

Un análisis para beneficiación del Plan de nivelación académica
Talentos en la Universidad de Valle y sus efectos sobre el
desempeño académico

Marco Antonio Lasso Muñoz

Universidad del Valle
Facultad de Ciencias Sociales y Económicas
Departamento de Economía
Santiago de Cali
2018

Un análisis para beneficiación del Plan de nivelación académica
Talentos en la Universidad de Valle y sus efectos sobre el
desempeño académico

Trabajo de Grado presentado para obtener el título de:
Economista

Marco Antonio Lasso Muñoz

Director:
Jaime Humberto Escobar Martínez

Universidad del Valle
Facultad de Ciencias Sociales y Económicas
Departamento de Economía
Santiago de Cali
2018

Agradecimientos

Agradezco primero que nada a Dios y a mi madre, por siempre estar conmigo; a todos mis familiares, amigos y compañeros, por aportar algo de cada uno en mi formación; y al profesor Jaime Escobar, por haberme guiado, no solo durante la escritura de este documento, sino durante toda mi carrera universitaria.

Tabla de Contenido

	Pág.
Agradecimientos	1
Resumen/Abstract	3
1. Introducción	4
2. Referentes Empíricos y Teóricos	6
2.1 El Método de <i>Propensity Score Matching (PSM)</i>	6
2.2 Aplicaciones del método de <i>PSM</i>	8
2.3 El Desempeño Académico, la Deserción Estudiantil, y sus Determinantes	11
3. Metodología	18
3.1 Aproximación empírica	18
3.2 Base de datos	23
3.3 Características de la muestra	24
4. Resultados	26
4.1 Desempeño Académico y Participación en el Plan Talentos	26
4.2 Probabilidad de Graduarse y Participación en el Plan Talentos	28
4.3 Efectos Marginales de los Modelos	31
4.4 Método de PSM	33
4.4.1 Cálculos del Propensity Score por los Modelos Logit y Probit	33
4.4.2 Elección entre los Modelos de PSM	35
4.4.3 Estimación del Efecto Promedio del Tratamiento por PSM	35
4.4.4 Pruebas de Balanceo del Método de PSM	38
5. Conclusiones	39
Referencias	41
Anexos	44

Resumen

En este trabajo se evalúa el impacto que trajo la participación en el Plan Talentos sobre el desempeño académico de los estudiantes de la Universidad del Valle que ingresaron en el año 2010. En este trabajo se utilizó el método de *propensity score matching* para obtener buenos estimadores de la relación entre la participación en el Plan Talentos y el desempeño académico. Se encontró que la participación en el Plan Talentos tiene una relación positiva con el desempeño académico de los estudiantes universitarios. Varios trabajos han encontrado que la principal causa de la deserción entre los estudiantes universitarios es el bajo desempeño académico. Este trabajo da evidencia de que el Plan Talentos podría ser una herramienta para evitar dicha deserción estudiantil.

Palabras Clave: Desempeño Académico, Plan Talentos, Deserción Estudiantil, Evaluación de Impacto, Propensity Score Matching.

Abstract

This work evaluates the impact of participation in the Plan Talentos on the academic performance of the students of the Universidad del Valle who entered the year 2010. In this work, a propensity score matching method was used to obtain good estimates of the relationship between participation in the Plan Talentos and academic performance. It was found that participation in the Plan Talentos has a positive relationship with students' academic performance. Several studies have found that the main cause of dropout among university students is low academic performance. This work gives evidence that the Plan Talentos could be a relevant approach to avoid such student desertion.

Keywords: Academic Performance, Plan Talentos, Student Dropout, Impact Evaluation, Propensity Score Matching.

1. Introducción

Este trabajo tiene como objetivo evaluar el impacto que tuvo la participación en el Plan de Nivelación Académica Talentos (Plan Talentos) sobre el desempeño académico de los estudiantes de la Universidad del Valle. El Plan Talentos fue un programa de nivelación académica ofrecido por la Alcaldía de Cali y la Universidad del Valle en el año 2009 dirigido a bachilleres menores de 23 años y pertenecientes a familias de los estratos 1 y 2 que no lograron acceder a la educación superior. Este programa seleccionó a los participantes con base en sus puntajes ICFES (importancia de 40%) y sus notas del grado 11 (importancia de 60%) para hacer parte del Plan. A los participantes del Plan se les otorgaron dos semestres nivelatorios gratuitos en la Universidad del Valle en los cuales cursaron las asignaturas obligatorias que presentan en la Prueba ICFES Saber11, siendo estas impartidas por personal docente capacitado de la Universidad. Además, se les otorgó un auxilio económico de 100 mil pesos para asegurar su continuidad y permanencia en el Plan (Alcaldía de Cali, 2009).

El Plan Talentos tuvo como objetivo que los estudiantes adquieran competencias académicas, ciudadanas y habilidades que les garantice un buen desempeño en la Prueba ICFES Saber11, y de esta forma poder acceder a la educación superior. Además, al mejorar el desempeño académico de los participantes, el Plan también pretende mejorar la permanencia en la Universidad de aquellos que hayan logrado acceder a ésta. El Plan Talentos busca que los estudiantes mejoren sus desempeños académicos en las áreas de lenguaje, matemáticas, inglés, ciencias naturales (física, biología y química), ciencias sociales y humanas (historia, geografía, filosofía, cultura ciudadana) y vida universitaria (Alcaldía de Cali, 2009). En este trabajo se busca evaluar el logro de estos objetivos midiendo el impacto que trajo el participar en el Plan Talentos sobre el desempeño académico de los estudiantes que finalmente lograron ingresar a la Universidad del Valle.

Son dos los principales objetivos del Plan Talentos: 1) mejorar el desempeño académico de sus participantes con el objetivo de que estos accedan a la educación superior, y 2) a partir de la mejora de su desempeño académico asegurar su permanencia en la universidad (Alcaldía de Cali, 2009). El estudio de Burbano (2013) ha encontrado que el Plan Talentos tuvo éxito en su meta de mejorar el desempeño académico de sus participantes en las pruebas

de estado ICFES y, de esta forma, mejorar sus probabilidades de acceder a la educación superior. Por otro lado, varios trabajos han encontrado que la principal causa de la deserción estudiantil en la Universidad del Valle está asociada con el bajo desempeño de los estudiantes (Escobar et al. 2008, 2006). A su vez, estos trabajos han encontrado que el bajo desempeño de los estudiantes está ligado a factores socioeconómicos de los estudiantes, como el nivel de ingresos del hogar, la educación de los padres, etc. Este trabajo pretende dar evidencia de como el Plan Talentos tiene un impacto positivo en el desempeño académico de los estudiantes de la Universidad del Valle. Por otro lado, está la mejora en el desempeño académico puede influir en una menor tasa de deserción estudiantil según Escobar et al. (2008, 2006).

Para evaluar el impacto del Plan Talentos sobre el desempeño estudiantil se siguen diferentes métodos. Los métodos de evaluación de impacto presentes en la literatura incluyen la realización de estadísticos descriptivos sencillos como los promedios condicionales, o métodos más sofisticadas como las regresiones por Mínimos Cuadrados Ordinarios (MCO) y el uso de modelos de probabilidad como las regresiones Logísticas o el modelo Probit (Bernal, 2011). Sin embargo, estos métodos pierden utilidad cuando las personas que participan de los programas a evaluar se autoseleccionan para participar. Este problema de autoselección ha llevado al desarrollo de métodos más complejos que permiten obtener estimadores y estadísticos con un menor sesgo relativo a los asociados a los métodos más simples y permiten realizar mejores comparaciones entre el grupo de individuos tratados y no tratados por un proyecto. El problema de autoselección se da porque los individuos que se presentan para participar del programa no es una muestra completamente aleatoria de la población de estudiantes. Los estudiantes que participan de estos planes suelen ser escogidos por criterios éticos o políticos como su nivel socioeconómico, sus bajos desempeños académicos, su pertenecía a una etnia, etc., lo cual implica que se tienen muestras de individuos que no son completamente aleatorias

En este trabajo se utiliza el método de *Propensity Score Matching (PSM)* para evaluar el impacto de participar en el Plan Talentos sobre el desempeño académico de los estudiantes de la Universidad del Valle. El método de PSM busca mejorar las comparaciones entre grupos de tratados (participantes del Plan) y no tratados (no participantes del Plan) para

reducir posibles problemas de autoselección que no pueden ser sorteados por los estimadores MCO, y los otros convencionales.

El trabajo se divide en cinco secciones. Esta es la primera sección y comprende una Introducción al documento de trabajo de grado y la problemática de investigación abordada. La segunda sección contiene una revisión de los principales Referentes Empíricos y Teóricos en la literatura sobre el método de *PSM* y los determinantes del desempeño académico en la educación universitaria. La tercera sección describe los Datos y Metodología a utilizar en este trabajo. La cuarta sección expone los principales Resultados de la investigación y en la quinta sección se desarrollan las Conclusiones.

2. Referentes Empíricos y Teóricos

2.1 El Método de *Propensity Score Matching* (PSM)

Es de gran interés para los investigadores en el área de la educación, el establecimiento de inferencias causales sobre el impacto que pueden traer al desempeño académico la gran variedad de planes de mejora de rendimiento académico (PMRA) que existen. En una visión menos rigurosa, los investigadores buscan evidencia a favor del impacto directo de los PMRA que se ofrecen sobre el aprendizaje, desempeño, tasa de deserción y graduación de los estudiantes. Sin embargo, en la literatura existente se reconoce que el principal inconveniente que deben afrontar los investigadores en este contexto es poder trabajar con el aspecto cuasi-experimental de los datos recopilados (Bernal, 2011; Holland, 1986). La naturaleza cuasi-experimental de los datos implica que las posibilidades de hacer inferencia causal en este contexto son muy restringidas.

En la literatura para la estimación de los impactos causales sobre las variables de interés es de gran importancia que los sujetos de tratamiento con los que se estén trabajando hayan sido seleccionados aleatoriamente. Sin embargo, en el contexto de la evaluación de PMRA la asignación por azar no es posible debido a problemas de autoselección de los individuos o por cuestiones éticas o políticas. Para ello, en la literatura se trabaja con un método de

estimación que busca simular lo mejor posible un diseño experimental verdadero, aun cuando los datos que se tengan no hayan surgido de uno, el método se conoce como *Propensity Score Matching* (Luellen, Shadish, & Clark, 2005).

El método de PSM busca controlar la autoselección de los individuos y como primer paso se tiene el cálculo de los *Propensity Score*. Un *Propensity Score* es un puntaje balanceado que se calcula para crear un grupo de comparación coincidente que es similar al grupo de intervención en un conjunto de características base (Abadie Imbens, 2016; Rosenbaum y Rubin, 1983). En otras palabras, los *Propensity Score* son la probabilidad de exposición a un tratamiento condicional en unas variables observadas (Joffe & Rosenbaum, 1999). El *Propensity Score* puede considerarse como la combinación de múltiples factores que le permiten al investigador comparar individuos que participaron de un programa de intervención con individuos similares a este que no participaron del programa.

En términos más formales, los *Propensity Score* son obtenidos a partir de una regresión logística y se corresponden con la probabilidad condicional predicha de que un estudiante haga parte del programa de intervención. La probabilidad de participación de los estudiantes es representada por un único puntaje calculado asociado con la probabilidad de participar en una intervención, dado un conjunto de variables. Condicional en el conjunto de variables, los estudiantes que participaron y no participaron en la intervención tienen distribuciones de probabilidad idéntica. Por ello, la elección del conjunto de variables en las que se condicionará la probabilidad de participación es de suma importancia para conseguir que las comparaciones entre estudiantes que participaron y no participaron sea la mejor posible y no contenga sesgo de autoselección.

Al tener individuos con distribuciones de probabilidad condicionales iguales en relación con la probabilidad de participar o no en el programa, se obtienen resultados que en promedio son similares a los que se obtendrían si se realizara la elección de los estudiantes al azar (Austin, 2011). Esto es posible dado que al condicionar en el conjunto de variables que se tienen de los estudiantes se obtienen resultados balanceados entre los grupos de tratamiento tal como se conseguiría si se obtuvieran estudiantes para la intervención a partir de un proceso completamente aleatorio.

La principal diferencia, como señala Austin (2011), entre PSM y la asignación aleatoria radica en que existen variables “no observadas” que no pueden ser tenidas en cuenta por el método de PSM por su naturaleza (ejemplo, no existe forma de medirlas). Esto último no quiere decir que el método de PSM no funcione porque existen variables que no son controladas. El método de PSM permite a los investigadores controlar el sesgo sistemático de las variables observadas que son medidas para hacer una estimación más precisa de los efectos del tratamiento (Rosenbaum y Rubin, 1983). Claramente, en caso de conocerse todos los factores relacionados con la autoselección de los estudiantes, el sesgo asociado podría llevarse a cero.

2.2 Aplicaciones del método de PSM

Dada su gran versatilidad y su funcionalidad el método de PSM ha ganado una gran fama entre los investigadores de una gran variedad de campos del conocimiento. Los campos en los que se suele aplicar PSM obedecen tanto a las Ciencias Sociales como a la Medicina, la Ingeniería, la Biología y otras disciplinas. Esto implica que existe una gigantesca literatura con relación a las aplicaciones del método de PSM. Para esta revisión tomaremos unos trabajos que representan bien la generalidad de las aplicaciones del método PSM, su alcance y limitaciones.

Por el lado de la medicina se destaca el trabajo de McIntire, Nelson, Macy, Seo, y Kolbe (2015). McIntire et. al. (2015) busca evidenciar como la exposición al humo de segunda mano (EHS) es responsable de numerosas enfermedades de los pulmones y otros sistemas corporales entre los niños. El estudio se centra en determinar una medida robusta de la correlación entre la EHS y la susceptibilidad al tabaquismo entre los adolescentes estadounidenses que no fuman. Este estudio utilizó datos de la Encuesta Nacional de Tabaco para Jóvenes de 2009 para Estados Unidos. Se utilizó el método de PSM para poder comparar adolescentes no fumadores que estuvieron expuestos a la EHS con los que no lo estuvieron, condicional en las características que hacen susceptibles a las personas de estar en condición de EHS. Entre los factores que se consideran influyen en estar en condición de EHS están género, raza/etnia, ingreso personal, normas al interior del hogar con relación a la ausencia de humo de cigarrillo, número de amigos que fuman, percepción del daño de EHS, beneficios percibidos del tabaquismo y exposición a mensajes de medios pro tabaco. Los resultados

mostraron una asociación significativa entre la exposición EHS y la susceptibilidad al tabaquismo entre adolescentes no fumadores.

Unos de los aspectos más atractivos del método de PSM es la notable reducción en la autoselección que resultan al momento de la recolección de los datos en un trabajo de investigación. Gundersen (2016) busca explotar esta ventaja del método de PSM para corregir los problemas que se presentan en las ecuaciones mincerianas tradicionales asociadas al rendimiento de la educación (Mincer, 1974). La ecuación de Mincer es estimada por MCO y busca encontrar una relación entre los salarios devengados por los individuos y su nivel de educación condicional en algunas características que determinan los salarios. Gundersen (2016) utiliza PSM para medir los beneficios de la educación de los trabajadores de Ghana entre 2004 y 2006, con el fin de corregir el posible sesgo por autoselección de MCO que resulta de no poder comparar grupos de individuos similares en el nivel de educación y otras características relevantes.

Gundersen (2016) encuentra que con el método de PSM la educación primaria no implica diferencias en los salarios de los individuos, los rendimientos de la educación intermedia son bastante altos pero los asociados a la educación secundaria no son muy altos, y las estimaciones puntuales bajo PSM son similares a las obtenidas bajo MCO. Estos resultados muestran que, si bien puede haber problemas de sesgo en los estimadores de MCO, estos no son muy diferentes de los resultados que surgen de corregir por el método de PSM. Por otro lado, se encuentra que para los trabajadores de Ghana es muy importante culminar sus estudios de la educación intermedia para obtener buenos salarios, pero estos no incrementarían de manera sustancial si prosiguen con sus estudios para culminar la escuela secundaria.

El método de PSM es muy popular entre los investigadores que evalúan el impacto de los diversos programas gubernamentales a los que un diseño experimental no sería política ni éticamente correcto. Dado que en muchos programas de apoyo gubernamental no es posible efectuar una escogencia al azar de los individuos se tienen que utilizar herramientas cuasi experimentales como el PSM para medir el impacto de estos programas. Sin embargo, existen trabajos como el de Díaz y Handa (2006) que tienen a disposición un programa gubernamental con diseño experimental que les permitirá evaluar qué tan sesgado resultan

los estimadores de PSM con relación al método experimental. Díaz y Handa (2006) aprovechan la disponibilidad de datos experimentales de la segunda fase del programa mexicano de lucha contra la pobreza PROGRESA (Programa de Educación, Salud y Alimentación). El programa en su segunda fase seleccionó aleatoriamente distintos hogares de distintas ciudades mexicanas con el fin de evaluar el impacto del programa, es decir, se dio paso a un diseño experimental en el programa gubernamental.

A partir de estos datos, Díaz y Handa (2006) buscan estimar el sesgo entre los estimadores de PSM y los del diseño experimental. Los resultados muestran que para datos de corte transversal existe un sesgo significativo para los resultados de los estimadores de PSM con relación a los experimentales. Por otro lado, se encuentra que los resultados de PSM mejoran cuando se utilizan instrumentos tomados de otras encuestas para completar el conjunto de variables observables. El sesgo en los estimadores no es indicativo de que el método de PSM no sea útil. Díaz y Handa (2006) indican que la evaluación de programas como PROGRESA utilizando PSM debe hacerse con cautela, con especial cuidado para garantizar la coherencia de los instrumentos de la encuesta y para justificar la suposición de la selección sobre observables.

Burga (2003) busca evaluar el impacto del programa de capacitación laboral juvenil PROJoven sobre las horas trabajadas a la semana y los ingresos laborales de esta población mediante el método de PSM. Entre los principales resultados de Burga (2003) se encuentra que el programa impacta de manera positiva los ingresos laborales mensuales, sin embargo, cuando se estima el efecto sobre los ingresos por hora no se muestra un impacto significativo. Por otro lado, Burga (2003) señala que el programa tiene un impacto positivo en las horas de trabajo a la semana, es decir, el programa también tiene efectos positivos en cuanto al subempleo por horas trabajadas ya que los beneficiarios terminan trabajando más horas en promedio que los individuos de control.

Para Colombia varios trabajos que aplican PSM se han enfocado en el tema de los cambios en la valorización de las viviendas en respuesta a los grandes proyectos que se llevan a cabo en las principales ciudades del país. Perdomo et al. (2007) utilizan el método de PSM para evaluar el impacto que ha tenido la construcción del sistema de transporte TransMilenio sobre el valor de las propiedades residenciales y comerciales en Bogotá. Los autores encuentran

que las viviendas ubicadas en zonas cercanas a las áreas correspondientes a TransMilenio se valorizan entre un 5,8% y un 17% más con relación a las que no están zonas cercanas al proyecto de TransMilenio.

En un estudio más reciente, Perdomo (2010) buscó evaluar cómo se comportan los métodos de PSM y de precio hedónico espacial (PHE) con relación al problema abordado por Perdomo et. al. (2007). El método de PHE consiste en la estimación (por ejemplo, MCO) del precio de las viviendas que refleja el valor de los atributos ambientales locales (entre estos la cercanía a las zonas asociadas al TransMilenio). Perdomo (2010) encuentra que con el PHE los predios se desvalorizan aproximadamente \$47 en promedio por cada metro que se distancian del sistema de TransMilenio. Con el método de PSM encuentran que esta desvalorización oscila entre aproximadamente \$46 y \$37.

En esta línea de investigación, Galvis y Carrillo (2012) buscan encontrar un índice de precios de las viviendas en el espacio que permita realizar comparaciones entre distintas regiones del país. De la mano del método de PSM, Galvis y Carrillo (2012) buscan emparejar viviendas base de la ciudad de Bogotá con viviendas estadísticamente similares en otras ciudades para de esta forma evaluar el diferencial de precios entre las viviendas de las regiones. Por otro lado, los autores también buscan evaluar si existe heterogeneidad en este diferencial de precios con relación a los distintos rangos de precios de las viviendas y sus características. Galvis y Carrillo (2012) encuentran que la ciudad con el precio de vivienda estándar más alto es Bogotá seguida de Cartagena y Villavicencio. La ciudad con el precio de vivienda estándar más bajo es Manizales seguida de Pasto y Pereira. Galvis y Carrillo (2012) encuentran que en promedio el diferencial de precios entre las regiones alcanza cifras cercanas al 30%, con una notable heterogeneidad asociada a las diferentes clases de viviendas y a los distintos rangos de precios.

2.3 El Desempeño Académico, la Deserción Estudiantil, y sus Determinantes

En esta sección abordaremos algunos artículos que tratan sobre los posibles determinantes del desempeño académico y la tasa de deserción de los estudiantes universitarios. Varios trabajos se aproximan al concepto de desempeño académico de distintas formas. En este

trabajo seguiremos el concepto de desempeño académico aproximado a partir de la variable instrumental de esfuerzo académico desarrollada por Escobar et al. (2008, 2006). El esfuerzo académico se obtiene como el promedio de las notas ponderado por los créditos de las materias, esta variable será discutida más extensamente en la sección de datos. El método de evaluación de impacto que utilizaremos en este trabajo es PSM¹. Algunos artículos de la literatura hacen uso del método de PSM mientras otros se quedan únicamente en establecer relaciones estadísticas descriptivas o estimar regresiones por MCO para obtener los posibles impactos que traen los determinantes sobre las variables de interés.

En el ámbito internacional se destacan los trabajos de evaluación de los Programas de Honor en Estados Unidos (Keller y Lacy, 2013; Slavin, Coladarci, y Pratt, 2008; Shushok, 2006; Cosgrove, 2004) y la evaluación del impacto de las escuelas con seguimiento académico para las escuelas de Hamburgo en Alemania (Guill, Lüdtke y Koller, 2017). También se destaca el trabajo de Porto y Di Gresia (2004) sobre el estudio del desempeño académico de los estudiantes universitarios para Argentina.

Los Programas de Honor (*Honors Programs*) en Estados Unidos son planes de enseñanza avanzada al interior de las universidades públicas y privadas. Estos planes avanzados incluyen estructuras curriculares suplementarias o alternativas, accesos especiales y otro tipo de privilegios que fomentan un mejor desarrollo del estudiante en el ámbito académico y personal (Keller y Lacy, 2013). Esta característica de los Programas de Honor ha motivado los trabajos de varios autores con relación a la determinación del impacto que traen estos programas en el desempeño de los estudiantes, sus tasas de deserción y de graduación.

Cosgrove (2004) busca determinar las tasas de graduación y el desempeño estudiantil de los estudiantes con relación a los Programas de Honor. Utilizó datos de los estudiantes de primer año de 1997 pertenecientes al Sistema Estatal de Educación Superior de Pensilvania. El estudio abarca un período de cinco años, desde el semestre de otoño de 1997 hasta el semestre de primavera de 2002, periodo en el cual se cumple el ciclo de cinco años de estudios. Cosgrove (2004) distinguió tres grupos de estudiantes: los estudiantes que completaron el Programa de Honores, los estudiantes que completaron parcialmente el

¹ En la sección de metodología se dará una explicación de cómo se abordará el problema de investigación.

Programa de Honores y los estudiantes que no participaron en los honores. Cosgrove (2004) calculo las tasas de graduación, tiempo para la graduación y las notas promedio al graduarse GPA (*Grade Point Average*, promedio de notas obtenidas por los estudiantes al momento de graduarse del *High School* o de la Universidad, variable proxy del desempeño académico). Entre sus resultados se encuentra que la tasa de graduación de cinco años para los que completaron el programa era del 100%², la tasa era del 82% para los estudiantes con honores parciales y del 76% para los no honores. Los GPA para los tres grupos fueron 3,71, 3,48 y 3,36, respectivamente.

Cosgrove (2004) encuentra que los estudiantes que terminaron el Programa de Honores tienen el más alto desempeño académico, tasas de graduación más altas, y un tiempo más corto para completar el grado, en comparación con otros estudiantes de alta habilidad, incluyendo estudiantes que participaron parcialmente del Programa de Honor. Por otro lado, la exposición parcial al Programa de Honor no mejora significativamente el desempeño académico, la tasa de graduación o el tiempo que a los estudiantes les toma completar el ciclo universitario y graduarse, con relación a sus pares que no participaron del Programa de Honor y que tienen altas habilidades.

Shushok (2006) recopiló datos para estudiantes de la Universidad de Maryland (College Park) para evaluar su desempeño académico y las tasas de deserción de los estudiantes. Posteriormente, aplicó el método de PSM para emparejar estudiantes con características similares y formar dos grupos de igual tamaño, un grupo se conforma de estudiantes que participaron en el Programa de Honor y el otro grupo de estudiantes que no lo hicieron. Shushok (2006) determinó que los estudiantes que participaron en Programas de Honor tenían un GPA significativamente más alto en comparación con los que no participaron en el Programa de Honor (honores 3.41 y no-honores 3.18). Además, estos también tenían menores tasas de deserción de un año (honores 3% y no honores 10%). Sin embargo, esta última diferencia perdía significancia estadística para el último año académico (cuarto año).

Slavin, Coladarci, y Pratt (2008) utilizaron datos de estudiantes de la Universidad de Maine identificando dos grupos comparables, uno para estudiantes de Programas de Honor

² Participar en el Programa de Honores implica que la tasa de graduación para los que terminaron fue 100% ya que un estudiante debe haber permanecido en la escuela para completar el Programa de Honor.

y otro para estudiantes de alta habilidad que no participaron de los Programas de Honor. Los criterios para definir que dos estudiantes de distintos grupos son comparables son los puntajes de SAT (examen de admisión universitaria estandarizado) y el rango de la escuela secundaria de la que se graduó. A partir de una regresión logística³ determinaron que las tasas de deserción de un año fueron significativamente más bajas para los estudiantes de honor que para los estudiantes no honor. Sus cálculos arrojan que un estudiante de honores tenía 3.1 veces más probabilidades de mantenerse como estudiante en la Universidad después de un año que un estudiante no honor, es decir, desertaban menos. Además, encontraron una tasa de graduación de cuatro años de 64% para los estudiantes de honores y 60% para los estudiantes no honor. Sin embargo, esta diferencia no es estadísticamente significativa.

Keller y Lacy (2013) utilizan datos de los estudiantes que ingresaron a la Universidad Estatal de Colorado (CSU) entre el otoño de 2005 y el otoño de 2008, y los resultados académicos registrados hasta el verano de 2012. Keller y Lacy (2013) utilizan el método de PSM para estimar el impacto que tiene la participación en el Programa de Honor de la universidad condicional en un conjunto de características de los estudiantes. Este conjunto de características comprende los logros académicos en la escuela secundaria, estatus étnico, sexo, una dummy para el nacimiento en el Estado de Colorado, y la unidad académica a la que entra. El logro académico fue medido usando el índice de admisión a la universidad (SAT). El estatus étnico es una dummy que toma el valor de 1 si es Blanco/americano y 0 en cualquier otro estatus, excluyendo a los estudiantes extranjeros del análisis. La unidad académica se refiere a la escuela a la cual el estudiante ingresa (Agricultura, Ingeniería, Ciencias Naturales, Artes Liberales, etc.) enumerados para el estudiante durante el primer semestre.

Entre los principales resultados de Keller y Lacy (2013) se encuentra que el Programa de Honor tiene un impacto positivo en la tasa graduación y reduce la tasa de deserción de los estudiantes. Los autores encontraron que las tasas de deserción para todos los años eran más bajas y la tasa de graduación eran mayores para un estudiante que participó en un Programa de Honor con relación a un estudiante de características similares que no participo del

³ Este método consiste en la estimación de modelo logit para la probabilidad de retención condicional en las características señaladas. Este método es una alternativa a PSM.

programa. Por ejemplo, la tasa de graduación de cinco años para los estudiantes de honores fue del 81,9% frente al 69,6% que se esperaba entre estos estudiantes si no hubieran participado en el programa.

Los trabajos encontrados en la literatura sobre el impacto de los Programas de Honor generalmente encuentran que estos tienen un impacto positivo sobre el desempeño académico de los estudiantes. Sin embargo, en el trabajo de Harris (2015), para un conjunto de estudiantes matriculados en las universidades públicas de tamaño mediano en la región del Atlántico Medio de los Estados Unidos, la estimación del impacto de los Programas de Honor mediante el método de PSM arrojó que estos no tienen un impacto significativo sobre el desempeño de los estudiantes. El trabajo de Harris (2015) mantiene abierto el debate sobre el impacto de los Programas de Honor en el desempeño académico, la deserción estudiantil y las tasas de graduación.

Otro trabajo en el campo del desempeño estudiantil que hace uso del método de PSM es el de Guill, Lüdtke y Koller (2017). Guill et al. (2017) utilizaron datos de estudiantes de las escuelas de Hamburgo en Alemania para comprobar cómo se diferenciaba el desarrollo cognitivo de los estudiantes que participaron en escuelas con *seguimiento académico* (*academic tracking*) con relación a estudiantes que participaron en escuelas que no tenían este formato de enseñanza. En Alemania las escuelas con seguimiento académico son aquellas que separan a los alumnos por la capacidad académica en grupos para todas las materias en su plan de estudios a nivel de la escuela. Para ello, toda la población escolar se asigna a las clases según si el rendimiento general de los estudiantes está por encima del promedio, normal o por debajo del promedio. Los estudiantes asisten a clases académicas solamente con estudiantes cuyo logro académico en general es el mismo que el suyo (Gamoran, 1992).

Guill et al. (2017) utilizan como variable dependiente un indicador de puntuación de inteligencia desarrollado por Lehmann, Gänsfuß y Peek (2011) en un estudio de rendimiento de las escuelas de Hamburgo. Guill et al. (2017) utilizan el método de PSM para comparar entre estudiantes con características similares. Como variables de control utilizan la capacidad de lectura de los estudiantes, el nivel de lenguaje (gramática y vocabulario) y ortografía, y la puntuación de matemáticas de la prueba base KS HAM 4/5 en el grado 5

(Lehmann, R. H., Gänßfuß, R., y Peek, 2011). Adicionalmente, se utiliza una medida del grado de intensidad con la que el estudiante realizó sus estudios primarios (el número de materias cursadas y el nivel de dificultad asociada a estas). Encuentran que, tras 4 años de seguimiento, los estudiantes de escuelas con seguimiento académico desarrollaron puntuaciones de inteligencia significativamente más altas que aquellos que no participaron de estas escuelas. Este resultado apoya la hipótesis de partida de los autores que indica que las escuelas con seguimiento académico permiten un mejor desarrollo cognitivo de los estudiantes.

En Latinoamérica se identificaron varios trabajos en esta línea de investigación sobre los determinantes del desempeño estudiantil. Porto y Di Gresia (2004) buscan encontrar los determinantes del desempeño académico de los estudiantes universitarios. Para ello, realizan una encuesta a estudiantes de la Universidad Nacional de La Plata en Argentina. Los autores definen desempeño académico estudiantil como una combinación de la cantidad de materias aprobadas por año y la nota promedio. Los autores realizan estimaciones del impacto de los determinantes sobre la variable desempeño a partir de una regresión por mínimos cuadrados ordinarios, el método de PSM no es utilizado por ellos en su trabajo.

Porto y Di Gresia (2004) encontraron que el desempeño depende de la carrera que cursan (algunas carreras ofrecen niveles de dificultad distintos a otras), el sexo (las mujeres tienen un rendimiento estudiantil promedio más alto), la edad (los estudiantes más jóvenes tienen mejor rendimiento), la educación de los padres, el tipo de escuela secundaria y la veracidad de las normativas que mantienen las facultades sobre su alumnado.

Para Colombia se destacan los trabajos realizados para la Universidad del Valle llevados a cabo por Escobar, Largo y Pérez (2008, 2006) y Burbano (2013). Escobar et al. (2008, 2006) utilizaron información de las características socioeconómicas de los estudiantes que ingresaron a los diferentes programas académicos durante el año 2001 en la Universidad del Valle. Además, utilizaron las calificaciones para todos los cursos que estos estudiantes matricularon hasta el segundo semestre de 2006 junto con información socioeconómica de los docentes que dictaron esos cursos. Encuentran una gran variedad de determinantes del desempeño estudiantil que está asociado con la caída en las tasas de deserción y las mejoras en las tasas de graduación. Entre estos se destaca que el nivel de formación de los docentes

no es un determinante significativo en el desempeño estudiantil. El tamaño del curso y la nota promedio de los estudiantes del curso es una variable que tiene un impacto positivo sobre el desempeño estudiantil. Esto refleja como el ambiente en el aula de clase influye en el resultado de los estudiantes. Las variables que tienen un impacto negativo sobre el desempeño estudiantil son ser menor de edad, residir en la zona de ladera de la ciudad o en el distrito Especial de Aguablanca, ser hombre y trabajar al momento de ingresar a la Universidad. Los autores encuentran que muchos de estos resultados son válidos incluso cuando se realizan las estimaciones a nivel de facultades exclusivamente.

En la Universidad del Valle, el Acuerdo 009⁴ dictamina que los estudiantes que incurran en bajo desempeño académico en más de dos ocasiones son retirados de la institución. Por bajo desempeño académico, el Acuerdo 009 estipula: 1) perder el 50% o más de los créditos matriculados en un semestre, y 2) perder dos veces una asignatura académica. Por ello, incurrir en bajo desempeño académico puede ser un causante de deserción estudiantil. Escobar et al. (2008, 2006) encuentran que el principal determinante de las altas tasas de deserción estudiantil son el bajo desempeño académico, y con menos importancia otros factores como las características socioeconómicas del hogar, o características individuales. Por ello, los autores proponen que buscar herramientas para mejorar el desempeño académico de los estudiantes es el objetivo principal para reducir la alta deserción estudiantil.

Burbano (2013) utilizó datos de 577 egresados del Plan Talentos de la cohorte 2010 para estimar la probabilidad de ingreso a la educación superior. Este trabajo encontró que el Plan Talentos tuvo un impacto positivo y significativo sobre la probabilidad de ingreso a la educación superior, controlando por características sociodemográficas de los estudiantes y de los hogares. A diferencia de Burbano (2013), en este trabajo se busca evaluar como el Plan Talentos influye en el desempeño de aquellos estudiantes que lograron acceder a la educación superior.

El aporte de este trabajo es la obtención de mejores estimadores del desempeño académico de los estudiantes universitarios a partir del método de PSM. Esto permite realizar mejores

⁴ El acuerdo que dictamina el reglamento estudiantil donde están estipuladas las normas que deben cumplir estudiantes y profesores miembros de la Universidad del Valle.

comparaciones entre grupos de individuos de las que se pueden realizar bajo otros métodos como MCO que fueron intensivamente usados en trabajos anteriores.

3. Metodología

3.1 Aproximación empírica

En este trabajo se siguen varios métodos para evaluar la relación entre el desempeño académico de los estudiantes y su participación en el Plan Talentos, igualmente, para la relación entre la probabilidad de graduación y la participación en el Plan Talentos. En primer lugar, se estima la relación entre el desempeño académico y la participación en el Plan Talentos con el método de MCO. En segundo lugar, se realiza la estimación de los modelos MCO, Logit y Probit para la relación entre la probabilidad de graduación y la participación en el Plan Talentos. En tercer lugar, se utiliza el método de PSM para obtener mejores comparaciones sobre el desempeño académico entre los grupos de tratados y no tratados por el Plan Talentos. Todas las estimaciones y cálculos son hechos a partir del software estadístico STATA.

La aplicación del método de PSM para medir el impacto del Plan Talentos del año 2010 sobre el desempeño académico de los estudiantes de la Universidad del Valle implica identificar y separar el cambio del desempeño académico debido al programa (el efecto causal), de aquel que de todas maneras se hubiera dado sin la ejecución del programa, por cambios exógenos en el entorno (Bernal, 2011). También se realizará un ejercicio para evaluar el impacto del Plan Talentos sobre la probabilidad de graduación, aunque, cabe aclarar que este no es uno de los objetivos principales del Plan Talentos como si lo es la reducir la tasa de deserción (Alcaldía de Cali, 2009).

La aplicación del método de PSM consiste en 3 pasos. El primer paso consiste en la estimación de los *propensity score* bajo los modelos Logit o Probit para poder definir el emparejamiento entre los grupos de comparación de tratados y no tratados. En el segundo paso se aplican criterios de elección de modelos para definir cuales *propensity score* se utilizan, los de Logit o Probit, y definir el conjunto de individuos tratados y no tratados que

serán emparejados por sus *propensity score*. En el tercer paso se estima el efecto promedio del tratamiento evaluando la diferencia entre el valor promedio de la variable objetivo entre los individuos tratados y los no tratados que fueron emparejados en el paso 2.

Los resultados de la aplicación del método de PSM pueden ser comparados con los que se derivan de la aplicación de los modelos de MCO para esfuerzo académico, y de probabilidad lineal, Logit y Probit para la probabilidad de graduación. Estas regresiones preliminares serán presentadas primero en la sección de resultados puesto que se conoce que estos métodos no permiten hacer buenas comparaciones entre grupos de tratamiento y por ende su interés en este trabajo sólo radica en su desempeño en comparación con PSM. La siguiente exposición presenta la variable esfuerzo académico como dependiente. Esta presentación aplica de igual forma en el caso en que tuviéramos la probabilidad de graduación como variable dependiente.

La elección del conjunto de variables de control en las que se condicionará la probabilidad de participación es de suma importancia para conseguir que las comparaciones entre estudiantes que participaron y no participaron sea la mejor posible y no contenga sesgo de autoselección. Con base en los resultados de Escobar et al. (2008,2006), se escogieron un conjunto amplio de posibles determinantes del desempeño académico como variables de control. Estos determinantes del desempeño académico son un conjunto de variables sociodemográficas de los individuos que son fundamentales al momento de decidir si participar o no en el Plan Talentos.

El esfuerzo académico se mide a partir de la nota promedio ponderada por créditos de las materias vistas por los individuos entre el primer semestre de 2010 y el primer semestre de 2016. La variable no incluye las notas de sus trabajos de grado (este aspecto es medido por la variable *grado*) y tampoco incluye aquellas materias cuyas notas no son medidas numéricamente. En el caso de que le individuo habilitara la materia se colocó la nota mayor entre la asignada en la habilitación y la nota del curso como exige el Acuerdo 009.

El objetivo de este ejercicio es evaluar cómo el Plan Talentos impacta en el esfuerzo académico de los estudiantes. Esta variable no se incluye en los determinantes de la participación en el Plan Talentos. Dado que los individuos no pueden observar su esfuerzo académico en la Universidad previo a su ingreso, esta variable no hace parte del conjunto de

variables de control. Las variables que se incluyan en las regresiones de PSM para Plan Talentos deben ser aquellas permitan realizar buenas comparaciones entre grupos de tratamiento.

Siguiendo a Bernal (2011), el impacto del programa a nivel de individuo se define como:

$$\alpha_i = esfuerzo_{1i} - esfuerzo_{0i}$$

Donde, *esfuerzo* es una variable que mide el nivel de esfuerzo académico que el estudiante a partir de sus notas académicas de los cursos. *esfuerzo_{1i}* es el resultado potencial de la variable *esfuerzo* si el individuo *i* fue sometido al tratamiento, y *esfuerzo_{0i}* es el resultado potencial si el individuo no fue sometido al tratamiento. Por desgracia, esta fórmula no es operativa, dado que para cada individuo se observa solo uno de los dos estados, lo que se conoce como el problema fundamental de identificación, lo que, a efectos prácticos, es un problema de observaciones perdidas.

Para sortear tal problema, se puede recurrir a medir el impacto esperado sobre los tratados (ATT por sus siglas en inglés, *Average Treatment effect among Treated*):

$$ATT = E[\alpha_i | D = 1] = E(esfuerzo_1 | D = 1) - E(esfuerzo_0 | D = 1) \quad (1)$$

Donde *D* es una variable dummy que toma el valor de 1 cuando el individuo hace parte del programa de tratamiento y cero en caso contrario. Los datos permiten aproximar este impacto promedio por medio de la diferencia entre $E(esfuerzo_1 | D = 1)$ y $E(esfuerzo_0 | D = 0)$, la cual sería igual a:

$$\begin{aligned} E(esfuerzo_1 | D = 1) - E(esfuerzo_0 | D = 0) \\ = ATT + E(esfuerzo_0 | D = 1) - E(esfuerzo_0 | D = 0) \end{aligned} \quad (2)$$

Por lo que, el parámetro *ATT* estaría identificado cuando:

$$E(esfuerzo_0 | D = 1) - E(esfuerzo_0 | D = 0) = 0 \quad (3)$$

Esta condición se cumple cuando la selección de los participantes del programa es aleatoria. Sin embargo, es probable que la selección de individuos responda a ciertos criterios de focalización, esto es, que exista autoselección, de modo que, dicho término no será igual

a cero. Por lo cual, el caso de que no se trate la autoselección, no se obtendrá una estimación confiable del impacto promedio (Bernal, 2011).

Una estrategia de identificación consiste en asumir que, dado un conjunto de variables observables $\mathbf{X}$, las cuales no responden al tratamiento, los resultados potenciales son independientes del tratamiento:

$$esfuerzo(0), esfuerzo(1) \perp D | \mathbf{X} \quad (4)$$

Lo cual implica que la selección está basada solamente en variables observables. Esto implica que, condicional en las características $\mathbf{X}$, los resultados de los no participantes en el programa se aproximan a los que los participantes hubieran obtenido de no haber participado en el programa. De modo que, para individuos con características similares:

$$E(esfuerzo_0 | D = 0) = E(esfuerzo_0 | D = 1) \quad (5)$$

Por lo que,

$$\begin{aligned} E(esfuerzo_1 | D = 1) - E(esfuerzo_0 | D = 0) \\ = ATT + E(esfuerzo_0 | D = 1) - E(esfuerzo_0 | D = 0) = ATT \end{aligned} \quad (6)$$

Con lo anterior surge un reto. En caso de que el conjunto de variables observables tenga un cardinal grande, encontrar individuos con características similares se vuelve un proceso complejo. Tal problema se sortea con la probabilidad de haber sido sujeto al tratamiento (PSM):

$$P(D = 1 | \mathbf{X})$$

$$esfuerzo(0), esfuerzo(1) \perp D | \mathbf{X} \rightarrow esfuerzo(0), esfuerzo(1) \perp D | P(\mathbf{X})$$

Lo que implica que un par de individuos serán similares en términos de sus características si sus probabilidades de ser sujetos al tratamiento son iguales (o muy cercanas), lo que reduce el problema a una sola dimensión.

Dado lo anterior, se sigue la siguiente secuencia para calcular el ATT

1) PSM

Se estima un modelo de probabilidad, Logit o Probit, que relacione la participación en el programa con las variables observables.

2) Vecino más cercano

Cada individuo del grupo de tratamiento se empareja con un individuo del grupo de control, de modo que, sea el más cercano en términos de $P(\mathbf{X})$ y que cumpla con cierta cercanía:

$$J(p_i) = \{j: \varepsilon > |p_i - p_j| = \min_{k \in \{D=0\}} \{|p_i - p_k|\}\}$$

Es decir, a cada individuo del grupo de tratamiento (el grupo i) se empareja con el individuo del grupo de control (el grupo j) más cercano en términos de $P(\mathbf{X})$ y además, que sea lo suficientemente cercano, $\varepsilon > |p_i - p_j|$, para un ε pequeño.

3) Calculo del efecto por individuo

El efecto se estima como se aprecia a continuación:

$$esfuerzo_{1i} - esfuerzo_{0j(i)}$$

Donde $j(i)$ la pareja del grupo de control para el individuo i.

4) ATT

$$ATT = \frac{1}{N_1} \sum_{i \in DD} [esfuerzo_{1i} - esfuerzo_{0j(i)}] \quad (8)$$

N_1 indica el número de parejas encontradas y DD indica el subconjunto del grupo de tratados a los cuales se les encontró pareja.

La anterior metodología puede utilizarse también para estimar el impacto de participar en el Plan Talentos sobre la probabilidad de graduación de los estudiantes. La variable dependiente *esfuerzo* puede ser reemplazada por *grado* sin derivar en grandes cambios en la metodología. *grado* será una variable dummy que tomara el valor de 1 si el estudiante se graduó y cero en caso contrario. El estimador del impacto del programa ATT indicara si la

participación en el tratamiento tiene algún impacto sobre la probabilidad de graduación de los estudiantes.

3.2 Base de datos

Los datos utilizados en este trabajo son suministrados por la Oficina de Registro Académico de la Universidad del Valle. En este trabajo se utilizaron tres bases de datos distintas que serán combinadas para obtener una base única con la que se podrá trabajar el problema de investigación. La base final permite empalmar las tres bases y poder tener identificados los individuos de forma correcta. La primera base de datos permite identificar los individuos que participaron del Plan Talentos en los años 2009 a 2011. La segunda base de datos contiene información sociodemográfica obtenida por la encuesta aplicada a todos los aspirantes a algún programa académico ofrecido por la Universidad del Valle en los años 2010 a 2012.

La tercera base de datos contiene información sobre las asignaturas matriculadas por los estudiantes admitidos, sus notas (tipo de nota, habilitación, etc.), el tipo de materias (obligatorias, electivas complementarias o profesionales, etc.), semestre matriculado o periodo académico, y otra información del historial académico de los estudiantes admitidos a la Universidad del Valle desde el primer semestre de 2010 hasta el primer semestre de 2016. Con esta base se crea la variable que utilizaremos para aproximarnos al desempeño académico del estudiante: el esfuerzo académico. Esta variable es definida como la nota promedio ponderada por créditos de las materias para los individuos entre el primer semestre de 2010 y el primer semestre de 2016 (Escobar et al., 2008, 2006). Las notas académicas capturan la destreza que el estudiante tuvo al momento de ver el curso, y los créditos de la materia capturan las horas que se deben dedicar al curso. Por ello, la nota promedio ponderada por créditos es una buena medida del desempeño académico. Esta variable no incluye las notas de sus trabajos de grado y tampoco incluye aquellas materias cuyas notas no son medidas numéricamente. En el caso de que el individuo habilitara la materia, se colocó la nota mayor entre la asignada en la habilitación y la nota del curso como exige el Acuerdo 009. Finalmente, se consolida todo en la base final que se utilizará en este trabajo.

La base inicialmente tenía disponibilidad de datos para estudiantes de 2010 a 2012. Dado que muchos de los estudiantes que ingresaron en los años 2011 y 2012 aún no se han

graduado para el primer semestre de 2016, se trabaja solamente con los estudiantes admitidos a la Universidad del Valle en el año 2010. En el Anexo 1 se presenta el diccionario de variables. Cabe resaltar que toda la información procesada no permite individualizar dato alguno, es decir, los datos son completamente anónimos.

3.3 Características de la muestra

En la Tabla 0.1 se tiene una tabla de frecuencias para la variable Talentos donde 1 son los individuos que participaron del Plan y cero son los individuos que no participaron del Plan, todos los individuos de la muestra ingresaron a la Universidad del Valle en el año 2010. Se observa que la población total de estudiantes es 2018, de los cuales, el 90.44% no participaron del Plan Talentos y el 9.56% restante si participó del Plan Talentos.

Tabla 0.1. Tabla de Frecuencias de la variable Talentos

Talentos	freq.	%	% acum
0	1825	90.44	90.44
1	193	9.56	100.00
Total	2018	100.00	

Fuente: Elaboración propia.

En la Tabla 0.2 se muestran algunos estadísticos descriptivos de las variables que se utilizan en este trabajo, diferenciados entre aquellos que participaron y no participaron en el Plan Talentos. Se observa que el esfuerzo académico es mayor en promedio entre los individuos que participaron del Plan Talentos con relación a los que no lo hicieron. Este resultado es una motivación a favor de la hipótesis de esta investigación. Por otro lado, se observa que la participación de las mujeres es mayor entre lo que participaron del Plan que entre los que no participaron. Así mismo, los individuos que participan del Plan suelen tener una edad menor. Además, la educación de la madre es, en promedio, menor para los individuos que participan del Plan con relación a los que no participan. La variable ingreso del hogar esta estandarizada para toda la muestra de tal forma que su media es cero y su varianza es uno (véase el Anexo 1). Por lo cual, la interpretación que se le debe dar a los valores promedio de esta variable

por los grupos de tratamiento es que el ingreso promedio de los hogares de los individuos que participan del Plan Talentos es menor que el de los que no participan del Plan.

Tabla 0.2. Estadísticos descriptivos de las variables por grupos de tratados y no tratados

Talentos	0	1
Esfuerzo académico	3.470983	3.743669
sexo	.3994521	.5906736
edad	20.3.014	18.91192
discapa	.2224658	.2124352
hijos	.0460274	.0207254
sisben	.1830137	.3937824
EPS	.6109589	.4093264
serv_publ_todos	.8367123	.8290155
vivienda_arrendada	.2608219	.3316062
soltero	.9627397	.9896373
blanco_mestizo	.8673973	.8134715
afrodescendiente	.0942466	.1658031
indigena	.0317808	.015544
trabaja	.1720548	.0725389
padre_vive	.5331507	.6787565
num_hermanos	1.542466	1.678756
estrato_residencia	2.392877	1.668394
madre_educ	2.42411	1.829016
ingreso_hogar	.0031719	-.0292969

Fuente: Elaboración propia.

4. Resultados

4.1 Desempeño Académico y Participación en el Plan Talentos

En la Tabla 1 se presentan los resultados de estimar el impacto que trae la participación en el Plan Talentos sobre la medida proxy del desempeño académico sugerida por Escobar et al.

(2008, 2006), el esfuerzo académico de los estudiantes en su ciclo universitario, sin incluir controles e incluyendo controles bajo el método de MCO. La columna 1 es una regresión simple entre Esfuerzo académico y Talentos, la columna 2 incluye todos los controles y la columna 3 excluye la variable educación del padre.

Tabla 1. Esfuerzo Académico y la participación en el Plan Talentos por MCO

	(1)	(2)	(3)
	Esfuerzo académico	Esfuerzo académico	Esfuerzo académico
Talentos	0.273*** (0.064)	0.188*** (0.069)	0.187*** (0.068)
sexo		0.287*** (0.039)	0.287*** (0.039)
edad		-0.024*** (0.007)	-0.024*** (0.007)
discapa		-0.085* (0.046)	-0.085* (0.046)
hijos		-0.308** (0.122)	-0.307** (0.122)
sisben		0.005 (0.063)	0.004 (0.062)
CALI		0.041 (0.050)	0.041 (0.050)
EPS		0.081 (0.051)	0.080 (0.051)
serv_publ_todos		-0.046 (0.054)	-0.046 (0.054)
vivienda_arrendada		0.024 (0.043)	0.024 (0.043)
soltero		-0.042 (0.141)	-0.042 (0.141)
blanco_mestizo		-0.041 (0.234)	-0.041 (0.234)
afrodescendiente		-0.125 (0.240)	-0.125 (0.240)
indígena		-0.128 (0.259)	-0.129 (0.259)
trabaja		-0.020 (0.061)	-0.020 (0.061)
padre_vive		0.092** (0.039)	0.093** (0.039)
num_hermanos		0.009	0.009

		(0.015)	(0.015)
estrato_residencia		0.028	0.028
		(0.022)	(0.021)
madre_educ		0.022	0.023
		(0.021)	(0.019)
padre_educ		0.003	
		(0.018)	
ingreso_hogar		-0.015	-0.015
		(0.019)	(0.019)
_cons	3.471***	3.782***	3.786***
	(0.020)	(0.330)	(0.329)
N	2018	1689	1689
r2_a	0.008	0.071	0.071
aic	5041.928	3936.434	3934.461
bic	5053.148	4055.936	4048.531

* p<0.10 ** p<0.05 *** p<0.01

Nota: N es el número de observaciones. r2_a es el estadístico R cuadrado ajustado. Aic es el criterio de información de Akaike. Bic es el criterio de información bayesiano de Schwarz.

Fuente: Elaboración propia.

Los resultados de la Tabla 1 indican que la participación en el Plan Talentos tiene un impacto positivo sobre el esfuerzo académico de los estudiantes de la Universidad del Valle admitidos en el año 2010. El valor promedio de la nota promedio ponderada de los individuos participantes en el Plan Talentos es 0.187 puntos mayor que la de los individuos no participantes, condicional en las variables de control (Columna 3). Entre los factores que influyen positivamente en el esfuerzo académico se encuentran el ser mujer y la condición de que el padre esté vivo. Entre los factores que impactan negativamente en el esfuerzo académico están la edad, la presencia de alguna discapacidad y el tener hijos al momento de ingresar a la universidad.

Las columnas 2 y 3 permiten comparar como se desempeña el modelo estimado cuando se incluye y no se incluye la educación del padre. Murnane, Maynard y Ohls (1981) y Escobar et al. (2008,2006) argumentan que principalmente es la educación de la madre la que más influye en el desempeño académico de los estudiantes. Los criterios de Akaike (aic) y el criterio bayesiano de Schwarz (bic) son criterios de selección de modelos. Para elegir un

modelo se deben comparar los estadísticos aic y bic en ambos y se elige aquel que arroje un valor del estadístico de prueba menor. En la Tabla 1 se puede observar que los criterios de información aic y bic indican que el modelo con mejor ajuste es el de la columna 3 que excluye la educación del padre. Esto implica que para tener un mejor ajuste del modelo, no es necesario incluir educación del padre.

4.2 Probabilidad de Graduarse y Participación en el Plan Talentos

En la Tabla 2 se presentan los resultados de la estimación del impacto que trae la participación en el Plan Talentos sobre la probabilidad de graduarse mediante los métodos de los modelos de probabilidad lineal, Probit y Logit, sin incluir variables de control.

Tabla 2. Probabilidad de graduación y la participación en el Plan Talentos

	(1)	(2)	(3)
	M. Lineal	M. Probit	M. Logit
Talentos	0.052 (0.032)	0.163 (0.102)	0.277 (0.171)
_cons	0.223*** (0.010)	-0.762*** (0.033)	-1.248*** (0.056)
N	2018.000	2018.000	2018.000
r2_a	0.001		
r2_p		0.001	0.001
aic	2222.242	2167.928	2167.928
bic	2233.462	2179.147	2179.147

* p<0.10 ** p<0.05 *** p<0.01

Nota: N es el número de observaciones. r2_a es el estadístico R cuadrado ajustado. r2_p es el pseudo R cuadrado de McFadden. Aic es el criterio de información de Akaike. Bic es el criterio de información bayesiano de Schwarz.

Fuente: Elaboración propia.

Los resultados de la Tabla 2 muestran que, sin controlar por otros factores asociados a la probabilidad de graduarse de la Universidad, la participación en el Plan Talentos no tiene impacto sobre la probabilidad de graduarse de la Universidad del Valle para los estudiantes

admitidos en el año 2010. Este resultado indica que la probabilidad de graduarse no es distinta entre estudiantes que participaron en el Plan Talentos y los que no lo hicieron. Sin embargo, este resultado no es muy relevante dado que no se están comparando estudiantes con características similares. Para ello, se deben incluir controles de forma que se reduzca el sesgo por omisión de variables relevantes.

En la Tabla 3 se presentan los resultados de la estimación del impacto que trae la participación en el Plan Talentos sobre la probabilidad de graduarse, añadiendo variables de control, mediante los métodos de los modelos de probabilidad lineal, Probit y Logit. Al controlar por otros factores asociados a la probabilidad de graduarse se puede esperar que se corrija el sesgo de omisión de variables relevantes que tienen los coeficientes estimados de la Tabla 2.

Tabla 3. Probabilidad de gradarse y la participación en el Plan Talentos incluyendo variables de control

	(1)	(2)	(3)
	M. Lineal	M. Probit	M. Logit
Talentos	0.026 (0.035)	0.085 (0.131)	0.150 (0.221)
esfuerzo_academico	0.172*** (0.013)	1.329*** (0.099)	2.260*** (0.175)
sexo	0.028 (0.020)	0.018 (0.078)	0.037 (0.133)
edad	-0.000 (0.004)	-0.001 (0.016)	-0.000 (0.028)
discapa	-0.003 (0.024)	-0.004 (0.093)	0.000 (0.158)
hijos	0.093 (0.063)	0.623** (0.288)	1.018** (0.515)
sisben	0.022 (0.032)	0.003 (0.130)	0.050 (0.225)
CALI	0.035 (0.026)	0.093 (0.099)	0.177 (0.168)
EPS	0.038 (0.026)	0.091 (0.105)	0.190 (0.182)
serv_publ_todos	0.038 (0.028)	0.122 (0.113)	0.222 (0.195)

vivienda_arrendada	-0.037*	-0.154*	-0.254*
	(0.022)	(0.087)	(0.148)
soltero	0.061	0.386	0.553
	(0.072)	(0.371)	(0.654)
blanco_mestizo	0.096	0.409	0.760
	(0.121)	(0.573)	(1.092)
afrodescendiente	0.089	0.489	0.923
	(0.124)	(0.583)	(1.108)
indigena	0.062	0.486	0.905
	(0.133)	(0.619)	(1.165)
trabaja	-0.138***	-0.760***	-1.365***
	(0.031)	(0.159)	(0.296)
padre_vive	0.040**	0.111	0.208
	(0.020)	(0.080)	(0.137)
num_hermanos	-0.007	-0.035	-0.065
	(0.008)	(0.032)	(0.055)
estrato_residencia	0.020*	0.048	0.092
	(0.011)	(0.043)	(0.073)
madre_educ	0.013	0.041	0.060
	(0.010)	(0.038)	(0.065)
ingreso_hogar	-0.004	-0.038	-0.099
	(0.010)	(0.287)	(0.733)
_cons	-0.659***	-6.852***	-11.663***
	(0.176)	(0.885)	(1.612)
N	1689	1689	1689
r2_a	0.145		
r2_p		0.214	0.211
aic	1691.453	1514.660	1521.495
bic	1810.954	1634.162	1640.996
ll	-823.726	-735.330	-738.747

* p<0.10 ** p<0.05 *** p<0.01

Nota: N es el número de observaciones. r2_a es el estadístico R cuadrado ajustado. r2_p es el pseudo R cuadrado de McFadden. Aic es el criterio de información de Akaike. Bic es el criterio de información bayesiano de Schwarz.

Fuente: Elaboración propia.

Los resultados de la Tabla 3 muestran que la participación en el Plan Talentos no tiene impacto significativo sobre la probabilidad de graduarse para los estudiantes de la

Universidad del Valle admitidos en el año 2010, controlando por un conjunto de factores que impactan en la probabilidad de graduación. Sin embargo, se puede evidenciar que el esfuerzo académico y el tener hijos tienen un impacto positivo sobre la probabilidad de graduarse de la universidad (para los tres modelos). Por otro lado, el trabajar al momento de ingresar a la Universidad o vivir en una casa arrendada impactan negativamente en la probabilidad de graduarse según los modelos Probit y Logit. Para el modelo Lineal las variables con un impacto negativo son vivir en una vivienda arrendada y trabajar al momento de ingresar a la universidad, mientras que el estrato socioeconómico y la condición de que el padre este vivo influyen positivamente.

4.3 Efectos Marginales de los Modelos

En la Tabla 4 se presentan los efectos marginales asociados a los tres modelos para cada una de las variables independientes estimadas.

Tabla 4. Efectos Marginales sobre la probabilidad de graduación

	Lineal			Probit			Logit		
	dy/dx	Std. Err.	P>t	dy/dx	Std. Err.	P>t	dy/dx	Std. Err.	P>t
Talentos	.0265377	.0353617	0.453	.0214839	.0318996	0.501	.0223774	.031903	0.483
esfuerzo_academico	.172156	.0126042	0.000	.3243602	.0200431	0.000	.3267022	.0205709	0.000
sexo	.0281003	.020412	0.169	.0044914	.019055	0.814	.0055405	.0191711	0.773
edad	-.0002134	.0037066	0.954	-6.28e-06	.0038842	0.999	.0000771	.0039885	0.985
discapa	-.0025578	.023551	0.914	-.0006549	.02275	0.977	.0002753	.022902	0.990
hijos	.092499	.0628474	0.141	.151501	.0698554	0.030	.1466564	.0740577	0.048
sisben	.0233572	.032234	0.469	.0025849	.0319998	0.936	.0090379	.0327533	0.783
CALI	.0353278	.0257113	0.170	.0227444	.0241403	0.346	.0256124	.024234	0.291
EPS	.0389894	.0262665	0.138	.0229334	.0256708	0.372	.0282112	.0263643	0.285
serv_publ_todos	.0374743	.0278798	0.179	.0297728	.0275175	0.279	.0320007	.0280685	0.254
vivienda_arrendada	-.037094	.0220062	0.092	-.0372172	.0211471	0.078	-.0364808	.0213627	0.088
soltero	.061592	.0724815	0.396	.0950955	.0901982	0.292	.0808411	.0942446	0.391
blanco_mestizo	.0956513	.1205443	0.428	.0995282	.1402622	0.478	.1073586	.1575174	0.496
afrodescendiente	.0887916	.1236578	0.473	.1192665	.1427169	0.403	.1311317	.1597764	0.412
indigena	.0625725	.1332661	0.639	.1197043	.1514575	0.429	.1294528	.168009	0.441

trabaja	-.1370956	.0314731	0.000	-.184006	.038081	0.000	-.195802	.0419701	0.000
padre_vive	.0395176	.0203248	0.052	.0259063	.0196818	0.188	.0289121	.0198591	0.145
num_hermanos	-.0075135	.0077397	0.332	-.0086579	.0076842	0.260	-.0094592	.0078895	0.231
estrato_residencia	.0191446	.0111712	0.087	.0107982	.0106064	0.309	.0123879	.0107296	0.248
madre_educ	.0113278	.010903	0.299	.0075682	.0103622	0.465	.0063467	.0104933	0.545
ingreso_hogar	-.0044456	.0096866	0.646	-.0102189	.0877437	0.907	-.0162697	.1391102	0.907

Fuente: Elaboración propia.

Los resultados de la Tabla 4 muestran los efectos marginales asociados a los modelos lineal, Probit y Logit. Estos resultados no tienen una diferencia cualitativa con relación a los expuestos en la Tabla 3, por lo cual, solo se interpretará el efecto marginal asociado al esfuerzo académico. Entre los resultados estadísticamente significativos se encuentra que para los modelos Probit y Logit un aumento de un punto en la nota promedio ponderada de los estudiantes conlleva un incremento en la probabilidad de graduarse de 32,4% y 32,6%, respectivamente. Para el modelo Lineal esta probabilidad se calcula en 17,2%. Esto indica que el esfuerzo académico que los estudiantes realizan en sus respectivos cursos o materias tiene un impacto significativo y alto sobre la probabilidad que estos tengan de graduarse.

Sin embargo, como se discutió en la sección de metodología, estas estimaciones pueden tener problemas dado que los grupos de tratamiento y de control pueden no estar bien definidos, implicando que las comparaciones de los efectos estimados sean erróneas. Además, también puede presentarse un problema de autoselección dado que los individuos que participan en el Plan Talentos son seleccionados por un método no experimental donde la muestra resultante no hace parte de un muestreo aleatorio. Esto implica que los individuos que no participaron del Plan Talentos pero que igual estaban en condiciones de hacerlo no son tenidos en cuenta por la muestra final de individuos. Este sesgo y las correctas comparaciones entre grupo de tratamiento y control puede ser corregido aplicando el método de PSM como se verá en la siguiente subsección. El método de PSM busca entre la muestra de individuos no participantes del Plan Talentos aquellos que se parecen bajo ciertas condiciones a los individuos que si participaron para poder realizar comparaciones más confiables entre los individuos y obtener mejores estimadores del impacto del Plan Talentos.

4.4 Método de PSM

4.4.1 Cálculos del *Propensity Score* por los Modelos Logit y Probit

En la Tabla 5 se presentan los resultados de la estimación de los *propensity score* para la variable Talentos mediante el método de PSM con los modelos Probit y Logit. A partir de los *propensity score* calculados se puede evaluar el impacto que trae la participación en el Plan Talentos sobre el esfuerzo académico y la probabilidad de graduarse de los individuos tratados con los no tratados que tienen *propensity score* similares.

Tabla 5. Probabilidad de participación en el Plan Talentos por el método de PSM

	(1)	(2)
Talentos	PSM Probit	PSM Logit
sexo	0.403*** (0.104)	0.760*** (0.191)
edad	-0.094*** (0.024)	-0.160*** (0.044)
discapa	-0.104 (0.126)	-0.179 (0.229)
hijos	0.101 (0.363)	0.171 (0.673)
sisben	0.234 (0.153)	0.422 (0.276)
EPS	-0.206 (0.142)	-0.349 (0.264)
serv_publ_todos	-0.071 (0.145)	-0.156 (0.255)
vivienda_arrendada	0.230** (0.113)	0.452** (0.203)
soltero	0.394 (0.472)	0.661 (0.875)
blanco_mestizo	-0.008 (0.548)	0.250 -1.166
afrodescendiente	0.302 (0.560)	0.808 -1.183
indigena	-0.221 (0.656)	-0.075 -1.336
trabaja	-0.279 (0.177)	-0.609* (0.348)
padre_vive	0.346*** (0.113)	0.604*** (0.207)

num_hermanos	-0.005 (0.039)	-0.016 (0.070)
estrato_residencia	-0.519*** (0.070)	-0.949*** (0.130)
madre_educ	-0.181*** (0.058)	-0.344*** (0.110)
ingreso_hogar	-5.844 -5.493	-11.627 -10.597
_cons	1.229 (0.958)	1.923 -1.849
N	1.343	1.343
r2_a		
r2_p	0.233	0.230
aic	805.841	808.930
bic	904.691	907.780
ll	-383.920	-385.465

* p<0.10 ** p<0.05 *** p<0.01

Nota: N es el número de observaciones. r2_a es el estadístico R cuadrado ajustado. r2_p es el pseudo R cuadrado de McFadden. Aic es el criterio de información de Akaike. Bic es el criterio de información bayesiano de Schwarz.

Fuente: Elaboración propia.

Los resultados de la Tabla 5 muestran los factores que influyen positivamente en la probabilidad de participar o no en el Plan Talentos. Estos son la condición de ser mujer, residir en una vivienda arrendada y la condición de que viva con el padre (para ambos modelos Probit y Logit). Según el modelo Probit, las variables que impactan negativamente en la probabilidad de participar en el Plan Talentos son la edad del individuo, el estrato socioeconómico y el nivel educativo de la madre. Según el modelo Logit, las variables que impactan de manera negativa en la probabilidad de participar en el Plan Talentos son la edad del individuo, el estrato socioeconómico, el nivel educativo de la madre y si trabaja al momento de ingresar a la universidad.

Los resultados de la Tabla 5 indican la probabilidad de participar en el Plan Talentos y sus determinantes. A partir de los valores predichos de las regresiones estimadas en la Tabla 5

se pueden construir los *propensity score*. Los *propensity score* se utilizan para poder emparejar a los individuos que participaron en el Plan Talentos con aquellos que no lo hicieron pero que tenían igual probabilidad de haberlo hecho (igual *propensity score*). Después de emparejar a los individuos se puede realizar una estimación del impacto de la participación en el Plan Talentos sobre el desempeño académico y la probabilidad de graduación.

4.4.2 Elección entre los Modelos de PSM

En esta subsección se exponen los resultados del cálculo de dos criterios de selección de modelos para definir cuál de los dos modelos expuestos por PSM (Probit o Logit) tiene un mejor ajuste a la información contenida en los datos. Los criterios de Akaike (*aic*) y el criterio bayesiano de Schwarz (*bic*) son criterios de selección de modelos. Para elegir un modelo se deben comparar los estadísticos *aic* y *bic* en ambos y se elige aquel que arroje un valor del estadístico de prueba menor. Los valores están en la Tabla 5 en las filas *aic* y *bic*. Como puede notarse son menores en ambos casos para el modelo Probit. Esto indica que el modelo que mejor calidad posee es el descrito por el modelo Probit y se utilizarán los *propensity score* calculados a partir de este método.

4.4.3 Estimación del Efecto Promedio del Tratamiento por PSM

En la Tabla 6 se muestra el parámetro ATT estimado. Dado que en las regresiones para encontrar el *propensity score* el modelo probit tiene un *aic* más bajo y un *bic* más bajo, el efecto promedio de la participación en el Plan Talentos sobre el esfuerzo académico se calcula con base en el modelo probit⁵:

Tabla 6. Efecto promedio de la participación en el Plan Talentos sobre el esfuerzo académico

Raw	Matched
-----	---------

⁵ Se utiliza el comando de STATA *teffects psmatch (Dependiente) (Controles, probit), atet*.

Number of obs =	1,689	330				
Treated obs =	165	165				
Control obs =	1,524	165				
Esfuerzo académico	AI Robust					
	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
ATT						
Talento (1 vs 0)	0.2166984	0.0938293	2.31	0.021	0.0327963	0.4006004

Fuente: Elaboración propia.

Como se puede observar, se emparejaron 330 observaciones, 165 tratados y 165 no tratados. Teniendo en cuenta que, del total de observaciones, 193 personas participaron en el programa, el número de observaciones emparejadas es bastante aceptable. Los resultados de la Tabla 6 muestran que las personas que participaron en el Plan Talentos tienen un esfuerzo académico mayor que las personas que no participaron. Este resultado es estadísticamente significativo a un nivel de significancia del 5%. El valor promedio de la nota promedio ponderada de los individuos participantes en el Plan Talentos es 0.2167 puntos mayor que la de los individuos no participantes. Comparando con los resultados de la Tabla 1 vemos que el efecto estimado por PSM es mayor al estimado por MCO, pero no son muy distintos el uno del otro (el valor estimado por MCO es de 0.187 y por PSM es 0.2167).

En la Tabla 7 se muestra el parámetro ATT estimado. Dado que en las regresiones para encontrar el *propensity score* el modelo probit tiene un *aic* más bajo y un *bic* más bajo, el efecto promedio de la participación en el Plan Talentos sobre la probabilidad de graduación se calcula con base en el modelo probit:

Tabla 7. Efecto promedio de la participación en el Plan Talentos sobre la probabilidad de graduación

grado	AI Robust					
	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]

ATT						
Talento	0.0484848	0.0549049	0.88	0.377	-0.0591268	0.1560965
(1 vs 0)						

Fuente: Elaboración propia.

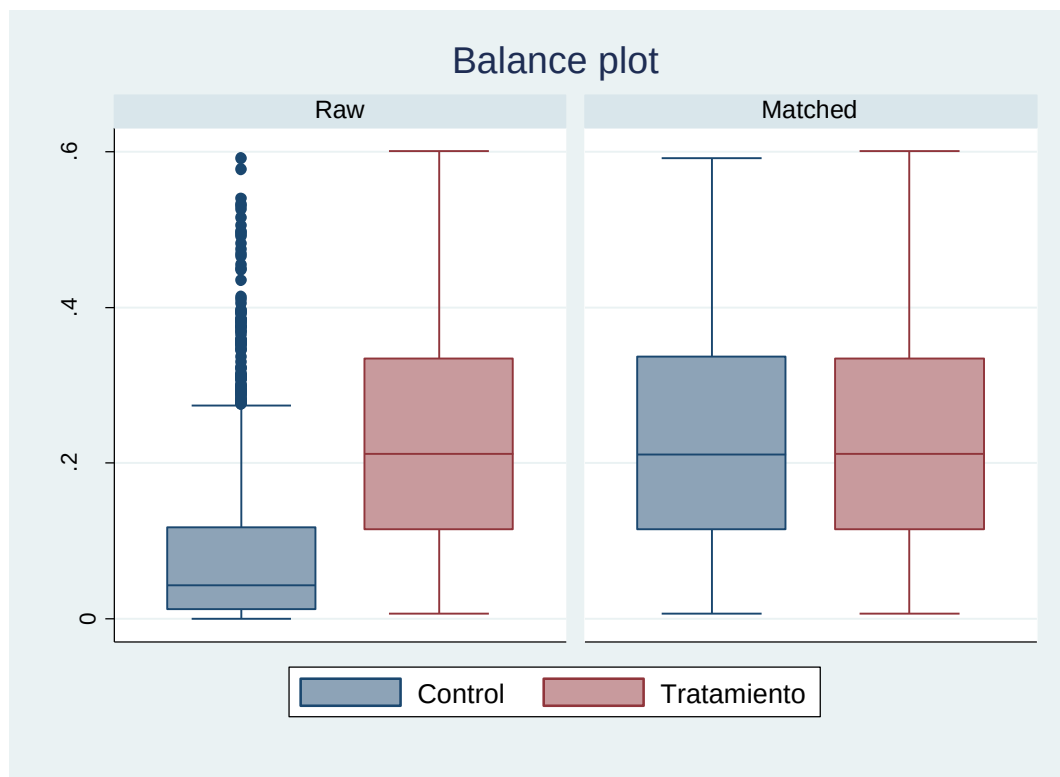
De la Tabla 6 sabemos que se emparejaron 330 observaciones, 165 tratados y 165 no tratados. Los resultados muestran que aquellas personas que participaron en el Plan Talentos, en promedio, aumentaron su probabilidad de graduarse un 4.8% aproximadamente. Sin embargo, tal efecto no es significativo. Por lo que se encuentra evidencia de que la participación en el Plan Talentos no influye en la probabilidad de graduarse, para el grupo de tratados.

Se pueden dar dos razones que podrían estar explicando la ausencia de una relación significativa entre la participación en el Plan Talentos y la probabilidad de graduarse. Es difícil poder evaluar el efecto que tiene el Plan Talentos sobre la probabilidad de graduarse de los estudiantes evaluando el caso de una sola cohorte como se está haciendo en este trabajo. Futuras investigaciones que tengan acceso a los datos de nuevas cohortes pueden obtener mejores estimaciones a partir de métodos de datos panel donde se podrán observar mejor las relaciones entre estas dos variables. Por otro lado, es posible que, al momento de ingresar a la universidad, los estudiantes que participaron del Plan Talentos ya no contaron con el mismo grado de acompañamiento con el que si contaron en sus estudios durante los semestres nivelatorios y que les permitió lograr con éxito entrar a la universidad. Esta ausencia de acompañamiento puede estar relacionada con la ausencia de un impacto sobre la probabilidad de graduarse de los estudiantes dado que el acompañamiento que se hace a los estudiantes es un determinante importante del éxito con el cual estos culminan sus estudios.

4.4.4 Pruebas de Balanceo del Método de PSM

A continuación, se muestra evidencia de que el proceso de pareo es confiable. Lo que nos dice que se ha sorteado el problema de autoselección satisfactoriamente⁶. Enseguida se muestra la distribución del *propensity score* discriminando por tratados y no tratados:

Figura 1. Prueba de Balanceo por Diagrama de Caja y Bigotes.



Fuente: Elaboración propia a partir de STATA.

Se puede observar que existen diferencias significativas entre tratados y no tratados para las observaciones no pareadas. Por otro lado, la distribución entre el grupo de control y el grupo de tratados es casi idéntica, por lo que, se cuenta con un buen soporte común, lo que sugiere que el proceso de emparejamiento que se llevó a cabo es confiable y los resultados obtenidos en el cálculo del efecto de tratamiento son estadísticamente válidos.

⁶ Claro está, bajo el supuesto de que la participación en el programa depende exclusivamente de variables observables.

5. Conclusiones

En este documento de trabajo se desarrollaron varios métodos de estimación para evaluar el impacto que trae participar en el Plan Talentos sobre el desempeño académico para los estudiantes de la Universidad del Valle admitidos el año 2010. Como proxy del desempeño académico se utilizó la variable esfuerzo académico siguiendo los trabajos de Escobar et al. (2008, 2006). Los métodos que se utilizaron fueron el modelo de probabilidad Lineal, Probit y Logit, y el método de *PSM*. Este método se propone dado que permite realizar mejores comparaciones entre grupos de tratamiento y de control. Además, también permite lidiar con posibles problemas de autoselección en la muestra dado que en este caso no se trabajan con datos completamente aleatorizados. Los datos son tomados de aquellos estudiantes que decidieron voluntariamente participar en el Plan Talentos, por lo cual, es posible que exista autoselección en la muestra. Esto implica que la utilización del método de *PSM* permite obtener mejores estimaciones de la relación entre el desempeño académico, sus determinantes y la participación en el Plan Talentos.

El principal resultado de esta investigación indica que el Plan Talentos tiene un efecto positivo sobre el desempeño académico de los estudiantes de la Universidad del Valle que ingresaron en el año 2010. Otros factores que influyen positivamente en el desempeño académico son el ser mujer y la condición de que viva con el padre. Entre los factores que impactan negativamente en el desempeño académico están la edad, la presencia de alguna discapacidad y el tener hijos al momento de ingresar a la universidad. Estos resultados son estadísticamente válidos dado que las pruebas de balanceo indican que la distribución entre el grupo de control y el grupo de tratados es casi idéntica, por lo que, se cuenta con un buen soporte común, lo que sugiere que el proceso de *matcheo* que se llevó a cabo es confiable.

Otro resultado encontrado en este trabajo indica que la participación en el Plan Talentos no tiene un impacto significativo sobre la probabilidad de graduación de los estudiantes. Se pueden dar tres razones que podrían estar explicando la ausencia de una relación significativa entre la participación en el Plan Talentos y la probabilidad de graduarse. Es difícil poder evaluar el efecto que tiene el Plan Talentos sobre la probabilidad de graduarse de los estudiantes evaluando el caso de una sola cohorte como se está haciendo en este trabajo.

Futuras investigaciones que tengan acceso a los datos de nuevas cohortes pueden obtener mejores estimaciones a partir de métodos de datos panel donde se podrán observar mejor las relaciones entre estas dos variables. Por otro lado, es posible que, al momento de ingresar a la universidad, los estudiantes que participaron del Plan Talentos ya no contaron con el mismo grado de acompañamiento con el que si contaron en sus estudios durante los semestres nivelatorios y que les permitió lograr con éxito entrar a la universidad. Esta ausencia de acompañamiento puede estar relacionada con la ausencia de un impacto sobre la probabilidad de graduarse de los estudiantes dado que el acompañamiento que se hace a los estudiantes es un determinante importante del éxito con el cual estos culminan sus estudios. Por último, debemos tener en cuenta que una de las mayores limitaciones de este trabajo es que se realiza un estudio del desempeño académico de estudiantes en el tiempo, pero con unas condiciones socioeconómicas iniciales, y tal vez, algunos efectos sobre el desempeño académico en el paso de los semestres, puede derivar de cambios en esas variables socioeconómicas.

Por otro lado, los trabajos de Escobar et al. (2008, 2006) encontraron que la deserción estudiantil entre los estudiantes de la Universidad del Valle podría ser explicada en su gran mayoría por el bajo desempeño académico. Esto implica que el Plan Talentos puede ser una buena herramienta para combatir la deserción estudiantil en la Universidad dado que este programa mejora significativamente el desempeño de los estudiantes.

Referencias

- Abadie, A., y Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, 84(2), 781-807.
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399-424.
- Bernal, R. P. (2011). Guía práctica para la evaluación de impacto (No. 361.25 B517). Universidad de Los Andes, Bogotá (Colombia).
- Burbano Espinosa, Carolina. (2013). Una aproximación a los determinantes del éxito del plan de nivelación académica Talentos. (*Tesis de pregrado*). Universidad del Valle.
- Burga, C. (2003). Re-evaluando Projovent: propensity score matching y una evaluación paramétrica. Consorcio de investigación económica y social. Recuperado el, 15 de agosto de 2017 de <http://cies.org.pe/sites/default/files/investigaciones/propensity-score-matching-re-evaluando-projovent.pdf>
- Alcaldía de Cali (2009). Plan Talentos, información al alcance de los jóvenes de Cali. [online] Nelson Pinzón G. Secretaría de Educación, 6 de Julio de 2009. Disponible en: http://www.cali.gov.co/planeacion/publicaciones/21280/plan_talentos_informacin_al_alcance_de_los_jvenes_de_cali/ [Recuperado el 22 ago. 2017].
- Cosgrove, J. (2004). The impact of honors programs on undergraduate academic performance, retention, and graduation. *Journal of the National Collegiate Honors Council--Online Archive*, 137.
- Díaz, J. J., y Handa, S. (2006). An assessment of propensity score matching as a nonexperimental impact estimator evidence from Mexico's PROGRESA program. *Journal of human resources*, 41(2), 319-345.
- Escobar, J., Largo, E., Pérez C. (2006). Factores Determinantes de la Deserción y la Permanencia en la Universidad del Valle para las Cohortes de 2001. Universidad del Valle (CIDSE – ViceAcadémica).
- Escobar, J., Largo, E., y Pérez, C. (2008). Rendimiento académico en la Universidad del Valle: determinantes y su relación con la deserción estudiantil. Cali, Colombia: Universidad del Valle.
- Galvis, L. A., y Carrillo Bermúdez, B. (2012). Un índice de precios espacial para la vivienda urbana en Colombia: Una aplicación con métodos de emparejamiento. *Revista de Economía del Rosario*. Vol. 16. No. 1. enero - junio 2013. 25 - 59

- Gamoran, A., 2000: Is Ability Grouping Equitable? p. 234-240 en: R. Arum / R. Beattie (Hrsg.), *The Structure of Schooling: Readings in the Sociology of Education*. Mountain View, CA: Mayfield Publishing Company.
- Guill, K., Lüdtke, O., y Köller, O. (2017). Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study. *Learning and instruction*, 47, 43-52.
- Gundersen, S. (2016). Disappointing returns to education in Ghana: A test of the robustness of OLS estimates using propensity score matching. *International Journal of Educational Development*, 50, 74-89.
- Harris, H. D. (2015). Propensity Score Matching in Higher Education Assessment. Masters Theses. James Madison University: Paper 55.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945-960.
- Joffe, M. M., y Rosenbaum, P. R. (1999). Propensity scores. *American Journal of Epidemiology*, 150, 327-333.
- Keller, R. R., y Lacy, M. G. (2013). Propensity Score Analysis of an Honors Program's Contribution to Students' Retention and Graduation Outcomes. *Journal of the National Collegiate Honors Council*, 14(2), 73-84.
- Lehmann, R. H., Gänsfuß, R., y Peek, R. (2011). LAU 7. Aspekte der Lernausgangslage und der Lernentwicklung–Klassenstufe 7 [An analysis of academic trajectories in secondary seventh-graders]. LAU–Aspekte der Lernausgangslage und der Lernentwicklung. Klassenstufen, 5(7), 121-285.
- Luellen, J. K., Shadish, W. R., y Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, 29, 530-558.
- McIntire, R. K., Nelson, A. A., Macy, J. T., Seo, D. C., y Kolbe, L. J. (2015). Secondhand smoke exposure and other correlates of susceptibility to smoking: a propensity score matching approach. *Addictive behaviors*, 48, 36-43.
- Mincer, J. (1974), *Schooling, experience, and earnings*, New York: National Bureau of Economic Research.
- Murnane, R., Maynard, R. y Ohls, J. (1981): Home Resources and children's achievement. *The Review of Economics and Statistics*, 63 (3), 369-377.

- Perdomo Calvo, J. A., Mendoza, C. A., Baquero-Ruiz, A. F., y Mendieta-López, J. C. (2007). Study of the effect of the TransMilenio mass transit project on the value of properties in Bogotá, Colombia. *Lincoln Institute of Land Policy* Working Paper No. WP07CA1.
- Perdomo, J. A. (2010). Una propuesta metodológica para estimar los cambios sobre el valor de la propiedad: estudio de caso para Bogotá aplicando propensity score matching y precios hedónicos espaciales. *Lecturas de Economía*, (73), 49-65.
- Porto, A., y Di Gresia, L. M. (2004). Rendimiento de estudiantes universitarios y sus determinantes. *Revista de Economía y Estadística*, Cuarta Época, 42 (1), 93-113.
- Rosenbaum, P. R., y Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Shushok Jr, F. (2006). Student outcomes and honors programs: A longitudinal study of 172 honors students 2000-2004. *Journal of the National Collegiate Honors Council*--Online Archive, 15.
- Slavin, C., Coladarci, T., y Pratt, P. A. (2008). Is Student Participation in an Honors Program Related to Retention and Graduation Rates?. *Journal of the National Collegiate Honors Council*--Online Archive, 66.
- Universidad del Valle. (2000). Acuerdo 009 de 2000. Consejo Superior. Cali: Universidad del Valle.

Anexos

Anexo 1. Diccionario de Variables

Nombre	Definición
<i>grado</i>	Variable dummy que toma el valor de 1 si el individuo está graduado para el primer semestre de 2016, cero si no está graduado.
<i>Talentos</i>	Variable dummy que toma el valor de 1 si el individuo participó en el Plan Talentos del año 2009, cero si no participó.
<i>Esfuerzo académico</i>	Variable que mide la nota promedio ponderada por créditos de las materias para los individuos entre el primer semestre de 2010 y el primer semestre de 2016. La variable no incluye las notas de sus trabajos de grado y tampoco incluye aquellas materias cuyas notas no son medidas numéricamente. En el caso de que le individuo habilitara la materia se colocó la nota mayor entre la asignada en la habilitación y la nota del curso como exige el Acuerdo 009.
<i>sexo</i>	Variable dummy que toma el valor de 1 si el individuo es mujer, y cero si es hombre.
<i>edad</i>	Variable que mide la edad de ingreso a la universidad del individuo.
<i>discapa</i>	Variable dummy que toma el valor de 1 si el individuo se auto-reconoce alguna discapacidad al momento de ingresar a la universidad, cero en caso contrario. Las discapacidades incluidas en la encuesta socio-demográfica son presencia de problemas en: 1) sistema nervioso, 2) los ojos, 3) los oídos, 4) los demás órganos de los sentidos (olfato, tacto y gusto), 5) la voz y el habla, 6) el sistema cardiorrespiratorio y las defensas, 7) la digestión, el metabolismo, las hormonas, 8) el sistema genital reproductivo, 9) el movimiento del cuerpo, manos, brazos, piernas, 10) la piel, y 11) partes del cuerpo por mutilaciones.
<i>hijos</i>	Variable dummy que toma el valor de 1 si el individuo tiene hijos al momento de ingresar a la universidad, cero en caso contrario.
<i>sisben</i>	Variable dummy que toma el valor de 1 si posee Sisben como servicio de salud, cero en caso contrario.
<i>EPS</i>	Variable dummy que toma el valor de 1 si el individuo posee un servicio de salud de EPS, cero en caso contrario.
<i>Cali</i>	Variable dummy que toma el valor de 1 si el individuo reside en Cali al momento de ser encuestado, cero en caso contrario.

<i>serv_publ_todos</i>	Variable dummy que toma el valor de 1 si la vivienda donde vive el individuo posee todos los servicios públicos (Acueducto, Alcantarillado, Energía, Teléfono y Recolección de Basuras), cero en caso contrario.
<i>vivienda_arrendada</i>	Variable dummy que toma el valor de 1 si la vivienda donde vive el individuo es arrendada, cero en caso contrario.
<i>soltero</i>	Variable dummy que toma el valor de 1 si el individuo es soltero, cero en caso contrario.
<i>blanco_mestizo</i>	Variable dummy que toma el valor de 1 si el individuo se auto reconoce como Blanco/Mestizo, cero en caso contrario.
<i>afrodescendiente</i>	Variable dummy que toma el valor de 1 si el individuo se auto reconoce como afro descendiente, cero en caso contrario.
<i>indígena</i>	Variable dummy que toma el valor de 1 si el individuo se auto reconoce como indígena, cero en caso contrario.
<i>trabaja</i>	Variable dummy que toma el valor de 1 si el individuo tiene algún trabajo remunerado al momento de ingresar a la universidad, cero en caso contrario.
<i>padre_vive</i>	Variable dummy que toma el valor de 1 si el padre del individuo está vivo al momento de ingresar a la universidad, cero en caso contrario.
<i>num_hermanos</i>	Variable que mide el número de hermanos que tiene el individuo al momento de ingresar a la universidad.
<i>estrato_residencia</i>	Variable que mide el estrato socioeconómico de la vivienda en la reside el individuo al momento de ingresar a la universidad.
<i>madre_educ</i>	Variable que mide el nivel educativo de la madre en 5 niveles, 0 si no tuvo escuela, 1 si tiene primaria completa o primaria incompleta, 2 si tiene bachillerato completo o bachillerato incompleto, 3 si tiene tecnólogo o técnico completo o tecnólogo o técnico incompleto, 4 si tiene universitario completo o universitario incompleto, y 5 si tiene postgrado
<i>padre_educ</i>	Variable que mide el nivel educativo del padre en 5 niveles, 0 si no tuvo escuela, 1 si tiene primaria completa o primaria incompleta, 2 si tiene bachillerato completo o bachillerato incompleto, 3 si tiene tecnólogo o técnico completo o tecnólogo o técnico incompleto, 4 si tiene universitario completo o universitario incompleto, y 5 si tiene postgrado
<i>ingreso_hogar</i>	Variable que mide el ingreso mensual de las personas que conforman el hogar del individuo. Principalmente, padre y madre. Esta variable fue estandarizada de tal forma que su media es cero, y su varianza es 1.

Fuente: Elaboración propia.