



Informe técnico



Medición de los efectos de la educación superior en Colombia sobre el aprendizaje estudiantil

Informe técnico
Medición de los efectos de la
educación superior en Colombia
sobre el aprendizaje estudiantil

Presidente de la República
Iván Duque Márquez

Ministra de Educación Nacional
María Victoria Angulo González

Viceministra de Educación
Preescolar, Básica y Media
Constanza Alarcón Párraga



Directora General
Mónica Ospina Londoño

Secretario General
Ciro González Ramírez

Directora de Evaluación
Natalia González Gómez

Director de Producción y Operaciones
Oscar Orlando Ortega Mantilla

Director de Tecnología e información
Sergio Andrés Soler Rosas

Subdirector de Diseño de Instrumentos
Luis Javier Toro Baquero

Subdirectora de Análisis y Divulgación
Mara Brigitte Bravo Osorio

Subdirectora de Estadísticas
Jeimy Paola Aristizabal Rodríguez

Subdirectora de Producción de Instrumentos
Nubia Rocío Sánchez Martínez

Subdirectora de Aplicación de Instrumentos
Yamile Ariza Luque

Subdirector de Información
Daniel Betancur Salazar

Subdirector de Abastecimiento
y Servicios Generales
Hans Ronald Niño García

Subdirector Financiero y Contable
Javier Santos Pacheco

Subdirector de Desarrollo de Aplicaciones

Subdirectora de Talento Humano
María Mercedes Corcho Caro

Jefe Oficina Asesora de
Comunicaciones y Mercadeo
María Paula Vernaza Díaz

Jefe Oficina de Gestión de
Proyectos de Investigación
Clara Lorena Trujillo Quintero

Jefe Oficina Asesora de Planeación
Luis Alberto Colorado Aldana

Jefe Oficina de Control Interno
Adriana Bello Cortés

Jefe Oficina Asesora Jurídica
Ana María Cristina de la Cuadra

Asesora Unidad Atención al Ciudadano
Alba Liliana Abril Daza

Elaboración
**Oficina de Gestión de Proyectos
de Investigación**

Revisión
Luis Eduardo Jaramillo Flechas

Procesamiento estadístico
Juan José Rubio Mesa

Diagramación e ilustración
Mónica Liliana López

ISBN: En trámite
Bogotá D.C., abril 2021
<http://www.icfes.gov.co/>
<https://www.facebook.com/icfescol>
<https://www.instagram.com/icfescol/>
<https://twitter.com/ICFEScol>

Todos los derechos de autor reservados ©..



Términos y condiciones de uso para las publicaciones y obras de propiedad del Icfes

El Instituto Colombiano para la Evaluación de la Educación (Icfes) pone a la disposición de la comunidad educativa y del público en general, de forma gratuita y libre de cualquier cargo, un conjunto de publicaciones a través de su portal www.icfes.gov.co. Dichos materiales y documentos están normados por la presente política y están protegidos por derechos de propiedad intelectual y derechos de autor a favor del Icfes. Si tiene conocimiento de alguna utilización contraria a lo establecido en estas condiciones de uso, por favor infórmenos al correo prensaicfes@icfes.gov.co

Queda prohibido el uso o publicación total o parcial de este material con fines de lucro. Únicamente está autorizado su uso para fines académicos e investigativos. Ninguna persona, natural o jurídica, nacional o internacional, podrá vender, distribuir, alquilar, reproducir, transformar*, promocionar o realizar acción alguna de la cual se lucre directa o

*La transformación es la modificación de la obra a través de la creación de adaptaciones, traducciones, compilaciones, actualizaciones, revisiones y, en general, cualquier modificación que de la obra se pueda realizar, de modo que la nueva obra resultante se constituya en una obra derivada protegida por el derecho de

indirectamente con este material. Esta publicación cuenta con el registro ISBN (International Standard Book Number, o Número Normalizado Internacional para Libros) que facilita la identificación no sólo de cada título, sino de la autoría, la edición, el editor y el país en donde se edita.

En todo caso, cuando se haga uso parcial o total de los contenidos de esta publicación del Icfes, el usuario deberá consignar o hacer referencia a los créditos institucionales del Icfes respetando los derechos de cita; es decir, se podrán utilizar con los fines aquí previstos transcribiendo los pasajes necesarios, citando siempre la fuente de autor, lo anterior siempre que estos no sean tantos y seguidos que razonadamente puedan considerarse como una reproducción simulada y sustancial, que redunde en perjuicio del Icfes.

autor, con la única diferencia respecto de las obras originales de que aquellas requieren para su realización de la autorización expresa del autor o propietario para adaptar, traducir, compilar, etcétera. En este caso, el Icfes prohíbe la transformación de esta publicación.

Así mismo, los logotipos institucionales son marcas registradas y de propiedad exclusiva del Instituto Colombiano para la Evaluación de la Educación (Icfes). Por tanto, los terceros no podrán usar las marcas de propiedad del Icfes con signos idénticos o similares respecto a cualquier producto o servicios prestados por esta entidad, cuando su uso pueda causar confusión. En todo caso queda prohibido su uso sin previa autorización expresa por Icfes. La infracción de estos derechos se perseguirá civil y, en su caso, penalmente, de acuerdo con las leyes nacionales y tratados internacionales aplicables.

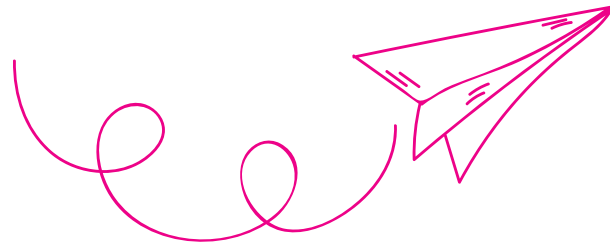
El Icfes adelantará las acciones legales pertinentes por cualquier violación a estas políticas y condiciones de uso.



Contenido

Introducción	5
1. Aspectos conceptuales	
1.1. Supuestos subyacentes a las estimaciones de medidas de valor agregado	7
1.2. Limitaciones y consideraciones en los análisis de valor agregado	8
1.2.1. Tratamientos	8
1.2.2. Unidades	9
1.2.3. Resultados	10
1.3. Desafíos técnicos: sesgo	12
2. Aspectos metodológicos y modelo de valor agregado y aporte relativo	
2.1. Estimación del modelo	14
2.2. Datos	16
2.2.1. Comparabilidad Saber 11	16
2.2.2. Relación entre las pruebas	17
2.2.3. Selección de la unidad de análisis	18
3. Resultados	
3.1. Relación entre modelos	24
Referencias	27
Anexos	30

Introducción



Los modelos de valor agregado (VA) son especificaciones estadísticas, a partir de las cuales se estima la efectividad educativa haciendo énfasis en el progreso de los estudiantes en el tiempo. En este sentido, el valor agregado de una institución se entiende como la contribución de esta al progreso neto de los estudiantes, hacia objetivos de aprendizaje establecidos, una vez se controla por la influencia de factores ajenos a la institución que pueden contribuir a ese progreso (Martínez R., Gaviria J., Castro M., 2009).

Los exámenes Saber 11° y Saber Pro ofrecen mediciones comparables sobre el nivel de competencias con el que los estudiantes inician y finalizan el ciclo de educación postsecundaria, lo cual permite evidenciar el valor agregado del sistema de educación superior en Colombia. Lo anterior contrasta con la situación que se presenta en la mayoría de los países, dado que es inusual encontrar exámenes estandarizados

y comparables entre la educación secundaria y superior. Existen excepciones importantes, como la Evaluación del Aprendizaje en la Educación Superior en Estados Unidos (Steedle, 2012). Sin embargo, al no ser una prueba periódica y censal, los datos de CLA⁽¹⁾ proveen análisis de alcance limitado.

Teniendo en cuenta el potencial de los datos que genera el sistema de evaluación de la educación en Colombia y la importancia de medir la efectividad educativa, a desde un enfoque que reconozca la trayectoria académica previa de los estudiantes, en 2012 el Icfes inició el estudio de valor agregado y aporte relativo⁽²⁾. Este proceso resultó en la publicación de los primeros reportes de aporte relativo en 2015. Dichos reportes, vinieron acompañados de un anexo técnico⁽³⁾ (Icfes, 2015), sin embargo, entre 2017 y 2020, se han incorporado algunos cambios que ameritan la publicación de este **nuevo anexo técnico**. Este que es una extensión del anterior, en el cual podrá encontrar todos los aspectos

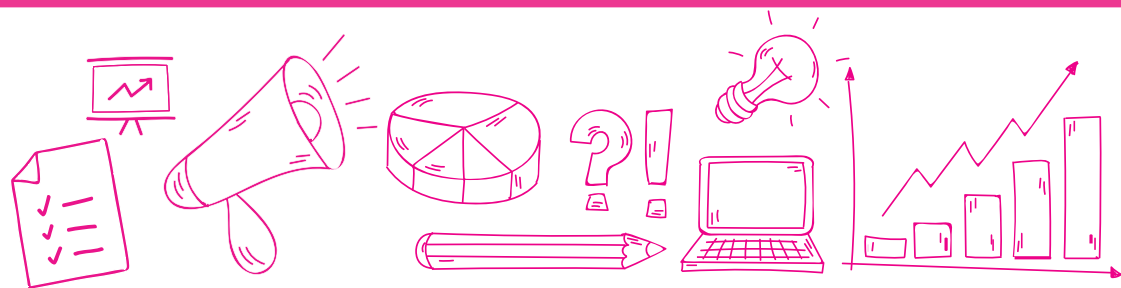
metodológicos concernientes a los reportes de aporte relativo, contruidos a partir de los resultados de los exámenes de Saber Pro Desde el año 2017.

Este anexo técnico, se compone de tres secciones. En la primera, se analizan los aspectos conceptuales del marco tradicional de inferencia causal para los modelos de valor agregado, los desafíos identificados y los aspectos relacionados con los sesgos que han motivado las decisiones tomadas en esta investigación. En la segunda, se analizan cuestiones metodológicas, la descripción del modelo de valor agregado y aporte relativo usado por el Icfes, haciendo énfasis en los aspectos relacionados con los datos y los modelos de estimación empleados en este estudio. Por último, en la tercera parte, se presentan algunos resultados del modelo descrito en la sección anterior.

¹ CLA se refiere al Collegiate Learning Assessment que es el examen estandarizado de educación superior que se hace en Estados Unidos. Unos renglones arriba sale como "Evaluación del Aprendizaje en la Educación Superior"

² El estudio de aporte relativo del Icfes es una modificación a los modelos de valor agregado, en el cual se controla no solo por las condiciones iniciales de los estudiantes, sino también por el tipo de población que atienden las IES, esto medido como el puntaje promedio de Saber 11 de los estudiantes de cada institución.

³ Los análisis y conclusiones del documento técnico anterior siguen siendo válidos y pueden ser consultados. Este documento es una actualización de el.



1

Aspectos conceptuales

1.1. Supuestos subyacentes a las estimaciones de medidas de valor agregado

En contextos observacionales, las inferencias causales válidas requieren diversos supuestos.

Reardon y Raudenbush (2009) identifican los siguientes supuestos para una inferencia causal válida con modelos de valor agregado:

a. Manipulabilidad: teóricamente, los estudiantes podrían estar expuestos a cualquier tratamiento.

b. No interferencia entre unidades: el resultado de un estudiante depende solamente de su asignación a un determinado tratamiento.

c. El supuesto métrico: los resultados del puntaje de la prueba están en una escala de intervalo.

d. Homogeneidad: el efecto causal no varía como función de una característica del estudiante.

e. Tratamiento altamente ignorable: la asignación al tratamiento es esencialmente aleatoria, después de condicionarse por las variables de control.

f. Forma funcional: la forma funcional empleada para el control de las características del estudiante es la correcta.

En el supuesto (a), se requiere que sea factible que un estudiante tenga un resultado potencial para todas las posibles instituciones de educación superior (IES). En este sentido, los estudiantes son “manipulables” (Estos podrían asistir a todas las IES disponibles). Este problema está presente en la educación superior dado que gran parte de las universidades solamente matriculan a los estudiantes seleccionados. ¿Si una universidad le niega la admisión a un estudiante, se puede afirmar razonablemente que este estudiante tiene un resultado potencial en esa institución?

Es por este supuesto que dentro del estudio realizado en el Icfes se emplea el concepto de vecindad de comparación, que se explicará más adelante.

En el supuesto (b) se requiere que los resultados potenciales de un estudiante no dependan de las IES a las que asisten otros estudiantes. Este supuesto, también conocido como el supuesto de estabilidad del valor de tratamiento unitario (SUTVA, por su sigla en inglés), no toma en cuenta los efectos del grupo par. El efecto de la escuela sobre el estudiante es independiente de las características de sus compañeros: si son estudiantes fuertes o débiles, ricos o pobres, etcétera. La investigación de los efectos del grupo par en la educación superior (Winston y Zimmerman, 2004; Sacerdote, 2011) sugiere que existen efectos moderados, de manera que el supuesto (b) no puede ser sostenible. Una alternativa para lidiar con este dilema es considerar que el tratamiento (pertenencia a una IES particular) se aplica a un grupo de estudiantes y no a un estudiante individualmente, como proponen Rubin, Stuart y Zanutto (2004). Esto lleva a cambios sustanciales en la interpretación de los resultados, pues las estimaciones solo tomarían en cuenta la asistencia a una institución.

El supuesto (c), establece que los puntajes de la prueba están en una escala de intervalo, que es el caso de los puntajes de los exámenes de Saber Pro y Saber 11°.

El supuesto (d), establece que los efectos de la escuela son una constante independiente de los antecedentes del estudiante. Este supuesto puede ser potencialmente erróneo con respecto a la educación superior, debido a que los antecedentes de los estudiantes pueden influenciar no solo el desempeño previo sino también el desempeño posterior. En Raudenbush y Willms (1995) se discute que este supuesto debe ser objeto de una cuidadosa investigación en los estudios del valor agregado.

En el supuesto (e) se requiere que la asignación a las instituciones pueda tratarse aleatoriamente después de controlar por las covariables. Esto pone un peso

sustancial sobre cómo elegir las covariables correctas. Por ejemplo, la inclusión de una variable como el nivel socioeconómico (NSE) podría tener sentido dada la asociación entre la riqueza y la asistencia a ciertos tipos de instituciones. Sin embargo, la inclusión de tal covariable puede ser poco atractiva puede ser poco pertinente en un escenario de rendición de cuentas, puesto que equivale a establecer expectativas diferentes para los estudiantes en función de su NSE.

Esta tensión entre la inclusión o exclusión de ciertas covariables puede corregirse con la inclusión de muchas covariables a nivel del estudiante, ya que las estimaciones suelen ser poco sensibles a su inclusión (Ballou, Sanders y Wright, 2004). Es decir, las estimaciones de los efectos de la institución no se afectan dramáticamente por la inclusión de covariables adicionales, después de incluir un control apropiado de los logros previos. Sin embargo, cuando se consideran covariables a nivel de institución, las estimaciones pueden cambiar dramáticamente, debido a que se estaría cambiando la interpretación de los resultados de los modelos.

Por último, el supuesto (f) afirma que la forma funcional de la ecuación de regresión es la correcta. Por ejemplo, los puntajes previos a menudo se incluyen

como predictores lineales y, por tanto, se asume que la verdadera relación entre los resultados y los puntajes previos es lineal.⁽⁴⁾ En relación con este supuesto, quizás lo más importante es que se supone que hay suficiente “soporte común”, es decir, que a cada universidad asiste un número suficiente de estudiantes a través de las distribuciones de los controles relevantes (NSE, aptitud académica, etcétera). Lo anterior se relaciona con el supuesto (a). Si muy pocos estudiantes con ciertas características (por ejemplo, NSE alto) asisten a una escuela determinada, entonces el efecto estimado de asistir a esa institución dependerá en gran medida de la especificación del modelo de regresión más que de las comparaciones entre unidades comparables en diferentes instituciones. Los problemas que acarrear los supuestos descritos cuando existen datos escasos son notables en la literatura sobre inferencia causal (Ho & Rubin, 2011).

1.2. Limitaciones y consideraciones en los análisis de valor agregado

Existen varias implicaciones sobre los análisis de valor agregado, principalmente, derivadas de las decisiones operativas sobre los supuestos expuestos anteriormente. En este sentido, se presenta un conjunto de consideraciones técnicas que deben ser tenidas en cuenta en el análisis de valor agregado.

1.2.1. Tratamientos

La especificación de una condición de tratamiento requiere que se seleccione una condición de control que la acompañe, es decir, una condición que sirva como punto de comparación. ¿Debe compararse a la institución educativa j con una institución “promedio”, con otra institución a la que el estudiante i podría razonablemente asistir, o con otra institución a la que definitivamente no podría asistir?, ¿deben compararse carreras académicas dentro de la misma institución?

La opción estándar en la literatura de valor agregado es permitir que la condición de control, o el contrafactual, se obtenga a partir de una institución “promedio”. Esta elección tiene implicaciones para la selección de las covariables, debido a que estas se utilizan para definir la institución “promedio” (Manzi, San Martín y Van Belleghem, 2014).

Una preocupación en la educación superior es que naturalmente, hay más variabilidad en la duración de la exposición al tratamiento, en comparación a la que hay en el ciclo de educación primaria y secundaria. En promedio, completar la educación universitaria puede tomar de 4 a 8 años (o incluso más). Esta variabilidad es inevitable, pero puede plantear desafíos. Por ejemplo, los estudiantes que tardan más tiempo para obtener un título

⁴ Ver Lockwoody McCaffrey (2015) para explorar la inclusión de términos de orden superior.

pueden haber estado trabajando, cambiado de carrera o haber interrumpido momentáneamente sus estudios. Por consiguiente, el cambio en las habilidades de los estudiantes no se puede atribuir exclusivamente a las instituciones, debido a que el progreso de aprendizaje pudo haber estado influenciado por experiencias exógenas a la institución.

Otra consideración, con respecto a la definición del tratamiento, es la especificación de los aspectos institucionales que se intentan evaluar. Raudenbush (2004), separa los efectos causales que las instituciones pueden tener sobre los estudiantes en efectos de práctica y contexto. Los primeros se relacionan con el aprendizaje debido a prácticas pedagógicas intencionales, mientras que los últimos se deben al contexto en el cual funciona la institución.

Raudenbush y Willms (1995) describen dos estrategias (denominadas modelos tipo A y B) que ayudan a analizar estos dos tipos de efectos causales. Los modelos tipo A, incluyen tanto el efecto del contexto como el de la práctica, pero no pretenden analizar la proporción en que cada uno de esos componentes contribuyen al aprendizaje de los estudiantes. Estos modelos pueden ser útiles para estudiantes y padres de familia, pues ofrecen una estimación del resultado final del aprendizaje que tendrían.

Por otro lado, los modelos tipo B solo miden el efecto de la práctica. Estos podrían ser de interés para los

encargados de las políticas públicas en educación, por ejemplo, debido a que describen los efectos sobre el aprendizaje derivados de los elementos que una escuela o institución educativa puede controlar como la calidad de la enseñanza y las expectativas para el éxito de los estudiantes. En estos modelos no se responsabiliza a las escuelas por los efectos de pares.

En el caso de la educación superior es posible que los modelos tipo A sean de interés tanto para los estudiantes como para los responsables de las políticas, ya que las IES controlan el ingreso de sus estudiantes y, por tanto, se podría afirmar que causan el aprendizaje contextual. En un sentido técnico, Raudenbush y Willms (1995) señalan que es difícil obtener estimaciones tipo B de manera adecuada debido a que típicamente la práctica escolar no es observable.

1.2.2. Unidades

La elección de la unidad de análisis se relaciona con el debate sobre el supuesto (b). Específicamente, responde a si el tratamiento debe aplicarse a estudiantes individuales o a grupos de estudiantes. Un enfoque útil para analizar esta problemática es considerar cuál sería la forma ideal de realizar un experimento aleatorio. Es decir, el tratamiento se aplicaría a grupos de estudiantes conformados aleatoriamente o a grupos de estudiantes que existen dentro de los datos observacionales. La

respuesta a esta pregunta depende del enfoque sobre los efectos de los pares o compañeros. Si pensamos que estos conforman un componente inevitable del aprendizaje de los estudiantes, solo la segunda opción es sostenible.

El supuesto de SUTVA tiene implicaciones en la selección de las unidades de análisis. Rubin et al. (2004) señalan que el SUTVA requiere que haya solamente una versión del tratamiento. En universidades, donde los estudiantes tienen muchas opciones diferentes para el enfoque académico, pueden surgir interrogantes sobre la existencia de un único tratamiento para todos los estudiantes. Si el tratamiento se define como la obtención de un grado en una institución particular, la implicación es que el tratamiento no varía entre los estudiantes de esa institución. Sin embargo, se podría contrargumentar que obtener títulos diferentes en una misma universidad, por ejemplo, un título de Ingeniería y a uno de Historia del Arte, implica realizar tratamientos muy distintos.

A pesar de las consideraciones anteriores, reducir la definición de unidades de tratamiento dificulta el análisis de valor agregado. En primer lugar, un análisis minucioso lleva a una reducción del tamaño de la muestra y de la precisión de las estimaciones. En segundo lugar, si una universidad está dividida en múltiples unidades, la atribución causal resulta

más compleja para los estudiantes que cambian de programa académico o que toman clases en múltiples facultades.

Considere esta simple ilustración: imagínesse un estudiante de literatura en una Facultad de Artes Liberales. Él cursa dos clases de matemáticas, ambas en la Facultad de Ciencias Naturales, mientras asiste a la universidad. Si la medición de los conocimientos matemáticos es un resultado relevante, ¿cuál debería ser la facultad responsable del aprendizaje de este estudiante?

1.2.3. Resultados

Establecer las medidas de resultados pertinentes en la educación superior es complejo, debido a la diversidad de aprendizajes que se desarrollan. Una posibilidad son las competencias genéricas, es decir, aquellas que son comunes entre diferentes programas académicos, como el caso de la Lectura Crítica. El uso de tales medidas posibilita las comparaciones entre diferentes campos de estudio. Sin embargo, ofrece poca información sobre aprendizajes propios del campo de formación al que se encuentran adscritos los estudiantes (competencias específicas), lo que representa un componente clave del tratamiento. Desarrollar las evaluaciones a través de un rango de campos podría ser costoso, pero benéfico cuando se considere crucial el énfasis en las competencias específicas dentro de un campo. Además de los desafíos relacionados con los tratamientos, las unidades y los resultados, se identifican tres aspectos

adicionales que merecen más atención cuando se pretende estimar la efectividad de las IES. Estos son la elección de covariables, datos faltantes y la selección de estudiantes. Estos aspectos están estrechamente relacionados con los supuestos de este tipo de modelos y constituyen la base para las preguntas de investigación que subyacen al análisis empírico.

a. Selección de covariables

La elección de las covariables, en un análisis de valor agregado, es una de las primeras preguntas en las investigaciones de este tipo (Ballou et al., 2004).

La principal consideración es cómo dar cuenta del logro previo a nivel del estudiante. En los estudios de valor agregado en educación primaria, los investigadores suelen utilizar el puntaje previo en una materia como único indicador de habilidades previas para evaluar el valor agregado logrado en la misma materia.

Pero, incluso en la configuración principal, hay buenas razones para considerar la inclusión de otros puntajes previos, pues reducen el efecto del error de medición (Lockwood y McCaffrey, 2014).

En educación superior es difícil encontrar una misma medida de habilidades previas disponible para todos los resultados a analizar y para todos los estudiantes, especialmente cuando son las pruebas específicas las que se usan como medidas de resultados. Incluso si se

1 Aspectos conceptuales



tiene disponible una prueba directamente comparable, el lapso entre la prueba previa y el resultado de las IES puede reducir la magnitud de las correlaciones entre los puntajes previos y los resultados (un mayor lapso entre las pruebas tiende a reducir la precisión de la estimación de los efectos (Leckie y Goldstein, 2009).

En general, se considera que la inclusión o exclusión de covariables adicionales a nivel del estudiante tiene efectos mínimos en las estimaciones de VA, una vez que los logros previos se han incluido de forma adecuada (Ballou et al., 2004).

Una segunda consideración es el uso de covariables contextuales o agregadas (por ejemplo, el promedio de la habilidad previa para todos los estudiantes que ingresan a una misma institución). Incluir



tales covariables, se cree, es una forma de producir estimaciones tipo B, aunque no necesariamente abordan las dificultades observadas por Raudenbush y Willms (1995), pues la práctica pedagógica que tendrá lugar durante el proceso de formación todavía no ha sido observada. En cambio, sí se sabe que la inclusión de tales covariables agregadas puede cambiar sustancialmente el rango de las estimaciones de VA resultantes, en comparación con los modelos sin covariables contextuales. Además, la inclusión de tales covariables tiene el efecto de cambiar la interpretación de las estimaciones del Valor Agregado, cambiando sutilmente la definición del contrafactual (Manzi et al., 2014).

b. Datos faltantes

En un gran porcentaje, los estudiantes universitarios abandonan sus carreras. En Colombia, la tasa promedio de deserción anual está alrededor del 10% (MEN, 2017). Esta pérdida de estudiantes puede sesgar las estimaciones. Aún si se hace caso omiso al sesgo, es importante considerar que algunos programas cuentan con tasas de deserción más altas que otros, de este modo, obtener un título puede ser tan significativo como el aprendizaje que obtiene un estudiante durante su formación (Spence, 1973).

c. Selección de estudiantes

La literatura de valor agregado en educación primaria y secundaria se ha enfocado en la selección de los estudiantes como una amenaza significativa para la inferencia causal (Rothstein, 2010; Briggs y Domingue, 2011; Chetty et al., 2011; Koedel y Betts, 2011; Goldhaber y Chaplin, 2012). Aunque el supuesto de grupos de estudiantes formados al azar dentro de las universidades es claramente insostenible, sigue siendo importante examinar la distribución del logro previo, a través de las unidades de análisis (IES, grupos de referencia, núcleos básicos del conocimiento, etc). para comprender el nivel de la selección no aleatoria. Si la distribución de habilidades previas de los estudiantes (capturadas por los puntajes obtenidos en pruebas anteriores) es relativamente equitativa entre unidades de análisis, entonces la inferencia causal se facilita, porque

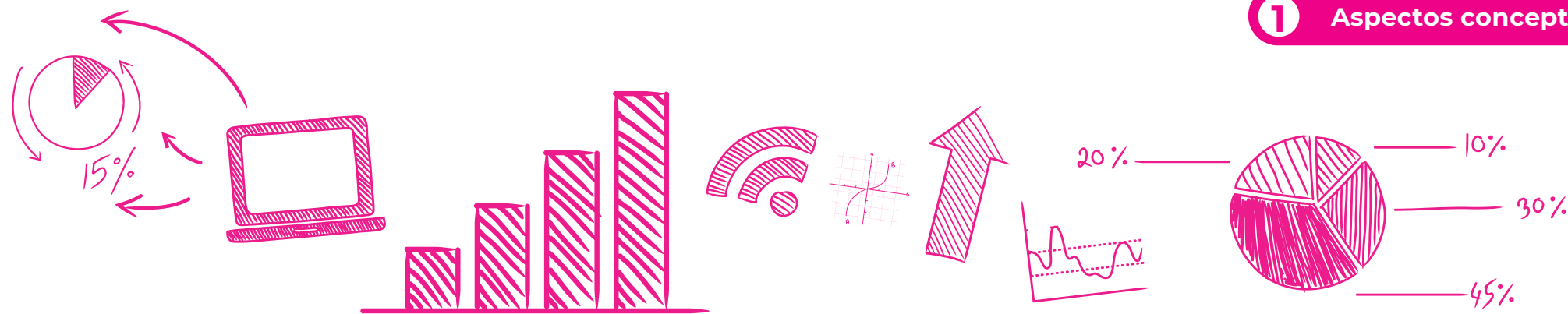
1 Aspectos conceptuales

se asigna menos peso a la forma funcional de la ecuación de regresión si se ha cumplido el supuesto (f) de “soporte común”. Para ilustrarlo, imagine un escenario en el que a cada institución asiste un grupo de estudiantes con un nivel idéntico de habilidades previas. En tal caso, no es necesario ningún ajuste de regresión y bastaría una comparación de los resultados promedio, pues no existirían diferencias importantes entre las habilidades de los estudiantes que entran a diferentes instituciones.

1.3. Desafíos técnicos: sesgo

El sesgo es un aspecto clave al hacer una investigación de valor agregado. En particular, este se puede observar en las estimaciones de los parámetros, debido a diversos factores. Esta sección se enfoca en tres tipos de sesgo:

- Sesgo en los coeficientes de los puntajes previos, debidos al error de medición.
- Sesgo en las estimaciones del efecto escolar para IES con características contextuales favorables, efectos tipo B.
- Sesgo en las estimaciones del efecto escolar para IES con características contextuales desfavorables.



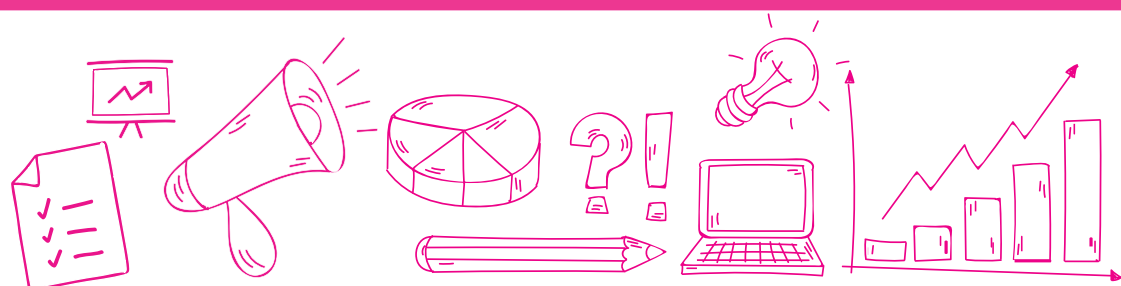
Un error de medición en una variable predictora es un problema bien conocido en la modelación estadística. En el contexto de los modelos de valor agregado, Lockwood y McCaffrey (2014) analizan los problemas relevantes. En general, la preocupación es que el error de medición en la prueba previa generará una estimación atenuada. En particular, podría preocupar que se sobreestimen los efectos escolares y que estos resulten altamente correlacionados, de manera artificial, con el resultado. Es posible realizar ajustes técnicos como SIMEX (Cook y Stefanski, 1994), pero su implementación en la educación superior podría ser desafiante, porque los puntajes de las pruebas previas pueden provenir de varias aplicaciones diferentes (las cuales pueden presentar diferentes niveles de errores de medición).

En general, si hay una correlación positiva entre el efecto escolar y la capacidad promedio en la prueba previa, entonces las estimaciones tipo B estarán sesgadas negativamente para las instituciones con un promedio alto de habilidades al entrar (Raudenbush y Willms, 1995). Una correlación positiva podría ser razonable en este caso. Por ejemplo, los estudiantes cuyo promedio de habilidades previas sea alto podrían venir de entornos socioeconómicos más aventajados y por tanto estar más dispuestos a invertir en su educación, lo que podría influenciar positivamente la calidad de los servicios ofrecidos por las instituciones a las que asistirían.

Es importante distinguir los efectos de la práctica escolar de los generados por los pares (compañeros), es decir, el aumento potencial del aprendizaje que los estudiantes pueden experimentar debido a sus compañeros. Esta diferencia la exploran Castellano, Rabe-Hesketh y Skrondal (2014).

En síntesis, los autores demuestran que las estimaciones del valor agregado de las instituciones estarán sesgadas si existe una correlación entre los pares y la variable de resultado, pero no hay un efecto contextual. Por otra parte, los estimadores (HT) de Hausman-Taylor (Hausman y Taylor, 1981) son consistentes. No obstante, si hay un efecto contextual, entonces las estimaciones HT estarán sesgadas, aunque sirven como límites para el verdadero efecto.

Finalmente, vale la pena hacer énfasis en que la mayor fuente potencial de problemas por sesgo proviene de subestimar sistemáticamente las mediciones del valor agregado para las instituciones que atienden a ciertas poblaciones, como los estudiantes de bajos ingresos. Esto podría llevar a una alta correlación entre las covariables y las medidas de VA, como se discutió en el contexto de errores de medición. La evidencia de que existe tal sesgo ha sido documentada en EE.UU. (Newton et al., 2010; Ehlert, Koedel, Parsons y Podgursky, 2012). Existe la preocupación de que tal sesgo podría existir incluso si se han incluido controles para las características de los antecedentes de los estudiantes.



2 Aspectos metodológicos y modelo de valor agregado y aporte relativo

Para desarrollar el modelo de valor agregado y aporte relativo se construye una base de datos con los registros de Saber Pro y Saber 11° de los estudiantes. Lo primero que se hace para construir esta base es tomar todos los individuos para los cuales coincide el número de documento de identidad registrado en el examen Saber 11° con aquel registrado en el examen Saber Pro. Debido al cambio en el documento de identidad que se da para muchos estudiantes entre la presentación de los dos exámenes (esto es, el paso de tarjeta de identidad a cédula de ciudadanía) y a que los estudiantes suelen escribir su nombre de manera diferente en ambos exámenes, se debe realizar un cruce probabilístico entre las dos bases de datos. Por tales razones, no es posible encontrar con exactitud los resultados de Saber 11° de todos aquellos individuos que presentan Saber Pro.

Para este cruce se tienen en cuenta todos los individuos que presentaron el Examen Saber 11° entre 4 y 8⁽⁵⁾ años antes de presentar el examen Saber Pro. Se limita el tiempo entre los dos exámenes para evitar atribuir al valor agregado de las instituciones experiencias externas de los estudiantes que no pueden ser observadas. Posterior a realizar el cruce, se excluyen del estudio los registros que tienen dos o más módulos de Saber Pro con un puntaje igual a 0 o sin calificación, ya que son personas que por alguna razón no respondieron la prueba y pueden sesgar las estimaciones.

Una vez se tienen en cuenta todos los filtros mencionados, se procede a hacer la estimación del modelo de valor agregado para aquellas instituciones dentro de un Núcleo Básico del Conocimiento (NBC), en las que existan al menos 25 estudiantes que hayan presentado Saber Pro, cuenten con su respectivo cruce en Saber 11°

y representen el 40% o más del total de la población de la institución en el NBC correspondiente. Esto último se hace con el fin de tener suficientes observaciones para hacer cálculos precisos y asegurar que los estudiantes sean una buena representación de la población atendida por los programas de una misma IES que pertenezcan a un NBC particular (INBC). Por último, solo se hacen las estimaciones de valor agregado en aquellos NBC que cuenten con al menos cinco instituciones y en los que se cumplan las situaciones anteriores para los dos años considerados en el reporte (ej. 2017-2018).

2.1. Estimación del modelo

En este estudio se definió como tratamiento que un estudiante i perteneciente al NBC k , dentro de la universidad j , sea elegible para presentar el examen Saber Pro, es decir, que haya cursado al menos el 75% de los créditos de su programa académico, lo que lo convierte en un estudiante que podría ser potencialmente evaluado). Es importante considerar que los modelos se estiman dentro de los NBC, es decir, cada modelo se estima teniendo en cuenta solo las instituciones que tienen programas pertenecientes al NBC k . Esto ayuda a garantizar que el tratamiento sea de una sola forma (en un intento por acercarse más al cumplimiento de la hipótesis del SUTVA) y que el enfoque se centre en la comparación que se considera más relevante para la toma de decisiones pedagógicas, administrativas y de política educativa.

⁵ Para medicina se extiende a 9 años.

Como consecuencia de este enfoque, los resultados no son comparables entre diferentes NBC. Para facilitar la comprensión del modelo, es importante tener en cuenta el modelo nulo, al que también se le conoce como especificación vacía, porque no tiene en cuenta variables explicativas y se usa para descomponer el resultado del estudiante en sus niveles de anidación. De esta forma, el modelo nulo se puede representar de la siguiente forma:

$$y_{ijk} = \mu_k + \gamma_{jk} + \varepsilon_{ijk}$$

Donde y_{ijk} es el resultado del estudiante i en la institución j dentro del NBC k en alguno de los módulos de Saber Pro (Razonamiento Cuantitativo, Lectura Crítica o Inglés); μ_k se refiere al promedio en el módulo de Saber Pro evaluado; γ_{jk} representa la desviación de la institución j en el NBC k del promedio general, es decir, cuánto más o cuánto menos es el promedio del INBC jk frente al promedio general; por último, ε_{ijk} es la desviación del estudiante i del promedio de su institución en el NBC k , es decir, cuanto más o cuanto menos es el puntaje del estudiante respecto al puntaje promedio de su institución. Estas dos últimas estimaciones se ven como variabilidad intergrupo (ε_{ijk}) e intragrupo (γ_{jk}).

Normalizando $\gamma_{jk} \sim N(0, \tau)$ y $\varepsilon_{ijk} \sim N(0, \sigma^2)$, se puede construir el índice de correlación intraclase (ICC) con los resultados del modelo nulo. Este indica qué parte de la varianza del resultado de los estudiantes es explicada por la varianza de las características de sus INBC, de tal forma que entre más alto sea este índice más influencia tiene el INBC en el resultado de los estudiantes, mientras que, si tiende a cero, el resultado puede atribuirse casi por completo a las características de los individuos. Así, el ICC se calcula de la siguiente forma:

$$ICC = \frac{\tau}{\tau + \sigma^2}$$

⁶El INSE es el índice de nivel socioeconómico construido por el Icfes, a partir de la información recogida de los estudiantes en los cuestionarios socioeconómicos de los exámenes (Icfes, 2019).

Para calcular el VA del INBC se utiliza la siguiente función para el estudiante i en la institución j , que ofrece programas académicos del NBC k :

$$y_{ijk} = N(\beta' X_{ijk} + \gamma_{jk} + \varepsilon_{ijk})$$

Donde y_{ijk} es el resultado en Saber Pro del individuo i y X_{ijk} es un vector de covariables. Los efectos de interés vienen dados por γ_{jk} , es decir un vector con el valor agregado de cada INBC jk (se usan técnicas de estimación de efectos aleatorios). Concretamente, X_{ijk} representa un vector de los puntajes obtenidos por el estudiante en el examen Saber 11° (Matemáticas, Lectura Crítica, Ciencias Naturales, Sociales y Ciudadanas e Inglés). Así, se aíslan del resultado del estudiante sus condiciones iniciales, permitiendo calcular la influencia del INBC en su desempeño. La ecuación a estimar es:

$$y_{ijk} = N(\beta_0 + \beta_1 \text{matemáticas}_{ijk} + \beta_2 \text{lectura crítica}_{ijk} + \beta_3 \text{ciencias naturales}_{ijk} + \beta_4 \text{sociales y ciudadanas}_{ijk} + \beta_5 \text{inglés}_{ijk} + \gamma_{jk} + \varepsilon_{ijk})$$

Este sería un modelo tipo A en la tipología de Raudenbush y Willms (1995), porque no están especificados efectos del contexto a nivel del INBC (covariables a nivel del promedio de los estudiantes del INBC). Es importante tener presente la advertencia de Raudenbush (2004) sobre los efectos de tipo B: la práctica pedagógica de las instituciones no está definida u observada, así que es difícil estimar su efecto. Aún más importante, con base en Raudenbush y Willms (1995), las estimaciones tipo B estarán sesgadas para las instituciones con contextos favorables si hay una correlación positiva entre práctica y el contexto medido de forma aproximada por el promedio del INSE⁽⁶⁾ o por los puntajes de Saber 11°. Es por ello que solo se presenta y analiza un modelo tipo A en este documento⁽⁷⁾.

⁷En el documento técnico construido en 2015 se presentan dos modelos tipo B, así como un análisis de sensibilidad y su comparación con las estimaciones tipo A (Icfes, 2015).

2.2. Datos

Este anexo aplica para los reportes construidos a partir de los resultados del Examen Saber Pro aplicado en los periodos 2017-2018 y posteriores periodos. Sin embargo, todos los análisis aquí presentados corresponden al periodo 2017-2018, al ser considerado el periodo de línea base de la nueva estructura de los reportes del estudio de aporte relativo.

2.2.1. Comparabilidad Saber 11°

El cruce entre Saber Pro y Saber 11° se extiende en un rango de tiempo de 4 a 9 años. Eso quiere decir que, para el caso del reporte correspondiente a los años 2017-2018, el periodo utilizado para Saber 11° abarca desde 2008 a 2014. Ahora bien, en el segundo semestre de 2014, el examen Saber 11° cambió su estructura y dejó de ser comparable con periodos anteriores. Sin embargo, el Icfes realizó una recalificación de los puntajes obtenidos por los estudiantes desde el primer semestre de 2012 hasta el primer semestre de 2014, de manera que la serie comprendida entre 2012-1 y 2014-1 es comparable con las series posteriores. Como consecuencia, para dicha serie se cuenta con dos tipos de puntajes del examen Saber 11°⁽⁸⁾: puntajes recalificados y puntajes originales.

Con el fin de hacer las estimaciones del modelo de aporte relativo, en las que se consideran puntajes de Saber 11 que originalmente no son comparables con la serie que inicia en 2012-1, se utilizaron los puntajes de la serie recalificada (2012-1 a 2014-1) para estimar los coeficientes asociados a los puntajes originales, por medio de un modelo de mínimos cuadrados ordinarios, de la siguiente forma:

⁸Tanto el puntaje original como el puntaje recalificado se encuentran reportados en la página de resultados individuales del Icfes y pueden ser consultados por los estudiantes que tomaron el examen entre 2012-1 y 2014-1. Los puntajes recalificados también están incluidos en las bases de datos de Saber 11° publicadas en el repositorio DataIcfes.

$$Sb11_{recalificado,i} = \beta_0 + \beta_1 matemáticas_i + \beta_2 lenguaje_i + \beta_3 ciencias\ sociales_i + \beta_4 filosofía_i + \beta_5 biología_i + \beta_6 química_i + \beta_7 física_i + \beta_8 inglés_i + \epsilon_i$$

donde $Sb11_{recalificado,i}$ corresponde a los puntajes recalificados de Saber 11° (Matemáticas, Lectura Crítica, Ciencias Naturales, Sociales y Ciudadanas o Inglés) para el periodo 2012-1 a 2014-1. Posteriormente, se utilizaron los coeficientes resultantes del modelo anterior para estimar los que serían los puntajes recalificados para el periodo que va de 2008-1 a 2011-2, con base en los puntajes originales y la ecuación de la regresión.

La Tabla 1 muestra los Coeficientes de determinación (R²) que se obtuvieron en las regresiones mencionadas, donde se puede resaltar que la medida de bondad de ajuste ⁽⁹⁾ es considerablemente alta, por lo que se podría confiar en la validez de las estimaciones.

Tabla 1.
R² serie recalificada

Nuevo puntaje	R ²
Matemáticas	0.8537
Lectura Crítica	0.7988
Ciencias Naturales	0.8858
Sociales y Ciudadanas	0.8655
Inglés	0.9892

⁹Las medidas de bondad de ajuste buscan determinar si los datos se ajustan a una determinada distribución".

2.2.2. Relación entre las pruebas

La Tabla 2 muestra el número de personas que presentaron el examen Saber Pro y a las que al mismo tiempo se les pudo cruzar los puntajes obtenidos en el examen Saber 11°.

Recuerde que para este ejercicio se considera una ventana de tiempo de presentación de 4 a 8 años entre los dos exámenes (9 para el caso de medicina).

Tabla 2.

Cantidad de puntajes cruzados entre el examen Saber Pro y Saber 11° por periodo de presentación

Periodo de presentación de Saber 11	Periodo de presentación de Saber Pro		
	2017	2018	Total
	2008-1	68	-
2008-2	267	-	267
2009-1	1.795	55	1.850
2009-2	10.271	236	10.507
2010-1	2.207	1.258	3.472
2010-2	17.072	11.124	28.196
2011-1	4.027	2.164	6.191
2011-2	25.535	15.896	41.431
2012-1	7.738	4.291	12.029
2012-2	41.366	25.368	67.004
2013-1	7.872	6.952	14.824
2013-2	12.708	39.742	52.450
2014-1	-	7.123	7.123
2014-2	-	13.102	1.302
Total	130.926	127.588	258.514

Las correlaciones a nivel de estudiantes entre los puntajes de las pruebas Saber 11 y Saber Pro en el análisis se muestran en la Tabla 3. Se puede apreciar que las correlaciones, en casi todos los casos, son superiores a 50%, lo que sugiere que hay una relación fuerte entre los puntajes previos (Saber 11°) y los resultados en el examen Saber Pro, de tal forma que se podrían considerar buenos predictores para el modelo.

Tabla 3.

Correlaciones entre puntajes

Saber 11°		Saber Pro		
		Razonamiento Cuantitativo	Lectura Crítica	Inglés
	Matemáticas	64,28%	50,54%	53,35%
	Lectura Crítica	50,82%	59,95%	53,89%
	Ciencias Naturales	62,08%	55,20%	54,56%
	Sociales y Ciudadanas	50,60%	58,54%	51,63%
	Inglés	44,54%	49,73%	75,95%

Al considerar las correlaciones mostradas y teniendo en cuenta lo encontrado por Ballou, Sanders y Wright (2004) y Lockwood y McCaffrey (2014), es factible justificar la inclusión de todos los puntajes de Saber 11° como predictores en las estimaciones de valor agregado, pues estas no son sensibles a la inclusión de variables a nivel de estudiante si se tienen identificados de forma precisa los puntajes previos. Además, los puntajes previos de diferentes áreas pueden ser transversales a diferentes competencias, ayudando a disminuir los posibles sesgos por errores de medición.

2.2.3. Selección de la unidad de análisis

Las decisiones de los estudiantes sobre a cuál institución educativa asistir y cuál programa educativo cursar inciden sustancialmente en los efectos de la clasificación y en las diferencias en el número de estudiantes disponibles para ciertos análisis. Asimismo, el nivel de agrupación que se elija tendrá mucha relevancia, pues cambiará la interpretación del modelo y su uso. La Tabla 4 muestra el número de instituciones, INBC e Instituciones - Programa que cumplirían con los filtros que se mencionaron en la sección del modelo.

Tabla 4.

Cantidad de puntajes cruzados entre el examen Saber Pro y Saber 11° según institución, INBC e Institución-Programa

	Totales	Cumplen filtros	%cumplimiento
Institución	255	197	77,2 %
INBC	2.309	1.687	73,0 %
Institución-Programa	3.668	2.309	62,9 %

Para complementar la información de la Tabla 4, la Tabla 5 muestra los coeficientes de correlación intraclase (ICC)⁽¹⁰⁾ por niveles de agrupación para los módulos de Razonamiento Cuantitativo, Lectura Crítica e Inglés para el examen Saber Pro. Analizar esto es importante, pues el ICC indica el grado de relación que existe entre los resultados observados a nivel de estudiantes y entre aquellos que se observan en el nivel de anidación correspondiente (NBC, institución, INBC).

¹⁰En la siguiente sección se explicará con mayor detalle el significado y construcción de este indicador.

Tabla 5.

ICC modelo vacío para diferentes niveles de agrupación

	Razonamiento Cuantitativo	Lectura Crítica	Inglés
NBC	0,2504	0,1367	0,1892
Institución	0,1937	0,1675	0,2921
Institución - NBC	0,3580	0,2601	0,3792
Institución - programa	0,3696	0,2801	0,4230

Por último, teniendo en cuenta el supuesto de SUTVA, el análisis a nivel de institución puede traer grandes sesgos para las estimaciones. Se puede apreciar que la mayor relación se ubica en los niveles de anidación de INBC e Institución. Además, teniendo en cuenta la pérdida de información que se podría tener al usar el nivel de anidación de programa académico que se observa, la elección de usar INBC como unidad de análisis parece ser la más indicada.

Tras exponer la pertinencia de la unidad de análisis seleccionada para los modelos de valor agregado, la Tabla 6 muestra el número de estudiantes por Núcleo Básico del Conocimiento, el número de estudiantes que se lograron cruzar con Saber 11° y la muestra final resultante después de adicionar todos los filtros requeridos por el estudio. De esta forma, se puede observar que existen seis NBC en los que el cruce entre los resultados de Saber 11° y Saber Pro fue menor al 60%, uno de ellos del 34%. Por otro lado, después de adicionar todos los filtros mencionados al principio de la sección, quedan tres NBC con cruces relativamente bajos, menores al 30%, entre los que llama la atención el NBC de “BIBLIOTECOLOGÍA, OTROS DE CIENCIAS SOCIALES Y HUMANAS”, pues su cruce es cercano al 10%.

Tabla 6.

Configuración de estudiantes por NBC

	Total	Cruce	% cruce	Aplicando filtros	% muestra final
ADMINISTRACIÓN	98.107	54.314	55,36%	30.973	31,57%
AGRONOMÍA	1.040	585	56,25%	291	27,98%
ANTROPOLOGÍA, ARTES LIBERALES	836	666	79,67%	566	67,70%
ARQUITECTURA	8.051	6.772	84,11%	6.234	77,43%
ARTES PLÁSTICAS, VISUALES Y AFINES	2.641	2.135	80,84%	1.806	68,38%
ARTES REPRESENTATIVAS	506	363	71,74%	209	41,30%
BACTERIOLOGÍA	1.433	1.239	86,46%	1.135	79,20%
BIBLIOTECOLOGÍA, OTROS DE CIENCIAS SOCIALES Y HUMANAS	965	333	34,51%	102	10,57%
BIOLOGÍA, MICROBIOLOGÍA Y AFINES	3.279	2.873	87,62%	2.483	75,72%
CIENCIA POLÍTICA, RELACIONES INTERNACIONALES	3.401	2.981	87,65%	2.721	80,01%
COMUNICACIÓN SOCIAL, PERIODISMO Y AFINES	12.524	10.665	85,16%	9.849	78,64%
CONTADURÍA PÚBLICA	39.199	21.595	55,09%	12.697	32,39%
DEPORTES, EDUCACIÓN FÍSICA Y RECREACIÓN	2.083	1.669	80,12%	1.460	70,09%
DERECHO Y AFINES	39.892	26.662	66,84%	22.806	57,17%
DISEÑO	7.500	6.181	82,41%	5.527	73,69%
ECONOMÍA	11.329	9.490	83,77%	8.629	76,17%
EDUCACIÓN	52.756	33.406	63,32%	26.720	50,65%

2 Aspectos metodológicos y modelo de valor agregado y aporte relativo

	Total	Cruce	% cruce	Aplicando filtros	% muestra final
ENFERMERÍA	8.536	6555	76,79%	5.803	67,98%
FILOSOFÍA, TEOLOGÍA Y AFINES	1.356	717	52,88%	273	20,13%
FORMACIÓN RELACIONADA CON EL CAMPO MILITAR O POLICIAL	1.236	847	68,53%	660	53,40%
FÍSICA	456	399	87,50%	255	55,92%
GEOGRAFÍA, HISTORIA	1.245	998	80,16%	800	64,26%
GEOLOGÍA, OTROS PROGRAMAS DE CIENCIAS NATURALES	725	665	91,72%	643	88,69%
INGENIERÍA ADMINISTRATIVA Y AFINES	1.918	1.287	67,10%	957	49,90%
INGENIERÍA AGROINDUSTRIAL, ALIMENTOS Y AFINES	2.402	1.842	76,69%	1.404	58,45%
INGENIERÍA AGRONÓMICA, PECUARIA Y AFINES	1.770	1.486	83,95%	1.303	73,62%
INGENIERÍA AGRÍCOLA, FORESTAL Y AFINES	1.420	1.219	85,85%	1.024	72,11%
INGENIERÍA AMBIENTAL, SANITARIA Y AFINES	11.859	9.851	83,07%	8.620	72,69%
INGENIERÍA BIOMÉDICA Y AFINES	1.089	927	85,12%	865	79,43%
INGENIERÍA CIVIL Y AFINES	15.323	12.898	84,17%	11.938	77,91%
INGENIERÍA DE MINAS, METALURGIA Y AFINES	1.883	1.752	93,04%	1.675	88,95%
INGENIERÍA DE SISTEMAS, TELEMÁTICA Y AFINES	15.061	9.039	60,02%	5.482	36,40%
INGENIERÍA ELECTRÓNICA, TELECOMUNICACIONES Y AFINES	6.990	4.802	68,70%	3.441	49,23%
INGENIERÍA ELÉCTRICA Y AFINES	3.475	2.255	64,89%	1.681	48,37%
INGENIERÍA INDUSTRIAL Y AFINES	26.329	18.487	70,22%	15.055	57,18%
INGENIERÍA MECÁNICA Y AFINES	7.642	5.825	76,22%	4.763	62,33%

	Total	Cruce	% cruce	Aplicando filtros	% muestra final
INGENIERÍA QUÍMICA Y AFINES	3.072	2869	93,39%	2781	90,53%
INSTRUMENTACIÓN QUIRÚRGICA	1.295	1167	90,12%	1101	85,02%
LENGUAS MODERNAS, LITERATURA, LINGÜÍSTICA Y AFINES	1.659	1376	82,94%	1269	76,49%
MATEMÁTICAS, ESTADÍSTICA Y AFINES	916	772	84,28%	466	50,87%
MEDICINA	13.417	11860	88,40%	11507	85,76%
MEDICINA VETERINARIA	3.185	2668	83,77%	2312	72,59%
MÚSICA	1.367	1095	80,10%	874	63,94%
NUTRICIÓN Y DIETÉTICA	1.265	1106	87,43%	951	75,18%
ODONTOLOGÍA	3.879	3217	82,93%	2919	75,25%
OPTOMETRÍA, OTROS PROGRAMAS DE CIENCIAS DE LA SALUD	477	391	81,97%	349	73,17%
OTRAS INGENIERÍAS	1.800	1491	82,83%	1271	70,61%
OTROS PROGRAMAS ASOCIADOS A BELLAS ARTES	404	268	66,34%	210	51,98%
PSICOLOGÍA	29.695	20239	68,16%	16368	55,12%
PUBLICIDAD Y AFINES	1.812	1393	76,88%	1151	63,52%
QUÍMICA Y AFINES	2.100	1713	81,57%	1449	69,00%
SALUD PUBLICA	3.458	1785	51,62%	1064	30,77%
SOCIOLOGÍA, TRABAJO SOCIAL Y AFINES	10.471	7521	71,83%	6389	61,02%
TERAPIAS	5.689	4881	85,80%	4422	77,73%
ZOOTECNIA	1.498	1098	73,30%	811	54,14%
TOTAL	48,3716	30,690	68,36%	25.8514	53,44%

Tabla 7

ICC por NBC

En la **Tabla 7** se pueden ver los ICC de cada uno de los NBC en los módulos de Razonamiento Cuantitativo, Lectura Crítica e Inglés, además del número de instituciones dentro de cada uno de ellos. En esta Tabla se puede ver que en los NBC de “ARTES REPRESENTATIVAS”, “BIBLIOTECOLOGÍA, OTROS DE CIENCIAS SOCIALES Y HUMANAS”, “FORMACIÓN RELACIONADA CON EL CAMPO MILITAR O POLICIAL” y “SALUD PÚBLICA” los ICC son muy pequeños e incluso llegan a ser 0.

Esto quiere decir que el modelo utilizado para calcular el valor agregado puede no ser el que mejor se adapte a los datos, es decir, que la anidación en los niveles de estudiantes e INBC no es relevante para explicar los puntajes de Saber Pro. Sin embargo, al mismo tiempo es de resaltar que esos NBC tienen menos de cinco instituciones, por lo que es posible que los ICC bajos estén explicados por la poca presencia de instituciones, más que por la no significancia de los niveles de anidación escogidos. Debido a que estos NBC tienen menos de 5 instituciones no es posible generarles los reportes de aporte relativo⁽¹⁾, lo que tampoco es posible para los NBC de “AGRONOMÍA” y “OTROS PROGRAMAS ASOCIADOS A BELLAS ARTES” pues no cumplen con este requisito.

Esta sección recoge los principales resultados de los

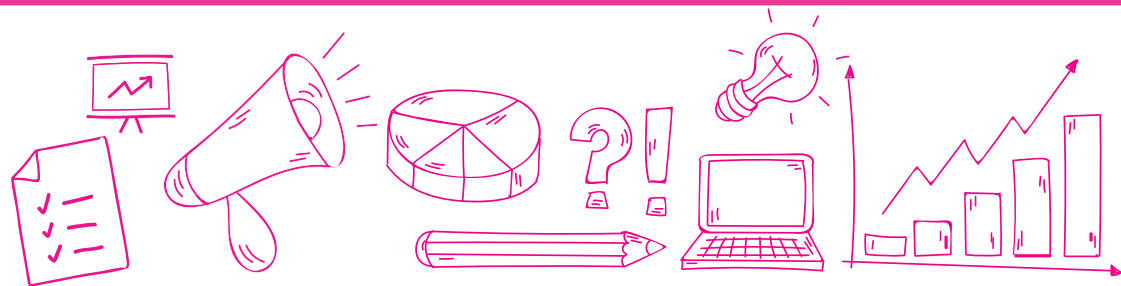
¹Este documento es construido a partir de la información de Saber Pro de los períodos 2017 y 2018. Por ello, aunque para este periodo los NBC mencionados no cumplen con los requisitos para que se realicen los reportes, nada impide que en periodos posteriores sí los cumplan y por tanto se generen los reportes.

	Razonamiento Cuantitativo	Lectura Crítica	Inglés	Número instituciones
ADMINISTRACIÓN	0,1818	0,1630	0,3470	138
	0,1130	0,0927	0,0917	4
ANTROPOLOGÍA, ARTES LIBERALES	0,1821	0,1977	0,3724	9
ARQUITECTURA	0,1927	0,2094	0,2988	48
ARTES PLÁSTICAS, VISUALES Y AFINES	0,1350	0,1334	0,2796	23
ARTES REPRESENTATIVAS	0,0996	0,0181	0,2570	3
BACTERIOLOGÍA	0,2244	0,2289	0,2189	14
BIBLIOTECOLOGÍA, OTROS DE CIENCIAS SOCIALES Y HUMANAS	0,0000	0,0000	0,0362	2
BIOLOGÍA, MICROBIOLOGÍA Y AFINES	0,2570	0,2351	0,3668	32
CIENCIA POLÍTICA, RELACIONES INTERNACIONALES	0,1136	0,1042	0,3340	24
COMUNICACIÓN SOCIAL, PERIODISMO Y AFINES	0,1187	0,1254	0,2400	52
CONTADURÍA PÚBLICA	0,1682	0,1554	0,2342	102
DEPORTES, EDUCACIÓN FÍSICA Y RECREACIÓN	0,1557	0,1517	0,1616	14
DERECHO Y AFINES	0,1802	0,2239	0,3317	93
DISEÑO	0,1771	0,1494	0,2466	45
ECONOMÍA	0,2680	0,2126	0,3427	68
EDUCACIÓN	0,1596	0,2007	0,2786	72

2 Aspectos metodológicos y modelo de valor agregado y aporte relativo

	Razonamiento Cuantitativo	Lectura Crítica	Inglés	Número instituciones
ENFERMERÍA	0,2473	0,2416	0,2170	54
FILOSOFÍA, TEOLOGÍA Y AFINES	0,2376	0,3371	0,2840	6
FORMACIÓN RELACIONADA CON EL CAMPO MILITAR O POLICIAL	0,0000	0,0225	0,1280	3
FÍSICA	0,1222	0,2782	0,3623	5
GEOGRAFÍA, HISTORIA	0,2428	0,3590	0,4390	13
GEOLOGÍA, OTROS PROGRAMAS DE CIENCIAS NATURALES	0,2350	0,2752	0,3084	6
INGENIERÍA ADMINISTRATIVA Y AFINES	0,3385	0,2280	0,3109	10
INGENIERÍA AGROINDUSTRIAL, ALIMENTOS Y AFINES	0,1149	0,1127	0,1628	22
INGENIERÍA AGRONÓMICA, PECUARIA Y AFINES	0,3418	0,2832	0,3002	16
INGENIERÍA AGRÍCOLA, FORESTAL Y AFINES	0,3331	0,2440	0,3082	13
INGENIERÍA AMBIENTAL, SANITARIA Y AFINES	0,2406	0,2160	0,3001	54
INGENIERÍA BIOMÉDICA Y AFINES	0,1720	0,2021	0,3180	11
INGENIERÍA CIVIL Y AFINES	0,2666	0,2267	0,3426	57
INGENIERÍA DE MINAS, METALURGIA Y AFINES	0,2383	0,0856	0,3397	9
INGENIERÍA DE SISTEMAS, TELEMÁTICA Y AFINES	0,2893	0,2424	0,4098	80
INGENIERÍA ELECTRÓNICA, TELECOMUNICACIONES Y AFINES	0,2383	0,2084	0,2839	51
INGENIERÍA ELÉCTRICA Y AFINES	0,2158	0,2065	0,2585	20

	Razonamiento Cuantitativo	Lectura Crítica	Inglés	Número instituciones
INGENIERÍA INDUSTRIAL Y AFINES	0,2410	0,1700	0,3266	83
INGENIERÍA MECÁNICA Y AFINES	0,2626	0,2049	0,3554	39
INGENIERÍA QUÍMICA Y AFINES	0,1798	0,1767	0,2664	17
INSTRUMENTACIÓN QUIRÚRGICA	0,1936	0,1747	0,2100	14
LENGUAS MODERNAS, LITERATURA, LINGÜÍSTICA Y AFINES	0,1405	0,1551	0,3311	10
MATEMÁTICAS, ESTADÍSTICA Y AFINES	0,2029	0,2767	0,4407	8
MEDICINA	0,2210	0,2266	0,3316	52
MEDICINA VETERINARIA	0,1884	0,2321	0,2821	25
MÚSICA	0,0508	0,0823	0,3011	16
NUTRICIÓN Y DIETÉTICA	0,3071	0,2411	0,2413	9
ODONTOLOGÍA	0,1872	0,1864	0,2060	21
OPTOMETRÍA, OTROS PROGRAMAS DE CIENCIAS DE LA SALUD	0,0680	0,0511	0,0612	7
OTRAS INGENIERÍAS	0,3169	0,3275	0,4611	16
OTROS PROGRAMAS ASOCIADOS A BELLAS ARTES	0,1288	0,1074	0,4499	3
PSICOLOGÍA	0,2211	0,2204	0,3566	80
PUBLICIDAD Y AFINES	0,0676	0,0550	0,1222	13
QUÍMICA Y AFINES	0,1648	0,1999	0,2729	15
SALUD PUBLICA	0,0141	0,0118	0,0204	4
SOCIOLOGÍA, TRABAJO SOCIAL Y AFINES	0,2851	0,3613	0,4318	39
TERAPIAS	0,1988	0,2107	0,2343	29
ZOOTECNIA	0,2177	0,1881	0,2794	14



3

Resultados

Esta sección recoge los principales resultados de los análisis de valor agregado descritos anteriormente para los INBC en las pruebas de Razonamiento Cuantitativo, Lectura Crítica e Inglés del examen Saber Pro. En el anexo, se relacionan las tablas que contienen la información de las estimaciones de los modelos que se presentan en este capítulo.

3.1. Relación entre modelos

Para este ejercicio se estimaron tres modelos de valor agregado. El modelo principal es el descrito en la sección 2.1 de este documento, es decir, se estima uno de los puntajes de los módulos de Saber Pro (Razonamiento Cuantitativo, Lectura Crítica o Inglés) con base en los cinco puntajes de los módulos de Saber 11º: Matemáticas, Lectura Crítica, Ciencias Naturales, Sociales y Ciudadanas e Inglés. Al segundo modelo se le agrega el puntaje del INSE del estudiante, para controlar por aspectos socioeconómicos. Por último, el tercer modelo solo se estima para las competencias de Lectura Crítica y Razonamiento Cuantitativo y su diferencia con el modelo principal es que no incluye inglés como covariable.

La elección del modelo principal se dio por las siguientes dos razones: i) el INSE utilizado en el segundo modelo

proviene del examen Saber Pro debido a que no es posible calcular este índice para todos los periodos de Saber 11º utilizados. Esto implica que el INSE observado no puede considerarse como una condición inicial del estudiante, previa al ingreso a la educación superior, por lo que teóricamente no debería ser incluida en el modelo. ii). Como se mencionó antes, una vez se controla por los puntajes previos correctos, la inclusión de variables a nivel de estudiante tiene efectos mínimos en las estimaciones de valor agregado (Ballou et al., 2004).

Las Figuras 1, 2 y 3 muestran la correlación de Spearman de los tres modelos estimados para los módulos de Lectura Crítica, Razonamiento Cuantitativo e Inglés, respectivamente, por NBC. En las tres figuras se puede observar una relación alta dentro de los modelos, siendo la mayoría superior a 95%. Esto implica que la sensibilidad del modelo es buena ante la inclusión o exclusión de alguna variable y que los resultados son muy robustos ante estos cambios en la especificación de la estimación.

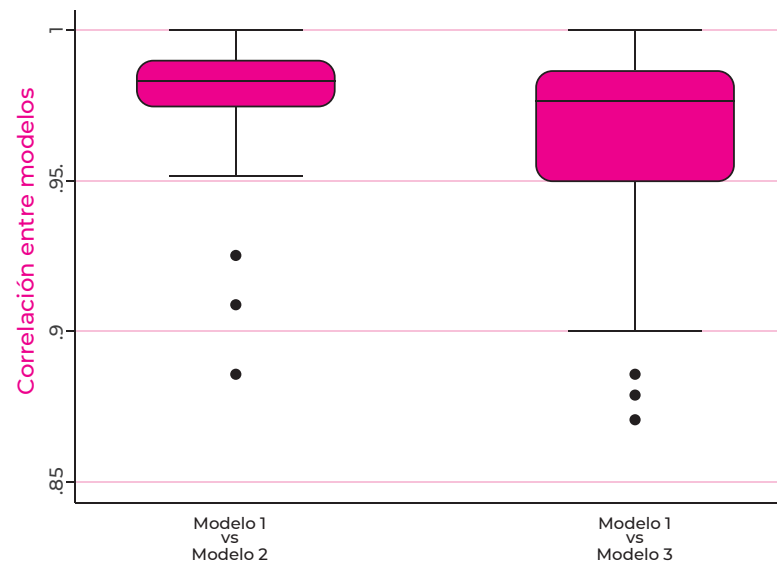


Figura 1. Correlación de Spearman entre modelos: Lectura Crítica

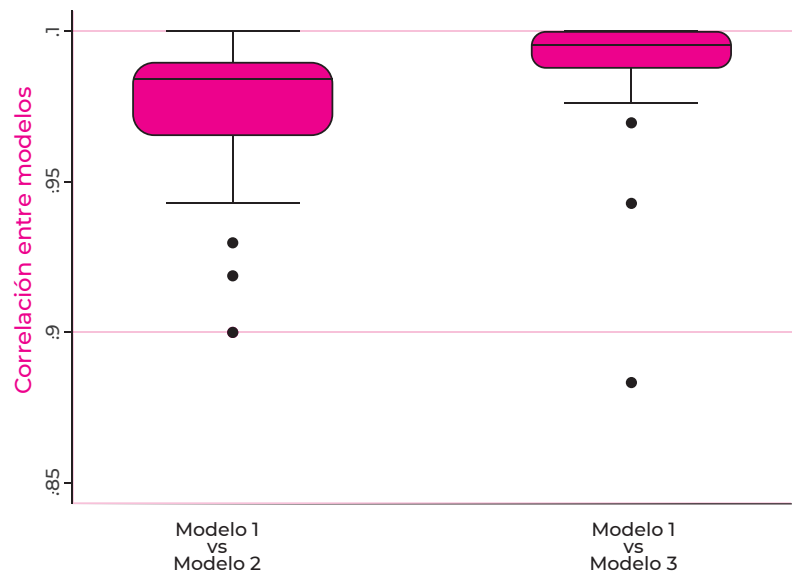


Figura 2. Correlación de Spearman entre modelos: Razonamiento Cuantitativo

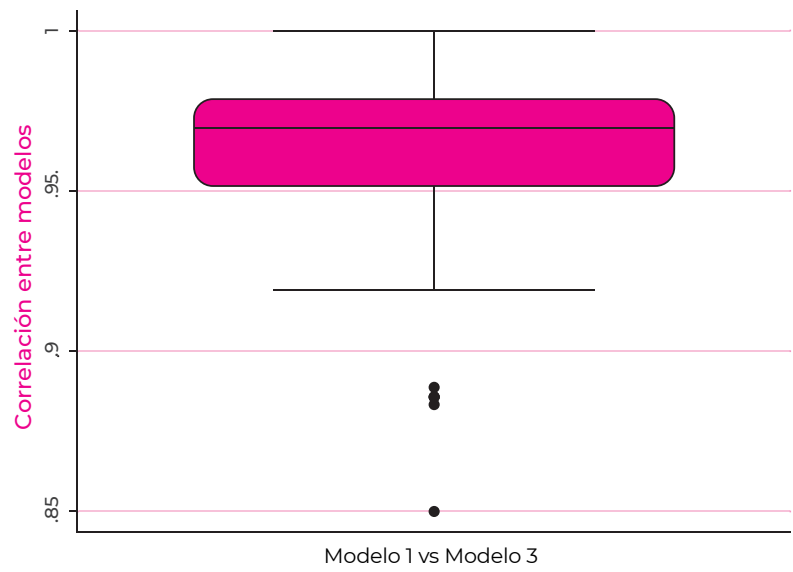


Figura 3. Correlación de Spearman entre modelos: Inglés

3.2. Relaciones con variables de interés

Las Figuras 4, 5 y 6 muestran las correlaciones de Spearman entre las estimaciones de valor agregado por INBC y cuatro variables de interés, promediadas por INBC: **i)** el puntaje promedio de Saber Pro del módulo evaluado, **ii)** el puntaje promedio de Saber 11° de la competencia más parecida al módulo evaluado, **iii)** el puntaje global promedio, y **iv)** el INSE de Saber Pro.

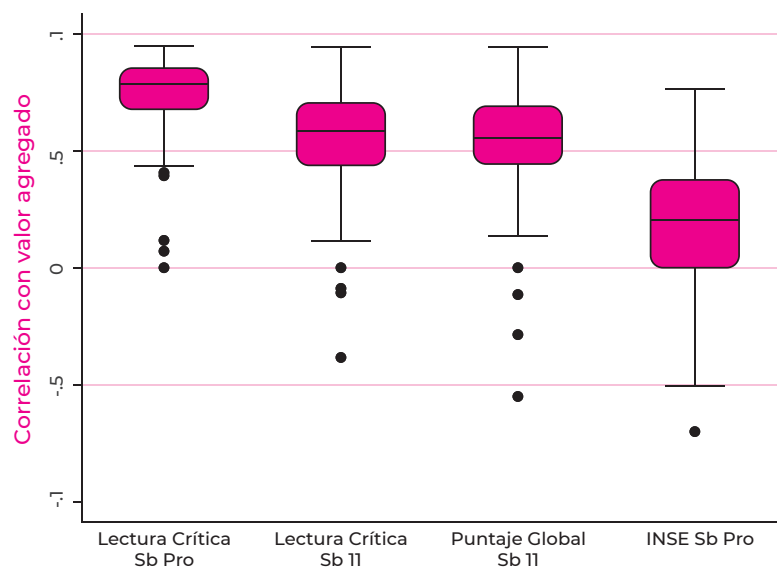


Figura 4. Correlación de Spearman con variables de interés: Lectura Crítica

¹⁰ Cada uno de los puntos de las figuras es una correlación de Spearman entre dos modelos estimados para un mismo NBC

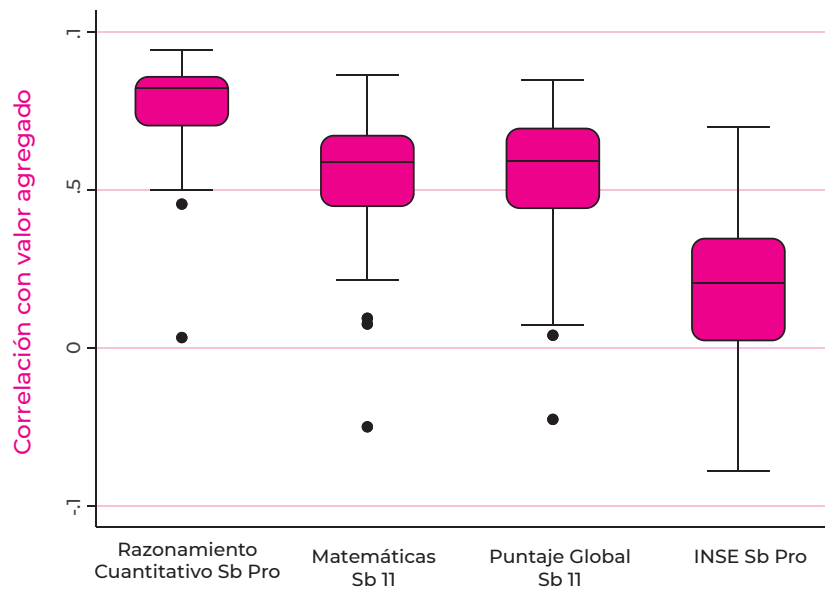


Figura 5. Correlación de Spearman con variables de interés: Razonamiento Cuantitativo

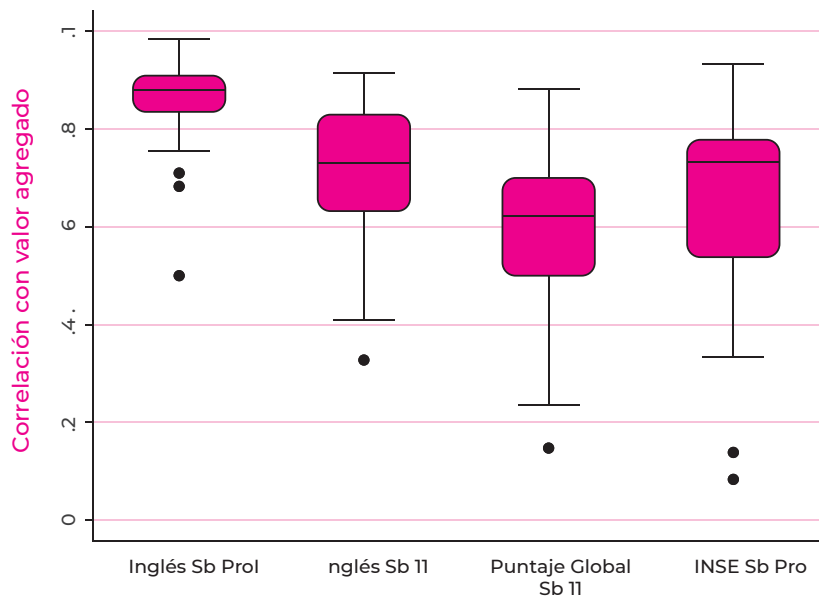


Figura 6. Correlación de Spearman con variables de interés: Inglés

Para los tres módulos analizados se puede ver que la correlación más alta es con su respectivo puntaje de Saber Pro. Aunque es alta, para ninguno de los NBC la correlación llega a 100%, indicando que, aunque similares, una ordenación de instituciones que tenga solo en cuenta el desempeño puede estar ignorando que las condiciones iniciales de los estudiantes sí tiene relación con esos resultados.

Dentro de estos datos, vale la pena resaltar que las correlaciones con el INSE en el valor agregado de Inglés son más altas que en los otros módulos. Esto puede ser un indicador que el aprendizaje de este idioma aún se relaciona mucho con el nivel socioeconómico de los estudiantes, por lo que le queda un trabajo pendiente al sistema educativo en cerrar este tipo de brecha.

Referencias

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37–65.

Briggs, D., & Domingue, B. (2011). Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles unified school district teachers by the "Los Angeles Times". National Education Policy Center.

Castellano, K., Rabe-Hesketh, S., & Skrondal, A. (2014). Composition, context, and endogeneity in school and teacher comparisons.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood (Working Paper No. 17699). Cambridge, MA: National Bureau of Economic Research.

Cook, J., & Stefanski, L. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association*, 89, 1314–1328.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2012, August). Selecting growth measures for school and teacher evaluations (Working Paper No. 80). Washington, D.C.: CALDER.

Goldhaber, D., & Chaplin, D. (2012, January). Assessing the "Rothstein falsification test": Does it really show teacher value-added models are biased (Working Paper No. 2012- 1.3). Seattle, WA: Center for Education Data and Research

Hausman, J. A., & Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric Society*, 1377–1398.

Ho, D. E., & Rubin, D. B. (2011). Credible causal inference for empirical legal studies. *Annual Review of Law and Social Science*, 7, 17–40.

Icfes, Instituto Colombiano para la Evaluación de la Educación (2015). Medición de los efectos de la educación superior en Colombia sobre el aprendizaje estudiantil - Informe Técnico. Bogotá D.C. ISBN: 978-958-11-0672-1

Icfes, Instituto Colombiano para la Evaluación de la Educación (2019). Saber al detalle ¿Cómo se calcula el índice de nivel socioeconómico (INSE) en el contexto de las pruebas Saber?. Bogotá D.C. ISBN: 2590-4663.

Koedel, C., & Betts, J. R. (2011). Does student sorting invalidate value-added models of teacher effectiveness? an extended analysis of the rothstein critique. *Education*, 6(1), 18–42

Leckie, G., & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society, A*, 172, 835–851.

Lockwood, J., & McCaffrey, D. (2015). Should nonlinear functions of test scores be used as covariates in a regression model?

Lockwood, J., & McCaffrey, D. (2014). Correcting for test score measurement error in ancova models for estimating treatment effects. *Journal of Educational and Behavioral Statistics*, 39(1), 22-52. Tomado de <http://jeb.sagepub.com/content/39/1/22.abstract>

Manzi, J., San Martin, E., & Van Belleghem, S. (2014). School system evaluation by value added analysis under endogeneity. *Psychometrika*, 79(1), 130-153. Tomado de <http://dx.doi.org/10.1007/s11336-013-9338-0>

Martínez, María del Rosario, Gaviria, José-Luis & Morera, María. (2009). Concepto y evolución de los modelos de valor añadido en educación. *Revista de educación*, ISSN 0034-8082, N° 348, 2009 (Ejemplar dedicado a: El valor añadido en educación), pags. 15-46.

MEN, Ministerio de Educación Nacional (2017). Reporte sobre deserción y graduación en educación superior año 2016. Tomado de https://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-357549_recurso_5.pdf

Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Education Policy Analysis Archives*, 18(23).

Raudenbush, S. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29, 121-130.

Raudenbush, S., & Willms, J. (1995). The estimation of

school effects. *Journal of Educational and Behavioral Statistics*, 20, 307-335.

Reardon, S. F., & Raudenbush, S. W. (2009). Assumptions of value-added models for estimating school effects. *Education*, 4(4), 492-519.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement*. *Quarterly Journal of Economics*, 125(1), 175-214. Tomado de <http://www.mitpressjournals.org/doi/abs/10.1162/qjec.2010.125.1.175>

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value added assessment in education. *Journal of Educational and Behavioral Statistics*, 29, 103-116.

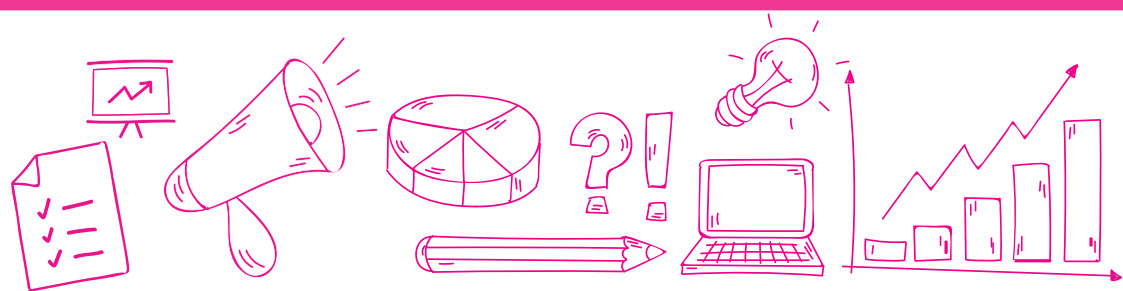
Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*, 3.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355-374.

Steedle, J. T. (2012). Selecting value-added models for postsecondary institutional assessment. *Assessment &*

Evaluation in Higher Education, 37(6), 637-652.

Winston, G., & Zimmerman, D. (2004). Peer effects in Higher Education. In *College Choices: The economics of where to go, when to go, and how to pay for it* (pp. 395-424). University of Chicago Press.



4 Anexo

Tabla A1. Coeficientes y desviaciones estándar estimados para Lectura Crítica

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
ADMINISTRACIÓN	2,022 (0,215)	6,107 (0,235)	4,736 (0,249)	5,217 (0,237)	4,131 (0,202)	150,930 (0,184)
ANTROPOLOGÍA, ARTES LIBERALES	2,086 (1,503)	5,776 (1,337)	4,303 (1,696)	1,392 (1,394)	5,894 (1,129)	183,504 (0,911)
ARQUITECTURA	2,149 (0,446)	5,499 (0,513)	5,251 (0,502)	5,425 (0,480)	3,885 (0,386)	156,113 (0,341)
ARTES PLÁSTICAS, VISUALES Y AFINES	1,410 (0,839)	6,851 (0,953)	2,939 (1,036)	5,470 (0,975)	4,761 (0,773)	166,070 (0,610)
BACTERIOLOGÍA	3,177 (0,938)	4,636 (1,042)	2,159 (1,164)	6,273 (1,079)	3,876 (0,799)	162,089 (0,792)
BIOLOGÍA, MICROBIOLOGÍA Y AFINES	2,016 (0,703)	6,045 (0,710)	3,718 (0,754)	4,296 (0,687)	4,282 (0,642)	168,467 (0,504)
CIENCIA POLÍTICA, RELACIONES INTERNACIONALES	2,177 (0,674)	5,437 (0,672)	3,512 (0,747)	5,086 (0,650)	3,238 (0,589)	181,442 (0,526)
COMUNICACIÓN SOCIAL, PERIODISMO Y AFINES	2,649 (0,340)	6,590 (0,373)	3,358 (0,399)	5,643 (0,372)	3,134 (0,314)	165,772 (0,285)
CONTADURÍA PÚBLICA	2,192 (0,308)	5,063 (0,312)	3,407 (0,322)	5,108 (0,320)	2,462 (0,238)	145,806 (0,245)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
DEPORTES, EDUCACIÓN FÍSICA Y RECREACIÓN	2,627 (0,873)	5,935 (1,063)	4,223 (0,929)	2,506 (1,021)	2,663 (0,802)	142,170 (0,657)
DERECHO Y AFINES	1,916 (0,250)	6,646 (0,281)	4,490 (0,280)	5,305 (0,271)	4,541 (0,219)	162,004 (0,192)
DISEÑO	2,013 (0,497)	6,061 (0,512)	4,846 (0,537)	4,396 (0,508)	5,029 (0,417)	160,020 (0,363)
ECONOMÍA	2,473 (0,388)	6,658 (0,409)	4,070 (0,433)	5,365 (0,389)	4,452 (0,359)	163,964 (0,304)
EDUCACIÓN	2,258 (0,206)	7,142 (0,257)	3,326 (0,234)	5,656 (0,258)	4,312 (0,183)	147,885 (0,222)
ENFERMERÍA	1,043 (0,435)	5,722 (0,480)	5,040 (0,499)	4,376 (0,439)	3,301 (0,373)	148,969 (0,315)
FILOSOFÍA, TEOLOGÍA Y AFINES	7,091 (2,243)	7,003 (2,696)	-1,614 (2,435)	6,085 (2,122)	0,384 (1,845)	179,129 (1,408)
FÍSICA	2,227 (2,808)	1,109 (3,625)	2,108 (2,705)	8,001 (2,994)	5,092 (2,395)	186,768 (1,717)
GEOGRAFÍA, HISTORIA	2,141 (1,193)	6,415 (1,455)	1,697 (1,313)	5,230 (1,298)	1,998 (1,182)	168,596 (0,900)
GEOLOGÍA, OTROS PROGRAMAS DE CIENCIAS NATURALES	1,470 (1,218)	4,767 (1,264)	2,027 (1,120)	2,656 (1,033)	4,674 (1,083)	178,823 (1,065)
INGENIERÍA ADMINISTRATIVA Y AFINES	2,914 (1,377)	6,195 (1,445)	3,619 (1,467)	3,919 (1,389)	4,350 (1,189)	163,737 (0,861)
INGENIERÍA AGROINDUSTRIAL, ALIMENTOS Y AFINES	3,315 (0,952)	6,035 (1,038)	1,139 (1,028)	5,010 (1,021)	3,651 (0,719)	155,143 (0,665)
INGENIERÍA AGRONÓMICA, PECUARIA Y AFINES	0,544 (1,039)	4,978 (1,149)	4,470 (1,117)	5,403 (1,124)	2,226 (0,758)	151,111 (0,694)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
INGENIERÍA AGRÍCOLA, FORESTAL Y AFINES	3,004 (1,077)	7,644 (1,366)	0,007 (1,152)	7,519 (1,229)	1,356 (0,965)	156,034 (0,780)
INGENIERÍA AMBIENTAL, SANITARIA Y AFINES	2,671 (0,377)	5,287 (0,386)	3,399 (0,429)	5,375 (0,400)	2,996 (0,318)	154,794 (0,276)
INGENIERÍA BIOMÉDICA Y AFINES	1,589 (1,178)	4,833 (1,323)	4,695 (1,371)	4,561 (1,266)	5,982 (1,059)	161,728 (0,941)
INGENIERÍA CIVIL Y AFINES	2,785 (0,364)	6,027 (0,358)	3,785 (0,388)	5,483 (0,354)	3,935 (0,307)	158,861 (0,261)
INGENIERÍA DE MINAS, METALURGIA Y AFINES	1,676 (0,943)	7,820 (0,944)	2,661 (0,925)	5,305 (0,907)	4,441 (0,810)	169,417 (0,795)
INGENIERÍA DE SISTEMAS, TELEMÁTICA Y AFINES	3,097 (0,571)	6,449 (0,580)	3,045 (0,648)	5,673 (0,580)	4,370 (0,476)	156,432 (0,372)
INGENIERÍA ELECTRÓNICA, TELECOMUNICACIONES Y AFINES	3,152 (0,655)	6,524 (0,693)	2,950 (0,701)	3,873 (0,626)	4,589 (0,486)	166,687 (0,421)
INGENIERÍA ELÉCTRICA Y AFINES	1,133 (0,925)	4,438 (1,047)	3,705 (1,004)	5,357 (0,938)	3,956 (0,774)	164,026 (0,666)
INGENIERÍA INDUSTRIAL Y AFINES	2,624 (0,325)	5,895 (0,331)	3,874 (0,368)	5,193 (0,330)	3,884 (0,297)	157,235 (0,226)
INGENIERÍA MECÁNICA Y AFINES	2,971 (0,563)	5,458 (0,597)	3,118 (0,656)	6,677 (0,587)	4,450 (0,490)	163,342 (0,396)
INGENIERÍA QUÍMICA Y AFINES	2,573 (0,628)	4,369 (0,622)	2,900 (0,695)	5,104 (0,617)	4,029 (0,515)	176,055 (0,482)
INSTRUMENTACIÓN QUIRÚRGICA	0,897 (0,836)	6,216 (0,944)	4,996 (0,955)	4,068 (0,960)	3,837 (0,775)	143,572 (0,716)
LENGUAS MODERNAS, LITERATURA, LINGÜÍSTICA Y AFINES	2,397 (0,913)	4,324 (0,987)	3,193 (1,073)	6,213 (1,112)	2,858 (0,868)	186,834 (0,824)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
MATEMÁTICAS, ESTADÍSTICA Y AFINES	3,087 (1,627)	5,529 (1,689)	1,956 (1,818)	3,990 (1,654)	3,449 (1,443)	175,318 (1,147)
MEDICINA	2,240 (0,386)	5,380 (0,386)	1,243 (0,426)	4,481 (0,368)	5,185 (0,312)	172,623 (0,261)
MEDICINA VETERINARIA	3,115 (0,721)	5,572 (0,793)	2,369 (0,771)	4,669 (0,809)	3,268 (0,645)	153,852 (0,547)
MÚSICA	1,827 (1,609)	8,035 (1,574)	0,029 (1,673)	5,178 (1,367)	4,879 (1,296)	166,533 (0,982)
NUTRICIÓN Y DIETÉTICA	-0,520 (1,060)	4,776 (1,106)	6,194 (1,153)	6,304 (1,251)	3,225 (0,902)	160,314 (0,779)
ODONTOLOGÍA	1,477 (0,640)	5,129 (0,708)	4,214 (0,715)	3,722 (0,724)	3,048 (0,534)	152,623 (0,497)
OPTOMETRÍA, OTROS PROGRAMAS DE CIENCIAS DE LA SALUD	3,696 (1,819)	6,219 (1,942)	3,738 (2,012)	4,958 (1,787)	3,284 (1,588)	146,015 (1,272)
OTRAS INGENIERÍAS	1,483 (1,156)	4,446 (1,103)	4,847 (1,210)	6,803 (1,091)	5,204 (1,018)	163,034 (0,715)
PSICOLOGÍA	1,863 (0,280)	6,869 (0,313)	3,803 (0,308)	5,448 (0,305)	3,800 (0,263)	157,624 (0,246)
PUBLICIDAD Y AFINES	0,935 (1,051)	5,881 (1,142)	3,599 (1,042)	4,637 (1,032)	3,784 (0,796)	155,245 (0,888)
QUÍMICA Y AFINES	3,415 (0,979)	5,148 (1,032)	-0,235 (1,017)	4,772 (1,020)	4,361 (0,723)	166,362 (0,773)
SOCIOLOGÍA, TRABAJO SOCIAL Y AFINES	1,945 (0,422)	6,176 (0,485)	4,274 (0,465)	5,978 (0,469)	2,744 (0,379)	152,058 (0,339)
TERAPIAS	1,560 (0,477)	7,184 (0,499)	4,650 (0,547)	4,121 (0,517)	3,190 (0,414)	150,697 (0,383)
ZOOTECNIA	0,722 (1,133)	6,238 (1,409)	4,164 (1,306)	4,107 (1,476)	1,879 (0,982)	148,667 (0,788)

Tabla A2. Coeficientes y desviaciones estándar estimados para Razonamiento Cuantitativo

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
ADMINISTRACIÓN	8,461 (0,198)	1,744 (0,208)	5,757 (0,231)	2,108 (0,217)	1,417 (0,193)	152,123 (0,169)
ANTROPOLOGÍA, ARTES LIBERALES	9,301 (1,461)	2,187 (1,406)	4,869 (1,602)	0,091 (1,534)	4,131 (1,198)	156,869 (0,979)
ARQUITECTURA	8,505 (0,417)	1,638 (0,451)	5,299 (0,481)	1,646 (0,423)	0,623 (0,361)	157,257 (0,295)
ARTES PLÁSTICAS, VISUALES Y AFINES	8,645 (0,871)	1,383 (0,929)	6,770 (1,021)	1,704 (0,898)	0,917 (0,722)	146,719 (0,665)
BACTERIOLOGÍA	7,351 (1,007)	0,724 (1,045)	4,304 (1,203)	2,406 (1,078)	1,542 (0,883)	156,797 (0,786)
BIOLOGÍA, MICROBIOLOGÍA Y AFINES	8,556 (0,657)	1,223 (0,686)	3,686 (0,676)	2,190 (0,668)	2,187 (0,592)	164,904 (0,468)
CIENCIA POLÍTICA, RELACIONES INTERNACIONALES	9,966 (0,652)	0,447 (0,670)	6,003 (0,729)	1,885 (0,679)	1,177 (0,540)	158,139 (0,491)
COMUNICACIÓN SOCIAL, PERIODISMO Y AFINES	7,421 (0,345)	1,624 (0,379)	6,226 (0,397)	1,749 (0,371)	1,146 (0,325)	143,297 (0,290)
CONTADURÍA PÚBLICA	7,926 (0,276)	1,585 (0,324)	4,868 (0,287)	1,537 (0,308)	0,623 (0,228)	153,893 (0,225)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
DEPORTES, EDUCACIÓN FÍSICA Y RECREACIÓN	8,160 (0,740)	2,008 (1,008)	5,582 (0,912)	1,473 (1,042)	1,091 (0,697)	146,848 (0,642)
DERECHO Y AFINES	7,995 (0,252)	1,480 (0,266)	6,547 (0,259)	1,669 (0,257)	1,175 (0,225)	146,001 (0,197)
DISEÑO	9,103 (0,444)	1,138 (0,460)	5,748 (0,530)	1,023 (0,474)	1,531 (0,402)	153,081 (0,354)
ECONOMÍA	9,286 (0,362)	1,987 (0,351)	4,941 (0,385)	2,007 (0,348)	1,391 (0,323)	165,075 (0,277)
EDUCACIÓN	7,701 (0,202)	1,393 (0,235)	6,397 (0,223)	1,223 (0,238)	0,157 (0,168)	136,855 (0,207)
ENFERMERÍA	5,682 (0,435)	1,762 (0,434)	6,301 (0,470)	1,071 (0,404)	1,438 (0,347)	146,121 (0,303)
FILOSOFÍA, TEOLOGÍA Y AFINES	10,919 (2,816)	-0,518 (2,670)	9,794 (2,983)	1,872 (2,767)	-1,644 (2,423)	145,648 (1,776)
FÍSICA	10,819 (2,946)	-2,165 (2,967)	0,705 (2,627)	5,526 (2,606)	0,810 (2,072)	201,804 (1,697)
GEOGRAFÍA, HISTORIA	9,398 (1,308)	1,638 (1,289)	6,328 (1,393)	1,854 (1,170)	-0,108 (1,102)	147,416 (0,763)
GEOLOGÍA, OTROS PROGRAMAS DE CIENCIAS NATURALES	9,157 (1,223)	1,919 (1,098)	2,136 (1,291)	1,437 (1,030)	2,470 (1,059)	186,217 (0,957)
INGENIERÍA ADMINISTRATIVA Y AFINES	9,802 (1,199)	1,898 (1,258)	2,441 (1,342)	2,648 (1,215)	3,325 (1,015)	172,823 (0,795)
INGENIERÍA AGROINDUSTRIAL, ALIMENTOS Y AFINES	7,056 (0,886)	1,465 (0,903)	3,553 (0,856)	1,132 (0,870)	1,930 (0,635)	165,952 (0,584)
INGENIERÍA AGRONÓMICA, PECUARIA Y AFINES	7,021 (0,896)	1,064 (1,044)	5,169 (1,051)	1,765 (0,979)	0,312 (0,778)	161,930 (0,755)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
INGENIERÍA AGRÍCOLA, FORESTAL Y AFINES	7,944 (0,891)	1,797 (1,083)	3,592 (0,989)	3,573 (1,120)	0,794 (0,745)	169,506 (0,679)
INGENIERÍA AMBIENTAL, SANITARIA Y AFINES	7,449 (0,351)	1,184 (0,395)	5,724 (0,386)	2,483 (0,376)	1,070 (0,297)	162,358 (0,264)
INGENIERÍA BIOMÉDICA Y AFINES	6,934 (1,188)	0,617 (1,229)	6,415 (1,372)	3,265 (1,185)	1,739 (1,067)	171,835 (0,812)
INGENIERÍA CIVIL Y AFINES	9,293 (0,356)	0,968 (0,354)	4,546 (0,365)	2,153 (0,334)	1,355 (0,282)	177,345 (0,248)
INGENIERÍA DE MINAS, METALURGIA Y AFINES	7,210 (0,859)	2,735 (0,958)	3,257 (0,858)	1,669 (0,741)	1,614 (0,684)	184,275 (0,789)
INGENIERÍA DE SISTEMAS, TELEMÁTICA Y AFINES	9,187 (0,504)	1,438 (0,527)	5,154 (0,501)	2,425 (0,558)	1,237 (0,421)	169,110 (0,323)
INGENIERÍA ELECTRÓNICA, TELECOMUNICACIONES Y AFINES	8,247 (0,561)	1,352 (0,581)	3,290 (0,632)	1,082 (0,560)	2,494 (0,516)	182,459 (0,397)
INGENIERÍA ELÉCTRICA Y AFINES	8,135 (0,825)	1,394 (0,899)	3,394 (0,834)	1,076 (0,861)	0,942 (0,759)	182,127 (0,603)
INGENIERÍA INDUSTRIAL Y AFINES	8,994 (0,284)	1,705 (0,295)	5,048 (0,307)	1,895 (0,277)	1,419 (0,251)	171,222 (0,209)
INGENIERÍA MECÁNICA Y AFINES	8,536 (0,523)	1,214 (0,567)	4,215 (0,560)	2,219 (0,560)	1,758 (0,462)	181,494 (0,345)
INGENIERÍA QUÍMICA Y AFINES	7,462 (0,681)	1,298 (0,648)	3,481 (0,701)	2,331 (0,643)	1,498 (0,548)	189,066 (0,481)
INSTRUMENTACIÓN QUIRÚRGICA	5,323 (0,768)	2,313 (0,841)	6,170 (0,794)	1,994 (0,956)	0,562 (0,733)	140,012 (0,660)
LENGUAS MODERNAS, LITERATURA, LINGÜÍSTICA Y AFINES	8,940 (1,115)	-0,314 (1,101)	8,130 (1,238)	1,271 (1,145)	0,377 (0,783)	158,648 (0,866)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
MATEMÁTICAS, ESTADÍSTICA Y AFINES	10,075 (2,179)	2,321 (2,035)	-2,476 (1,945)	0,019 (1,861)	3,696 (1,509)	191,461 (1,279)
MEDICINA	9,094 (0,405)	1,061 (0,384)	2,309 (0,428)	1,703 (0,375)	3,078 (0,300)	169,792 (0,254)
MEDICINA VETERINARIA	8,901 (0,652)	1,368 (0,741)	4,121 (0,760)	1,717 (0,760)	1,137 (0,560)	158,652 (0,469)
MÚSICA	12,000 (1,526)	5,203 (1,530)	3,632 (1,553)	-1,102 (1,399)	0,845 (1,026)	156,048 (0,848)
NUTRICIÓN Y DIETÉTICA	5,277 (1,002)	1,804 (0,979)	4,718 (1,093)	3,423 (1,133)	1,465 (0,881)	155,389 (0,703)
ODONTOLOGÍA	6,985 (0,574)	1,844 (0,676)	4,114 (0,675)	1,408 (0,653)	1,477 (0,529)	147,770 (0,500)
OPTOMETRÍA, OTROS PROGRAMAS DE CIENCIAS DE LA SALUD	4,225 (1,836)	3,318 (1,740)	8,136 (1,890)	-0,988 (1,697)	-0,549 (1,325)	143,580 (1,090)
OTRAS INGENIERÍAS	11,434 (1,093)	0,401 (1,340)	3,700 (1,190)	3,711 (1,317)	1,549 (0,980)	177,700 (0,735)
PSICOLOGÍA	6,694 (0,244)	2,329 (0,292)	5,452 (0,277)	1,503 (0,275)	1,539 (0,224)	141,109 (0,217)
PUBLICIDAD Y AFINES	6,718 (1,092)	3,761 (1,146)	3,287 (1,188)	0,886 (1,193)	1,437 (0,846)	143,059 (0,851)
QUÍMICA Y AFINES	8,753 (0,983)	-0,244 (0,896)	4,036 (0,978)	1,970 (0,909)	2,149 (0,798)	177,186 (0,593)
SOCIOLOGÍA, TRABAJO SOCIAL Y AFINES	5,863 (0,389)	0,898 (0,498)	6,395 (0,435)	2,410 (0,463)	1,431 (0,374)	135,109 (0,324)
TERAPIAS	6,759 (0,459)	1,380 (0,461)	5,372 (0,532)	1,845 (0,480)	0,768 (0,393)	144,623 (0,376)
ZOOTECNIA	8,483 (1,181)	1,778 (1,411)	3,384 (1,590)	0,388 (1,392)	-0,038 (0,956)	158,450 (0,766)

Tabla A3. Coeficientes y desviaciones estándar estimados para Inglés

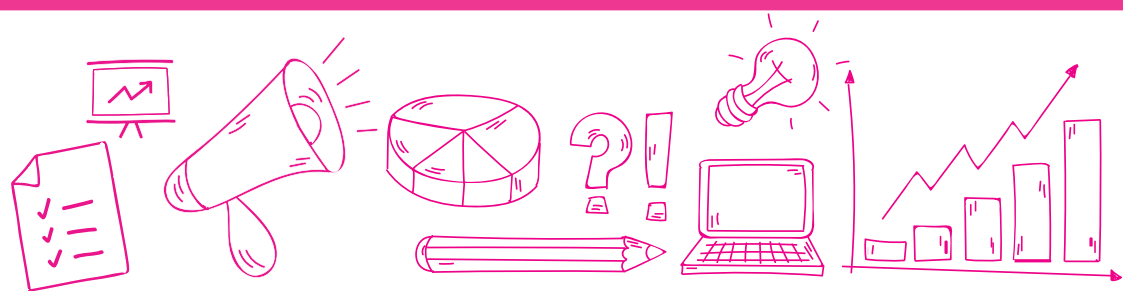
	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
ADMINISTRACIÓN	1,944 (0,179)	1,679 (0,186)	1,256 (0,202)	1,078 (0,192)	16,467 (0,178)	158,967 (0,154)
ANTROPOLOGÍA, ARTES LIBERALES	2,942 (1,360)	0,079 (1,046)	2,318 (1,467)	0,527 (1,274)	19,699 (1,254)	176,265 (0,807)
ARQUITECTURA	2,014 (0,360)	1,696 (0,407)	0,535 (0,401)	1,259 (0,384)	17,920 (0,337)	163,903 (0,285)
ARTES PLÁSTICAS, VISUALES Y AFINES	1,125 (0,732)	2,120 (0,782)	0,592 (0,816)	1,061 (0,844)	20,081 (0,725)	174,771 (0,550)
BACTERIOLOGÍA	1,353 (0,796)	1,972 (0,952)	0,557 (1,007)	2,014 (0,928)	13,457 (0,629)	154,499 (0,688)
BIOLOGÍA, MICROBIOLOGÍA Y AFINES	1,780 (0,686)	2,079 (0,672)	1,217 (0,710)	0,327 (0,708)	18,810 (0,610)	170,861 (0,462)
CIENCIA POLÍTICA, RELACIONES INTERNACIONALES	1,280 (0,569)	1,935 (0,570)	0,756 (0,648)	0,285 (0,573)	16,324 (0,498)	181,975 (0,452)
COMUNICACIÓN SOCIAL, PERIODISMO Y AFINES	1,395 (0,306)	1,028 (0,304)	1,487 (0,344)	1,560 (0,304)	18,190 (0,274)	164,785 (0,252)
CONTADURÍA PÚBLICA	1,773 (0,232)	1,480 (0,259)	0,457 (0,253)	0,958 (0,267)	10,783 (0,208)	145,393 (0,199)
DEPORTES, EDUCACIÓN FÍSICA Y RECREACIÓN	2,551 (0,626)	1,748 (0,754)	1,185 (0,746)	-0,892 (0,757)	13,112 (0,547)	145,839 (0,520)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
DERECHO Y AFINES	1,245 (0,205)	1,823 (0,222)	1,019 (0,232)	1,515 (0,226)	18,315 (0,189)	153,670 (0,160)
DISEÑO	0,962 (0,389)	1,589 (0,432)	0,318 (0,415)	1,372 (0,403)	19,242 (0,343)	173,442 (0,301)
ECONOMÍA	0,956 (0,333)	2,215 (0,323)	1,091 (0,372)	1,195 (0,337)	16,845 (0,303)	169,584 (0,259)
EDUCACIÓN	1,536 (0,204)	3,824 (0,250)	0,433 (0,223)	0,848 (0,236)	16,776 (0,177)	151,576 (0,214)
ENFERMERÍA	1,101 (0,350)	1,050 (0,364)	1,319 (0,401)	0,999 (0,349)	11,480 (0,263)	144,301 (0,254)
FILOSOFÍA, TEOLOGÍA Y AFINES	5,022 (2,385)	2,759 (2,865)	-3,203 (2,770)	5,804 (2,530)	20,854 (2,140)	164,823 (1,473)
FÍSICA	2,206 (2,467)	2,593 (2,702)	0,945 (2,901)	3,118 (1,878)	15,692 (1,842)	195,616 (1,665)
GEOGRAFÍA, HISTORIA	1,667 (1,025)	2,033 (1,150)	1,013 (1,240)	-1,006 (1,205)	20,372 (1,119)	158,820 (0,709)
GEOLOGÍA, OTROS PROGRAMAS DE CIENCIAS NATURALES	0,094 (1,283)	2,448 (1,241)	0,104 (1,132)	0,845 (1,061)	17,658 (1,063)	184,244 (0,960)
INGENIERÍA ADMINISTRATIVA Y AFINES	0,142 (0,921)	0,257 (1,062)	3,108 (1,071)	1,814 (1,083)	17,964 (0,915)	173,204 (0,729)
INGENIERÍA AGROINDUSTRIAL, ALIMENTOS Y AFINES	1,293 (0,754)	0,928 (0,880)	0,569 (0,842)	-0,067 (0,805)	15,696 (0,722)	153,788 (0,567)
INGENIERÍA AGRONÓMICA, PECUARIA Y AFINES	2,337 (0,732)	0,749 (0,819)	0,604 (0,860)	1,959 (0,809)	12,471 (0,734)	147,350 (0,588)
INGENIERÍA AGRÍCOLA, FORESTAL Y AFINES	1,795 (0,833)	1,253 (1,005)	-1,124 (0,930)	1,011 (1,060)	15,436 (0,919)	155,479 (0,656)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
INGENIERÍA AMBIENTAL, SANITARIA Y AFINES	2,287 (0,282)	1,333 (0,334)	0,853 (0,340)	1,193 (0,335)	14,108 (0,267)	154,692 (0,226)
INGENIERÍA BIOMÉDICA Y AFINES	2,257 (1,171)	2,483 (1,065)	-0,426 (1,088)	0,966 (1,036)	17,045 (0,825)	171,308 (0,673)
INGENIERÍA CIVIL Y AFINES	1,558 (0,274)	2,085 (0,303)	0,723 (0,284)	1,457 (0,299)	16,449 (0,260)	160,646 (0,198)
INGENIERÍA DE MINAS, METALURGIA Y AFINES	1,088 (0,722)	2,405 (0,719)	1,166 (0,809)	0,687 (0,676)	15,925 (0,641)	172,394 (0,670)
INGENIERÍA DE SISTEMAS, TELEMÁTICA Y AFINES	1,914 (0,444)	2,471 (0,471)	1,098 (0,473)	1,436 (0,460)	16,996 (0,412)	166,817 (0,281)
INGENIERÍA ELECTRÓNICA, TELECOMUNICACIONES Y AFINES	2,185 (0,505)	1,789 (0,561)	0,077 (0,554)	1,441 (0,551)	17,139 (0,476)	172,887 (0,358)
INGENIERÍA ELÉCTRICA Y AFINES	1,454 (0,736)	1,171 (0,794)	0,168 (0,721)	2,917 (0,771)	16,205 (0,654)	168,762 (0,530)
INGENIERÍA INDUSTRIAL Y AFINES	1,813 (0,245)	2,090 (0,248)	0,648 (0,267)	1,517 (0,246)	16,408 (0,226)	163,658 (0,197)
INGENIERÍA MECÁNICA Y AFINES	1,484 (0,459)	1,414 (0,471)	0,748 (0,542)	2,469 (0,478)	17,536 (0,387)	171,756 (0,336)
INGENIERÍA QUÍMICA Y AFINES	1,079 (0,531)	1,045 (0,479)	1,875 (0,537)	1,875 (0,468)	13,832 (0,449)	180,897 (0,442)
INSTRUMENTACIÓN QUIRÚRGICA	0,964 (0,686)	2,246 (0,834)	1,394 (0,841)	0,037 (0,796)	11,105 (0,739)	144,269 (0,615)
LENGUAS MODERNAS, LITERATURA, LINGÜÍSTICA Y AFINES	0,796 (0,971)	1,311 (1,000)	2,863 (1,037)	0,706 (1,027)	15,302 (0,832)	193,778 (0,987)
MATEMÁTICAS, ESTADÍSTICA Y AFINES	2,031 (1,282)	0,746 (1,564)	-2,408 (1,615)	1,680 (1,196)	19,508 (1,392)	177,122 (1,110)

	Matemáticas	Lectura Crítica	Ciencias Naturales	Sociales y Ciudadanas	Inglés	Constante
MEDICINA	1,845 (0,351)	1,588 (0,349)	-1,859 (0,368)	2,129 (0,326)	18,722 (0,273)	174,113 (0,215)
MEDICINA VETERINARIA	2,218 (0,571)	1,521 (0,648)	0,607 (0,659)	-0,005 (0,650)	17,626 (0,538)	156,004 (0,449)
MÚSICA	2,367 (1,174)	2,132 (1,144)	-2,524 (1,176)	2,039 (1,175)	18,488 (1,131)	178,554 (0,758)
NUTRICIÓN Y DIETÉTICA	-0,469 (0,884)	1,112 (0,970)	2,159 (1,053)	2,734 (0,961)	16,972 (0,887)	155,716 (0,632)
ODONTOLOGÍA	2,410 (0,469)	1,697 (0,572)	-0,485 (0,569)	0,569 (0,521)	15,602 (0,467)	155,509 (0,426)
OPTOMETRÍA, OTROS PROGRAMAS DE CIENCIAS DE LA SALUD	1,100 (1,303)	1,397 (1,266)	1,886 (1,310)	2,978 (1,206)	12,078 (1,030)	151,573 (0,877)
OTRAS INGENIERÍAS	2,639 (0,945)	1,331 (0,925)	-0,852 (0,864)	1,505 (0,911)	19,608 (0,770)	169,982 (0,590)
PSICOLOGÍA	1,337 (0,216)	2,202 (0,251)	1,032 (0,239)	1,171 (0,245)	17,011 (0,243)	154,046 (0,199)
PUBLICIDAD Y AFINES	2,446 (0,889)	1,649 (0,934)	-1,052 (0,899)	0,777 (0,770)	15,845 (0,605)	164,779 (0,663)
QUÍMICA Y AFINES	3,789 (0,800)	2,319 (0,782)	-1,455 (0,809)	0,818 (0,796)	15,639 (0,665)	167,213 (0,562)
SOCIOLOGÍA, TRABAJO SOCIAL Y AFINES	1,352 (0,326)	1,237 (0,421)	0,968 (0,353)	1,351 (0,381)	12,421 (0,355)	141,932 (0,296)
TERAPIAS	2,026 (0,381)	1,653 (0,387)	0,635 (0,421)	1,052 (0,394)	12,314 (0,353)	149,038 (0,321)
ZOOTECNIA	0,351 (0,907)	2,610 (1,133)	0,241 (1,253)	0,001 (1,142)	13,153 (0,877)	145,148 (0,670)

Desviaciones estándar en paréntesis



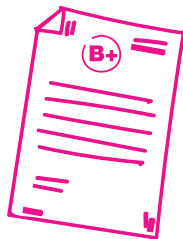
5 Anexo

Supuestos del Modelo

Modelo base del aporte relativo, el valor agregado:

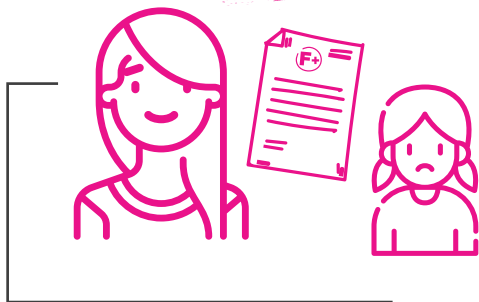
Especificaciones estadísticas a partir de las cuales se estima la contribución de una Institución Educación Superior (IES), diferenciando los NBC de cada programa al progreso de los estudiantes, una vez se controla por la influencia de factores ajenos a ésta.

Los exámenes Saber 11° y Saber Pro ubican a Colombia en una posición privilegiada para reportar el valor agregado de su sistema de educación superior, debido a que ofrecen mediciones comparables sobre el nivel de competencias con el que los estudiantes inician y finalizan el ciclo de educación postsecundaria.



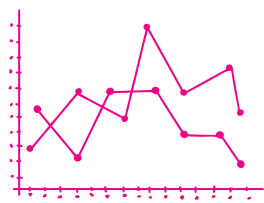
Manipulabilidad

Requiere que sea teóricamente factible que un estudiante tenga un resultado potencial para todas las posibles IES.



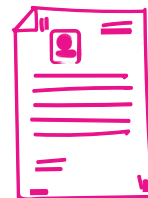
No interferencia entre unidades:

Requiere que los resultados potenciales de un estudiante no dependan de las IES a las que asisten otros estudiantes. El efecto de la IES sobre el estudiante es independiente de las características de sus compañeros.



El supuesto métrico:

Los resultados del puntaje de la prueba están en una escala de intervalo.



Homogeneidad:

Establece que los efectos de la IES son una constante independiente de los antecedentes del estudiante.



Tratamiento altamente ignorable:

La asignación al tratamiento es esencialmente aleatoria después de condicionarse por las variables de control.



Forma funcional:

La forma funcional empleada para el control de las características del estudiante es la correcta. Se supone que a cada IES asiste un número suficiente de estudiantes a través de las distribuciones de los controles relevantes.

Limitaciones y consideraciones

Análisis de los desafíos ocasionados y los aspectos relacionados con los sesgos que han motivado las decisiones tomadas en esta investigación. Existen varias implicaciones sobre los análisis de valor agregado derivados de las decisiones operativas sobre los supuestos del modelo de valor agregado.

Desafíos



Tratamientos:

¿Cómo escoger el tratamiento para cada control, dadas las características propias de cada contrafactual?

Unidades:

¿Cuáles deberían ser las unidades de análisis? Teniendo en cuenta, por ejemplo, que un estudiante puede pertenecer a dos programas en una misma universidad.



Resultados:

¿Cuáles deberían ser las medidas de los resultados? Se podrían medir competencias comunes a todos los programas, pero esto brindaría poca información sobre los aprendizajes propios de cada uno de éstos.



Aspectos adicionales a tener en cuenta

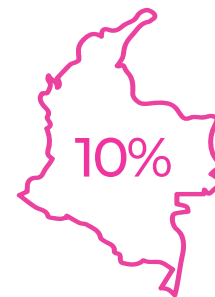


Covariables:

La definición de las covariables tiene el efecto de cambiar las interpretaciones de la estimación del valor agregado.

Datos faltantes:

En Colombia, la tasa promedio de deserción anual en educación superior está alrededor del 10% (MEN, 2017) ¿Cómo manejar esto para no sesgar los resultados?



Selección de estudiantes:

Los estudiantes se autoseleccionan en los programas e IES, aspecto que en estudios internacionales se evidencia como una amenaza significativa para la inferencia causal.



Desafíos técnicos - sesgos

En el coeficiente de los puntajes previos de los estudiantes, por causa de errores de medición.

En las estimaciones del efecto educativo para IES con características favorables.

En las estimaciones del efecto escolar para IES con características desfavorables.

Aspectos Metodológicos

Modelo de valor agregado y aporte relativo

El estudio de aporte relativo del Icfes es una modificación a los modelos de valor agregado, en el cual se controla no sólo por las condiciones iniciales de los estudiantes, sino también por el tipo de población que atienden las IES (medido como el puntaje promedio de Saber 11° de los estudiantes de cada institución).

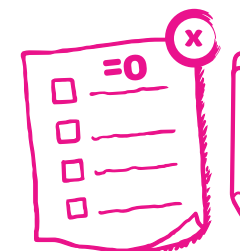
Filtros para la estimación



Se toman todos los individuos para los cuales se cruza el documento de identidad entre Saber 11° y Saber Pro. También hay un cruce fonético de los nombres.



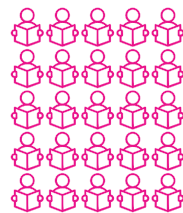
Se limita el tiempo entre los dos exámenes (entre 4 y 8 años, y 9 para Medicina) para evitar atribuir al valor agregado de las instituciones experiencias externas de los estudiantes.



Se excluyen del estudio los estudiantes que tienen dos o más módulos de Saber Pro con un puntaje igual a 0 o sin calificación.

Estimación

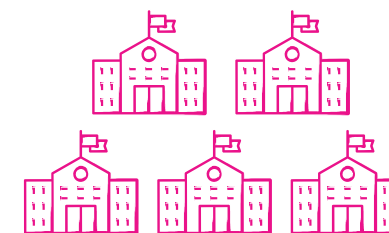
La estimación del modelo de Valor Agregado se da entonces en cada IES dentro de un Núcleo Básico del Conocimiento (INBC).



Cada INBC debe tener al menos 25 estudiantes que hayan presentado Saber Pro, con su respectivo cruce en Saber 11°.

40%

Los estudiantes cruzados deben representar mínimo el 40% de la población total de la IES en ese NBC



Sólo se hacen las estimaciones de valor agregado a aquellos NBC que cuenten con al menos 5 instituciones.

Aspectos Metodológicos

Modelo de valor agregado y aporte relativo

El estudio de aporte relativo del Icfes es una modificación a los modelos de valor agregado, en el cual se controla no sólo por las condiciones iniciales de los estudiantes, sino también por el tipo de población que atienden las IES (medido como el puntaje promedio de Saber 11° de los estudiantes de cada institución).

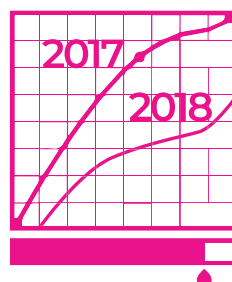
Este estudio se realiza desde 2015, en esta figura ponemos como ejemplo la información del último informe realizado.

Estimación del Modelo

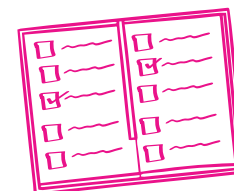
El modelo se construye buscando el puntaje que obtendría un estudiante en el Examen Saber Pro, y luego comparando con el puntaje obtenido, teniendo en cuenta características que provienen de sus resultados en el Examen Saber 11°.



Filtros para la estimación

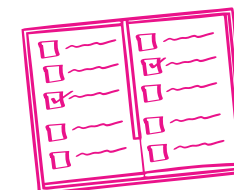


El estudio toma dos años de Saber Pro. Por ejemplo para Saber Pro 2018 y 2017 se buscan los resultados de Saber 11° en los años 2008 a 2015. Esto quiere decir que se cruzan los exámenes con un rango de 4 a 9 años.



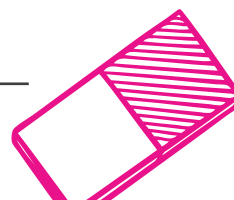
2012-1/2014-1

Los coeficientes asociados a los puntajes recalificados, se obtienen por medio de un modelo de Mínimos Cuadrados Ordinarios, que incluye como controles todas las pruebas (2012-1 a 2014-1).



2018-1/2011-2

Usando los coeficiente recalificados, se estiman los puntajes de 2008-1 a 2011-2 por medio de un modelo de mínimo cuadrado ordinarios.



2012-1/2014-1

Debido a un cambio en la estructura del Examen Saber 11° en 2014, el Icfes hace una recalificación desde 2012-1 hasta 2014-1 para que los resultados comparables. Estos se conocen en el modelo como resultados "recalificados".

Correlación entre puntaje

		Saber Pro		
		Razonamiento Cuantitativo	Lectura Crítica	Inglés
Saber 11°	Matemáticas	642,8%	0,5054	0,5335
	Lectura Crítica	508,2%	0,5995	0,5389
	Ciencias Naturales	620,8%	0,5520	0,5456
	Sociales y Ciudadanas	506,0%	0,5854	0,5163
	Inglés	445,4%	0,4973	0,7595

Hallazgos del desarrollo del modelo de valor agregado y aporte relativo

0,5

En casi todos los casos, las correlaciones son superiores a 0,5, lo que sugiere que hay una relación fuerte entre los puntajes previos (Saber 11°) y los resultados en Saber Pro, de tal forma que se podrían considerar buenos predictores para el modelo.

Es factible justificar la inclusión de todos los puntajes de Saber 11° como predictores en las estimaciones de valor agregado, pues estas no son sensibles a la inclusión de variables a nivel de estudiante si se tienen identificados de forma precisa los puntajes previos.



Instituciones por nivel de agrupación

	Razonamiento Cuantitativo	Lectura Crítica	Inglés
NBC	250,4%	0.1367	0.1892
Institución	193,7%	0.1675	0.2921
Institución - NBC	358,0%	0.2601	0.3792
Institución - programa	369,6%	0.2801	0.4230

Hallazgos del desarrollo del modelo de valor agregado y aporte relativo



El nivel de agrupación que se elija tendrá mucha relevancia, pues cambiará la interpretación del modelo y su uso.

INBC

Se puede apreciar que la mayor relación se ubica en los niveles de anidación de INBC e Institución-Programa.



Teniendo en cuenta la pérdida de información que se podría tener al usar el nivel de anidación de programa académico, en el estudio se utiliza la INBC.



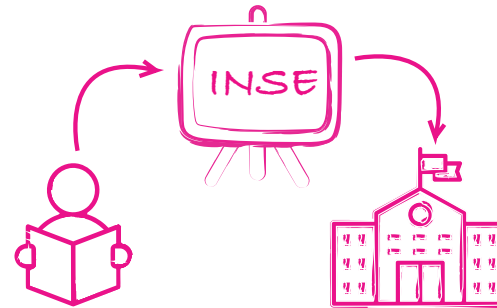
Dicho comportamiento es esperable ya que en estos niveles se presentan la mayoría de las estrategias de enseñanza que influyen las competencias de los estudiantes.

¿Cuántos puntajes se pudieron cruzar para el estudio de aporte relativo?

Hallazgos del desarrollo del modelo de valor agregado y aporte relativo

		Periodo de presentación de Saber Pro		
		2017	2018	Total
Periodo de presentación de Saber 11	2008-1	68	-	68
	2008-2	267	-	267
	2009-1	1795	55	1850
	2009-2	10271	236	10507
	2010-1	2207	1258	3472
	2010-2	17072	11124	28196
	2011-1	4027	2164	6191
	2011-2	25535	15896	41431
	2012-1	7738	4291	12029
	2012-2	41366	25368	67004
	2013-1	7872	6952	14824
	2013-2	12708	39742	52450
	2014-1	-	7123	7123
	2014-2	-	13102	1302
	Total		130926	127588

La elección del modelo principal se dio por las siguientes dos razones:



Hallazgos del desarrollo del modelo de valor agregado y aporte relativo

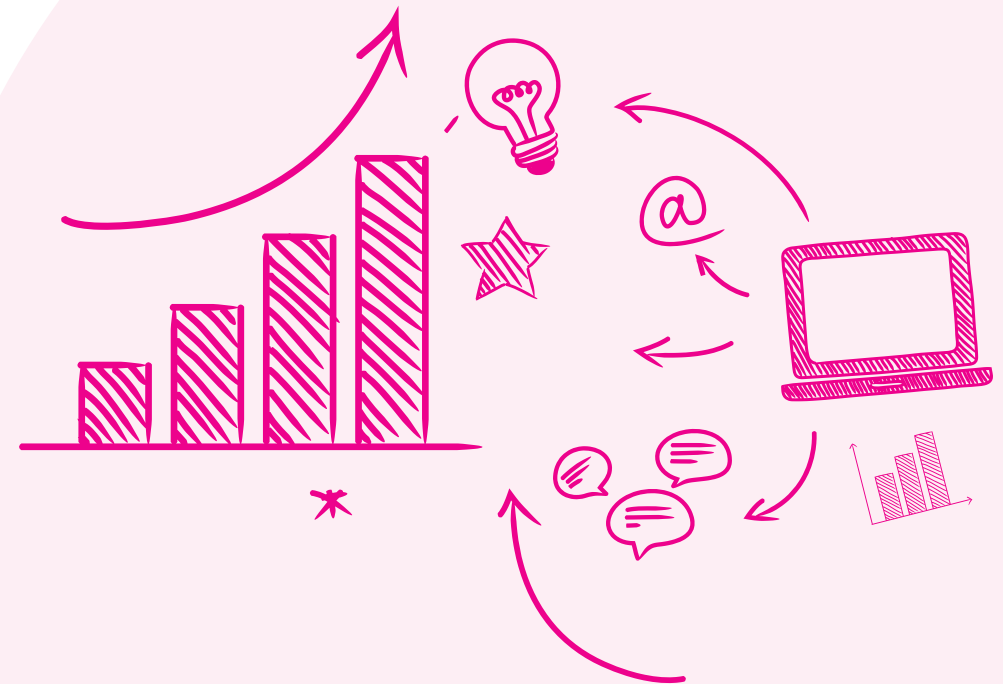
1

Uno de los modelos analizados incluía el Índice de Nivel Socioeconómico. Sin embargo, en este modelo el INSE utilizado proviene del Examen Saber Pro debido a que no es posible calcular este índice para todos los periodos de Saber 11° utilizados. Esto implica que el INSE observado no puede considerarse como una condición inicial del estudiante, previa al ingreso a la educación superior, por lo que teóricamente no se incluye en el modelo.

Una vez se controla por los puntajes previos correctos, la inclusión de variables a nivel de estudiante tiene efectos mínimos en las estimaciones de valor agregado.



2



icfes.gov.co



facebook.com/icfescol



twitter.com/ICFEScol



instagram.com/icfescol



youtube.com/c/icfescol