muestra se presenta en promedio un estudiante por hogar, estos son considerados parte del mismo nivel.

institución dadas las covariables. Esta dependencia se puede cuantificar mediante el *Coefficiente de Correlación Intraclase (ICC)* o *Coefficiente de Partición de Varianza*, ρ , así:

$$\rho \equiv \text{Cor}(Y_{ij}^*, Y_{i'j}^* | X_{ij}, X_{i'j}) = \text{Cor}(\xi_{ij}, \xi_{i'j}) = \frac{\sigma_u^2}{\sigma_u^2 + 3, 29}, \quad i' \neq i \quad (8)$$

Por tanto, la correlación intraclase, en este caso intracolegios, representa la proporción de la varianza residual total ($\sigma_u^2 + 3, 29$) que se debe a la varianza residual entre colegios σ_u^2 . La existencia de un ICC diferente de cero significa que el método de estimación de mínimos cuadrados no es adecuado ya que conduce a inferencias incorrectas.

5. Datos

Los datos utilizados en esta investigación provienen del Instituto Colombiano para la Evaluación de la Educación (ICFES), la institución estatal a cargo de las pruebas estandarizadas en materia educativa en Colombia. Este conjunto de datos contiene información de todos los estudiantes que tomaron el examen SABER 11 en el segundo semestre de cada uno de los cuatro años de duración del programa SPP, 2014 a 2017. En cada cohorte hay aproximadamente 545,000 estudiantes. Se debe tener en cuenta que, para la realización de las estadísticas descriptivas y las estimaciones econométricas, se hizo una depuración de la base de datos eliminando registros para los cuales no se tiene información. En la subsección 5.2, se describirá esta información y se caracterizarán las covariables utilizadas en las estimaciones.

5.1. Identificación del Umbral de Necesidad

Identificar con precisión el criterio de necesidad ha sido una limitante importante en esta investigación. La información que proporciona el ICFES en sus bases de datos, permite conocer el nivel de Sisbén en que está clasificada la familia del estudiante hasta el periodo 2016-2. Dada la necesidad de identificar, dentro de los evaluados en SABER 11 para el periodo 2017, aquellos que cumplieran con el criterio de necesidad, fueron consideradas dos estrategias: Pronosticar el nivel de Sisbén y utilizar una variable proxy.

■ Pronóstico del nivel de Sisbén.

Como se expuso en la subsección 2.2, determinar el nivel de Sisbén no es tarea sencilla. Para su cálculo, se requieren de un número considerable de variables relacionadas con necesidades básicas a nivel de hogar. Actualmente se han desarrollado técnicas de Inteligencia Artificial que le permiten a un software aprender y dar un diagnóstico del NSE de los individuos, incluso predecir de acuerdo a un conjunto de datos. Dentro del contexto de la minería de datos, dichas técnicas se conocen como *Aprendizaje de Máquina* (ML, por sus siglas en inglés).

Al contar con datos correctamente identificados sobre los niveles de Sisbén en las tres primeras cohortes de SPP, fue posible utilizar métodos de ML supervisado. Para el conjunto de datos de entrenamiento, fueron seleccionados aleatoriamente 1.067.197 estudiantes (equivalente al 80 %) del total de evaluados en SABER 11 entre 2014 y 2016. El 20 % restante (266.795 estudiantes) conformó el conjunto de datos de prueba. Los parámetros empleados en la predicción fueron: El índice de nivel socioeconómico del estudiante (continua), el estrato de residencia del estudiante (categórica),

el máximo nivel educativo del padre y la madre del estudiante (categóricas), el índice del nivel socioeconómico del colegio¹⁰ (continua) y el municipio de residencia del estudiante (categórica). La cantidad reducida de parámetros, tiene su justificación en las pocas variables comunes en todos los periodos (2014 - 2017), ya que es necesario emplear los mismos parámetros, tanto en la etapa de entrenamiento (datos cohortes 2014 - 2016), como en la etapa de clasificación (cohorte 2017).

Los clasificadores considerados para este ejercicio fueron: Naïve Bayes, Random Forest, Decision Tree y Gradient boosting; utilizando el paquete “sklearn” de *Python*. Las métricas de evaluación de los métodos escogidos se presentan en la tabla a continuación:

Tabla 1: Resultados de las métricas de evaluación de los algoritmos de ML considerados

Clasificadores	Métricas de Evaluación		
	Exactitud	Precisión	Retirada
Gradient Boosting	0,7509	0,5534	0,4969
Decision Tree	0,6924	0,4680	0,4686
Extra Trees	0,7087	0,4832	0,4756
Random Forest	0,7163	0,4912	0,4794
Naïve Bayes (Bernulli)	0,4963	0,0993	0,2000

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014-2016.

En cuanto a la evaluación de las técnicas escogidas, el clasificador *Gradient Boosting* presentó un desempeño superior en todas las métricas durante la etapa de entrenamiento. De acuerdo con el criterio de *exactitud*, los resultados indican que, en el mejor de los casos, se tendrá una clasificación errónea 25 de cada 100 intentos de clasificar. A continuación, se presentan los resultados obtenidos con el método de *Gradient Boosting* en la etapa de clasificación.

Tabla 2: Resultados clasificación nivel de Sisbén para el periodo 2017

Nivel de Sisbén	Observado			Pronosticado
	2014	2015	2016	2017
Nivel 1	251.903	267.244	205.487	173.892
Nivel 2	120.148	119.555	115.833	167.476
Nivel 3	21.435	19.427	18.997	9.954
Clasificado en otro nivel	4.267	3.589	4.743	1.213
No está clasificado por el Sisbén	144.263	128.083	117.892	143.270
Total	542.016	537.898	545.018	495.805

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014-2017.

¹⁰Variable construida a partir del promedio por colegio del Índice Nivel Socioeconómico (INSE) de los estudiantes. Se explica con mas detalle en la subsección siguiente.

- **Utilizar una variable proxy**

La otra estrategia considerada es la utilización de una variable proxy. Como se introdujo en la subsección 2.2, el *estrato socioeconómico* y el *nivel de Sisbén* son los instrumentos utilizados por el gobierno colombiano en la focalización de sus políticas públicas, dado que ambas permiten identificar el nivel socioeconómico de los individuos. A fin de considerar la utilización del estrato socioeconómico como una variable proxy del nivel de Sisbén, resulta conveniente comparar el desempeño de estos dos instrumentos.

En el caso, en que el estrato socioeconómico (niveles 1, 2 y 3) se utilizara como criterio de necesidad para determinar los beneficiarios del programa SPP, 333.562 estudiantes no elegibles (ubicados en las categorías *Nivel superior* y *Sin clasificación Sisbén* pero residentes en estratos 1, 2 y 3) serían potenciales beneficiarios del programa (ver tabla 3). En este contexto, la proporción de estudiantes que “erróneamente” podrían beneficiarse del programa por residir en los estratos 1, 2 y 3, el error de inclusión, sería del 21,5%. De igual manera, la fracción de personas que serían “erróneamente” excluidas de los potenciales elegibles, estaría dada por el cociente de los estudiantes en los niveles 1, 2 y 3 del Sisbén que viven en estratos 4, 5 y 6 (6.453 estudiantes no elegibles basados en el estrato), y el total de estudiantes clasificados en Sisbén 1, 2 y 3 (1.222.095 estudiantes). Así el error de exclusión sería del 0,5%.

Tabla 3: Distribución de los estudiantes evaluados en SABER 11 entre 2014 y 2016 según estrato socioeconómico y nivel de Sisbén

Estrato	Nivel de Sisbén					Total
	Nivel 1	Nivel 2	Nivel 3	Nivel sup.	Sin registro	
1	645.129	29.228	2.735	2.981	37.643	717.716
2	130.957	286.691	13.473	5.102	124.611	560.834
3	28.482	37.531	41.416	3.540	159.685	270.654
4	1.656	1.757	1.893	771	47.819	53.896
5	332	255	282	168	14.726	15.763
6	144	74	60	37	5.754	6.069
Total	806.700	355.536	59.859	12.599	390.238	1.624.932

Nota: El coeficiente de correlación de Pearson entre estas dos variables es 0,713.

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014-2016.

En concordancia con los hallazgos de [Bottia et al. \(2012\)](#), la alta correlación entre el Sisbén y la estratificación socioeconómica (coeficiente de correlación de Pearson de 0,713 en esta muestra) hace que el error de exclusión sea prácticamente nulo, mientras que el de inclusión podría alcanzar el 21,5%.

Como se pudo notar, el utilizar la variable estrato como proxy del nivel de Sisbén para la identificación del criterio de necesidad para los evaluados en 2017, reduce en 3,5 p.p. la probabilidad de cometer error tipo II en la clasificación. Por lo expuesto anteriormente, en lo que resta de esta investigación, se determina la utilización del estrato como el criterio de necesidad en la elegibilidad del programa SPP.

5.2. Variables y Muestra

Las variables utilizadas en este estudio, consisten en información a nivel de estudiante y colegio, las cuales fueron seleccionadas a luz de la revisión de literatura y la teoría expuesta en la sección anterior. A continuación se realiza una descripción de estas variables:

- **SPP:** Es la variable binaria a explicar, creada con valor uno si el evaluado es beneficiario del programa SPP y cero en caso contrario.
- **Puntaje Global:** Variable que relaciona el puntaje global obtenido por el estudiante en la prueba SABER 11. Valores altos de esta variable indican una mayor capacidad académica demostrada por el estudiante.
- **Nivel Educativo del Hogar (NEH):** Es una variable categórica ordenada que toma el valor del mayor nivel educativo del padre o la madre del estudiante, la cual recoge el el *background familiar*. Los niveles de la variable son los siguientes: Inferior, Bachiller, Técnico y Superior.
- **Estrato:** Variable categórica que representa el estrato socioeconómico del estudiante. La variable toma valor uno si el estudiante reside en el estrato 1; dos si reside en el estrato 2 y tres si reside en el estrato 3.
- **Oficial:** Variable dicotómica que define la naturaleza pública o privada del colegio, toma el valor uno si el colegio es oficial y cero en caso contrario.
- **Índice Compañeros de Colegio (ICE):** Es un indicador del nivel académico medio de los pares escolares del estudiante i . Construida a partir del promedio por colegio del puntaje global en la prueba SABER 11 de los estudiantes menos el efecto del estudiante i . Esta variable intenta recoger el *efecto par*.
- **Área Geográfica (Área):** Variable categórica que intenta recoger los efectos diferenciales de los criterios de elegibilidad en el espacio geográfico. La variable toma valor uno, si el municipio de residencia del estudiante está en áreas rurales; dos, si el municipio pertenece a una de las ciudades intermedias (demás áreas urbanas); tres, si el municipio pertenece a una de las 14 áreas metropolitanas principales.

Por supuesto, ningún conjunto de covariables podría capturar todas las posibles razones por las cuales el gobierno colombiano podría proporcionar ayuda a un estudiante. Sin embargo, el propósito de este estudio no es predecir la cantidad de ayuda financiera brindada a cualquier estudiante con un alto nivel de precisión, sino estimar la respuesta del gobierno a las necesidades o al mérito de los estudiantes elegibles dentro de su programa de ayuda gubernamental. Dado este objetivo, el anterior conjunto de covariables debería proporcionar una estabilidad razonable en la estimación de los coeficientes para la variable de interés ([Angrist and Pischke, 2008](#)).

En concordancia con el interés principal de esta investigación, limito la muestra a estudiantes que cumplieron simultáneamente con los criterios de elegibilidad del programa SPP en los periodos 2014 a 2017. Así mismo, elimino aquellas observaciones con características individuales o institucionales incompletas. Lo anterior proporcionó los siguientes tamaños de muestra para los periodos en consideración: 39.224 estudiantes en 2014, 35.854 en 2015, 18.283 en 2016 y 12.877 en 2017. Los

diferentes procedimientos realizados para la obtención de estas muestras se detallan en el Anexo A.1.

5.3. Análisis Descriptivo

En la Tabla 4, se presentan las estadísticas descriptivas de la muestra y las variables utilizadas para el análisis. Se puede observar que en la población de elegibles, el 44 % residen en el estrato 3, alrededor del 38 % tienen al menos uno de sus padres con educación superior y egresan mayoritariamente de colegios oficiales con un 57 %. En relación con los beneficiarios del programa SPP, cerca del 26 % estudian en colegios privados, el 21 % viven en hogares con cuyos padres no lograron completar el título de bachiller, y alrededor del 32 % residen en el nivel socioeconómico más bajo. Por otro lado, la capacidad académica demostrada, el puntaje global en la prueba SABER 11, es alrededor de 2,4 puntos superior en el grupo de beneficiarios, siendo esta diferencia estadísticamente significativa.

Así mismo, se observan un efectos par mayor en la población de elegibles, con una brecha de 7,3 puntos. Una prueba de diferencia de medias confirma estadísticamente esta brecha con respecto a la población de beneficiarios. Finalmente, en cuanto a la ubicación geográfica se puede resaltar, que el mayor peso de los beneficiarios la tiene la categoría de ciudades principales con 53 %, situación que se invierte en la población de beneficiarios, donde la mayor proporción se concentra en zonas rurales con 51 %.

Tabla 4: Estadísticas descriptivas de elegibles y beneficiarios del programa SPP para el agregado de los periodos 2014 - 2017

Variable	Muestra Completa: Elegibles					Beneficiarios				
	Obs.	Media	Des. Est.	Min.	Max.	Obs.	Media	Des. Est.	Min.	Max.
Puntaje Global	106.238	344,86	23,42	310	492	39.941	347,23	22,20	310	492
NEH (Base: Inferior)										
Bachiller	106.238	0,2715	0,44	0	1	39.941	0,3227	0,47	0	1
Técnica	106.238	0,1980	0,40	0	1	39.941	0,1927	0,39	0	1
Superior	106.238	0,3828	0,49	0	1	39.941	0,2745	0,45	0	1
Estrato (Base: Estrato 1)										
Estrato 2	106.238	0,3916	0,49	0	1	39.941	0,4667	0,50	0	1
Estrato 3	106.238	0,4360	0,50	0	1	39.941	0,2216	0,42	0	1
Oficial	102.767	0,5656	0,50	0	1	39.941	0,7433	0,44	0	1
ICE	106.238	284,20	30,60	183,5	419,1	39.941	276,93	28,69	183,5	417,1
Área (Base: Zona rural)										
Intermedia	106.238	0,0693	0,25	0	1	39.941	0,0644	0,25	0	1
Principal	106.238	0,5278	0,50	0	1	39.941	0,4319	0,50	0	1

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014 - 2017.

6. Resultados Empíricos

La estimación del modelo logístico multinivel descrito en la sección 4 se realizó bajo una estructura secuencial iniciando con la estimación del modelo nulo, con el fin de calcular la variación total en cada una de las partes estocásticas del modelo y de evaluar la pertinencia de este en el

análisis. El método de estimación empleado para cada una de las etapas es máxima verosimilitud utilizando cuadratura adaptativa con 15 puntos de integración. Las estimaciones se realizaron bajo el comando *xtnlogit*, principal rutina de Stata para ajustar modelos multinivel con variables de respuesta binaria (Rabe-Hesketh and Skrondal, 2012).

En la Tabla 5 recoge los resultados de las estimaciones del modelo logístico de interceptos aleatorios agrupando las distintas covariables según nivel individual o de colegio. El primer modelo, modelo nulo, no incluye variables explicativas y solo permite analizar la influencia del colegio. El siguiente modelo (modelo 3), incluye las variables a nivel de estudiante y hogar. El modelo final (modelo 6), agrega al modelo anterior las características del nivel superior.

En relación con la influencia del contexto, el modelo nulo presenta solida evidencia que la varianza entre colegios es diferente de cero. Los hallazgos sugieren que un (27,7 %, 36,8 %, 48,2 % y 51,6 %) de la varianza en la probabilidad favorable a ser beneficiario del programa SPP es atribuible a características del nivel superior, para cada periodo del programa respectivamente. Al controlar por características a nivel individual (Modelo 3), se observa que la varianza estimada de los residuos se reduce, en promedio, 16,2 puntos porcentuales (p.p.) para cada periodo, con lo cual el ICC podría atribuir un (16 %, 19,3 %, 30 % y 34,3 %) de la misma a las características del nivel superior que no son observadas en este modelo. Por su parte, cuando se incluyen factores del nivel superior (Modelo 6), el ICC disminuye a 11,5 %, 15,1 %, 26,5 % y 28,9 % para los periodos 2014 a 2017, respectivamente; lo cual permitiría inferir que en la probabilidad de elegibilidad del programa SPP, tienen cierta influencia el rol que desempeña la naturaleza público/privada del colegio y el municipio de residencia.

A continuación, se realiza un análisis de la capacidad explicativa de las regresoras consideradas en esta investigación bajo una estructura paso a paso. Los resultados de las estimaciones de los distintos modelos para cada uno de los periodos en consideración, se presentan en los Anexos A.3, A.4, A.5 y A.6, los cuales permiten evidenciar la correspondencia de los signos esperados con los observados, así como una fuerte significancia estadística de los parámetros y varianzas estimadas.

Tabla 5: Estimación modelo logístico de interceptos aleatorios según características individuales y de comunidad

Parámetro	Modelo Nulo				Modelo 3: Variables Nivel 1				Modelo 6: Variables Nivel 1 + 2			
	2014	2015	2016	2017	2014	2015	2016	2017	2014	2015	2016	2017
	Coef.	Coef.	Coef.	Coef.	OR	OR	OR	OR	OR	OR	OR	OR
Parte Fija												
Factores del Primer Nivel												
Puntaje Global					1,004*** (0,001)	1 (0,001)	1 (0,001)	1 (0,001)	1,006*** (0,001)	1,001* (0,001)	1,001* (0,001)	0,999 (0,001)
NEH (Base: Inferior)												
Bachiller					0,840*** (0,032)	0,756*** (0,031)	0,745*** (0,057)	0,764*** (0,076)	0,894*** (0,034)	0,824*** (0,034)	0,808*** (0,061)	0,831* (0,082)
Técnica					0,700*** (0,031)	0,587*** (0,027)	0,520*** (0,041)	0,627*** (0,061)	0,805*** (0,035)	0,689*** (0,032)	0,611*** (0,049)	0,734*** (0,072)
Superior					0,488*** (0,021)	0,410*** (0,018)	0,370*** (0,028)	0,429*** (0,04)	0,603*** (0,026)	0,513*** (0,023)	0,463*** (0,035)	0,551*** (0,051)
Estrato (Base: Estrato 1)												
Estrato 2					0,527*** (0,02)	0,415*** (0,016)	0,264*** (0,019)	0,227*** (0,022)	0,618*** (0,023)	0,489*** (0,019)	0,319*** (0,022)	0,282*** (0,027)
Estrato 3					0,237*** (0,01)	0,150*** (0,007)	0,074*** (0,006)	0,052*** (0,005)	0,336*** (0,015)	0,211*** (0,009)	0,107*** (0,008)	0,080*** (0,008)
Factores del Nivel Superior												
Oficial									1,789*** (0,083)	2,305*** (0,109)	2,518*** (0,185)	2,845*** (0,25)
ICE (Efecto Par)									0,988*** (0,001)	0,994*** (0,001)	0,995*** (0,001)	0,993*** (0,002)
Áreas (Ref: Zonas rurales)												
Ciudades Interm.									0,538*** (0,045)	0,635*** (0,053)	0,718*** (0,088)	0,778* (0,108)
Ciudades Ppales.									0,889*** (0,034)	0,695*** (0,028)	0,614*** (0,039)	0,486*** (0,037)
Constante	-1,251*** (0,229)	-0,278*** (0,001)	0,465*** (0,039)	0,519*** (0,047)								
Parte Aleatoria:												
Varianza Nivel-2 (σ_u^2)	1,258*** (0,063)	1,912*** (0,084)	3,061*** (0,168)	3,506*** (0,226)	0,639*** (0,041)	0,784*** (0,044)	1,413*** (0,093)	1,721*** (0,13)	0,429*** (0,032)	0,583*** (0,036)	1,188*** (0,083)	1,335*** (0,111)
Estadísticos Derivados												
Función Verosimilitud	-19.895	-21.426	-10.681	-7.473	-18.997	-19.933	-9.626	-6.659	-18.631	-19.528	-9.423	-6.439
ICC (ρ)	27,7 %	36,8 %	48,2 %	51,6 %	16,3 %	19,2 %	30,0 %	34,3 %	11,5 %	15,1 %	26,5 %	28,9 %
Criterio Schuartz	39.811,0	42.872,2	21.381,8	42.872,2	38.078,6	39.949,8	19.330,7	39.949,8	37.389,9	39.182,2	18.964,7	39.182,2
Núm. Observaciones (i)	39.224	35.854	18.283	12.877	39.224	35.854	18.283	12.877	39.224	35.854	18.283	12.877
Núm. Grupos (j)	5.580	5.422	3.920	3.276	5.580	5.422	3.920	3.276	5.580	5.422	3.920	3.276

Nota: OR (Odds Ratio). Errores estándar entre parentesis. Niveles de significancia: * <0.1 ; ** <0.05 ; *** <0.01

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014 - 2017.

Modelo Nulo. Mide el logaritmo de las posibilidades de ser beneficiario del programa SPP para un estudiante de un colegio promedio (sin características distintivas). Aunque no son directamente interpretables, los coeficientes estimados son negativos para los dos primeros periodos del programa y positivos para los posteriores (ver Tabla 5), lo cual podría sugerir una mayor opción del colegio promedio de tener Pilos conforme avanza el programa. Estos resultados permiten afirmar que las variaciones en la propensión en favor de ser beneficiario del programa SPP a nivel de colegio son significativas, indicando que es relevante realizar un estudio multinivel a partir de estos datos. Por su parte, el índice de correlación intraclase (ICC) ha presentado un comportamiento creciente durante el programa, pasando de 28 % en 2014 a 52 % en 2017, lo cual indica que a medida que avanza el programa, en mayor medida las diferencias entre colegios están explicando los distintos Pilos.

Modelo 1. Se obtiene luego de agregar la variable puntaje global al modelo nulo. Esta variable intenta capturar los cambios en la probabilidad en favor de ser beneficiario de programa SPP en relación con la capacidad académica demostrada del estudiante. La variable permaneció insignificante estadísticamente durante los cuatro periodos analizados. Sin embargo, solo en la primera cohorte del programa, la variable logró significancia estadística al incluirse otras covariables al modelo, evidenciándose un aumento en la probabilidad de ser beneficiario del programa conforme aumenta el puntaje global obtenido en la prueba SABER 11 de 2014. Para las cohortes restantes, la variable continuó insignificante estadísticamente. Lo anterior permitiría inferir que el mérito académico no fue una característica diferenciable en la propensión a obtener la ayuda financiera dentro de la población elegible, situación que se hace aún más visible al no genera ningún cambio en el ICC.

Modelo 2. Se obtiene de agregar el *background familiar* al modelo 1. Los resultados insinúan que los estudiantes pertenecientes a hogares de menor capital humano tienen mayores opciones de ser beneficiarios del programa SPP. Así, para el primer periodo del programa, se observa con un alto nivel de significancia estadística que la probabilidad de ser Pilo disminuye gradualmente (26,9 %, 44,4 % y 65 %) a medida que aumenta el máximo nivel educativo del hogar respecto a un hogar donde ninguno de los padres lograron completar la educación media. Para los periodos siguientes el comportamiento ha sido similar, sin embargo, para 2016 los jóvenes cuyos padres tenían como mínimo un título universitario disminuían sus probabilidades en cerca del 79 % frente sus similares de menor capital cultural en el hogar. Pareciera que el entorno familiar explica con cierta amplitud la variación en la elegibilidad del programa SPP en el segundo nivel del modelo, situación que se refleja en la disminución que presenta el ICC al caer 4,3 p.p. durante los cuatro periodos.

Modelo 3. Se logra al incluir el estrato socioeconómico al modelo anterior. El estrato explica la variabilidad de los errores del nivel agregado en mayor proporción, aportando el 7,3 %, 12,2 %, 14,7 % y 14,1 % de esta, para cada uno de los periodos del programa, respectivamente. Los resultados sugieren, a cualquier nivel de significancia, una disminución en la probabilidad de ser becario del programa SPP a medida que mejoran las condiciones económicas del estudiante, hipótesis respaldada quizás por la regla de selección del programa. Para la ultima cohorte del programa, los datos en esta investigación indican que los residentes del estrato 2 tenían 77,2 % menos probabilidades de ser elegidos frente a quienes residen en el estrato 1; mientras que para los

residentes del estrato 3 esta brecha fue 95 % aproximadamente.

Modelo 4. La especificación de este modelo se completa incluyendo la naturaleza del colegio al modelo 3, intentando medir el diferencial en la propensión a ser Pilo según el carácter público o privado de la institución educativa. Por su parte, los resultados sostienen que los estudiantes de instituciones oficiales tenían como mínimo 2,5 veces más opciones de ser favorecidos con el crédito-beca que sus contrapartes de colegios privados. La proporción de variación entre escuelas explicada por la variable naturaleza tomando como referencia el modelo inmediatamente anterior es 3,3 p.p.p. menos para el total de periodos en el estudio.

Modelo 5. Una medida del efecto compañeros de clase (ICE) se adiciona a la especificación del modelo anterior. Esta variable intenta capturar el *efecto par* a partir de la influencia de los compañeros de clase sobre la probabilidad en favor a ser Pilo, para el estudiante i . El modelo predice, manteniendo la restante información constante, una disminución promedio del 1,2 % en la probabilidad a ser beneficiario del programa SPP por cada unidad que se incremente la variable ICE. Vale la pena señalar que aunque la correlación del índice con la obtención del crédito-beca es negativa, este efecto ha ido reduciendo con el paso del tiempo, alcanzando el mínimo en 2016 con 0,5 %. En relación con la influencia del contexto del colegio, al controlar por la variable ICE, no se observa un patrón de comportamiento en la varianza de los residuos, ya que en los primeros dos periodos del programa el ICC disminuye, mientras que aumenta para los dos últimos.

Modelo 6. El modelo final se completa con la inclusión de la variable Área Geográfica al modelo 5. La inclusión de esta variable se justifica en la modificación de los umbrales de elegibilidad según el espacio geográfico, con lo cual esta variable intenta capturar dicho efecto diferencial en la probabilidad de ser beneficiario del programa. Los resultados sugieren que, manteniendo la demás información constante, los estudiantes residentes zonas urbanas (14 áreas principales o ciudades intermedias), tenían en promedio 33 % menos probabilidad de obtener la ayuda que sus similares de zonas rurales, siendo este efecto marginal creciente para los estudiantes de las principales ciudades del país y decreciente para los de las ciudades intermedias. Tomando como referencia el modelo inmediatamente anterior, la proporción de la variación entre colegios explicada por el área geográfica es, en promedio, 2,6 p.p. menor para cada periodo en el estudio, permitiendo atribuirle un 20,5 % de esta varianza a las características del nivel superior que no son observadas en este modelo.

7. Discusión Final

En Colombia, el sistema de educación superior se ha caracterizado por una segregación debido a los altos costos de matrícula y el escaso apoyo financiero para los estudiantes de bajo NSE. Como respuesta a esta problemática, el Ministerio de Educación Nacional lanzó en 2014 el programa Ser Pilo Paga, otorgándole a la población vulnerable y destacada en las pruebas SABER 11, un crédito-beca que cubre los costos de matrícula del programa académico elegido por el beneficiario y además un subsidio de manutención diferencial. En este contexto, se utilizó un modelo lineal jerárquico para abordar la estructura multinivel existente en los datos, encontrando algunos predictores significativos a nivel de estudiante y colegio que se correlacionaron fuertemente con la probabilidad de obtención de la ayuda financiera del programa SPP. Para ello, se consideró sólo

aquellos estudiantes que cumplieron simultáneamente con los criterios de elegibilidad (necesidad y mérito académico) del programa SPP a partir criterios entorno al nivel socioeconómico del hogar y el puntaje global en la prueba SABAER 11, y posteriormente se cuantificó el nivel de asociación las distintas covariables en probabilidad de obtención del crédito-beca mediante la estimación del modelo.

Los hallazgos reunidos en esta investigación sugieren que el gobierno colombiano fue menos receptivo a las características académicas y más sensible a las necesidades de los estudiantes elegibles, al momento de distribuir la ayuda financiera. En ese sentido, la principal variable que relaciona el mérito académico, la capacidad académica demostrada, no configuró un buen predictor de la probabilidad de obtención de la ayuda financiera dentro del grupo de elegibles; mientras que el nivel académico medio del colegio, el efecto par, presentó una asociación inversa con la probabilidad de ser becario del programa SPP. Estos resultados cobran sentido si se tiene en cuenta que los estudiantes analizados se encuentran en el decil superior de la distribución de los puntajes SABER 11, dificultando la posible existencia de diferencias en la capacidad académica entre estos estudiantes. Así mismo, un mayor nivel académico del colegio estaría correlacionado con un mayor nivel socioeconómico de sus estudiantes, lo que estaría explicando la dirección inversa con la probabilidad de obtención de la ayuda.

En relación con las necesidades, los resultados reafirman la importancia de la educación en el contexto familiar, así como la fuerza de la relación entre la condición económicamente del hogar y la educación de los padres. El modelo estaría validando el objetivo principal del programa SPP, el cual estaría dirigido a los estudiantes de escasos recursos económicos, ya que en la medida que aumenta el nivel educativo de los padres y por ende el nivel social, económico y cultural en el hogar, la probabilidad de ser favorecido con la beca disminuye, lo que garantiza una vez más que el criterio de necesidad se privilegia cumpliendo de esta manera con su vocación primaria.

Quizás, el resultado más sobresaliente reportado aquí es la reducción en la brecha entre los estudiantes de bajo NSE, particularmente de las partes más pobres y más lejanas de Colombia, en comparación con los estudiantes de bajo NSE que viven en las 14 ciudades más grandes del país. Este resultado sugiere una reducción en la brecha de desigualdad en el acceso a la educación superior, lo que podría producir una reducción significativa en la disparidad socioeconómica en Colombia en el futuro cercano. Estos resultados estarían en línea con los obtenidos por [Medina et al. \(2018\)](#), quien proporciona evidencia de una reducción de esta brecha entre los estudiantes de SISBEN residentes en las principales ciudades de Colombia y aquellos que habitan en zonas rurales.

Finalmente, dados sus datos y limitaciones metodológicas, este estudio puede mejorarse de varias maneras. Por una parte, la varianza entre colegios que asoma incluso después de controlar por características individuales y escuelas hace evidente la necesidad de enfocar futuras investigaciones a los efectos contextuales que estas presentan sobre la distribución de la ayuda financiera para la educación en Colombia, las cuales podrían estar asociadas a otras variables sociales, económicas y psicológicas no contempladas en las bases de datos del ICFES. Así mismo, sería importante observar los cambios de probabilidad a mayor nivel de análisis, por ejemplo ciudades o departamentos, lo que incrementaría la comprensión de la descomposición de la varianza asociada a cada uno de los grupos

y la identificación de los aportes de cada una de las características, consideradas a la explicación de dicha varianza.

Por otra parte, estudios posteriores deben encaminarse a producir mediciones causales de eficiencia sobre este tipo de intervenciones para comparar los beneficios económicos de los programas que se centran en la población estudiantil y las universidades públicas, y contribuir al debate sobre la elección que hacen los gobiernos entre otorgar recursos a universidades públicas o ayuda financiera basada en la necesidad a los estudiantes.

Referencias

- Agasisti, T. and Murtinu, S. (2016), ‘Grants in italian university: a look at the heterogeneity of their impact on students’ performances’, *Studies in Higher Education* **41**(6), 1106–1132.
- Álvarez, M. J., Castro, C., Corredor, J., Londoño, J., Maldonado, C., Rodríguez, C., Sánchez, F., Pulido, X. et al. (2017), El programa ser pilo paga: impactos iniciales en equidad en el acceso a la educación superior y el desempeño académico, Technical report, Universidad de los Andes-CEDE.
- Angrist, J. D. and Lavy, V. (2002), The effect of high school matriculation awards: Evidence from randomized trials, Technical report, National Bureau of Economic Research.
- Angrist, J. D. and Pischke, J.-S. (2008), *Mostly harmless econometrics: An empiricist’s companion*, Princeton university press.
- Angrist, J., Lang, D. and Oreopoulos, P. (2009), ‘Incentives and services for college achievement: Evidence from a randomized trial’, *American Economic Journal: Applied Economics* **1**(1), 136–63.
- Arango, S. (2017), ‘Externalidades en el desempeño académico de la provisión de incentivos: evidencia del programa ser pilo paga’, *Revista de Economía del Rosario* **20**(2), 175–212.
- Arango-Thomas, L. E., Bonilla, G. and Arango, L. E. (2015), ‘Human capital agglomeration and social returns to education in colombia’, *Borradores de Economía; No. 883* .
- Barnes, G. T. and Neufeld, J. L. (1980), ‘The predictability of college financial aid offers: Evidence from the class of 1972’, *Economic Inquiry* **18**(4), 667–691.
- Barrera-Osorio, F., Bertrand, M., Linden, L. L. and Perez-Calle, F. (2011), ‘Improving the design of conditional transfer programs: Evidence from a randomized education experiment in colombia’, *American Economic Journal: Applied Economics* **3**(2), 167–95.
- Barrera-Osorio, F. and Filmer, D. (2016), ‘Incentivizing schooling for learning: Evidence on the impact of alternative targeting approaches’, *Journal of Human Resources* **51**(2), 461–499.
- Basto Aguirre, N. et al. (2019), The Unexpected Effects of a Merit-Based Scholarship: Evidence from Ser Pilo Paga, PhD thesis, Universidad del Rosario.
- Bernal, G. L. and Penney, J. (2019), ‘Scholarships and student effort: Evidence from colombia’s ser pilo paga program’, *Economics of Education Review* .

- Bettinger, E. (2004), How financial aid affects persistence, *in* 'College choices: The economics of where to go, when to go, and how to pay for it', University of Chicago Press, pp. 207–238.
- Boatman, A. and Long, B. T. (2016), 'Does financial aid impact college student engagement?', *Research in Higher Education* **57**(6), 653–681.
- Boschung, M. D., Sharpe, D. L. and Abdel-Ghany, M. (1998), 'Racial, ethnic, and gender differences in postsecondary financial aid awards', *Economics of Education Review* **17**(2), 219–222.
- Bottia, M., Sosa, L. C. and Medina, C. (2012), 'El sisben como mecanismo de focalización individual del régimen subsidiado en salud en colombia: ventajas y limitaciones', *Revista de Economía del Rosario* **15**(2), 137–177.
- Chen, R. and DesJardins, S. L. (2010), 'Investigating the impact of financial aid on student dropout risks: Racial and ethnic differences', *The Journal of Higher Education* **81**(2), 179–208.
- Cornwell, C., Lee, K. H. and Mustard, D. B. (2003), 'The effects of merit-based financial aid on course enrollment, withdrawal and completion in college'.
- De Wit, H., Jaramillo, I. C., Knight, J. and Gacel-Ávila, J. (2005), *Higher education in Latin America: The international dimension*, The World Bank.
- DNP (2015), 'Departamento nacional de planeación - todos por un país'.
- Doyle, W. R. (2010), 'Changes in institutional aid, 1992–2003: The evolving role of merit aid', *Research in Higher Education* **51**(8), 789–810.
- Duffy, E. A. and Goldberg, I. (2014), *Crafting a class: College admissions and financial aid, 1955-1994*, Vol. 77, Princeton University Press.
- Dynarski, S. and Scott-Clayton, J. (2013), Financial aid policy: Lessons from research, Technical report, National Bureau of Economic Research.
- Epple, D., Romano, R. and Sieg, H. (2003), 'Peer effects, financial aid and selection of students into colleges and universities: An empirical analysis', *Journal of Applied Econometrics* **18**(5), 501–525.
- Fack, G. and Grenet, J. (2015), 'Improving college access and success for low-income students: Evidence from a large need-based grant program', *American Economic Journal: Applied Economics* **7**(2), 1–34.
- Ferreira, M. M. (2017), 'The demand side of the higher education expansion', *Higher Education in Latin America and the Caribbean* **112**(1), 148.
- Goldrick-Rab, S., Kelchen, R., Harris, D. N. and Benson, J. (2016), 'Reducing income inequality in educational attainment: Experimental evidence on the impact of financial aid on college completion', *American Journal of Sociology* **121**(6), 1762–1817.
- Grubb, W. N. and Tuma, J. (1991), 'Who gets student aid?: Variations in access to aid', *The Review of Higher Education* **14**(3), 359–381.

- Heller, D. E. (2001), 'Race, gender, and institutional financial aid awards.', *Journal of Student Financial Aid* **31**(1), 7–24.
- Hox, J. J., Moerbeek, M. and Van de Schoot, R. (2017), *Multilevel analysis: Techniques and applications*, Routledge.
- Kane, J. and Spizman, L. M. (1994), 'Race, financial aid awards and college attendance: Parents and geography matter', *American Journal of Economics and sociology* **53**(1), 85–96.
- Kim, J. (2012), 'Exploring the relationship between state financial aid policy and postsecondary enrollment choices: A focus on income and race differences', *Research in Higher Education* **53**(2), 123–151.
- Laajaj, R., Moya, A. and Sánchez, F. (2018), 'Equality of opportunity and human capital accumulation: Motivational effect of a nationwide scholarship in colombia (igualdad de oportunidades y acumulación de capital humano: Efectos motivacionales de una beca de gran escala en colombia)'.
(2018).
- Leuven, E., Oosterbeek, H. and Van der Klaauw, B. (2010), 'The effect of financial rewards on students' achievement: Evidence from a randomized experiment', *Journal of the European Economic Association* **8**(6), 1243–1265.
- Levitt, S. D., List, J. A., Neckermann, S. and Sadoff, S. (2016), 'The behavioralist goes to school: Leveraging behavioral economics to improve educational performance', *American Economic Journal: Economic Policy* **8**(4), 183–219.
- Londono-Velez, J., Rodriguez, C. and Sánchez, F. (2017), 'The intended and unintended impacts of a merit-based financial aid program for the poor: The case of ser pilo paga', *Documento CEDE* (2017-24).
- McPherson, M. and Schapiro, M. (2006), 'Watch what we do (and not what we say)', *College access: Opportunity or privilege* pp. 50–73.
- Medina, P., Ariza, N., Navas, P., Rojas, F., Parody, G., Valdivia, J. A., Zarama, R. and Penagos, J. F. (2018), 'An unintended effect of financing the university education of the most brilliant and poorest colombian students: The case of the intervention of the ser pilo paga program', *Complexity* **2018**.
- Melguizo, T., Sanchez, F. and Velasco, T. (2016), 'Credit for low-income students and access to and academic performance in higher education in colombia: A regression discontinuity approach', *World development* **80**, 61–77.
- OCDE, B. M. (2016), 'Revisión de políticas nacionales de educación: La educación superior en colombia', *La Educación Superior en Colombia* .
- OECD and Bank, W. (2013), 'Reviews of national policies for education: Tertiary education in colombia 2012', *OECD Publishing* .

- Rabe-Hesketh, S. and Skrondal, A. (2012), *Multilevel and longitudinal modeling using Stata*, STATA press.
- Scott-Clayton, J. (2011), ‘On money and motivation a quasi-experimental analysis of financial incentives for college achievement’, *Journal of Human resources* **46**(3), 614–646.
- Scott-Clayton, J. and Zafar, B. (2019), ‘Financial aid, debt management, and socioeconomic outcomes: Post-college effects of merit-based aid’, *Journal of Public Economics* **170**, 68–82.
- Singell Jr, L. D. (2002), ‘Merit, need, and student self selection: is there discretion in the packaging of aid at a large public university?’, *Economics of Education Review* **21**(5), 445–454.
- Steele, F. (2009), ‘Module 7: Multilevel models for binary responses: Concepts’, *Multilevel models for binary responses. Bristol: Center for Multilevel Modelling* .
- Toutkoushian, R. K. and Shafiq, M. N. (2010), ‘A conceptual analysis of state support for higher education: Appropriations versus need-based financial aid’, *Research in Higher Education* **51**(1), 40.
- Yang, P. (2010), ‘Who gets more financial aid in china? a multilevel analysis’, *International Journal of Educational Development* **30**(6), 560–569.
- Ziderman, A. (2003), *Student loans in Thailand: are they effective, equitable, sustainable?*, UNESCO Bangkok.

8. Anexos

Tabla A.1: Distribución de los evaluados en SABER 11 según criterios de elegibilidad del programa SPP

	Periodo 2014		Periodo 2015		Periodo 2016		Periodo 2017	
	Frec.	Porc.	Frec.	Porc.	Frec.	Porc.	Frec.	Porc.
Panel A: Cumple criterio académico (Puntaje global 9% sup)								
No	491,274	90.64	491,928	91.45	520,259	95.46	478,026	96.41
Si	50,742	9.36	45,970	8.55	24,760	4.54	17,779	3.59
Total	542,016		537,898		545,019		495,805	
Panel B: Cumple criterio académico y elegible de acuerdo al NSE								
No	11,518	22.70	10,116	22.01	6,477	26.16	4,902	27.57
Si	39,224	77.30	35,854	77.99	18,283	73.84	12,877	72.43
Total	50,742		45,970		24,760		17,779	
Panel C: Elegibles y beneficiarios del programa SPP								
No	30,243	77.10	21,673	60.45	8,504	46.51	5,877	45.64
Si	8,981	22.90	14,181	39.55	9,779	53.49	7,000	54.36
Total	39,224		35,854		18,283		12,877	

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014-2017.

Tabla A.2: Matriz de correlaciones de las variables consideradas en el modelo

Covariables	Mujer	NEH	Estrato	Oficial	ICE	Área
Puntaje Global	1					
NEH	0,2022	1				
Estrato	0,0899	0,3509	1			
Oficial	-0,1564	-0,3503	-0,3755	1		
ICE	0,3580	0,3831	0,3758	-0,5500	1	
Área	0,0229	0,0917	0,2651	-0,2021	0,1861	1

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014 - 2017.

Tabla A.3: Estimación modelo logístico de efectos aleatorios: Modelo nulo y extensiones periodo 2014

Parámetro	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
	Coef.	OR	OR	OR	OR	OR	OR
Parte Fija							
Factores del Primer Nivel							
Puntaje Global		1,001 (0,001)	1,003*** (0,001)	1,004*** (0,001)	1,005*** (0,001)	1,006*** (0,001)	1,006*** (0,001)
NCH (Base: Inferior)			0,728*** (0,028)	0,840*** (0,032)	0,873*** (0,033)	0,887*** (0,033)	0,894*** (0,034)
Bachiller			0,550*** (0,024)	0,700*** (0,031)	0,767*** (0,034)	0,797*** (0,035)	0,805*** (0,035)
Técnica			0,342*** (0,015)	0,488*** (0,021)	0,560*** (0,024)	0,596*** (0,026)	0,603*** (0,026)
Superior				0,527*** (0,02)	0,581*** (0,022)	0,617*** (0,023)	0,618*** (0,023)
NSEH (Base: Estrato 1)				0,237*** (0,01)	0,293*** (0,013)	0,335*** (0,015)	0,336*** (0,015)
Estrato 2					2,518*** (0,11)	1,811*** (0,084)	1,789*** (0,083)
Estrato 3						0,987*** (0,001)	0,988*** (0,001)
Factores del Primer Superior							
Oficial							
ICE (Efecto Par)							
Áreas (Ref: Zonas rurales)							
Ciudades Intern.							0,538*** (0,045)
Ciudades Ppales.							0,889*** (0,034)
Constante	-1,251*** (0,229)						
Parte Aleatoria:							
Varianza Nivel-2 (σ_u^2)	1,258*** (0,063)	1,269*** (0,064)	1,010*** (0,055)	0,639*** (0,041)	0,523*** (0,037)	0,449*** (0,033)	0,429*** (0,032)
Estadísticos Derivados							
Función Verosimilitud	-19.895	-19.894	-19.543	-18.997	-18.768	-18.661	-18.631
ICC (ρ)	27,7 %	27,8 %	23,5 %	16,3 %	13,7 %	12,0 %	11,5 %
Criterio Schuartz	39.811,0	39.820,0	39.149,4	38.078,6	37.630,7	37.428,5	37.389,9
Núm. Observaciones (i)	39.224	39.224	39.224	39.224	39.224	39.224	39.224
Núm. Grupos (j)	5.580	5.580	5.580	5.580	5.580	5.580	5.580

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2014.

Tabla A.4: Estimación modelo logístico de efectos aleatorios: Modelo nulo y extensiones periodo 2015

Parámetro	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
	Coef.	OR	OR	OR	OR	OR	OR
Parte Fija							
Factores del Primer Nivel							
Puntaje Global		0,997*** (0,001)	0,999 (0,001)	1 (0,001)	1 (0,001)	1,001* (0,001)	1,001* (0,001)
NCH (Base: Inferior)			0,636*** (0,026)	0,756*** (0,031)	0,803*** (0,033)	0,814*** (0,033)	0,824*** (0,034)
Bachiller			0,434*** (0,02)	0,587*** (0,027)	0,663*** (0,03)	0,680*** (0,031)	0,689*** (0,032)
Técnica			0,267*** (0,012)	0,410*** (0,018)	0,490*** (0,022)	0,513*** (0,023)	0,513*** (0,023)
Superior				0,415*** (0,016)	0,464*** (0,018)	0,477*** (0,018)	0,489*** (0,019)
NSEH (Base: Estrato 1)				0,150*** (0,007)	0,189*** (0,008)	0,201*** (0,009)	0,211*** (0,009)
Estrato 2					2,935*** (0,126)	2,426*** (0,115)	2,305*** (0,109)
Estrato 3						0,993*** (0,001)	0,994*** (0,001)
Factores del Primer Superior							
Oficial							
ICE (Efecto Par)							
Áreas (Ref: Zonas rurales)							
Ciudades Interm.							0,635*** (0,053)
Ciudades Ppales.							0,695*** (0,028)
Constante	-0,278*** (0,001)						
Parte Aleatoria:							
Varianza Nivel-2 (σ_u^2)	1,912*** (0,084)	1,877*** (0,083)	1,500*** (0,069)	0,784*** (0,044)	0,619*** (0,038)	0,609*** (0,037)	0,583*** (0,036)
Estadísticos Derivados							
Función Verosimilitud	-21.426	-21.415	-20.886	-19.933	-19.611	-19.574	-19.528
ICC (ρ)	36,8 %	36,3 %	31,3 %	19,2 %	15,8 %	15,6 %	15,1 %
Criterio Schuartz	42.872,2	42.861,6	41.835,4	39.949,8	39.317,3	39.251,9	39.182,2
Núm. Observaciones (i)	35.854	35.854	35.854	35.854	35.854	35.854	35.854
Núm. Grupos (j)	5.422	5.422	5.422	5.422	5.422	5.422	5.422

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2015.

Tabla A.5: Estimación modelo logístico de efectos aleatorios: Modelo nulo y extensiones periodo 2016

Parámetro	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
	Coef.	OR	OR	OR	OR	OR	OR
Parte Fija							
Factores del Primer Nivel							
Puntaje Global		1,002 (0,001)	0,999 (0,001)	1 (0,001)	1 (0,001)	1,001* (0,001)	1,001* (0,001)
NCH (Base: Inferior)			0,588*** (0,045)	0,745*** (0,057)	0,800*** (0,061)	0,812*** (0,062)	0,808*** (0,061)
Bachiller			0,358*** (0,029)	0,520*** (0,041)	0,600*** (0,048)	0,614*** (0,049)	0,611*** (0,049)
Técnica			0,211*** (0,016)	0,370*** (0,028)	0,458*** (0,035)	0,474*** (0,036)	0,463*** (0,035)
Superior				0,264*** (0,019)	0,300*** (0,021)	0,306*** (0,022)	0,319*** (0,022)
NSEH (Base: Estrato 1)				0,074*** (0,006)	0,095*** (0,007)	0,099*** (0,007)	0,107*** (0,008)
Estrato 2							
Estrato 3							
Factores del Primer Superior							
Oficial					3,130*** (0,201)	2,654*** (0,196)	2,518*** (0,185)
ICE (Efecto Par)						0,994*** (0,001)	0,995*** (0,001)
Áreas (Ref: Zonas rurales)							
Ciudades Intern.							0,718*** (0,088)
Ciudades Ppales.							0,614*** (0,039)
Constante	0,465*** (0,039)						
Parte Aleatoria:							
Varianza Nivel-2 (σ_u^2)	3,061*** (0,168)	3,075*** (0,169)	2,654*** (0,149)	1,413*** (0,093)	1,213*** (0,084)	1,221*** (0,085)	1,188*** (0,083)
Estadísticos Derivados							
Función Verosimilitud	-10.681	-10.680	-10.374	-9.626	-9.463	-9.453	-9.423
ICC (ρ)	48,2 %	48,3 %	44,7 %	30,0 %	26,9 %	27,1 %	26,5 %
Criterio Schuartz	21.381,8	21.389,1	20.807,2	19.330,7	19.013,5	19.003,9	18.964,7
Núm. Observaciones (i)	18.283	18.283	18.283	18.283	18.283	18.283	18.283
Núm. Grupos (j)	3.920	3.920	3.920	3.920	3.920	3.920	3.920

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2016.

Tabla A.6: Estimación modelo logístico de efectos aleatorios: Modelo nulo y extensiones periodo 2017

Parámetro	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
	Coef.	OR	OR	OR	OR	OR	OR
Parte Fija							
Factores del Primer Nivel							
Puntaje Global		1,002 (0,001)	0,999 (0,001)	1 (0,001)	1 (0,001)	1,001* (0,001)	0,999 (0,001)
NCH (Base: Inferior)							
Bachiller			0,647*** (0,064)	0,764*** (0,076)	0,817** (0,08)	0,827* (0,082)	0,831* (0,082)
Técnica			0,460*** (0,045)	0,627*** (0,061)	0,717*** (0,07)	0,726*** (0,071)	0,734*** (0,072)
Superior			0,244*** (0,023)	0,429*** (0,04)	0,546*** (0,051)	0,563*** (0,053)	0,551*** (0,051)
NSEH (Base: Estrato 1)							
Estrato 2				0,227*** (0,022)	0,261*** (0,025)	0,267*** (0,025)	0,282*** (0,027)
Estrato 3				0,052*** (0,005)	0,069*** (0,007)	0,072*** (0,007)	0,080*** (0,008)
Factores del Primer Superior							
Oficial					3,841*** (0,293)	3,071*** (0,273)	2,845*** (0,25)
ICE (Efecto Par)						0,993*** (0,002)	0,993*** (0,002)
Áreas (Ref: Zonas rurales)							
Ciudades Intern.							0,778* (0,108)
Ciudades Ppales.							0,486*** (0,037)
Constante	0,519*** (0,047)						
Parte Aleatoria:							
Varianza Nivel-2 (σ_u^2)	3,506*** (0,226)	3,517*** (0,227)	3,090*** (0,202)	1,721*** (0,13)	1,407*** (0,115)	1,422*** (0,117)	1,335*** (0,111)
Estadísticos Derivados							
Función Verosimilitud	-7.473	-7.472	-7.293	-6.659	-6.496	-6.485	-6.439
ICC (ρ)	51,6 %	51,7 %	48,4 %	34,3 %	30,0 %	30,2 %	28,9 %
Criterio Schuartz	42.872,2	42.861,6	41.835,4	39.949,8	39.317,3	39.251,9	39.182,2
Núm. Observaciones (i)	12.877	12.877	12.877	12.877	12.877	12.877	12.877
Núm. Grupos (j)	3.276	3.276	3.276	3.276	3.276	3.276	3.276

Fuente: Cálculos del autor basados en ICFES - SABER 11, 2017.