











Valor agregado en la educación superior colombiana: una revisión de la contribución de los programas y las instituciones a la calidad educativa

Recibido: 30 de agosto de 2024
Evaluado: 08 de abril de 2025
Publicado: 01 de abril de 2026

Frederick Andrés Mendoza-Lozano*  
Claudia Milena Pico-Bonilla**  
David Bacca-Morales***  
Mauricio Alejandro Cano-Niño****  

Tipología: artículo de investigación

Resumen

El valor agregado es una medida de la contribución de las instituciones al logro académico de los estudiantes. Sin embargo, las medidas de valor agregado pueden ser imprecisas porque los mejores estudiantes intentan entrar a las universidades con mejor reputación y, por tanto, es más probable que estas instituciones obtengan mejores resultados. Para corregir la imprecisión de medida se requiere controlar por el historial del estudiante y su entorno socioeconómico para evaluar el efecto neto que está generando la institución sobre el logro; esto es, corregir por sesgo de selección. Este artículo presenta una propuesta de medición del valor agregado de las instituciones y programas de educación superior en Colombia a partir de la estimación de un modelo de regresión jerárquica y una metodología de evaluación de impacto conocida como *Propensity Score Matching*. Con este objetivo, compara el desempeño de los estudiantes en pruebas Icfes Saber II y Saber Pro presentadas en 2020. Los resultados indican que existen disparidades en la educación superior y hay evidencia de sesgo de selección; no obstante, al aplicar procedimientos de corrección, se demuestra que la mayor contribución institucional se concentra en la Corporación Universitaria Adventista, la Universidad la Gran Colombia en Armenia y la Universidad de la Sabana. El estudio amplía el debate a nivel de programa académico y presenta un listado de los que tienen mayor valor agregado a nivel nacional.

Palabras clave

análisis multivariado; calidad de la educación; evaluación continua; modelo estadístico

* Institución Universitaria Politécnico Grancolombiano. Departamento de Ciencias Básicas. famendoza@poligran.edu.co

** Universidad Militar Nueva Granada. Programa académico: Facultad de Ciencias Económicas. claudia.pico@unimilitar.edu.co

*** Institución Universitaria Politécnico Grancolombiano. Departamento de Ciencias Básicas. hbaccamo@poligran.edu.co

**** Institución Universitaria Politécnico Grancolombiano. Ingeniería en Ciencia de Datos. alejandrocannon@gmail.com

Value Added in Colombian Higher Education: A Review of the Contribution of Programs and Institutions to Educational Quality

Abstract

Value added is a measure of the contribution of institutions to students' academic achievement. However, value-added measures can be imprecise, as the best students tend to apply to the most reputable universities, making it more likely that these institutions will achieve better results. To correct this measurement bias, it is necessary to control for students' prior academic history and socioeconomic background in order to evaluate the net effect of the institution on achievement, that is, to adjust for selection bias. This article presents a proposal for measuring the value added of Colombian higher education institutions and programs by estimating a hierarchical regression model and applying an impact evaluation methodology known as *Propensity Score Matching*. For this purpose, it compares student performance in the Icfes, Saber 11 and Saber Pro tests administered in 2020. The results indicate disparities in higher education and evidence of selection bias. Nonetheless, after applying correction procedures, the greatest institutional contribution is shown to be concentrated in Corporación Universitaria Adventista, Universidad La Gran Colombia in Armenia, and Universidad de La Sabana. The study further extends the debate to the level of academic programs and provides a list of those with the highest value added at the national level.

Keywords

multivariate analysis; educational quality; continuous assessment; statistical model

Valor agregado no ensino superior colombiano: uma revisão da contribuição dos programas e instituições para a qualidade educacional

Resumo

O valor agregado é uma medida da contribuição das instituições para o desempenho acadêmico dos estudantes. No entanto, as medidas de valor agregado podem ser imprecisas, pois os melhores estudantes tendem a ingressar nas universidades de maior reputação, o que aumenta a probabilidade de que essas instituições obtenham melhores resultados. Para corrigir essa imprecisão de medida, é necessário controlar o histórico acadêmico e o contexto socioeconômico dos estudantes, de modo a avaliar o efeito líquido gerado pela instituição sobre o desempenho, ou seja, corrigir o viés de seleção. Este artigo apresenta uma proposta de medição do valor agregado das instituições e programas de ensino superior na Colômbia a partir da estimação de um modelo de regressão hierárquica e da aplicação de uma metodologia de avaliação de impacto conhecida como *Propensity Score Matching*. Para isso, compara o desempenho dos estudantes nas provas Icfes, Saber 11 e Saber Pro aplicadas em 2020. Os resultados indicam disparidades no ensino superior e evidências de viés de seleção. Contudo, após a aplicação dos procedimentos de correção, demonstra-se que a maior contribuição institucional se concentra na Corporación Universitaria Adventista, na Universidad La Gran Colombia (em Armenia) e na Universidad de La Sabana. O estudo amplia o debate para o nível dos programas acadêmicos e apresenta uma lista daqueles com maior valor agregado em nível nacional.

Palavras-chave

análise multivariada; qualidade da educação; avaliação contínua; modelo estatístico

Para citar este artículo:

Mendoza-Lozano, F. A., Pico-Bonilla, C. M., Bacca-Morales, D. y Cano-Niño, M. A. (2026). Valor agregado en la educación superior colombiana: una revisión de la contribución de los programas y las instituciones a la calidad educativa, *Revista Colombiana de Educación*, (99), e22105, <https://doi.org/10.17227/rce.num99-22105>

Introducción

La educación superior colombiana enfrenta retos de diversa naturaleza. Recientemente se ha puesto en evidencia por parte del Ministerio de Educación Nacional la reducción de la matrícula en este nivel de formación, con cerca de 37 000 estudiantes menos entre 2020 y 2021, lo que se atribuye al cambio demográfico (Meisel y Granger, 2022), a variaciones en los intereses de la población joven (Zerpa y Rodríguez-Montoya, 2024) y a la falta de oportunidades en los mercados de trabajo (Rincón y Espitia, 2022). Tras la pandemia, en 2023, las cifras de cobertura se recuperaron y se alcanzaron niveles similares a los que tenía el país en 2017 (cuando la matrícula fue de 2 280 327 y en 2023, 2 259 970). Aunque la cobertura de la educación superior en el país es importante, hay otras dos problemáticas que son de atención urgente, pues se asocian con la movilidad social y la reducción de brechas de desigualdad: la calidad y la equidad en las instituciones de educación superior.

Como lo advierte Tam (2010) en su revisión, la medición de la calidad en la educación no es una tarea sencilla, pues, aunque hay claros indicios de determinantes individuales, del hogar e institucionales, que favorecen las mejoras en calidad, en ocasiones resulta difícil determinar de manera precisa las causas del buen rendimiento de los alumnos. Los estudios sobre el valor agregado aportan en esa dirección, pues permiten examinar cómo un entorno institucional específico contribuye a mejorar el logro académico de un estudiante, al observar su trayectoria desde una evaluación de entrada, hasta una evaluación posterior.

El valor agregado en el contexto educativo corresponde a la variación en el logro de los estudiantes, entendido como el crecimiento en conocimiento, habilidades, capacidades y otros atributos, que ellos han ganado como resultado de sus experiencias en un sistema educativo en el tiempo (OCDE, 2008). Estas mediciones comportan retos, pues, como advierten Perna y Thomas (2006), la evaluación del éxito de los estudiantes debe contextualizarse en múltiples capas que incluyan el contexto interno de los estudiantes, el familiar, el de la escuela y el social, económico y político. Rodríguez-Revilla (2022), por su parte, afirma que el valor agregado es “la contribución de las instituciones educativas o los profesores al progreso académico neto de los estudiantes, una vez eliminada la influencia de otros factores ajenos a la institución” (p. 45).

La medición del progreso académico neto implica debates metodológicos, toda vez que si se quiere capturar el aporte institucional en la formación de un estudiante, es preciso aislar el efecto de otros factores fuera del control de la escuela (Duarte *et al.*, 2016). Además de los retos metodológicos, el estudio del valor agregado es importante para identificar los diferenciales de calidad entre instituciones educativas, pues sus resultados pueden convertirse en una parte

importante en la construcción de reputación, que eleve la matrícula, y en una fuente para focalizar el otorgamiento de recursos públicos (Liu *et al.*, 2012).

Son numerosos los estudios realizados sobre la medición del valor agregado. Algunas aproximaciones se centran en medir la diferencia neta entre los niveles de entrada y de salida de los estudiantes en un programa académico. En este grupo de estudios, se inscribe el realizado por Campos *et al.* (2017), quienes se centran en la competencia de lectura crítica y cuantifican la diferencia entre la prueba de ingreso a la universidad y la prueba Saber Pro. Por su parte, el trabajo de Ramírez (2014) da cuenta de la importancia de la trayectoria académica previa para predecir el buen rendimiento académico; así, un buen desempeño en la prueba Saber 11, por lo general, se convierte en un buen predictor del desempeño en la prueba Saber Pro.

En esta misma línea, Owusu y Larbi (2017) adoptan un enfoque constructivista y comparan el rendimiento promedio ponderado acumulado al inicio y final del proceso de formación de un grupo de 127 estudiantes de Ingeniería de Telecomunicaciones. Los autores concluyen que el nivel de entrada de los estudiantes es un predictor de su desempeño al final de su formación, lo que supone que la selección de los mejores estudiantes otorga ventajas en el rendimiento a lo largo de los procesos de formación.

Aunque los trabajos que se centran en la medición de la diferencia entre niveles de entrada y salida verifican la existencia de variaciones en el rendimiento como producto del proceso de formación, no logran demostrar qué proporción de estas variaciones se debe a las instituciones educativas, es decir, producen mediciones sesgadas, pues no corrigen por sesgo de selección, al no introducir controles de otros factores diferentes a la institución.

En términos generales, el sesgo de selección se produce cuando se utilizan procedimientos inadecuados en la selección de la muestra, lo que lleva a que el grupo seleccionado no sea un reflejo exacto de la población (Alexander *et al.*, 2015). En ciencias sociales, el problema puede surgir por reglas de selección distorsionantes y problemas de autoselección de los participantes (Heckmann, 1990).

En el caso de la educación superior, se presenta el problema de la autoselección, pues los estudiantes eligen una universidad basados en la reputación de la institución, en su capacidad de pago y en sus calidades académicas, hecho que hace más probable que las universidades agrupen estudiantes con características comunes y que no sea posible obtener una muestra aleatoria representativa de toda la población en cada universidad. En palabras de Rodríguez y López (2016): “estudiantes con un desempeño académico notable en su educación básica y media tienen una alta probabilidad

de continuar con un elevado rendimiento en niveles superiores independientemente de la institución a la que asisten” (p. 120).

La literatura sobre valor agregado, en su mayoría, parte de la necesidad de corregir el sesgo de selección y, por tanto, suele tener aproximaciones a partir del método de regresión lineal jerárquica con variables instrumentales, propuesto por (Manzi *et al.*, 2013). Usualmente, los estudios utilizan métodos econométricos de regresiones jerárquicas, en los que se compara la variación entre el rendimiento promedio de los estudiantes en áreas como lectura crítica, razonamiento cuantitativo, entre otras (Monroy *et al.*, 2018).

Rodríguez-Revilla (2015) propuso una evaluación de la calidad del sistema educativo para la educación superior en Bogotá, a partir del uso de modelos lineales jerárquicos con variables instrumentales, en los que se controla por características familiares, sociales y económicas y algunas institucionales, con el fin de reducir el sesgo de selección. El trabajo logra identificar que el 71 % del desempeño académico de los estudiantes se explica por las universidades, hecho que justifica la realización de diagnósticos permanentes sobre valor agregado, pues, en este caso, constituyen una porción importante de la explicación sobre la calidad educativa en el nivel superior.

Bogoya y Bogoya (2013) estiman un modelo lineal que considera el rendimiento académico en pruebas estandarizadas y el nivel socioeconómico. Los autores encontraron que el historial académico, representado en los resultados Saber 11, es un buen predictor del desempeño en las Saber Pro, por encima del nivel socioeconómico. En un trabajo posterior, Bogoya *et al.* (2017) usaron una aproximación a la medición del valor agregado en pruebas estandarizadas, con estimaciones a través de mínimos cuadrados ordinarios y regresión cuantílica. La modelación con mínimos cuadrados muestra que el desempeño en pruebas Saber 11, en conjunto con otras características socioeconómicas son buenos predictores del desempeño en las Saber Pro; sin embargo, debido a los problemas de sesgo referidos previamente, esta metodología no permite cuantificar adecuadamente el valor agregado de una institución. En esa medida, la regresión cuantílica reduce la suma de errores al cuadrado, pero no logra aislar el efecto de la institución sobre el rendimiento en la Saber Pro.

Shavelson *et al.* (2016) discuten los retos que se presentan en la medición de valor agregado para el caso colombiano y advierten la existencia de supuestos en los ejercicios de modelación, que pueden ser vistos como limitaciones o imprecisiones en las mediciones de valor agregado. Dichos supuestos son, entre otros, la no interferencia, en la que se presume que no hay efectos de grupo par asociados con los tratamientos; la homogeneidad del efecto causal, que no varía en función de las características individuales; y el de la presunción de correctitud

de las formas funcionales lineales. Cada uno de estos supuestos revela limitaciones de los métodos para medir el valor agregado, pues la literatura ha documentado la existencia de efectos de grupo par (Arnott y Rowse, 1987), y se ha demostrado la efectividad de las aproximaciones no lineales (Schiltz *et al.*, 2018).

Los estudios reseñados hasta este punto usan en su mayoría las pruebas estandarizadas como fuente de información; sin embargo, como advierten Cunha y Miller (2014), el uso de datos administrativos para medir el valor agregado en la educación superior reviste algunos problemas. En primer lugar, el valor agregado no solo deriva de las pruebas estandarizadas, sino que puede verse reflejado en las tasas de permanencia, tasas de graduación y ganancias de los egresados, lo que implica que las medidas que se concentran en los resultados de las pruebas desconocen otros efectos que puede tener una institución educativa sobre la calidad de sus egresados. En segundo lugar, los autores reconocen que las medidas estandarizadas no dan cuenta precisa de la agregación de valor en los campos disciplinares, pues suelen concentrarse en la observación de competencias generales.

Por su parte, Liu *et al.* (2012) critican el uso de pruebas estandarizadas como un instrumento de diagnóstico para la educación superior, pues advierten que en estas no se captura el nivel de motivación de los estudiantes que presentan la prueba. Esto implica que, aunque se trata de un instrumento importante para medir la calidad educativa, no se constituye en una fuente confiable, toda vez que induce sesgos. En esta misma línea, Sparks (2011) y Tam (2004) argumentan que la calidad de la institución educativa y su generación de valor agregado no se reduce al rendimiento en pruebas, sino a los resultados en términos de empleo, la motivación de los estudiantes, el tiempo de contacto y la experiencia con la institución. Por su parte, Díaz-Palacios (2015) reconoce la importancia de las pruebas estandarizadas, pero advierte que su aplicación desconoce otras acciones que se adelantan en las instituciones en procura de mejorar la calidad educativa.

Ante la presencia de múltiples fuentes de sesgos, se ha optado recientemente por el uso de métodos alternativos para minimizar los errores de predicción. Blankenberger *et al.* (2021) señalan que en la evaluación de la calidad educativa es recomendable emplear métodos cuasiexperimentales que reduzcan las amenazas de pérdida de validez interna y externa. En esa medida, consideran que el método Propensity Score Matching podría servir para emular la aleatorización y, por tanto, sería efectivo en la reducción del sesgo que se deriva de la autoselección. Esta observación es confirmada por Austin (2011), quien afirma que este método logra eliminar una alta proporción de las diferencias que

hay en las características básicas entre quienes reciben el tratamiento y aquellos que hacen parte del grupo de control.

En el campo de la investigación experimental y cuasiexperimental, destaca el trabajo de Schiltz *et al.* (2018), quienes construyeron un ranking de escuelas basado en estimaciones de valor agregado con el método de bosques aleatorios. Los autores demostraron, con datos de pruebas estandarizadas para Italia en el 2013, que el método de bosques aleatorios permite minimizar los errores de predicción si se le compara con los métodos tradicionales de regresión lineal, pues, al capturar no linealidades, describe de forma más ajustada las relaciones más complejas.

En definitiva, la existencia de sesgo de selección, defectos en las variables que se usan en la medición y no linealidades, obliga a una revisión permanente de los métodos usados para la medición de valor agregado. Estos métodos no solo deben atender a la necesidad de obtener el efecto neto de la contribución institucional, sino que además pueden convertirse en una herramienta de diagnóstico e identificación de instituciones y programas que contribuyen a la mejora de la calidad educativa.

En esta medida, este trabajo aporta a la literatura existente sobre valor agregado en educación superior, porque avanza en la corrección del sesgo de selección a partir del uso de un método cuasiexperimental de Propensity Score Matching, usualmente utilizado en estudios de evaluación de impacto, al tiempo que evalúa el desempeño de los estudiantes de todos los programas en todas las instituciones de Colombia, y desarrolla una herramienta de consulta y comparación, que sirve como insumo para el diseño de estrategias de mejora a nivel de programa e instituciones.

Metodología

Este artículo se desarrolló con métodos cuantitativos. Tras el proceso de depuración de las bases de datos, fue ejecutado en tres fases: en la primera fase, se construyó un análisis descriptivo, en el que se identificaron las trayectorias y el desempeño en competencias generales de los estudiantes de cada universidad y cada programa académico, desde la presentación de la prueba Saber 11, hasta la presentación de la prueba Saber Pro, previas a la corrección por sesgo de selección.

Posteriormente, se evaluó el desempeño y se corrigió por sesgo de selección, para lo cual se estimó el modelo de regresión lineal jerárquica, que se comparó con el método de Matching. Con la aplicación tanto del método de regresión jerárquica como del método de Matching, se construyó un ranking por institución y, para el último caso, también se construyó un ranking por programa. Por último, en la tercera fase, para sistematizar los resultados y permitir la identificación y

caracterización de los programas y sus resultados, en términos de valor agregado, se elaboró una aplicación de visualización de la información a través de Shiny Apps.

Datos

Los microdatos que se usaron como referencia para la construcción del estudio provienen de las bases de datos del Icfes (2021), que, para el año de consulta, contiene 247 834 observaciones a nivel nacional. Un rasgo diferencial de esta investigación tiene que ver con el reconocimiento de las modalidades en las IES. El Icfes provee una llave para enlazar los resultados de la prueba Saber 11 con la prueba Saber Pro; así, se buscaron, por medio de una llave provista por el Icfes, todos los estudiantes del periodo 2020 en los repositorios de la prueba Saber 11 entre 2010-1 y 2017-2.

Conviene señalar que los datos usados corresponden a estudiantes egresados en un periodo previo a los confinamientos por covid-19, por lo que este estudio no avanza en la evaluación de los posibles efectos de la pandemia sobre los resultados en las pruebas Saber Pro.

Análisis descriptivo

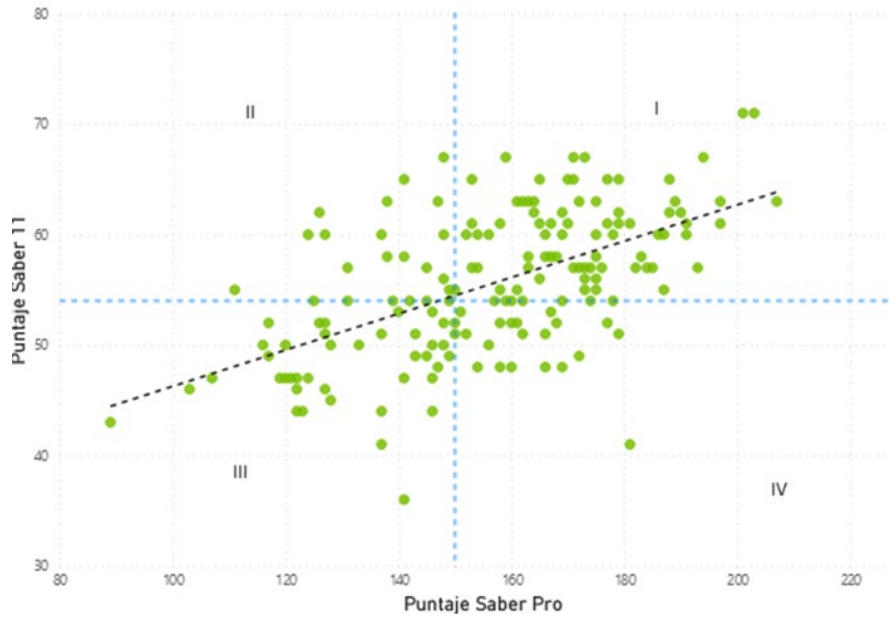
Con la información depurada, se construyó un análisis de estadística descriptiva de la prueba Saber Pro por cada una de las covariables que se definen en el modelo de valor agregado. Dado que la prueba Saber 11 tuvo un cambio estructural a partir del periodo 2014-2, que hizo cambiar significativamente los resultados promedio, es posible hacer las comparaciones según los resultados de uno y otro periodo, esto es, antes de 2014-2 y después.

Se ejecutó un análisis de puntajes globales, a partir de un diagrama de dispersión de la prueba Saber Pro versus las pruebas Saber 11, que se divide en cuatro cuadrantes (ver figura 1). En el marco de este análisis, los valores de referencia corresponderán a promedios nacionales (representados en las líneas punteadas). En el cuadrante I, se ubican aquellos estudiantes que estuvieron por encima del promedio nacional tanto en las pruebas Saber 11 como en las Saber Pro; en el cuadrante II, están los estudiantes que tenían buen rendimiento en las pruebas Saber 11 y obtuvieron un rendimiento desfavorable en las Saber Pro; en el cuadrante III, se encuentran aquellos estudiantes que tenían un desempeño por debajo del promedio en las Saber 11 y que mantuvieron un desempeño por debajo de la media en las Saber Pro; por último, la descriptiva más importante corresponde al cálculo del número de personas en el cuadrante IV, que equivale a los estudiantes que ingresaron con un puntaje de la prueba Saber 11 por debajo

de la media y lograron obtener puntajes por encima del promedio en la prueba Saber Pro.

Figura 1.

Modelo de análisis de dispersión entre los resultados de las pruebas Saber 11 y Saber Pro



Fuente: Informe de valor agregado, Universidad de Cundinamarca.

Modelos de referencia

Modelo de regresión jerárquica

Con base en Velasco (2006), los modelos de regresión jerárquicos se usan cuando el conjunto de datos que se pretende analizar tiene una estructura jerárquica, por ejemplo, se pretende estudiar las relaciones entre alumnos (primer nivel) y el grupo de estudiantes del que hacen parte (segundo nivel).

A partir de la técnica del modelo de regresión lineal jerárquico, y para confirmar los efectos heterogéneos en los resultados de los estudiantes, se trabajó un modelo en dos niveles: estudiantes e IES. La definición de variables y covariables de cada nivel sigue de cerca el trabajo de Rodríguez- Revilla (2015); así, la variable dependiente fue el puntaje global prueba Saber Pro, entretanto, para las variables independientes, se hicieron dos ajustes (ver tabla 1): en el primer nivel, se incluyeron los resultados promedio de la prueba Saber 11, y en el segundo nivel, la modalidad de las IES. El modelo incluyó la totalidad de observaciones del país en la prueba Saber Pro y descartó los valores perdidos en

los resultados de puntaje global en la Saber Pro, o los que no se pudieron asociar con un registro en la Saber 11, al usar la llave que provee el Icfes para tal efecto.¹

Tabla 1.

Covariables en el modelo de regresión multinivel

Covariables de primer nivel: estudiante	Covariables de segundo nivel: IES
Puntaje saber 11	Sector IES (pública o privada)
Género	Cantidad de docentes/total de estudiantes
Años de educación del padre	Proporción de docentes con doctorado
Años de educación de la madre	Modalidad: virtual, distancia, presencial
Nivel socioeconómico	
Situación laboral	

Fuente: elaboración propia, con base en Rodríguez-Revilla (2015) y datos del Icfes (2021).

El modelo lineal multinivel implica la estimación de efectos fijos y aleatorios. La especificación del primer nivel corresponde a esta expresión:

$$Saberpro_{ij} = \beta_{0j} + \beta_{1j} + \mu_{ij} \quad (1)$$

Donde β_{0j} representa el intercepto de cada IES, β_{1j} representa el vector de parámetros beta de cada IES, X_{ij} representa las covariables de primer nivel y μ_{ij} corresponde al error aleatorio del primer nivel.

La especificación de segundo nivel sigue esta formulación:

$$\beta_{0j} = \alpha_{00} + \omega_{1j} + v_{0j}\beta_{0j} \quad (2)$$

$$\beta_{1j} = \alpha_{10} + \omega_{1j} + v_{1j}\beta_{1j} \quad (3)$$

Donde α_{00} corresponde a la media global del desempeño en el puntaje en la prueba Saber Pro en toda la población, ω_{1j} representa las covariables de segundo nivel y v_{0j} da cuenta del error que mide el valor agregado de la IES.

La validez del modelo multinivel se verifica al comparar la anterior especificación contra un modelo nulo, en el que no hay covariables, solo los interceptos y los términos del error. Además, la estrategia multinivel se evalúa calculando la proporción de varianza en cada nivel. El modelo se ejecutó con ayuda de la librería lme4 de R. El procedimiento y su interpretación se fundamentaron en Roback y Legler (2021).

¹ De los 246 436 estudiantes del país que presentaron la prueba Saber Pro en el 2020, se encontraron 160 439 que presentaron la prueba Saber 11 entre los años 2010 y 2017. De esta forma, se tiene una muestra seleccionada para 2020, que corresponde al 65 % del total de evaluados en Saber Pro.

Propensity Score Matching

En este trabajo se propone un modelo alternativo al jerárquico, inspirado en la metodología Propensity Score Matching de amplia aplicación en el ámbito de la evaluación de impacto en ciencias sociales (Bernal y Peña, 2011). El método de Propensity Score fue presentado originalmente por Rosenbaum y Rubin (1983) y se entiende como la probabilidad de asignación a un tratamiento particular, dado un vector de covariables observadas. Por su parte, la técnica de Matching empareja uno o más casos de control con un “vecino cercano” que ha recibido el tratamiento (Benedetto *et al.*, 2018). El Matching intenta eliminar el sesgo de selección al hacer comparaciones en el valor agregado entre estudiantes, que por sus características generales son comparables, esto es, se empareja a cada estudiante con su vecino más cercano.

El procedimiento que se plantea implica que cada IES o programa académico sea un tratamiento. Los estudiantes que ingresan tienen unas condiciones previas como su rendimiento en la prueba Saber 11, el ingreso individual y familiar, el área de residencia (urbana o rural), el acceso a internet, el nivel de formación del padre, el grupo de referencia y su estado civil.

Para eliminar el sesgo de selección, cada estudiante del tratamiento (IES) se compara con una muestra aleatoria de diez mil estudiantes de otras IES y se encuentran parejas o grupos de estudiantes muy similares. Esa similitud se mide como la probabilidad de seleccionar la universidad (tratamiento). Las probabilidades se calculan por medio de una regresión logística en la que $X_1 \dots X_k$ son las variables que definen la probabilidad de pertenecer a una IES o programa académico. Este proceso se conoce como emparejamiento:

$$\text{logit}(P(T = 1 | X)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (4)$$

Una vez obtenidos los emparejamientos, se accede a un subconjunto de datos en el que los estudiantes de cada IES están agrupados con estudiantes de otras IES similares, es decir, aunque escogieron otra institución tenían una alta probabilidad de pertenecer a la IES que se está analizando. Posteriormente, se define una regresión lineal en el grupo en la que la variable dependiente es la diferencia entre los puntajes estandarizados de la prueba Saber Pro y la Saber 11. Además, se controla por las mismas variables del emparejamiento, más una *dummy* de tratamiento, que toma el valor de 1, si el estudiante pertenece a la IES, y de cero, si no pertenece (grupo de control). El valor del coeficiente de la variable tratamiento indica el efecto promedio sobre los tratados (ATT, por sus siglas en inglés), esto es, el valor agregado de la institución.

Dado que los puntajes de las pruebas Saber 11 y Saber Pro y sus diferencias están estandarizados, la interpretación del modelo puede hacerse en términos de

desviaciones estándar. Así, por ejemplo, un valor agregado de 0,5 indica un efecto promedio de 0,5 desviaciones estándar en el valor agregado de una IES o programa, con respecto a otros espacios de formación que reciben estudiantes similares o comparables. Este estándar de comparación es el que se usa tanto en el ranking que resulta del modelo de regresión jerárquica como en el que se obtiene del Propensity Score Matching.

Resultados

El análisis descriptivo realizado demuestra la importancia del sesgo de selección, pues, como se puede verificar en la figura 2, las universidades de alto logro como la Universidad de los Andes y la Universidad del Norte en Barranquilla, concentran la mayor proporción de sus estudiantes (98 %, para la primera, y 95 %, para la segunda) en el primer cuadrante, esto es, reciben estudiantes de alto logro académico y gradúan estudiantes con rendimiento en Saber pro por encima del promedio. En los casos en los que reciben estudiantes por debajo del promedio en Saber 11, logran que su rendimiento esté por encima del promedio en la prueba Saber Pro.

Figura 2.

Relación de los puntajes de las pruebas Saber pro y Saber 11, Universidades de los Andes (izquierda) y Universidad del Norte (derecha)



Fuente: elaboración propia con los datos del Icfes (2021).

Los patrones observados muestran una tendencia a la diferenciación en calidad y en procesos de selección cuando se comparan con instituciones de educación superior que reciben mayores proporciones de estudiantes con menores niveles socioeconómicos y menor rendimiento en pruebas Saber 11. La figura 3 da cuenta de los patrones observados para dos universidades en la ciudad de Barranquilla, la Universidad de la Costa y la Universidad Simón Bolívar. De esta comparación se deduce que, a diferencia de las universidades de alto logro, la proporción de estudiantes con buen rendimiento en las pruebas Saber 11 se ubica en 79 % para la Universidad de la Costa y en 72 % para la Universidad

Simón Bolívar. En las Saber Pro, 68 % de los estudiantes logra estar por encima del promedio en la primera universidad y 58 % lo hace en la segunda.

Figura 3.

Relación de los puntajes de las pruebas Saber Pro y Saber 11, Universidad de la Costa (izquierda) y Universidad Simón Bolívar (derecha)



Fuente: elaboración propia con los datos del Icfes (2021).

Resultados del modelo lineal jerárquico

Una vez corregido el sesgo de selección mediante el modelo lineal jerárquico, las diferencias socioeconómicas que son evidentes en el análisis descriptivo se reducen. La tabla 2 muestra el valor agregado por institución educativa en las doce instituciones que logran mejor rendimiento en sus estudiantes tras corregir por sesgo de selección. El ranking revela que el CESA logra que sus estudiantes estén 0,5 desviaciones estándar por encima del promedio. También se reconoce el desempeño de instituciones regionales como la Universidad de Antioquia, la sede Armenia de la Gran Colombia, la Corporación Adventista y la Universidad de la Sabana en el municipio de Chía.

Tabla 2.

Ranking universitario usando el modelo lineal jerárquico

Posición	Nombre	Sede	Puntaje
1	Colegio de Estudios Superiores de Administración CESA	Bogotá D. C.	0,53
2	Universidad de Antioquia	Carmen de Viboral	0,43
3	Fundación Universitaria Cafam	Bogotá D. C.	0,36
4	Universidad La Gran Colombia	Armenia	0,36
5	Universidad de la Sabana	Chía	0,35
6	Institución Universitaria Latina Unilatina	Bogotá D. C.	0,35
7	Universidad Nacional de Colombia	Bogotá D. C.	0,34
8	Fundación Universidad de América	Bogotá D. C.	0,32
9	Corporación Universitaria Adventista	Medellín	0,32
10	Universidad Colegio Mayor de Cundinamarca	Bogotá D. C.	0,27
11	Corporación Escuela de Artes y Letras	Bogotá D. C.	0,27
12	Universidad EIA	Medellín	0,26

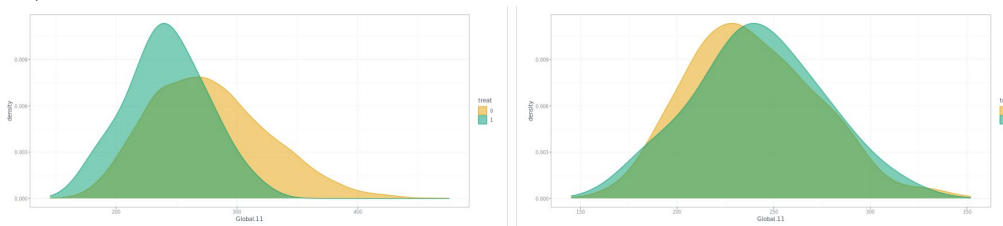
Fuente: elaboración propia con los datos del Icfes (2021).

Resultados del Propensity Score Matching

Al aplicar el procedimiento de Propensity Score Matching se busca lograr un ajuste mayor entre cada estudiante y su vecino cercano, esto es, un estudiante de similares características en términos de nivel socioeconómico, programa académico, puntaje de la prueba Saber 11, entre otros, para diferenciar sus resultados tras pasar por las instituciones y programas que están siendo evaluados. Posterior al emparejamiento, se observa un mejor ajuste, como ejemplo se muestra la distribución de probabilidad para los estudiantes de la Corporación Universitaria Adventista (ver figura 4), que se sitúa como novena en generación de valor agregado con el modelo lineal jerárquico, con un promedio institucional que está 0,32 desviaciones estándar por encima del promedio nacional. Los patrones de distribución presentados en las Figuras 4 y 5 ponen en evidencia que los resultados en términos de valor agregado suelen ser más bajos para la Corporación Universitaria Adventista, cuando se compara a los estudiantes con la totalidad de la muestra, y tiende a ser más ajustados cuando se consideran sus pares cercanos en términos de resultados de Saber 11 y nivel socioeconómico. Asimismo, la información de la tabla 3 corrobora esta información, pues es evidente que la agregación de valor de esta institución es superior cuando sus estudiantes son comparados con sus pares cercanos, ya que la desviación estándar pasa de estar 0,32 por encima a 0,47.

Figura 4.

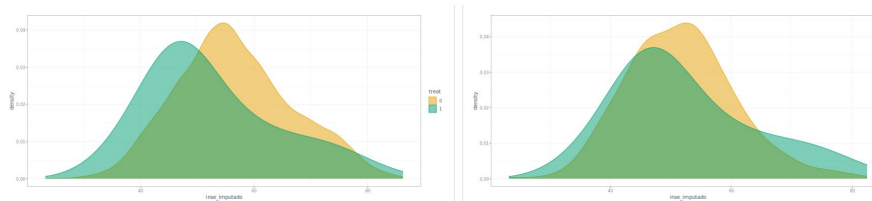
Ajuste con emparejamiento (derecha) y sin emparejamiento (izquierda) para pruebas Saber 11, Corporación Universitaria Adventista



Fuente: elaboración propia con los datos del Icfes (2021).

Figura 5.

Ajuste con emparejamiento (derecha) y sin emparejamiento (izquierda) por nivel socioeconómico, Corporación Universitaria Adventista



Fuente: elaboración propia con los datos del Icfes (2021).

En cuanto al ranking de uso del Matching (ver tabla 3), se observa que algunas instituciones siguen registrando buen desempeño en términos de valor agregado, es el caso de la Universidad La Gran Colombia en Armenia, Unilatina, la Universidad de Antioquia y la Corporación Universitaria Adventista. Instituciones como el CESA (0,14) y la Fundación Universitaria Cafam (0,19) tienen un aporte más modesto al valor agregado cuando se compara a sus estudiantes con otros de similares características. En el caso de la Universidad Nacional de Colombia (-0,03), se observa un deterioro en la generación de valor agregado cuando se aplica el procedimiento de Matching. Al aplicarlo, destaca la presencia de Universidades como la Sergio Arboleda en Santa Marta, la Universidad de Cundinamarca, la Manuela Beltrán y la Konrad Lorenz, las cuales muestran agregación de valor para sus estudiantes al compararlos con sus pares.

Tabla 3.

Ranking universitario al usar el modelo Propensity Score Matching

Ranking	Nombre	Ciudad	Puntaje
1	Corporación Universitaria Adventista	Medellín	0,47
2	Universidad Sergio Arboleda	Santa Marta	0,36
3	Institución Universitaria Latina Unilatina	Bogotá D. C.	0,35
4	Universidad La Gran Colombia	Armenia	0,32
5	Universidad de Antioquia	Carmen de Viboral	0,32
6	Universidad Libre	Cúcuta	0,28
7	Universidad de La Sabana	Chía	0,28
8	Fundación Universitaria de Ciencias de la Salud	Bogotá D. C.	0,24
9	Fundación Universitaria Juan N. Corpas	Bogotá D. C.	0,23
10	Universidad de Cundinamarca UDEC	Ubaté	0,23
11	Universidad Manuela Beltrán UMB	Bogotá D. C.	0,22
12	Fundación Universitaria Konrad Lorenz	Bogotá D. C.	0,22

Fuente: elaboración propia con los datos del Icfes (2021).

Por último, en términos de programas académicos, se obtuvieron los resultados expuestos en la tabla 4. El programa de Cine y Televisión de la Universidad Nacional de Colombia logra situar a sus estudiantes 2,97 desviaciones estándar por encima de sus pares en las pruebas Saber Pro. Entretanto, programas como Microbiología Industrial de la Javeriana, Física y Arquitectura de la Nacional, Economía de La Salle y Medicina Veterinaria, logran superar a sus pares en más de una desviación estándar. Estos resultados por programa revelan dos hechos interesantes, el primero es que universidades como la Nacional y la Sabana aportan diferenciales importantes a nivel de programa, pues logran posicionar varios de estos entre los que mayor valor agregado generan. Otro hecho destacable es que la UNAD, una universidad cuya modalidad

es a distancia, posiciona tres programas entre los veinticinco que producen mayor valor agregado en el país.

Tabla 4.

Ranking de programas al usar el modelo Propensity Score Matching

Ranking	Universidad	Programa	Puntaje
1	Universidad Nacional de Colombia	Cine y Televisión	2,97
2	Pontificia Universidad Javeriana	Microbiología Industrial	1,64
3	Universidad Nacional de Colombia	Física	1,34
4	Universidad de La Salle	Economía	1,25
5	Universidad Nacional de Colombia	Arquitectura	1,2
6	Universidad de Antioquia	Medicina Veterinaria	1,2
7	Universidad Militar Nueva Granada	Ingeniería Civil	0,9
8	Universidad Nacional Abierta y a Distancia UNAD	Zootecnia	0,85
9	Universidad Distrital Francisco José de Caldas	Matemáticas	0,83
10	Universidad Nacional Abierta y a Distancia UNAD	Licenciatura en Inglés Como Lengua Extranjera	0,83
11	Universidad de La Sabana	Comunicación Audiovisual y Multimedios	0,82
12	Universidad Nacional Abierta y a Distancia UNAD	Ingeniería Electrónica	0,81
13	Universidad de La Sabana	Licenciatura En Pedagogía Infantil	0,77
14	Universidad de La Sabana	Enfermería	0,76
15	Universidad Eafit	Psicología	0,75
16	Universidad Nacional de Colombia	Ciencia Política	0,73
17	Universidad Pontificia Bolivariana	Trabajo Social	0,71
18	Universidad del Magdalena UniMagdalena	Antropología	0,71
19	Pontificia Universidad Javeriana	Ciencias de la Información- Bibliotecología	0,7
20	Universidad Pedagógica Nacional	Licenciatura en Educación Básica con énfasis en Humanidades: Español e Inglés	0,69
21	Universidad El Bosque	Diseño Industrial	0,69
22	Universidad del Tolima	Licenciatura en Inglés	0,68
23	Universidad Externado de Colombia	Gobierno y Relaciones Internacionales	0,67
24	Corporación Universitaria Taller Cinco Centro de Diseño	Profesional en Diseño Grafico	0,67
25	Universidad de La Sabana	Fisioterapia	0,67

Fuente: elaboración propia con los datos del Icfes (2021).

Los hallazgos de este trabajo incluyen un comparativo en el nivel descriptivo y para los dos tipos de modelos que se estimaron para aproximarse al valor agregado. En esta sección, se tomaron algunos ejemplos que ilustran las técnicas utilizadas y las principales observaciones derivadas de la aplicación de metodologías estadísticas. Sin embargo, como se anotó previamente, el estudio incluyó la totalidad de instituciones de educación superior en Colombia y sus programas, por lo cual se desarrolló un aplicativo anexo para consulta en el que se pueden visualizar los resultados de cada fase del estudio.²

Comparación entre los resultados de ambas aproximaciones metodológicas

La comparación de las técnicas estadísticas utilizadas en esta investigación demuestra que, si bien a nivel general hay correspondencia entre los hallazgos del modelo de regresión lineal jerárquico y del Propensity Score Matching (ver figura 6), pues hay una correlación positiva entre ambas técnicas, a nivel individual se observan diferencias que conviene tener en cuenta en los debates sobre corrección del sesgo de selección. En la tabla 5 se presentan las IES que obtuvieron los cinco mejores y los cinco peores resultados con la técnica de PSM, cuya posición y resultados fueron comparados en la aplicación del método de regresión jerárquico. Las tendencias observadas muestran que, entre las entidades que lideran, solo la Universidad la Gran Colombia, sede Armenia, se mantuvo en la misma posición. Este resultado puede sugerir que en el método de regresión lineal jerárquica persiste el sesgo de selección. No obstante, la información observada demuestra que solo en un caso, el de la Universidad Sergio Arboleda sede Santa Marta, la variación en la posición es importante, pues con la técnica de PSM está en la segunda posición y en el modelo lineal jerárquico en la posición 22. En este caso, el ajuste de la comparación con vecinos cercanos demuestra el efecto positivo que tiene esta institución en sus estudiantes. En las IES de peor desempeño, la dispersión entre los resultados de los cálculos con PSM y Modelo lineal jerárquico es menos evidente y se observa que las posiciones y puntajes en términos de desviación estándar son cercanos, lo que evidencia el poco aporte de estas instituciones al proceso de formación de sus estudiantes. Estos casos merecen especial atención, pues requerirían del rediseño de estrategias orientadas a la mejora en la calidad educativa para que la institución contribuya al proceso de su estudiante y genere valor agregado.

² Dicho aplicativo está disponible en este enlace: <https://app.myeduapp.cloud/>

Tabla 5.

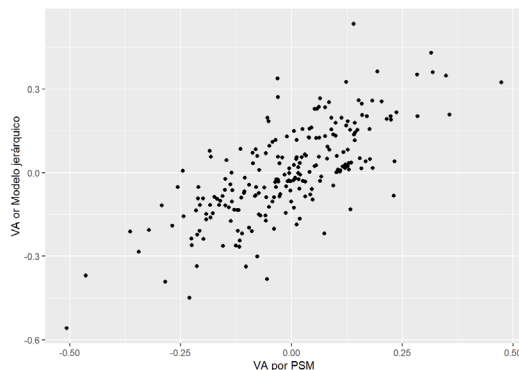
Comparativo de desempeño PSM y lineal jerárquico de instituciones con mejor y peor desempeño

IES	Resultado PSM	Ranking PSM	Resultado lineal	Ranking lineal jerárquico
Corporación Universitaria Adventista	0,47	1	0,32	9
Universidad Sergio Arboleda	0,36	2	0,21	22
Institución Universitaria Latina Unilatina	0,35	3	0,35	6
Universidad La Gran Colombia	0,32	4	0,36	4
Universidad de Antioquia	0,32	5	0,43	2
Fundación Universitaria para el desarrollo humano Unimpahu	-0,32	217	-0,21	198
Institución Universitaria Escolme	-0,34	218	-0,28	213
Escuela Militar de Cadetes general José María Córdoba"	-0,36	219	-0,21	202
Corporación Universitaria Centro superior Unicuces	-0,46	220	-0,37	217
Dirección Nacional de Escuelas	-0,51	221	-0,56	221

Fuente: elaboración propia con datos del Icfes (2021).

Figura 6.

Correlación entre valor agregado por el método de regresión jerárquica y por Propensity Score Matching.



Fuente: elaboración propia.

Conclusiones

Este artículo se aproximó a la discusión metodológica sobre valor agregado y optó por la comparación entre un diseño no experimental, el de regresión lineal jerárquica y uno cuasiexperimental, el Propensity Score Matching. Los resultados muestran que el Propensity Score Matching es efectivo en la reducción del sesgo de selección, porque compara los resultados de cada estudiante en las pruebas Saber Pro con los de sus pares más cercanos. Este ajuste metodológico deriva en resultados en valor agregado que difieren de los obtenidos mediante el método de regresión lineal jerárquica.

Es importante anotar que la comparación de métodos que se llevó a cabo en este trabajo permite tener una aproximación a la medición de la calidad a partir de la observación de los resultados en las pruebas estandarizadas y, aunque se presentaron los casos más destacados, también se observan instituciones y programas con resultados negativos en la medida de valor agregado. El análisis a profundidad de los resultados reportados en este estudio puede constituir un punto de partida para la revisión de las estrategias institucionales frente a las pruebas estandarizadas y lograr que los problemas que revelan las medidas de valor agregado sean corregidos.

En el análisis descriptivo, es posible reconocer la existencia de disparidades entre instituciones de educación superior que se evidencian en el nivel de entrada de los estudiantes a los programas de formación. El análisis por cuadrantes revela que las universidades de alto logro reciben más del 90 % de estudiantes con buen rendimiento, producto de lo cual cuentan con resultados por encima de la media en las Saber Pro. En contraste, la proporción de buenos puntajes se reduce a alrededor del 70 % en instituciones de nivel mediano; en este caso, el ejercicio de garantizar mejor rendimiento en las Saber Pro resulta más exigente, sobre todo, en casos en los que la proporción de estudiantes resilientes es baja (Mendoza *et al.*, 2024).

En cuanto a los resultados del análisis descriptivo, merece especial atención el caso de los estudiantes que entraron con buen rendimiento en las Saber 11 y obtuvieron un resultado desfavorable en las Saber Pro. El estudio de las condiciones en las que estos estudiantes empeoraron su desempeño en el puntaje global de Saber Pro puede revelar que los enfoques educativos que no consideran las características propias del estudiante son insuficientes para garantizar su buen desempeño. En razón a esto, la utilización de enfoques personalizados mediados por tecnología puede contribuir a reducir la incidencia de dichos casos (Iyer *et al.*, 2022).

A pesar de la corrección de posibles sesgos, el artículo tiene limitaciones en su aproximación conceptual al valor agregado, pues parte de pruebas estandarizadas y deja de lado aspectos como la empleabilidad (Cunha y Miller, 2014) o los tiempos de interacción con el estudiante (Sparks, 2011), que también han sido identificados como parte del valor que generan las instituciones educativas. Otro elemento que conviene revisar a futuro tiene que ver con el estudio de los campos disciplinares y su contribución al valor agregado siguiendo metodologías como la propuesta por Balcázar y Ñopo (2016), quienes se aproximaron al valor agregado en los campos relacionados con las licenciaturas o carreras enfocadas en pedagogía.

Otra posible limitación se asocia con que el periodo considerado en la elaboración de este estudio corresponde al año 2020, momento en el cual se produjo un cambio en la educación superior provocado por las políticas de confinamiento implementadas para detener el avance de la pandemia por covid-19 (Organización Mundial de la Salud, 2020). Sin embargo, el momento de la evaluación no permite medir el efecto del confinamiento sobre el rendimiento de los estudiantes en la educación superior, por lo que puede considerarse como un estudio que ofrece un panorama sobre la calidad educativa en una etapa previa al covid-19. Estudios previos han demostrado que los confinamientos llevaron a variaciones en el logro académico de los estudiantes (Ehsan, 2024) y pusieron en evidencia la importancia de atributos individuales y socioeconómicos de los estudiantes resilientes (Cheung *et al.*, 2024), por lo que es preciso que el método presentado en este estudio se aplique a los resultados obtenidos en 2022 y 2023 para evaluar las diferencias de desempeño de los estudiantes que avanzaron en su proceso de formación en educación superior durante los confinamientos.

También resulta importante la validación de las fuentes de información y la observación en el tiempo de las medidas de valor agregado. Frente a esto, Steedle (2012) llama la atención sobre la confiabilidad de los datos, la consistencia en el tiempo y la información acerca de la precisión de los puntajes. Esto acarrea que los procesos de evaluación constante de las pruebas estandarizadas sean clave para garantizar la fiabilidad de este tipo de estudios.

Este estudio se concentró en el desempeño de estudiantes que presentaron la prueba Saber Pro en un contexto en el que la pandemia por covid-19 aún no mostraba sus efectos sobre la calidad. Dado que hallazgos previos (Bartholo *et al.*, 2023; Engzell *et al.*, 2020) han dado cuenta de pérdidas en calidad asociadas con las dinámicas que siguieron las instituciones para cumplir con las políticas de aislamiento, se requiere el desarrollo de estudios comparativos que evalúen el efecto por institución y programa en una prueba posterior al covid.

Una lección importante de este trabajo es que fue posible identificar experiencias exitosas de programas en el desarrollo de competencias genéricas.

Una observación cualitativa de estos programas y un mapeo de sus buenas prácticas puede contribuir a identificar estrategias de intervención que redunden en la mejora de capacidades para la presentación de las pruebas estandarizadas en el nivel superior.

Referencias

- Alexander, L., Lopez, B., Richetti-Masterson, K. y Yeatts, K. (2015). *Selection Bias*. https://sph.unc.edu/wp-content/uploads/sites/112/2015/07/nciph_ERIC13.pdf
- Arnott, R. y Rowse, J. (1987). Peer group effects and educational attainment. *Journal of Public Economics*, 32(3), 287-305. doi: [https://doi.org/10.1016/0047-2727\(87\)90034-X](https://doi.org/10.1016/0047-2727(87)90034-X)
- Austin, P. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46(3), 399-424. doi: <https://doi.org/10.1080/00273171.2011.568786>
- Balcázar, C. y Ñopo, H. (2016). Broken Gears: The Value Added of Higher Education on Teachers' Academic Achievement. *Higher Education*, 72, 341-361. doi: <https://doi.org/10.1007/s10734-015-9960-0>
- Bartholo, T., Kolinsky, M., Tymms, P. y Lopez, D. (2023). Learning loss and learning inequality during the Covid-19 pandemic. *Ensaio: Avaliação e Políticas Públicas em Educação*, 31(119), 1-24. doi: <https://doi.org/10.1590/S0104-40362022003003776>
- Benedetto, U., Head, S., Angelini, G. y Blackstone, E. (2018). Statistical primer: propensity score matching and its alternatives. *European Journal of Cardio-Thoracic Surgery*, 53, 1112-1117. doi: <https://doi.org/10.1093/ejcts/ezy167>
- Bernal, R. y Peña, X. (2011). *Guía práctica para la evaluación de impacto*. Ediciones Uniandes.
- Blankenberger, B., Gelhausen, S. y Lichtenberger, E. (2021). Improving Institutional Evaluation Methods: Comparing Three Evaluations Using PSM, Exact and Coarsened Exact Matching. *Research in Higher Education*, 62, 1248-1275. doi: <https://doi.org/10.1007/s11162-021-09632-0>
- Bogoya, J. y Bogoya, J. (2013). An academic value-added mathematical model for higher education in Colombia. *Ingeniería e investigación*, 33(2), 76-81. doi: <https://doi.org/10.15446/ing.investig.v33n2.39521>
- Bogoya, J., Bogoya, J. y Peñuela, A. (2017). Value-added in higher education: ordinary least squares and quantile regression for a Colombian case.

- Ingengería e Investigación*, 37(3), 30-36. doi:
<https://doi.org/10.15446/ing.investig.v37n3.61729>
- Campos Torres, J., Carrillo Cárdenas, C., Otero Ceballos, M. y Risueño Rueda, M. (2017). Valor agregado en la educación de la Universidad Nacional de Colombia. *Revista papeles*, 9(17), 107-118.
- Cheung, K., Sit, P., Zheng, J., Lam, C. y Mak, S. (2024). A machine-learning model of academic resilience in the times of the Covid-19 pandemic: Evidence drawn from 79 countries/economies in the PISA 2022 mathematics study. *British Journal of Educational Psychology*, 94(4), 1224-1244. doi:
<https://doi.org/10.1111/bjep.12715>
- Cunha, J. y Miller, T. (2014). Measuring value-added in higher education: Possibilities and limitations in the use of administrative data. *Economics of Education Review*, 42, 64-77. doi:
<https://doi.org/10.1016/j.econedurev.2014.06.001>
- Díaz-Palacios, J. (2015). Re-significación y re-conceptualización de la categoría calidad educativa: una nueva mirada, una nueva perspectiva. *Revista Colombiana de Educación*, (68), 173-194. doi:
<https://doi.org/10.17227/01203916.68rce173.194>
- Duarte, J., Godoy, S. y Dueñas, X. (2016). Valor agregado en la educación primaria y secundaria: siguiendo cohortes en el tiempo. *Comunicaciones en Estadística*, 9(1), 11-41. <https://doi.org/10.15332/s2027-3355.2016.0001.01>
- Ehsan, M. (2024). *Impact of Covid-19 Lockdowns on Learning Losses: Structural and Individual Factors*. <https://www.proquest.com/openview/23eb1a24342ef4de064a5168621eafb9/1?pq-origsite=gscholar&cbl=2026366&diss=y>
- Engzell, P., Frey, A. y Verhagen, M. (2020). *Learning Inequality During the Covid-19 Pandemic*. <https://ideas.repec.org/p/osf/socarx/ve4z7.html#download>
- Heckmann, J. (1990). Selection Bias and Self-selection. En: J. Eatwell, M. Milgate y P. Newman, *Econometrics* (pp. 201-224). Palgrave Macmillan.
- Icfes. (2021). *Resultados únicos Saber Pro*. <https://www.datos.gov.co/d/u37r-hjmu>
- Iyer, L., Bharadwaj, S., Shetty, S., Verma, V. y Devanathan, M. (2022). Advancing Equity in Digital Classrooms: A Personalized Learning Framework for Higher Education Institutions. En: M. García, *Socioeconomic Inclusion During an Era of Online Education* (pp. 225-245). IGI Global.
- Liu, L., Bridgeman, B. y Adler, R. (2012). Measuring learning outcomes in higher education: motivation matters. *Educational Researcher*, 41(9), 352-362. doi:
<https://doi.org/10.3102/0013189X12459679>

- Manzi, J., San Martín, E. y Van Belleghem, S. (2013). School System Evaluation by Value Added Analysis Under Endogeneity. *Psychometrika*, 79, 130-153. doi: <http://dx.doi.org/10.1007/s11336-013-9338-0>
- Meisel, A. y Granger, A. (2022). Transición demográfica y sus consecuencias en la matrícula universitaria en Colombia. *Economía y Región*, 14(1), 1-34. doi: <http://dx.doi.org/10.32397/er.vol14.n1.1>
- Mendoza, F., Pico, C. y Arias, N. (2024). Entorno socioeconómico y resiliencia académica en la educación superior colombiana. *Revista Iberoamericana de Educación Superior*, 15(43), 123-140. doi: <https://doi.org/10.22201/iisue.20072872e.2024.43.1781>
- Monroy, A., Aguirre, C. y Espitia, A. (2018). Propuesta metodológica para identificar el valor agregado de programas de ingeniería a partir del análisis de resultados de pruebas estandarizadas. *Revista Educación en Ingeniería*, 13(25), 102-107. doi: <http://dx.doi.org/10.26507/rei.v13n25.868>
- OCDE. (2008). *Measuring improvements in Learning Outcomes: Best practices to assess the value added of schools*. París: OCDE.
- Organización Mundial de la Salud. (19 de mayo de 2020). *Concluye una histórica Asamblea de la Salud con el compromiso mundial de respuesta a la covid-19*. <https://www.who.int/es/news/item/19-05-2020-historic-health-assembly-ends-with-global-commitment-to-covid-19-response>
- Owusu, Y. y Larbi, O. (2017). Measuring Students' Learning using a Value Added Approach. *Africa Education Review*, 15(4), 99-117. doi: <https://doi.org/10.1080/18146627.2016.1224582>
- Perna, L. y Thomas, S. (2006). *A Framework for Reducing the College Success Gap and Promoting Success for All*. National Postsecondary Education Cooperative. https://nces.ed.gov/npec/pdf/Perna_Thomas_Report.pdf
- Ramírez, C. (2014). Factores asociados al desempeño académico según nivel de formación pregrado y género de los estudiantes de educación superior Colombia. *Revista Colombiana de Educación*, (66), 201-222. doi: <https://doi.org/10.17227/01203916.66rce201.222>
- Rincón, C. y Espitia, A. (2022). La Educación Superior de Colombia en riesgo: ¿Dónde están los estudiantes? *Ecos de Economía*, 24(51), 4-28. doi: <https://doi.org/10.17230/ecos.2020.51.1>
- Roback, P. y Legler, J. (2021). *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*. Taylor & Francis Group.
- Rodríguez Revilla, R. (2015). *Medición de valor agregado para la educación superior en Bogotá*. <https://repository.usta.edu.co/handle/11634/301>

- Rodríguez, R. y López, D. (2016). El valor agregado de la educación superior en la formación en segunda lengua en Colombia. *Civilizar Ciencias Sociales y Humanas*, 16(30), 119-136. doi: <https://doi.org/10.22518/16578953.538>
- Rodríguez-Revilla, R. (2022). Valor agregado y las competencias genéricas de los estudiantes de educación superior en Colombia. *Revista Iberoamericana de Educación Superior*, 13(36), 44-62. doi: <https://doi.org/10.7440/res64.2018.03>
- Rosenbaum, P. y Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. doi: <https://doi.org/10.2307/2335942>
- Schiltz, F., Sestito, P., Agasistit, T. y De Witte, K. (2018). The added value of more accurate predictions for school rankings. *Economics of Education Review*, 67, 207-215. doi: <https://doi.org/10.1016/j.econedurev.2018.10.011>
- Shavelson, R., Domingue, B., Mariño, J., Molina, A., Morales, A. y Wiley, E. (2016). On the practices and challenges of measuring higher education value-added: the case of Colombia. *Assesment and Evaluation in Higher Education*, 41(5), 695-720. doi: <https://doi.org/10.1080/02602938.2016.1168772>
- Sparks, R. (2011). A Value-Added Model to Measure Higher Education Returns On Gvernment Investment. *Contemporary Issues in Education Research*, 4(2), 15-22. doi: <https://doi.org/10.19030/cier.v4i2.4078>
- Steedle, J. (2012). Selecting value-added models for postsecondary institutional assesment. *Assesment and Evaluation in Higher Education*, 37(6), 637-652. doi: <https://doi.org/10.1080/02602938.2011.560720>
- Tam, M. (2004). Using Students' Self-reported Gains as a Measure of Value-Added. *Quality in Higher Education*, 10(3), 253-260. doi: <https://doi.org/10.1080/1353832042000299531>
- Tam, M. (2010). Measuring Quality and Performance in Higher Education. *Quality in Higher Education*, 7(1), 47-54. doi: <https://doi.org/10.1080/13538320120045076>
- Velasco, F. (2006). Modelo lineal general jerárquico. *Revista de ciencias básicas UJAT*, 4(2), 20-28. <https://dialnet.unirioja.es/descarga/articulo/9493621.pdf>
- Zerpa, C. y Rodríguez-Montoya, C. (2024). Disminución de la matrícula universitaria por deserción: una revisión estructurada. *Ciencia y Educación*, 8(1), 59-78. doi: <https://doi.org/10.22206/cyed.2024.v8i1.2920>