

**VALOR AGREGADO EN LA EDUCACIÓN SUPERIOR. APLICACIÓN PARA
LAS COMPETENCIAS GENÉRICAS EN LAS PRUEBAS SABER PRO 2012.**

CARLOS ANDRÉS OREJUELA RÍOS

Universidad del Valle

Facultad de Ciencias Sociales y Económicas

Maestría en Economía Aplicada

SANTIAGO DE CALI

2015

**VALOR AGREGADO EN LA EDUCACIÓN SUPERIOR. APLICACIÓN PARA
LAS COMPETENCIAS GENÉRICAS EN LAS PRUEBAS SABER PRO 2012.**

CARLOS ANDRÉS OREJUELA RÍOS

Trabajo de grado para optar al título de:

Magister en Economía Aplicada

Director

Juan Byron Correa Fonnegra

Magister en Economía

Universidad del Valle

Facultad de Ciencias Sociales y Económicas

Maestría en Economía Aplicada

SANTIAGO DE CALI

2015

CONTENIDO

	Pag.
1. INTRODUCCIÓN	7
2. REVISIÓN DE LITERATURA	11
3. MARCO TEÓRICO	18
4. METODOLOGÍA	20
5. DATOS	24
6. RESULTADOS	29
7. DISCUSIÓN	35
REFERENCIAS	37
ANEXOS	39

LISTA DE TABLAS

	Pag.
Tabla 5.1 Información de Evaluados Saber Pro 2012.	39
Tabla 5.2 Información de registros cruzados Saber 11 - Saber Pro 2012.	40
Tabla 5.3 Evaluados por grupo de referencia: total y porcentaje con cruce.	41
Tabla 5.4 Promedios competencias genéricas por grupo de referencia.	42
Tabla 5.5 Razonamiento cuantitativo por sexo y sector.	42
Tabla 5.6 Variables de control Grupo Ingenierías.	43
Tabla 5.7 Evidencia de clasificación cruzada en los datos.	44
Tabla 5.8 Relación colegio estudiantes para una IES.	44
Tabla 6.1 Estimación modelos de valor agregado. Nivel 1: estudiantes. Nivel 2: IES	45
Tabla 6.2 Estimación modelos de valor agregado con clasificación cruzada IES - IEM	46
Tabla 6.3 Ranking de universidades por Valor Agregado en Razonamiento Cuantitativo. Grupo Ingenierías. Mejores veinte clasificados.	47

LISTA DE FIGURAS

	Pag.
Figura 3.1 Metodología de estimación valor agregado – modelo de crecimiento del estudiante.	48
Figura 4.1 Modelo de clasificación cruzada en el nivel 2 (IES y Colegio) con estudiantes en el nivel 1.	49
Figura 5.1 Desempeños Razonamiento Cuantitativo Ingenierías Hombres-Mujeres.	50
Figura 5.2 Desempeños Razonamiento Cuantitativo Ingenierías Público-Privado.	50
Figura 6.1 Mejores veinte, modelo clasificación cruzada, control Saber 11–Ingreso.	51
Figura 6.2 Mejores veinte, modelo dos niveles, control Saber 11–Ingreso.	51
Figura 6.3 Mejores veinte, modelo nulo.	52

VALOR AGREGADO EN LA EDUCACIÓN SUPERIOR. APLICACIÓN PARA LAS COMPETENCIAS GENÉRICAS EN LAS PRUEBAS SABER PRO 2012.

Resumen

Este trabajo explora la medición de valor agregado en la educación superior para Colombia aplicado a las competencias genéricas en razonamiento cuantitativo. Se usan datos publicados por el ICFES como resultado de cruzar las bases de datos de las evaluaciones presentadas por los estudiantes al salir de la educación media, Saber 11, y al salir de la educación superior, Saber Pro, para el año 2012. El principal aporte del trabajo está en la validación de la hipótesis de clasificación cruzada de los evaluados en IES (universidad) y en IEM (colegio). Se encuentra que la varianza del colegio es significativa, cuando se controla simultáneamente por el cruce con la varianza de la universidad, y por lo tanto la estimación de valor agregado debe descontar este efecto para aproximarse a una cuantificación insesgada del aporte de las IES a sus egresados en las competencias analizadas.

Palabras claves: Valor agregado, modelos jerárquicos lineales, modelos de varianza con clasificación cruzada.

Clasificación JEL: J24, I20.

VALUE ADDED IN HIGHER EDUCATION. AN APPLICATION TO GENERIC SKILLS IN SABER PRO 2012 TEST.

Abstract

This paper explores the value-added's measure in Colombia's higher education, applied to generic skills in quantitative analysis. Data are public from ICFES by merging information from evaluations to students at the end of high school, Saber 11, and ending higher education, Saber Pro, in 2012. The main finding relates the hypothesis testing of cross classification by students between college and high school. The high school variance is significant when controlling by cross with college variance, so value added estimations must discount this effect to reach an unbiased measure of college's effect in the skills evaluated to their students.

Keywords: value-added, hierarchical linear models, cross classification variance models.

JEL Codes: J24, I20.

1. INTRODUCCIÓN

La literatura económica se interesa por la calidad de los sistemas de educación en las naciones del mundo actual. Es reconocido su papel relevante en la construcción, y sobre todo en la transformación, de las estructuras productivas de los países que deseen alejarse de sendas de crecimiento dependientes de la explotación de los recursos naturales en busca del desarrollo basado en la ciencia y la tecnología. Colombia no es la excepción, los políticos y especialmente los académicos e investigadores se presentan interesados en desarrollar propuestas que permitan mejorar la calidad de nuestro sistema educativo. Ocupa un lugar relevante en el análisis la necesidad de igualar el desempeño de los mejores sistemas de educación del mundo, de tal forma que podamos integrarnos completamente a la nueva sociedad del conocimiento.

Pero, ¿Cómo evaluamos si una institución del sistema educativo está haciendo bien su trabajo? La primera respuesta sugiere revisar el desempeño observado de sus alumnos, y la medida evidente es el promedio de esos desempeños, o alguna variante del mismo. El problema es que este tipo de miradas estáticas no nos dice nada sobre el desempeño de la institución en el tiempo, y ante todo corresponde a una comparación injusta que no puede cuantificar el progreso de los estudiantes con respecto a su desempeño al ingreso. Más preocupante en ese tipo de estudios es que no se considera la desventaja que tiene para alcanzar altos desempeños promedios una institución que recibe niños en condiciones de desventaja social económica y cultural. Su trabajo puede estar bien hecho, el progreso de sus estudiantes puede ser destacado, pero siempre van a estar mal clasificados porque el entorno y las condiciones iniciales los condenan a tener desempeños inferiores al promedio. Surge entonces la idea de evaluar el progreso de los estudiantes dentro de una institución, y por lo tanto evaluar el aporte de esa institución a la mejora o deterioro del rendimiento de los estudiantes. Este tipo de evaluación dinámica, más próxima a la realidad del proceso educativo, es conocida como análisis de valor agregado. Lo más importante es que en este tipo de análisis se puede identificar con relativa precisión a las instituciones que hacen bien su labor de mejorar la competencia de sus alumnos para desenvolverse en el mundo moderno.

Surge entonces la pregunta ¿Cómo se mide este valor agregado? El ICFES ha hecho público el trabajo de su equipo de investigación en el manejo de los métodos de estimación más usados en la literatura actual (Cuellar, 2014). En general, dicho trabajo corresponde a la implementación de modelos multinivel que son ampliamente aceptados porque coinciden en especificación con la naturaleza jerárquica de los datos. Lo que no se evidencia en los trabajos actuales es la consideración de controlar por el posible efecto que tiene el desempeño observado del colegio (en adelante IEM) de egreso en la educación media, en el valor que se adjudica como agregado por la universidad o equivalente (en adelante IES) en las competencias evaluadas. Esta situación se conoce como clasificación cruzada, ya que los individuos evaluados están asociados simultáneamente a dos grupos o jerarquías. Así una IES recibe estudiantes de muchas IEM, y simultáneamente una IEM envía estudiantes a muchas IES. Este trabajo explora entonces el efecto de esta clasificación cruzada en la medida de valor agregado, estimada con los datos publicados por el ICFES resultado de cruzar las pruebas Saber 11 y Saber Pro. Se trata entonces de identificar si este efecto cruzado es significativo, y en caso de serlo presentar el impacto en la medida de valor agregado obtenida.

¿Por qué es importante esta medida y especialmente la precisión que se tenga en su cálculo? Porque el objetivo fundamental de la educación superior, y de hecho de la educación en todos los niveles, es construir el capital humano de los estudiantes bien sea en forma de ganancias en competencias genéricas o como ganancias en las competencias específicas propias de cada profesión. Se habla de ganancias porque se toman como referencia las condiciones de ingreso del individuo a la institución evaluada. En el caso en que efectivamente se evidencien ganancias en el proceso educativo se dice que el análisis es de valor agregado, y se afirma que la institución evaluada agrega valor en la formación de sus estudiantes. No es justo ni preciso evaluar el desempeño de las instituciones sólo con base en el resultado obtenido por sus estudiantes en una prueba estandarizada, sin tener en cuenta las condiciones de ingreso ni el entorno socio económico del que vienen dichos individuos. Sólo los modelos que consideren el progreso controlando por estas condiciones se clasifican como modelos de valor agregado (OECD, 2013). En los casos en los que se

estime el aporte de las instituciones al desempeño sin dichos controles no es posible garantizar que la medida corresponda razonablemente a lo agregado por la institución.

Si el objetivo principal del sistema educativo es agregar valor (en forma de ganancias en competencias) en el proceso formativo, es razonable medir dicho aporte con el menor sesgo posible. De esta forma, con el empleo de modelos de valor agregado se puede entregar información adecuada a los interesados en usarla para tomar decisiones. Estos pueden ser individuos eligiendo sobre su futuro profesional; o los encargados de la política educativa para diseñar, con información más precisa y confiable, los planes y proyectos que permitan mejorar la calidad de los resultados obtenidos.

Se encuentra que para el año 2012, en competencias genéricas de razonamiento cuantitativo, la varianza asociada al colegio es significativa cuando se controla simultáneamente por el cruce con la varianza de la universidad. Por lo tanto, la estimación de valor agregado debe descontar este efecto para aproximarse a una cuantificación insesgada del aporte de las IES a sus egresados en las competencias analizadas. Se presenta la comparación de las medidas de valor agregado calculadas con los distintos modelos estimados.

El resto del trabajo se organiza en seis secciones adicionales que contienen en su orden los siguientes temas: la revisión de literatura relevante para el estudio, la presentación del modelo teórico de referencia, la descripción de la metodología a usar en el proceso de inferencia causal, la descripción de los datos disponibles, los resultados obtenidos en la estimación del valor agregado en la educación superior, y finalmente una breve discusión de los principales resultados.

Objetivos

El objetivo general del trabajo es la identificación del valor agregado por las universidades colombianas en competencias genéricas a los estudiantes egresados de los programas profesionales. En particular se va a aplicar la metodología a los estudiantes de los programas de ingeniería que presentaron la prueba Saber Pro en el año 2012, los datos son

tomados del sitio web del ICFES encargado del diseño, administración y calificación de las pruebas estandarizadas de educación en Colombia.

Para cumplir con el objetivo general se realizan los siguientes objetivos específicos:

- Cuantificar el valor agregado por las universidades en competencias genéricas condicionado por el desempeño previo de los estudiantes en las competencias relacionadas evaluadas en las pruebas Saber 11.
- Cuantificar el valor agregado por las universidades en competencias genéricas condicionado por las características socio económicas del estudiante.
- Identificar la posible existencia de estructuras de varianza de clasificación cruzada para descontar el valor agregado por los colegios de origen de los estudiantes en las competencias genéricas, buscando una identificación más precisa del valor agregado por las universidades en dichas competencias.
- Usar los resultados obtenidos en la medición de valor agregado para identificar las instituciones de educación superior que más agregan valor en competencias genéricas en los programas de ingeniería.
- Comparar los resultados obtenidos con los diferentes modelos en la clasificación, y por tanto cuantificación, del valor agregado por la universidad.

2. REVISIÓN DE LITERATURA

El uso de sistemas de contabilidad educativa se ha generalizado en casi todos los países del mundo. Se entiende por sistemas de contabilidad al conjunto de evaluaciones del desempeño de las instituciones de educación a partir de medidas del desempeño de los estudiantes de esas instituciones. En la mayoría de los casos los estudiantes son evaluados con pruebas estandarizadas, estas pueden tener alcance local, regional, nacional o internacional. La literatura destaca dos tipos de enfoque para evaluar el desempeño de las instituciones a partir de los resultados en los test aplicados a sus estudiantes: las medidas de prestigio y las medidas de crecimiento. En el caso de las medidas de prestigio la evaluación a la institución se juzga con base en los niveles de desempeño de los estudiantes. Ejemplo de esas medidas son el porcentaje de estudiantes que obtienen un rango de puntajes en la evaluación o el valor promedio de los resultados para los individuos que presentan el test. Las medidas de crecimiento, también conocidas como medidas de ganancia o de valor agregado, clasifican las escuelas-colegios-universidades por la evidencia de mejora de sus estudiantes en el logro observado de un año a otro, o en general en evaluaciones realizadas entre dos periodos diferentes del tiempo en los que al menos una variable a observar cambie, jugando el papel de tratamiento (Figlio y Loeb, 2011).

En relación con el valor agregado hay consenso sobre el resultado principal que debe evaluarse para la educación en los niveles de básica y media. Los estudiantes asisten a la escuela y al colegio para mejorar sus competencias, en el sentido de potenciar el aprovechamiento de sus actividades, en áreas del conocimiento que resultan esenciales para la convivencia en la sociedad moderna. Aquí se adopta el concepto de competencias genéricas que resultan transversales en la vida cotidiana de un adulto, independiente de su decisión de participar en actividades laborales, académicas o investigativas, además de su necesidad de interacción con otros individuos en aspectos de la vida como el social, emocional, familiar o político. De esta forma el objeto de medida sobre el que se agrega valor en los dos primeros niveles de educación no genera controversia en la mayoría de las investigaciones sobre el tema.

Para la educación superior el problema es muy complejo. En este nivel de los sistemas educativos el resultado esperado se puede evaluar en varias dimensiones. En el caso de Colombia se ha decidido por parte del gobierno identificar las ganancias en las competencias genéricas, medida que no parece muy acertada ya que se espera que los individuos busquen desarrollar esencialmente competencias específicas asociadas a la profesión que eligen. Esas competencias específicas pueden ser el segundo resultado a tener en cuenta para identificar una institución como exitosa o no – en este contexto como agregadora de valor o no. El problema importante es que no se tiene una medida previa del desempeño de los individuos en esas competencias específicas, lo que hace difícil y poco confiable la implementación de medidas del valor que se agrega en el proceso formativo. El otro resultado relevante a considerar como producto de la educación superior es el retorno monetario que percibe el estudiante al ejercer la profesión para la que fue formado. Este tipo de análisis entrega información – a hogares, firmas, gobiernos- sobre esos retornos monetarios. En consecuencia permite identificar las instituciones que más agregan valor, en este caso en forma de mayores ganancias con respecto al salario esperado para los individuos egresados de otras instituciones que estudien profesiones comparables (Cuhna y Miller, 2014).

El término valor agregado es tomado de la literatura sobre la teoría económica de la producción. En ese campo, el valor agregado se refiere a la cantidad de valor que se incrementa a una mercancía en cada etapa del proceso productivo. En el caso de la educación se supone que el capital humano acumulado se puede identificar en cada instante del tiempo. Por lo tanto, será posible medir el cambio en ese capital humano asociado a los factores de producción usados en el periodo evaluado (Koedel, Mihaly y Rockoff, 2015).

Pero, no se trata solo de un préstamo del término. El uso y medida del constructo de valor agregado está asociado al supuesto de existencia de una tecnología de producción de educación, o la posibilidad de especificar una función de producción educativa. La discusión es planteada en un artículo seminal por Hanushek (1979) en el que el autor no sólo describe los requisitos que debe cumplir tal especificación, sino que incluye de manera precisa las limitaciones teóricas y empíricas que este enfoque puede tener. En el área

específica del valor agregado Raudenbush y Willms (1995) describen la forma funcional que resulta más usada en los trabajos recientes. En este primer trabajo presentan como grupos de factores asociados al desempeño de los estudiantes los siguientes: práctica escolar, efecto contexto asociado a la institución y características del individuo y su entorno familiar. Lo más importante de este trabajo es la identificación de los requerimientos para obtener medidas de valor agregado tanto a nivel de profesor como de escuela. Este análisis es completado con los trabajos posteriores de Raudenbush (2004) y, Reardon y Raudenbush (2009) que discuten los supuestos sobre los que se basa la estimación insesgada y consistente de los efectos causales del valor agregado. Los trabajos presentan un análisis sobre la viabilidad de esos supuestos y las consecuencias de su no cumplimiento.

El elemento central en esta discusión es que la medida de valor agregado corresponde a un ejercicio de inferencia causal. Esta medición enfrenta dos retos enormes cuando se aplica el modelo al sistema educativo. El primero es que no existen experimentos controlados, y ni siquiera son éticamente aceptables, en los que se someta el sistema educativo a un tipo específico de controles y tratamientos que permitan identificar de manera inequívoca el efecto causal de cada elemento –factor- del sistema. El segundo es que la evidencia indica que las selecciones de escuela, colegio y universidad distan mucho de ser aleatorias. Hay un claro proceso de selección de los padres en las primeras etapas, y del estudiante en las superiores que está altamente correlacionado con los desempeños observados en las evaluaciones estandarizadas. El problema se acrecienta en la medida que el individuo pasa entre los diferentes niveles de educación. En el nivel básico es posible identificar razonablemente el efecto del profesor en las evaluaciones de competencias genéricas, esto porque generalmente un profesor se encarga de un curso completo. De hecho el objetivo de ese primer nivel de formación es el desarrollo de competencias genéricas con énfasis en lecto-escritura y pensamiento lógico-matemático. Al terminar el nivel básico y en el nivel medio el individuo tiene un profesor para cada área de conocimiento. En esta etapa, con las competencias genéricas desarrolladas es posible presentar a los estudiantes saberes específicos. Ahora es posible incluir como unidad de análisis la escuela –o el colegio-, pero

el efecto estimado se puede separar por las ganancias atribuibles a los profesores y sus prácticas, al contexto escolar o al arreglo institucional (Reardon y Raudenbush, 2009).

En el caso de la educación superior, la identificación de efectos causales atribuibles a los actores del sistema es menos probable. No sólo por la diversidad de interacciones posibles dentro y fuera de la universidad que pueden afectar los desempeños evaluados, sino como se ha mencionado por los posibles resultados esperados en este nivel de formación (Yunker, 2005; Rothstein, 2007). Lo anterior se suma a la dificultad de construcción de la información relevante para obtener estimaciones razonables. En general se encuentran tres tipos de medidas denominadas cohorte a cohorte, cuasi-longitudinal y longitudinal según la muestra analizada. En el caso de los estudios entre cohortes se evalúan las ganancias de los estudiantes de un año y nivel con respecto a los estudiantes del mismo nivel pero de un periodo anterior, los estudios cuasi longitudinales comparan el desempeño de un grupo de estudiantes para un año y nivel con los de el nivel anterior en el periodo anterior, finalmente los estudios longitudinales evalúan las ganancias promedio de los individuos de un año y nivel con el de ellos mismos en el periodo y nivel anteriores (Koretz, 2008).

Una vez definido el grupo a estudiar debe elegirse la metodología empírica a utilizar para estimar el valor agregado. Debe destacarse que los trabajos se concentran en mayor cantidad en la identificación de los efectos asociados al profesor en educación básica, esto claramente obedece a la mayor disponibilidad de información ya que en algunos países desarrollados los estudiantes son evaluados cada año. Estados Unidos y el Reino Unido fueron pioneros en la implementación de medidas de valor agregado obligatorias para este nivel de educación (Houng y Justman, 2013) (Tekwe, y otros 2004). En relación con la medición de valor agregado en la educación superior hay tres grupos de modelos dominantes. El primer modelo es un modelo de diferencias en residuales que usa OLS y tiene como unidad de análisis las instituciones. En este caso se compara el valor de una prueba promedio para dos grupos de estudiantes ubicados en cohortes extremas (principiantes y candidatos a grado) que han presentado una prueba común antes de ingresar a la educación superior. Se estima el residual para cada grupo y el valor agregado es la diferencia entre los residuales de los dos grupos. El segundo modelo también estima

diferencias en residuales, pero ahora usa una especificación jerárquica lineal o multinivel. Los sujetos comparados son los mismos del modelo de diferencias en OLS, sólo que ahora se utiliza la información disponible en componentes de varianza agrupando por instituciones -en el nivel 2- y por individuos -en el nivel 1. El valor agregado corresponde a la diferencia en los residuales estimados para el nivel 2 en los dos grupos. Los dos modelos tienen como problema fundamental la validez de la hipótesis de comparabilidad de los residuales teniendo en cuenta que se analizan dos cohortes diferentes de estudiantes que no necesariamente serán objeto de las mismas prácticas pedagógicas ni del mismo ambiente institucional. El tercer modelo estudia la misma cohorte al principio y al final del ciclo educativo, se trata entonces de un análisis longitudinal, y compara el resultado obtenido en la evaluación a la salida con el resultado esperado dados los desempeños a la entrada. La estructura es multinivel y el valor agregado se estima con el residual del nivel dos que corresponde a las instituciones. Estos valores agregados se pueden calcular en forma de intervalos de confianza con mayor precisión de la que pueden entregar los dos modelos previos (OECD, 2013).

En relación con el grupo de covariantes a incluir en el modelo, hay consenso sobre la obligatoriedad de controlar por el desempeño del estudiante previo al ingreso a la educación superior. Claramente se busca que las escalas sean comparables, es decir que las evaluaciones deben estar encaminadas a medir el desempeño en las mismas características de los estudiantes. Esta condición de comparabilidad es especialmente difícil de lograr para la educación superior, más si se tiene en cuenta que el resultado para evaluar el logro no necesariamente es el desempeño en unas competencias medidas en test estandarizados. De hecho un grupo importante de estudios se dedican a identificar el efecto atribuible a las universidades en los mercados laborales, en este caso en forma de salarios que remuneran el capital humano acumulado en todo el ciclo educativo. La consideración de otras variables representa un reto en términos del sesgo que su inclusión y su omisión pueden introducir al modelo. La discusión más relevante está relacionada con los índices de estatus socio económico del entorno familiar del estudiante. Las variables socio económicas deben ser incluidas porque ellas están asociadas a las ganancias que los estudiantes tienen en su

ciclo académico. Si la idea es identificar el aporte de la institución, omitir el control por el entorno familiar le asignará erróneamente toda la ganancia al establecimiento cuando una parte es aportada por su núcleo social y familiar. De otra parte, si se incluyen los indicadores del estatus socio económico estos se asocian directamente con la calidad observada de las instituciones evaluadas. Los estudiantes hijos de las familias con más ingreso -generalmente los hijos de padres con mayor capital humano- están en las escuelas con mayor desempeño observado en las pruebas estandarizadas, escuelas que a su vez tienden a contratar mejores profesores y por lo tanto se puede presentar un sesgo en los valores estimados atribuibles a las escuelas. La cuestión clave es decidir cuál es el sesgo aceptable en cada estudio, ya que este problema tiene diferentes dimensiones en las distintas metodologías propuestas. (Ballou, Sanders y Wright 2004)

Son muchos los estudios realizados en el mundo en relación con el efecto profesor o el efecto escuela en los niveles de educación básica y secundaria. Es importante mencionar que el volumen de trabajos realizados en los Estados Unidos aumentó de manera significativa a partir de la ley conocida como “No Child Left Behind”, promulgada por la administración Bush en el año 2001, que obliga a los estados de la unión a fijar estándares medibles que permitan determinar mediante evaluaciones, entre los grados tercero a octavo, las escuelas que están mejorando y las que empeoran en relación con el desempeño de sus estudiantes. Estos resultados permiten decidir la continuidad de financiación para las instituciones que mejoran y el cierre de las que empeoran. La mayoría de estados ha diseñado estrategias de evaluación alrededor del concepto de valor agregado, pero lo asignan a los docentes aprovechando la evaluación continua en esos cinco años del ciclo educativo (Chetty, Friedman y Rockoff, 2013; Kelly y Downey, 2010; Fagioli, 2014; Manzi, SanMartin y van Belleghem, 2010).

En el caso de la educación superior apenas está en crecimiento el número de estudios, esto lleva a tener pocos resultados comparables relacionados con la formación en este nivel, especialmente por la diversidad de información usada como producto esperado de las universidades. Es relevante destacar que son pocos los países que tienen implementadas pruebas estandarizadas para los egresados de programas profesionales, así que en muchos

casos se recurre a estimar el valor agregado a partir de los retornos en los mercados laborales o en otras medidas dentro del ciclo educativo, entre las que destaca la tasa de graduación (Cuhna y Miller, 2014; Best y Keppo, 2014; Breslin y Samantha, 2014; Su-ran, Li-hiu, Guang-li y Hui-chan, 2014; Pedrosa, Amaral y Knobel, 2013).

En Colombia, no obstante la información disponible, son pocos los trabajos desarrollados en esta área, entre los que destaca el estudio de Balcazar y Ñopo (2014) que comparan el valor agregado en competencias genéricas a los estudiantes de programas de docencia con el valor agregado a estudiantes del resto de programas evaluados en 2011. Como variable de control usan el desempeño en Saber 11 en un modelo lineal, aunque controlan por sesgo de selección teniendo en cuenta que comparan estudiantes de diferentes programas. Encuentran evidencia de menor desempeño de los programas de licenciatura tanto al ingreso como a la salida de su formación profesional, pero como agravante son también el grupo que menos valor agregado recibe en la educación superior. Saavedra y Saavedra (2011) realizan un estudio entre cohortes con la prueba piloto que el ICFES realiza en 2008 en el proceso de construcción de los entonces ECAES. Usan estimación de paneles con efectos fijos y controlan la selección por un score de propensión a estar en la muestra por programas y universidades que pondera las observaciones de manera diferencial. Su resultado indica que los estudiantes ganan en todas las competencias evaluadas como resultado de asistir a la educación superior.

La mayor parte de los trabajos en el país se han ocupado de las medidas de eficacia escolar en cortes transversales, en estos modelos se identifican los factores asociados al desempeño de los evaluados en cada nivel de información incluido (pueden ser de estudiantes, colegios, ciudades, departamentos). No obstante no pueden clasificarse como de valor agregado ya que no controlan por el desempeño previo de los estudiantes evaluados, esto porque en la mayoría de los casos esa información no está disponible.

3. MARCO TEÓRICO

En la literatura económica se supone la existencia de una función de producción educativa (Pritchett y Filmer 1999) de tal forma que el producto del proceso educativo está directamente relacionado con un vector de insumos y factores disponibles en el sistema. Dicha función es presentada en (Checchi 2005) en su forma implícita como:

$$Y_i^t = f(\mathbf{X}) \quad (3.1)$$

En los estudios de contabilidad educativa el resultado relevante para evaluar la calidad de una institución es el desempeño de los estudiantes, Y_i^t . Si el colegio se modela como una firma, el desempeño del estudiante será el resultado de aprovechar un vector de entradas – factores- (\mathbf{X}) con una tecnología disponible $f(\cdot)$ en todo el sistema educativo.

En el caso específico del valor agregado, Raudenbush (2004) propone que el vector de insumos y factores –INPUTS- está compuesto por:

- Habilidad no observada del estudiante, A_{it} .
- El background del estudiante, S_{it} , que contempla una medida previa de aptitud y/o del estatus social y económico en el hogar, o en general de su capital social.
- La práctica académica, P_{it} , que incluye -por ejemplo- los recursos escolares, la estructura organizacional, la calidad y efectividad de la planta docente.
- La contribución del contexto escolar, C_{it} , que puede incluir el promedio del nivel social y económico para la escuela, la tasa de desempleo de la comunidad, las tasas de incidencia de la violencia, entre otras.

De esta forma (3.1) se puede presentar como:

$$Y_i^t = f(A_{it}, S_{it}, P_{it}, C_{it}) \quad (3.2)$$

En esta formulación pueden definirse dos tipos de valor agregado (Raudenbush y Willms 1995). El valor agregado tipo A, corresponde a la diferencia entre el desempeño actual del individuo y el desempeño esperado si él hubiese asistido a una escuela con características promedio tanto en práctica como en contexto. En este caso la efectividad de la escuela

puede generarse por su equipo de trabajo, por la composición de estudiantes, o por la influencia social y económica de la comunidad en la que se encuentra la escuela. El valor agregado tipo B, definido como la diferencia entre el desempeño de un individuo en una escuela dada y el desempeño esperado en caso de asistir a una escuela con un contexto semejante pero con prácticas escolares de diferente efectividad.

La figura 3.1 corresponde a una adaptación del modelo de valor agregado basado en el crecimiento del desempeño para estudiantes presentado en Raudenbush (2004). La línea punteada representa la trayectoria del valor esperado en el desempeño de los estudiantes con características promedio que se encuentran en una institución con prácticas y contexto promedio. Se supone que en el momento t , dos estudiantes $i(1)$ e $i(2)$ se encuentran exactamente con desempeños observados iguales al promedio ($Y_{i,j}^t$). Entre el periodo t y $t+1$ cada uno es asignado a una escuela diferente, identificadas como $j(+)$ y $j(-)$. Los cambios en los desempeños observados hasta $t+1$ están representados por las dos curvas continuas, una por encima (trazo más grueso) y otra por debajo de la línea de desempeño esperado ($Y_{i,j(0)}^{t+1}$).

Para el estudiante $i(1)$ el proceso ha sido virtuoso: su desempeño en $t+1$ es superior al esperado. Se concluye que la escuela $j(+)$ le agrega valor exactamente en la cantidad que separa verticalmente las dos trayectorias ($Y_{i(1),j(+)}^{t+1} - Y_{i,j(0)}^{t+1}$). Para $i(2)$ la situación es inversa: el desempeño observado en $t+1$ está por debajo de lo que se esperaba. En consecuencia la escuela $j(-)$ no le ha agregado valor en ese periodo. De hecho el total restado es la diferencia vertical entre las trayectorias observada y esperada ($Y_{i(2),j(-)}^{t+1} - Y_{i,j(0)}^{t+1}$). Como no se discrimina entre efectos de práctica ni de contexto, el valor agregado que se estima corresponde al tipo A. En este trabajo se estima el valor agregado tipo A por las instituciones de educación superior (IES) en Colombia en competencias genéricas de razonamiento cuantitativo.

4. METODOLOGÍA

El sistema educativo está organizado en grupos, esto genera información jerarquizada, en la que es posible identificar componentes de varianza. Esto implica que el modelo estadístico adecuado para realizar inferencia con esta información es el modelo jerárquico lineal o modelo multinivel. Dado que se va a realizar la estimación de valor agregado entre dos instantes del tiempo, el subíndice temporal no es necesario en el modelo. Dado que se incluye como unidad de análisis en el segundo nivel las IES, ellas se identifican con el subíndice j que toma valores $1, 2, \dots, J$. De esta forma para cada individuo, i , que pertenece a la IES j , tendremos una observación de su desempeño, Y_{ij} , y de su vector de covariantes, X_{ij} .

Siguiendo las exposiciones de (Raudenbush y Willms 1995) y (S. Raudenbush 2004) el valor agregado tipo A corresponde a la diferencia entre el desempeño potencial del individuo i en la escuela j , $Y_{ij}(S_i, P_j, C_j, \varepsilon_{ij})$, y su desempeño potencial en la escuela j^* , $Y_{ij^*}(S_i, P_{j^*}, C_{j^*}, \varepsilon_{ij^*})$. En cambio, el valor agregado tipo B será la diferencia entre el desempeño potencial del individuo i en la escuela j cuando la práctica escolar es P_j , $Y_{ij}(S_i, P_j, C_j, \varepsilon_{ij})$, y su desempeño potencial cuando en la escuela j se usa la práctica P_j^* , $Y_{ij}(S_i, P_j^*, C_j, \varepsilon_{ij}^*)$. Estimar el valor agregado tipo B requiere información detallada de los elementos que constituyen la práctica escolar. Generalmente esos datos no están disponibles y este caso no es la excepción. Cabe resaltar que el objeto de estudio en esta investigación es la IES como un todo, lo que incluye tanto sus prácticas educativas como el contexto en el que desarrolla su labor formativa. Esta situación nos lleva a identificar las condiciones necesarias para estimar el valor agregado tipo A.

Un tratamiento debe ser manipulable y sus efectos concebidos en relación con otros tratamientos para los que el sujeto estudiado pueda ser asignado. El estatus socioeconómico de un estudiante es un atributo, mientras que los posibles métodos de instrucción son tratamientos. Para el estudiante de educación superior la suma total de influencias que recibe en la IES, lo que incluye los atributos de sus pares, puede ser visto como un tratamiento. Se puede someter el estudiante a diferentes efectos par, cambiando de IES. En

este caso hablamos de valor agregado tipo A que resulta de una mezcla compleja de tratamientos que incluye normas de los pares y de la comunidad académica, además de las políticas y prácticas aplicadas en la enseñanza. Dadas J IES, el valor agregado tipo A corresponde a la diferencia entre el desempeño observado del estudiante y el promedio de las J posibles medidas para ese estudiante. Pero, sólo hay un desempeño observado por estudiante, por lo que la solución consiste en realizar un estudio aleatorio que garantice una asignación de tratamientos independiente de los J resultados por estudiante. El problema es que en educación no se dispone de experimentos de esa naturaleza. Los autores (Raudenbush y Willms 1995) proponen la condición de asignación de tratamientos fuertemente ignorables, lo que implica que los J desempeños del estudiante sean condicionalmente independientes de la asignación del tratamiento dado un conjunto de covariantes.

La pertenencia a una IES no es un proceso aleatorio, de hecho está determinado entre otros elementos por la segregación socioeconómica o racial, lo que garantiza que las IES serán diferentes en las características de ingreso de sus estudiantes. Se considera la presencia de tratamientos fuertemente ignorables cuando se incluyen covariantes relacionadas con el desempeño que afectan la propensión de un estudiante a asistir a una IES dada. Estas covariantes pueden ser una medida previa del desempeño o la aptitud que, junto a indicadores de entorno social permiten estimar razonablemente el valor agregado tipo A.

El modelo estadístico propuesto es:

$$Y_{ij} = \beta_0 + \beta_s S_{ij} + \beta_p P_{ij} + \beta_c C_{ij} + \varepsilon_{ij} \quad (4.1)$$

Que corresponde a un modelo jerárquico lineal en el que las unidades del primer nivel son estudiantes, i , y las del segundo nivel son IES, j . Las variables se corresponden a los factores de la función de producción educativa presentada en el marco teórico. El valor agregado tipo A, estará dado por:

$$VA_{Aij} = \beta_p P_{ij} + \beta_c C_{ij} \quad (4.2)$$

que incluye los efectos causados por la práctica y también los del contexto social. Estos efectos pueden ser calculados por adición, lo que implica tener estimadores insesgados y consistentes de β_p y β_c , así como la información relevante de P_{ij} y C_{ij} . Dada la dificultad para obtener información comparable de las prácticas educativas, es más viable obtener los valores agregados estimados por substracción:

$$VA_{Aij} = Y_{ij} - \beta_0 - \beta_s S_{ij} - \varepsilon_{ij} \quad (4.3)$$

Que será estimado sin sesgo, cuando β_s es estimado sin sesgo, lo que a su vez requiere tener la información relevante del estudiante y un método de estimación insesgado.

Si se define un modelo jerárquico lineal de la forma:

$$Y_{ij} = \alpha + \beta_s S_{ij} + \mu_j + \varepsilon_{ij} \quad (4.4)$$

En el que se puede obtener para cada escuela:

$$\mu_j = Y_{ij} - \alpha - \beta_s S_{ij} - \varepsilon_{ij} \quad (4.5)$$

Se puede ver que μ_j en (4.5) tiene la misma forma de VA_{Aij} en (4.3). El problema es que los valores estimados serán sesgados en presencia de efectos contextuales (omitidos), aunque es consistente en la medida que aumente el tamaño de la muestra en la IES.

Cabe destacar en este punto que, tal como anotan Cunha y Miller (2014), este modelo identifica las diferencias promedio en los resultados condicionales entre universidades, o medidas relativas de valor agregado. Estas medidas pueden ser estimaciones de valor agregado causal al considerar que las características observables del estudiante, previas al ingreso, controlan por las diferencias en la selección de la IES.

La pregunta central de este trabajo está relacionada con la identificación de la influencia de los colegios de los que egresan los estudiantes (con sus efectos contextuales y de práctica pedagógica), que en caso de resultar significativa estaría siendo asignada equivocadamente como valor agregado a la IES, sobreestimando la contribución de esta última. Goldstein (2011) define este modelo como multinivel de clasificación cruzada, en el que las unidades

de nivel 1 (estudiantes) pueden estar clasificadas en diferentes conjuntos del nivel 2. El modelo ahora toma la forma:

$$Y_{i(j_1j_2)} = \alpha + \beta_s S_{i(j_1j_2)} + \mu_{j_1} + \mu_{j_2} + \varepsilon_{i(j_1j_2)} \quad (4.6)$$

En el que μ_{j_1} corresponde a la IES actual, y μ_{j_2} corresponde al colegio del que egresó el estudiante. Si la varianza del error de nivel 2 para el colegio $-\mu_{j_2}$ resulta significativa tendremos evidencia de valor agregado que corresponde a los colegios de origen que por lo tanto no debe ser asignado a la IES.

En la figura 4.1 se presenta una posible configuración de clasificación cruzada, hay cinco estudiantes ($i1, \dots, i5$) que asisten a dos IES ($j1-1$ y $j1-2$) y terminaron la educación media en dos colegios ($j2-1$ y $j2-2$). Se puede observar que la IES identificada como $j1-1$ recibe un estudiante de cada uno de los colegios, mientras la IES $j1-2$ recibe un estudiante del colegio $j2-1$ y dos del otro colegio, $j2-2$. Con este ejemplo es claro por qué el desempeño de los estudiantes se identifica con $Y_{i(j_1j_2)}$. Desde el punto de vista de los estudiantes se puede observar que en el caso del estudiante $i1$ asiste a la IES $j1-1$ y es egresado del colegio $j2-1$. El mismo análisis se puede repetir para los otros estudiantes. Se espera que los dos grupos de instituciones tengan efectos en el aprendizaje del estudiante, y este efecto será mayor en el caso de las competencias genéricas que son objeto de formación en los dos niveles de educación. El objetivo del trabajo es identificar la varianza que sea atribuible a cada unidad del segundo nivel y no suponer que todo el valor es agregado exclusivamente por las IES.

Debe anotarse que en el modelo no se incluyen variables del segundo nivel porque el modelo de valor agregado tipo A busca identificar todos los efectos que pueden atribuirse a la institución por su práctica académica y por el contexto en el que la misma se desarrolla. La inclusión de variables de universidad afectarán el valor agregado estimado ya que se espera que ellas estén altamente correlacionadas con la práctica académica y el efecto contextual que se busca estimar.

5. DATOS

Evaluación de la educación media y superior en Colombia

Colombia tiene una oportunidad interesante y escasa en el mundo. Todos los estudiantes que van a egresar de la educación superior deben presentar una evaluación que incluye tanto competencias genéricas como específicas. En la actualidad esas pruebas son denominadas Saber Pro por el ICFES. Prácticamente todos esos estudiantes han presentado el examen de estado al terminar el ciclo de educación media (en una IEM). En la actualidad estas pruebas son denominadas Saber 11. Es decir, tenemos individuos evaluados en dos momentos de su proceso educativo. Por lo tanto, podemos hacer el ejercicio de medir la evolución en el tiempo de su desempeño académico en varias de las competencias incluidas en las dos pruebas. Aunque es necesario advertir que los datos disponibles hoy no corresponden a evaluaciones estrictamente comparables de competencias. Alineado con esta situación, el decreto 3963 fija como objetivo de los exámenes de estado de calidad de la educación superior la producción de indicadores de valor agregado para este nivel de formación. Siguiendo los lineamientos del decreto mencionado, el ICFES ha definido para el año 2014 el inicio de un ciclo de 12 años de pruebas comparables en los dos niveles de educación. En la búsqueda de ese objetivo fue necesario rediseñar las pruebas Saber 11 de tal forma que estén completamente alineadas con la cuantificación de las competencias genéricas evaluadas en Saber Pro. Es decir, que en el año 2019-2020 se puede contar con el primer conjunto completo de información coherente con la nueva metodología de medición.

Uno de los principales problemas que enfrenta el ICFES está relacionado con el cruce de información entre las bases de las dos evaluaciones. Dicho cruce de información se debe hacer porque generalmente el estudiante cambia de tipo y número de documento de identidad en los dos periodos. Generalmente al presentar Saber 11 el documento es la tarjeta de identidad, al presentar Saber Pro el documento es la cédula de ciudadanía. Los dos números son diferentes y es necesario agotar una serie de procedimientos tendientes a identificar los cruces de las dos bases de datos (ICFES 2012). Con el nuevo siglo, Colombia completó un cambio en el proceso de identificación de los ciudadanos y el

número de identidad será único para la mayoría de evaluados en los siguientes años. Por ahora, el ICFES nos entrega la base que resulta de su proceso de cruce de información, que por razones de protección de la información personal ya son de carácter anónimo.

Información de las pruebas Saber Pro

El examen Saber Pro lo presentan los estudiantes antes de egresar del sistema de educación superior formal, que incluye los niveles técnico, tecnológico y profesional. Presentarlo es un requisito obligatorio para el grado a partir del año 2010, pero no es una prueba de tipo aprobatorio, sólo se exige al estudiante evidencia de la presentación. Entre los datos publicados por el ICFES con cruce de información los que corresponden al año 2012 son los que tienen más registros por lo tanto serán usados en este estudio. La tabla 5.1 presenta las cifras de estudiantes evaluados, además del detalle por sexo, tipo de institución y nivel de formación para los registros de ese año. Entre los hechos que se pueden destacar está que el número de mujeres que presenta el examen es mayor que el de hombres en las dos aplicaciones; la mayor proporción de evaluados provienen de instituciones educativas clasificadas como Universidades, correspondientes en cada caso al cincuenta y seis por ciento de los evaluados (56%); la mayoría de estudiantes están en formación de tipo universitaria, seguida por tecnológica, técnica y por último normalistas; el primer grupo representa el sesenta y ocho por ciento del total (68%).

Sólo los registros hallados en las dos bases de datos (Saber 11 y Saber Pro) pueden ser usados para estimar el valor que agregan las IES en competencias genéricas a sus estudiantes. Se tiene que del total de evaluados en nivel universitario, el veintinueve (29%) y el treinta y uno (31%) por ciento tienen información identificada con cruce en las dos pruebas. La tabla 5.2 presenta las cifras de registros cruzados, además del detalle por sexo y tipo de institución para los registros de ese año. De estos el sesenta y dos por ciento (62) son mujeres y la mayoría (82%) provienen de universidades.

En todos los casos, los evaluados por lo menos presentan el examen de competencias genéricas que son la base del presente estudio de valor agregado. Aunque las competencias específicas son la razón de ser de la educación superior, debe recordarse que no se cuenta

con pruebas comparables de esta naturaleza al ingreso de los estudiantes a las IES o a la salida de la educación media. En la evaluación es necesario tener en cuenta: las especificidades de las diferentes áreas del conocimiento, la existencia de métodos y prácticas académicas apropiadas para cada saber y, los énfasis que cada disciplina puede implicar en el desarrollo diferencial de algunas de las competencias clasificadas como genéricas. Lo anterior da origen a la clasificación de los evaluados en grupos de referencia, de tal forma que cada grupo recoge programas afines en los criterios mencionados.

Para el año 2012 el ICFES definió un total de 23 grupos. La tabla 5.3 muestra el total de evaluados por grupo y el porcentaje para los que hay cruce exitoso de información en este periodo en las bases de Saber 11 y saber Pro. Los grupos que se omiten corresponden a niveles de formación diferente de la universitaria o profesional. Se puede destacar que los grupos más grandes de evaluados corresponden a estudiantes a egresar de las facultades de administración y afines, ingeniería, educación y salud en orden descendente. Pero, los grupos que presentan mayores casos exitosos de cruce son los que corresponden a las facultades/programas con menor número de estudiantes evaluados.

Variable de respuesta

Las competencias genéricas que se evalúan en Saber Pro incluyen: lectura crítica, razonamiento cuantitativo, comunicación escrita e inglés. En todos los casos los resultados se normalizan con media diez (10) y desviación estándar uno (1). Aunque todos los programas presentan pruebas de competencias genéricas, es necesario recordar que la forma funcional especificada incluye como factor de la tecnología de producción educativa el término P_{it} que representa la práctica académica. Esto coincide con la separación que hace el ICFES en grupos de referencia, se trata de programas afines, por lo tanto con prácticas académicas semejantes que harán énfasis en unas competencias más que en otras. Se espera que el efecto causal atribuible a una IES en una competencia específica sea comparable con los de otra IES dentro de cada grupo de referencia. Así el modelo elegido se usa para identificar el valor agregado en razonamiento cuantitativo para los programas de pregrado en ingeniería.

Como resultado de la consideración anterior se pueden estimar modelos de valor agregado para los quince (15) grupos de referencia con información por cada una de las cuatro (4) competencias genéricas evaluadas. La tabla 5.4 presenta para los grupos de referencia los promedios, con su error estándar, de los resultados observados en las competencias genéricas para la población total.

Destacan entre esos resultados los siguientes hechos: los grupos de economía y medicina tienen resultados superiores en promedio en todas las competencias genéricas, los grupos de ingeniería y de ciencias naturales destacan en razonamiento cuantitativo, el grupo de humanidades no se destaca en razonamiento cuantitativo, el grupo de bellas artes tiene el mejor desempeño en inglés, se destacan en comunicación escrita los grupos de ciencias sociales, comunicación-periodismo y derecho. Los peores resultados se encuentran en el grupo de educación y en el de salud (diferente a medicina).

El presente ejercicio usa como respuesta los resultados en competencias genéricas de razonamiento cuantitativo para el grupo de referencia en ingenierías. La figura 5.1 muestra la relación -dentro de las IES- entre el desempeño de hombres y mujeres, en el grupo de ingenierías, para la prueba de razonamiento cuantitativo. Se puede observar claramente que los hombres obtienen desempeños superiores en esta prueba.

La figura 5.2 muestra la relación entre el desempeño de instituciones públicas y privadas, en el grupo de ingenierías, para la prueba de razonamiento cuantitativo. Hay una pequeña ventaja a favor de las públicas.

La tabla 5.5 presenta las descriptivas de los dos casos comparados: sector de la IES y sexo del estudiante. Quedan claras las ventajas en desempeños promedios a favor de los hombres y de las instituciones públicas, en ambos casos acompañados de mayor dispersión del grupo con mayor media. Las diferencias son estadísticamente significativas.

Variables de control

La especificación presentada en el modelo teórico y en la metodología implica el uso de variables que permitan capturar efectos de selección inobservables que puedan estar

correlacionados con el resultado y la elección de la IES, esto requiere como mínimo una medida de desempeño previo del individuo y una medida de su background familiar al momento del ingreso a la educación superior. Como medida del desempeño previo en la competencia genérica al momento de ingreso a la IES se usa el desempeño en las pruebas evaluadas en Saber 11 (los resultados están estandarizados), con interés en el resultado de la prueba de matemáticas. Es relevante mencionar que en relación con el contenido las dos evaluaciones corresponden a la misma área del conocimiento, es decir que interpretamos esta prueba de matemática como la prueba relevante a ser tomada en cuenta como control para la medida de valor agregado.

En relación con las características del estudiante observadas previo al ingreso a la IES, se cuenta con algunas preguntas de la encuesta socioeconómica diligenciada al presentar Saber 11 (entre las que destacan por su uso frecuente en la literatura el ingreso de la familia, la educación y la ocupación de los padres), además está disponible un índice social y económico calculado por el ICFES con base en esas respuestas. La tabla 5.6 muestra las descriptivas para las variables a utilizar, corresponden a los resultados de saber 11 para los años 2006, 2007 y 2008.

Los datos usados en los modelos estadísticos estimados que permiten identificar los niveles de información son:

- Nivel 1: Código del estudiante en Saber Pro y Saber 11.
- Nivel 2: Código IES en la que presenta Saber Pro.
- Nivel 2 cruce: Código IEM en la que presenta Saber 11.

Las tablas 5.7 presentan la evidencia de información con clasificación cruzada. Las IES reciben estudiantes de varios colegios, y los colegios tienen egresados estudiando en varias IES. Por facilidad de lectura se muestra sólo una porción de la tabla de cruce. En el caso de la tabla 5.8 se relaciona el código del colegio y de estudiantes para una IES. La conclusión es la misma, la información presenta doble clasificación en el segundo nivel.

6. RESULTADOS

El primer paso en la construcción de la validación de la hipótesis planteada consiste en identificar la existencia de información con estructura jerárquica. Esto se hace mediante una prueba de identificación de componentes de varianza en un modelo nulo multinivel.

El modelo de partida es:

$$Y_{ij} = \beta_0 + \varepsilon_{ij} \quad \text{con } \beta_0 = \beta_{00} + \mu_j \quad (6.1)$$

Los resultados obtenidos en la estimación de los modelos en dos niveles se presentan en la tabla 6.1. En la columna (I) se observa que hay evidencia estadística bastante robusta en favor de un modelo de componentes de varianza frente a un modelo lineal. Destaca en este caso una correlación intraclase de aproximadamente veintitrés por ciento (23%), relativamente baja comparada con la misma medida en las pruebas Saber 11 o PISA que - en cálculos del autor- están cercanas al cuarenta y tres por ciento (43%). Esa correlación indica que las diferencias en el desempeño observado son atribuibles en veintitrés por ciento a las IES y el restante setenta y siete por ciento (77%) corresponde a diferencias entre individuos.

En el modelo presentado en la columna (II) se agrega el control por desempeño previo que se puede expresar como:

$$Y_{ij} = \beta_0 + \beta_Y Y_{ij,PRE} + \varepsilon_{ij} \quad \text{con } \beta_0 = \beta_{00} + \mu_j \quad (6.2)$$

Se observa que el coeficiente estimado para esta covariante es estadísticamente significativo, con un efecto marginal importante: un incremento en una desviación estándar en el desempeño en matemáticas en la prueba Saber 11 está asociado a un incremento de 0.438 desviaciones estándar en el desempeño en razonamiento cuantitativo en la prueba Saber Pro (es preciso recordar que ambas medidas están en una escala con varianza unitaria). Se evidencia una reducción en las varianzas del sesenta y ocho por ciento (68%) para el nivel dos -IES- y del veintitrés por ciento (23%) para el nivel uno -estudiante-.

Estas varianzas resultan significativas tanto en la aproximación normal como en una prueba de razón de verosimilitud.

Los dos modelos finales se presentan en las columnas (III) y (IV) e incluyen un control para la selección por nivel socio económico de la familia del estudiante. El modelo en general tiene la especificación:

$$Y_{ij} = \beta_0 + \beta_Y Y_{ij,PRE} + \beta_S X_{ij,S} + \varepsilon_{ij} \quad \text{con } \beta_0 = \beta_{00} + \mu_j \quad (6.3)$$

Al incluir como control el índice socio económico construido por el ICFES se obtiene un coeficiente significativo, con un efecto marginal muy pequeño, que se relaciona con un cambio muy pequeño en el efecto marginal y la varianza asociadas al desempeño previo. La parte aleatoria del modelo permanece estable con una disminución del siete por ciento (7%) en la varianza para los interceptos de las IES y sin cambios a nivel de estudiante.

Al considerar por separado las variables de caracterización socio económica, disponibles en la encuesta diligenciada por los estudiantes, se encuentra que la educación de la madre y del padre, la ocupación de los padres y la clasificación en el SISBEN no resultan significativas en un modelo junto al desempeño previo en Saber 11 prueba de matemáticas. En todo caso se incluyeron en el modelo como variables categóricas. Tampoco resultaron significativas las interacciones entre las categorías de esas variables.

La única variable que resultó significativa en el modelo fue el nivel de ingresos de la familia. Esta variable toma siete valores que identifican rangos de ingresos en salarios mínimos (menos de 1, 1-2, 2-3, 3-5, 5-7, 7-10, más de 10). La categoría de referencia corresponde a ingresos familiares de menos de un (1) salario mínimo mensual. Todas las categorías obtienen diferencias positivas y significativas en el desempeño promedio con respecto a la categoría base. La mayor diferencia se presenta con el grupo de mayores ingresos que obtiene en promedio 0.294 puntos por encima del grupo de menores ingresos. La parte aleatoria se comporta estable (siguen siendo significativas las varianzas) y el resto de los coeficientes de la parte fija no presenta cambios importantes.

El principal aporte de este trabajo está relacionado con la mejor identificación posible, dados los datos disponibles, del valor agregado que se atribuye a la IES al no controlar por las varianzas asociadas al colegio del que egresa el estudiante, lo que puede llevar sobre o sub estimar el aporte de la IES. En este caso el modelo es de clasificación cruzada porque el estudiante pertenece simultáneamente a dos grupos relacionados con su desempeño académico. Hay consenso en que la influencia del colegio es menor en el caso de las competencias específicas, pero en el caso de las competencias genéricas evaluadas en Saber Pro la relación es directa. Lo que pasa es que una parte importante de esa influencia ha sido capturada en el desempeño en Saber 11, ahora nos interesa identificar si existe un efecto asociado a ese colegio en la variabilidad de los desempeños observados al egresar de la educación superior.

La tabla 6.2 presenta los resultados para los modelos estimados en la tabla 6.1, pero ahora se incluye en el nivel 2 la varianza cruzada de las IEM. La parte fija de los modelos es la misma. El hallazgo más importante es que la varianza cruzada en el segundo nivel es significativa en todos los casos examinados: en el modelo nulo (V), en el modelo con control por desempeño previo en Saber 11 (VI), y en los modelos con control por estatus socio económico (VII y VIII). Lo que evidencia el hallazgo es que el efecto es pequeño: el colegio explica sólo cuatro punto cincuenta y nueve por ciento de la variabilidad en los desempeños no condicionados (ICC IEM). Las pruebas de razón de verosimilitud entre pares de modelos comparables en todos los casos favorecen la elección de los modelos con la clasificación cruzada. Esto implica que el modelo que mejor estima el valor agregado relativo (condicionado) entre las instituciones es el VIII, en él se controla por desempeño previo, determinantes socio económicos y la varianza del nivel se descompone en clasificación cruzada. El ordenamiento derivado de la estimación con este modelo (VIII) será el más justo (además robusto) para establecer una comparación de los efectos atribuibles como valor agregado tipo A en competencias genéricas -razonamiento cuantitativo- para los programas de ingeniería en las universidades colombianas.

Es importante destacar que la varianza de los residuales de individuos está siendo sobre estimada, siendo mayor el sesgo en los modelos con covariantes. En estos últimos se

observa un cambio mayor en valor absoluto en la varianza del nivel estudiante, pero en términos porcentuales la disminución en los dos casos es ligeramente superior al cuatro por ciento. Al incluir las variables de control se explica en promedio la varianza del nivel dos en un setenta por ciento en universidades (disminuye de 0.235 a 0.071 en el modelo VII), en sesenta por ciento en colegios (disminuye de 0.050 a 0.020 en el modelo VII), mientras sólo disminuye en veinte por ciento la varianza no explicada de estudiantes en el nivel uno. Los cambios porcentuales se calculan con respecto a los valores estimados en el modelo nulo (V).

En relación con la parte fija los cambios no son significativos en los coeficientes estimados ni en sus errores estándar. Sólo en el modelo VIII se cambia el orden de diferencias entre las categorías, ahora la mayor diferencia en los desempeños promedio se observa entre el grupo de estudiantes en familias con ingresos entre siete y diez salarios mínimos con respecto a la categoría de referencia, que sigue siendo el grupo de familias con ingresos inferiores a un salario mínimo. En ese mismo caso (control por ingreso familiar) disminuye la significancia para la diferencia en desempeños de estudiantes en familias con ingresos entre uno y dos salarios mínimos y la categoría de referencia.

Uno de los objetivos que se persigue al elaborar medidas de valor agregado consiste en entregar información objetiva a los grupos de interés acerca del desempeño de las instituciones evaluadas. En el caso de padres y estudiantes, esta información les permite tomar una mejor decisión a la hora de invertir en educación superior. Con este fin es posible construir un ordenamiento de las instituciones por su desempeño medido por el valor agregado estimado en los modelos. La tabla 6.3 presenta la lista de las mejores veinte instituciones de educación superior en el grupo de ingenierías, para las pruebas de razonamiento cuantitativo clasificadas a partir de los modelos estimados con (IV) y sin (VIII) clasificación cruzada controlando por Saber 11 e ingreso familiar. Se incluye el valor estimado del valor agregado en promedio a sus evaluados en el modelo nulo (I) y su diferencia con el mejor modelo (con controles y clasificación cruzada). Las instituciones están ordenadas por su clasificación en el modelo seleccionado (VIII), los valores mostrados corresponden al valor agregado relativo estimado en cada uno de los modelos.

Las figuras 6.1 a 6.3 presentan el ranking de la tabla 6.2, pero en un gráfico que además muestra un intervalo de confianza para el valor agregado estimado, en el eje horizontal se muestra la posición que ocupa la universidad en función del valor agregado y en el eje vertical se muestra el valor agregado condicionado (relativo) a las variables de control usadas en el modelo seleccionado. Se destaca la diferencia en los valores estimados entre el modelo con y sin clasificación cruzada: los sesgos por omitir la estructura de la varianza no son uniformes. En la parte alta de la tabla puede destacarse el caso de las universidades Escuela Colombiana de Ingeniería y Javeriana que mejoran su resultado estimado al considerar la varianza con controles y clasificación cruzada (VIII) con respecto al modelo sin clasificación cruzada (IV). Caso contrario sucede con Unillanos que empeora su ubicación en el rango final, tanto la Universidad Autónoma de Occidente como la Universidad el Quindío salen del grupo de las veinte mejores al considerar la varianza de los colegios de origen.

Con respecto al modelo nulo –I- (relacionado con los modelos de prestigio) destaca la pérdida de posiciones importante que afecta a las universidades ICESI y EAFIT, una vez se controla por el desempeño previo y los factores socio económicos (IV y VIII) salen del grupo de las diez mejores universidades, perdiendo nueve posiciones en el ranking estimado.

Si se piensa en usar los resultados de este trabajo para entregar información a los interesados será recomendable construir una interfaz (software) que le permita ingresar los datos de control (resultado Saber 11, variables socio económicas y colegio de egreso), y así poder presentarle la evolución de los resultados históricos para individuos semejantes en las universidades del país. No obstante el ranking general le da una idea razonable de las universidades que aportan valor a sus estudiantes en el proceso formativo, recordando que la estimación se hace para competencias genéricas que están muy relacionadas con el proceso académico y se consideran útiles para desempeñarse en las sociedades actuales.

Debe destacarse que el resultado encontrado tiene implicaciones importantes en la política educativa. Por una parte, se debe tener en cuenta la trayectoria de los individuos en el

sistema académico para definir sus mejores alternativas al pasar al nivel superior de educación, el colegio incide en el desempeño esperado en el futuro, no es suficiente la selección por el puntaje en las pruebas de selección. La realidad social es más compleja, y la estructura de varianza cruzada captura mejor esta intrincada red de relaciones que afectan el desempeño de los estudiantes. Aquí se supone que individuos mejor informados pueden tomar mejores decisiones, y el modelo precisamente puede entregar mejor información de la que aporta un ranking de prestigio. Por otra parte, sería razonable que los programas de intervención, tipo “Ser pilo paga”, se planearan con base en la maximización de los resultados esperados para los individuos beneficiarios. Si este fuese el caso los estudiantes deberían ser dirigidos al grupo de instituciones que mejor trayectoria esperada ofrecen para los estudiantes dados los valores de sus variables de control al momento de la selección.

7. DISCUSIÓN

Este trabajo explora la medición de valor agregado en la educación superior para Colombia aplicado a las competencias genéricas en razonamiento cuantitativo. Se usaron datos generados por el ICFES como resultado de cruzar las bases de datos de las evaluaciones presentadas por los estudiantes al salir de la educación media, Saber 11, y al salir de la educación superior, Saber Pro. El ejercicio se aplica a las cohortes evaluadas en Saber Pro para el año 2012, año en el que se inicia la implementación definitiva del nuevo examen de evaluación de la educación superior.

La información disponible aún no es la mejor que se puede construir para modelo de valor agregado, en el que se elabora un ejercicio de inferencia causal. Las medidas del desempeño previo y del estatus socio económico del entorno familiar del evaluado deben ser razonablemente insesgadas. Para resolver esto, El ICFES está cambiando la metodología de evaluar el desempeño previo, alineando los dos exámenes en competencias genéricas lo que permitirá una mejor comparación vertical de los resultados. Las primeras evaluaciones comparables para los programas universitarios estarán disponibles a partir del año 2018. Esto no debe ser un impedimento para avanzar en la especificación de los modelos a usar en la cuantificación del valor agregado de cada IES a sus egresados.

El resultado principal del trabajo corresponde a la validación de la hipótesis de clasificación cruzada de los estudiantes. El desempeño observado en competencias genéricas puede estar influenciado por el aporte del colegio (IEM) de egreso de la educación media, además del aporte que la universidad (IES) realiza en la medida en que construye las competencias específicas que responden a su objetivo principal. Se encuentra evidencia estadística significativa de clasificación cruzada IES- IEM para las varianzas de los desempeños promedio de los establecimientos en los dos niveles de educación estudiados. Esto implica que la cuantificación de los valores agregados en la educación superior debe descontar el aporte asociado al desempeño del colegio de origen. No tenerlo en cuenta lleva a estimar con sesgo (en dirección no definible) el valor agregado por la universidad y por lo tanto a

entregar información equivocada a los organismos de control y al público interesado en la decisión racional de invertir en educación superior.

Debe destacarse que los valores agregados estimados son bastante sensibles a la especificación tanto de la parte fija como de la parte aleatoria del modelo, lo que se convierte en la mayor debilidad referenciada por la literatura y que se confirma en el trabajo realizado. El método de estimación elegido obedece a la naturaleza jerárquica de la información. Pero, el método tiene un problema potencial en la estimación del valor agregado. Los grupos pequeños son castigados al calcular el residual del nivel dos (IES-IEM) llevándolos a cero y aumentando su error estándar. Esto es bastante frecuente en la educación superior colombiana por lo que el efecto puede no ser despreciable en grupos de referencia que tengan menor demanda por los estudiantes, o en las universidades con altas tasas de deserción en los programas analizados.

Es pertinente advertir que en general la clasificación de valor agregado es sensible a los cambios en las covariantes y en la especificación de la estructura de varianza. Esto es precisamente una debilidad que pueden aprovechar por los encargados de divulgar la información. Si tenemos que elegir un modelo en este caso, los mejores corresponden a aquellos que controlan la selección por desempeño previo y estatus socioeconómico, incluyendo además clasificación cruzada para controlar por varianza en la escuela de origen. Pero, una IES puede hacer público el resultado del modelo en el que resulta mejor posicionada. Y no es razonable exigirle al público general, parte interesada en la información, que se haga experto en modelos de valor agregado.

Es necesario replicar el ejercicio para las otras competencias genéricas y para los otros grupos de referencia con el fin de tener un panorama completo de la situación de la educación superior en nuestro país.

REFERENCIAS

- Balcázar, Carlos, y Hugo Nopo. «Broken Gears: The Value Added of Higher Education on Teachers' Academic Achievement.» *IZA Discussion Papers*, 2014.
- Ballou, Dale, William Sanders, y Paul Wright. «Controlling for Student Background in Value-Added Assessment of Teachers.» *Journal of Educational and Behavioral Statistics*, 2004: 37-65.
- Best, Katharina, y Jussi Keppo. «A major choice: An examination of higher education and ability-adjusted income.» *Working Paper*, 2014.
- Breslin, Thomas, y Subarna Samanta. «Value Added in Undergraduate Business Education: An Empirical Analysis.» *Modern Economy*, 2014: 1138-1145.
- Checchi, Daniele. *The Economics of Education*. New York: Cambridge University Press, 2005.
- Chetty, Raj, John Friedman, y Jonah Rockoff. «Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates.» *NBER Working Paper Series*, 2013.
- Cuellar, Edwin. «ICFES.» 2014. www2.icfes.gov.co/seminario/seminarios-antiguos/programa-academico/conferencias-principales-2014/item/128.html (último acceso: 25 de Enero de 2015).
- Cunha, Jesse, y Trey Miller. «Measuring value-added in higher education: Possibilities and limitations in the use of administrative data.» *Economics of Education Review*, 2014: 64-77.
- Fagioli, Loris. «A comparison between value-added school estimates and currently used metrics of school accountability in California.» *Educational Assessment, Evaluation and Accountability*, 2014: 203–222.
- Figlio, David, y Susanna Loeb. «School Accountability.» *Handbook of the Economics of Education Volume 3* (North-Holland), 2011: 383-422.
- Goldstein, Harvey. *Multilevel statistical models*. West Sussex (UK): John Wiley & Sons Ltd, 2011.
- Hanushek, Eric. «Conceptual and Empirical Issues in the Estimation of Educational Production Functions.» *The Journal of Human Resources*, 1979: 351-388.
- Houng, Brendan, y Moshe Justman. «Comparing Least-Squares Value-Added Analysis and Student Growth Percentile Analysis for Evaluating Student Progress and Estimating School Effects.» *Melbourne Institute Working Paper Series*, 2013.
- Kelly, Anthony, y Christopher Downey. «Value-added measures for schools in England: looking inside the ‘black box’ of complex metrics.» *Educational Assessment, Evaluation and Accountability*, 2010: 181-198.

- Koedel, Cory, Kata Mihaly, y Johan Rockoff. «Value-Added Modeling: A Review.» *Economics of Education Review*, 2015: 180-195.
- Koretz, Daniel. «A Measured Approach.» *American Educator*, 2008: 18-39.
- Manzi, Jorge, Ernesto San Martin, y Sebastien Van Belleghem. «School system evaluation by value-added analysis under endogeneity.» *Toulouse School of Economics Working Paper*, 2010.
- OECD. «Literature Review on the Value-Added Measurement in Higher Education.» *Organisation for Economic Co-operation and Development*, 2013.
- Pedrosa, Renato, Elian Amaral, y Marcelo Knobel. «Assessing higher education learning outcomes in Brazil.» *Higher Education Management and Policy*, 2013: Vol. 24/2.
- Pritchett, Lant, y Deon Filmer. «What education production function really show: a positive theory of education expenditures.» *Economics of Education Review*, 1999: 223-239.
- Raudenbush, Stephen W, y J. Douglas Willms. «The Estimation of School Effects.» *Journal of Educational and Behavioral Statistics*, 1995: 307-335.
- Raudenbush, Stephen. «What are value-added models estimating and what does this imply for statistical practice.» *Journal of Educational and Behavioral Statistical*, 2004: 121-129.
- Reardon, Sean, y Stephen Raudenbush. «Assumptions of Value-Added Models for Estimating School Effects.» *Education Finance and Policy*, 2009: 492–519.
- Rothstein, Jesse. «Do Value-Added Models Add Value? Tracking, Fixed Effects, and Causal Inference.» *NBER Working Paper*, 2007.
- Saavedra, Anna, y Juan Esteban Saavedra. «Do colleges cultivate critical thinking, problem solving, writing and interpersonal skills?» *Economics of Education Review*, 2011: 1516-1526.
- Su-ran, Kong, Jiang Li-hiu, Li Guang-li, y Wang Hui-chuan. «Value-added Evaluation of College Students' English learning efficiency based on Hierarchical linear model.» *International Conference on Advanced Information and Communication Technology for Education*, 2014.
- Tekwe, Carmen, y otros. «An Empirical Comparison of Statistical Models for Value-Added Assessment of School Performance.» *Journal of Educational and Behavioral Statistics*, 2004: 11-36.
- Yunker, James. «The dubious utility of the value-added concept in higher education: the case of accounting.» *Economics of Education Review*, 2005: 355-367.

ANEXOS

Tabla 5.1 Información de Evaluados Saber Pro 2012.

	Aplicada 2012-1	Aplicada 2012-2
Total evaluados	102,258	150,675
Detalle por sexo del evaluado⁺⁺		
Mujeres	58,050	81,707
Hombres	44,189	68,952
Detalle por tipo de institución –IES-⁺⁺		
Escuela Tecnológica	1,418	1,710
Institución Tecnológica	6,232	9,916
Institución Universitaria	28,008	43,937
Técnica Profesional	7,948	9,207
Universidad	57,602	81,320
Detalle por nivel de formación⁺⁺		
Normalista	1,050	4,585
Técnica profesional	8,771	20,985
Tecnológica	20,729	25,186
Universitaria	71,708	99,919

Fuente: cálculos propios con las bases públicas del ICFES.

⁺⁺ Las discrepancias con los totales se deben a datos missing.

Tabla 5.2 Información de registros cruzados Saber 11 - Saber Pro 2012.

	Aplicada 2012-1	Aplicada 2012-2
Total evaluados	102,258	150,675
Total universitarios evaluados	71,708	99,919
Total registros con cruce	20,528	31,422
	29%	31%
Detalle por sexo del evaluado⁺⁺		
Mujeres	12,941	19,319
Hombres	7,587	12,103
Detalle por tipo de institución –IES-⁺⁺		
Escuela Tecnológica	21	54
Institución Tecnológica	48	117
Institución Universitaria	3,540	5,501
Técnica Profesional	122	140
Universidad	16,797	25,610

Fuente: cálculos propios con las bases públicas del ICFES.

⁺⁺ Las discrepancias con los totales se deben a datos missing.

Tabla 5.3 Evaluados por grupo de referencia: total y porcentaje con cruce.

Grupo de Referencia	Evaluados 2012	% Cruce SaberPro Saber 11
ADMINISTRACIÓN Y AFINES	50.957	25.34%
ARQUITECTURA Y URBANISMO	2.266	44.88%
BELLAS ARTES Y DISEÑO	4.519	42.86%
CIENCIAS AGROPECUARIAS	3.372	17.94%
CIENCIAS MILITARES Y NAVALES	245	31.02%
CIENCIAS NATURALES Y EXACTAS	2.520	37.58%
CIENCIAS SOCIALES	10.352	37.83%
COMUNICACIÓN, PERIODISMO Y PUBLICIDAD	6.276	41.33%
DERECHO	14.717	33.82%
ECONOMÍA	2.791	45.40%
EDUCACIÓN	21.017	20.89%
HUMANIDADES	1.233	17.11%
INGENIERÍA	33.939	33.36%
MEDICINA	5,088	23.25%
SALUD	11,727	24.27%

Fuente: cálculos propios con las bases públicas del ICFES.

Tabla 5.4 Promedios competencias genéricas por grupo de referencia.

Grupo de Referencia	Razona	Lectura	Escritura	Inglés
ADMINISTRACIÓN Y AFINES	10.0(.91)	9.9(.91)	10.1(1.1)	10.0(1.1)
ARQUITECTURA Y URBANISMO	10.1(.93)	10.1(.96)	10.1(1.2)	10.6(1.4)
BELLAS ARTES Y DISEÑO	10.0(.94)	10.4(.95)	10.2(1.2)	11.2(1.5)
CIENCIAS AGROPECUARIAS	10.1(.96)	10.0(.95)	9.9(1.1)	10.0(1.1)
CIENCIAS MILITARES Y NAVALES	10.1(.91)	9.7(1.0)	10.0(1.0)	10.3(1.1)
CIENCIAS NATURALES Y EXACTAS	10.6(1.1)	10.4(1.0)	10.3(1.1)	10.9(1.4)
CIENCIAS SOCIALES	9.7(.90)	10.2(1.0)	10.4(1.2)	10.2(1.3)
COMUNICACIÓN, PERIODISMO Y PUBLICIDAD	9.8(.86)	10.3(.93)	10.4(1.1)	10.7(1.3)
DERECHO	9.8(.87)	10.2(.99)	10.4(1.1)	10.1(1.1)
ECONOMÍA	10.8(1.1)	10.5(.96)	10.6(1.1)	11.0(1.5)
EDUCACIÓN	9.5(.84)	9.7(1.0)	10.1(1.2)	9.9(1.1)
HUMANIDADES	9.9(1.0)	10.6(1.0)	10.8(1.2)	11.1(1.6)
INGENIERÍA	10.6(1.0)	10.2(.95)	10.1(1.1)	10.6(1.3)
MEDICINA	10.5(1.0)	10.7(.94)	10.4(1.1)	11.0(1.5)
SALUD	9.7(.85)	9.9(.94)	10.0(1.1)	9.9(.92)

Fuente: cálculos propios con las bases públicas por el ICFES.

Tabla 5.5 Razonamiento cuantitativo por sexo y sector.

	Promedio	Desviación típica
Detalle por sexo del evaluado⁺⁺		
Mujeres	10.40	0.99
Hombres	10.77*	1.11
Detalle por sector IES-⁺⁺		
Oficial (pública)	10.78**	1.12
No oficial (privada)	10.53	1.05

Fuente: cálculos propios con las bases públicas del ICFES.

Diferencias significativas al 1% *; al 5% **.

Tabla 5.6 Variables de control Grupo Ingenierías.

	Promedio	Correlación con Razonamiento
Puntajes Saber 11		
Matemáticas	1.12(1.21)	0.61
Lenguaje	0.84(0.94)	0.43
Física	0.82(1.15)	0.49
Sociales	0.93(0.97)	0.47
Filosofía	0.57(0.93)	0.30
Socioeconómicas		
INSE	58.53(17.6)	
Ingreso Familiar 1-2SM ⁺	24.62%	
Ingreso Familiar 2-3SM ⁺	23.35%	
Ingreso Familiar 3-5SM ⁺	22.32%	
Sisben 1/2/3	32.62%	
Sisben No clasificado	66.56%	
	Educación Madre	Educación Padre
Primaria	11.28%	14.84%
Secundaria	36.23%	31.44%
Técnica o tecnológica	17.44%	14.93%
Profesional	34.71%	38.76%
	Ocupación Madre	Ocupación Padre
Trabaja cuenta propia	10.24%	20.64%
Empleado profesional	11.78%	13.06%
Pequeño empresario	5.57%	11.72%
Pensionado	5.23%	9.99%
Ama de casa	36.73%	1.06%

Fuente: cálculos propios con las bases públicas por el ICFES.

Tabla 5.7 Evidencia de clasificación cruzada en los datos.

Código Colegio/Código IES	1101	1102	1204	1209
12005	-	-	2	-
12047	1	-	1	2
12054	-	-	-	2
12088	1	-	-	-
12252	1	-	-	3
12260	-	-	1	8
12294	-	1	-	2
12583	1	-	3	-

Fuente: cálculos propios con las bases públicas por el ICFES.

Tabla 5.8 Relación colegio estudiantes para una IES.

Código Estudiante	Código IES	Código Colegio
SBPRO201220021495	1101	3608
SBPRO201220018188	1101	3608
SBPRO201210059199	1101	4754
SBPRO201210002044	1101	4754
SBPRO201210050681	1101	4754
SBPRO201220042855	1101	4762
SBPRO201210059241	1101	4762
SBPRO201220029881	1101	5090
SBPRO201210002089	1101	5108
SBPRO201210065504	1101	5108
SBPRO201220023043	1101	5108

Fuente: cálculos propios con las bases públicas por el ICFES.

Tabla 6.1 *Estimación modelos de valor agregado. Nivel 1: estudiantes. Nivel 2: IES*

	Ingenierías Razonamiento Cuantitativo – Saber Pro 2012			
	I	II	III	IV
Parte Fija				
Intercepto	10.50 (0.043)*	10.19 (0.026)*	10.06 (0.04)*	10.05 (0.041)*
Saber11 matemática		0.438 (0.007)*	0.435 (0.007)*	0.43 (0.007)*
INSE			0.002 (0.0005)*	
Ingreso Familiar Menos de 1SM				Categoría de referencia
1-2 SM				0.075 (0.036)**
2-3 SM				0.12 (0.037)*
3-5 SM				0.185 (0.038)*
5-7 SM				0.257 (0.042)*
7-10 SM				0.27 (0.047)*
Más de 10 SM				0.294 (0.048)*
Parte aleatoria				
Nivel 2: IES				
Varianza(Intercepto)	0.256 (0.032)*	0.08 (0.011)*	0.074 (0.01)*	0.072 (0.01)*
Nivel 1: estudiantes				
Varianza(residual)	0.858 (0.011)*	0.664 (0.008)*	0.664 (0.008)*	0.66 (0.008)*
Correlación intraclase	22.98%			
Deviance	30802	27795	27776	27698

Fuente: cálculos propios con las bases públicas por el ICFES.

Valores p: * Significativo al 1 %; ** Significativo al 5 %; *** Significativo al 10 %

Errores estándar entre paréntesis.

Tabla 6.2 Estimación modelos de valor agregado con clasificación cruzada IES - IEM

	Ingenierías Razonamiento Cuantitativo – Saber Pro 2012			
	V	VI	VII	VIII
Parte Fija				
Intercepto	10.49 (0.042)*	10.18 (0.026)*	10.06 (0.04)*	10.05 (0.04)*
Saber11 matemática		0.441 (0.007)*	0.438 (0.007)*	0.433 (0.007)*
INSE			0.002 (0.0005)*	
Ingreso Familiar Menos de 1SM				Categoría de referencia
1-2 SM				0.072 (0.037)***
2-3 SM				0.123 (0.038)*
3-5 SM				0.179 (0.038)*
5-7 SM				0.275 (0.042)*
7-10 SM				0.287 (0.048)*
Más de 10 SM				0.267 (0.05)*
Parte aleatoria				
Nivel 2.				
Varianza IES (Intercepto)	0.235 (0.03)*	0.076 (0.011)*	0.071 (0.01)*	0.069 (0.01)*
Varianza IEM (Intercepto)	0.05 (0.006)*	0.021 (0.004)*	0.020 (0.004)*	0.021 (0.004)*
Nivel 1: estudiantes				
Varianza(residual)	0.804 (0.012)*	0.637 (0.009)*	0.637 (0.009)*	0.632 (0.009)*
Correlación intraclase IES	21.57%			
Correlación intraclase IEM	4.59%			
Deviance	29104	26266	26250	26170

Fuente: cálculos propios con las bases públicas por el ICFES.

Valores p: * Significativo al 1 %; ** Significativo al 5 %; *** Significativo al 10 %

Errores estándar entre paréntesis.

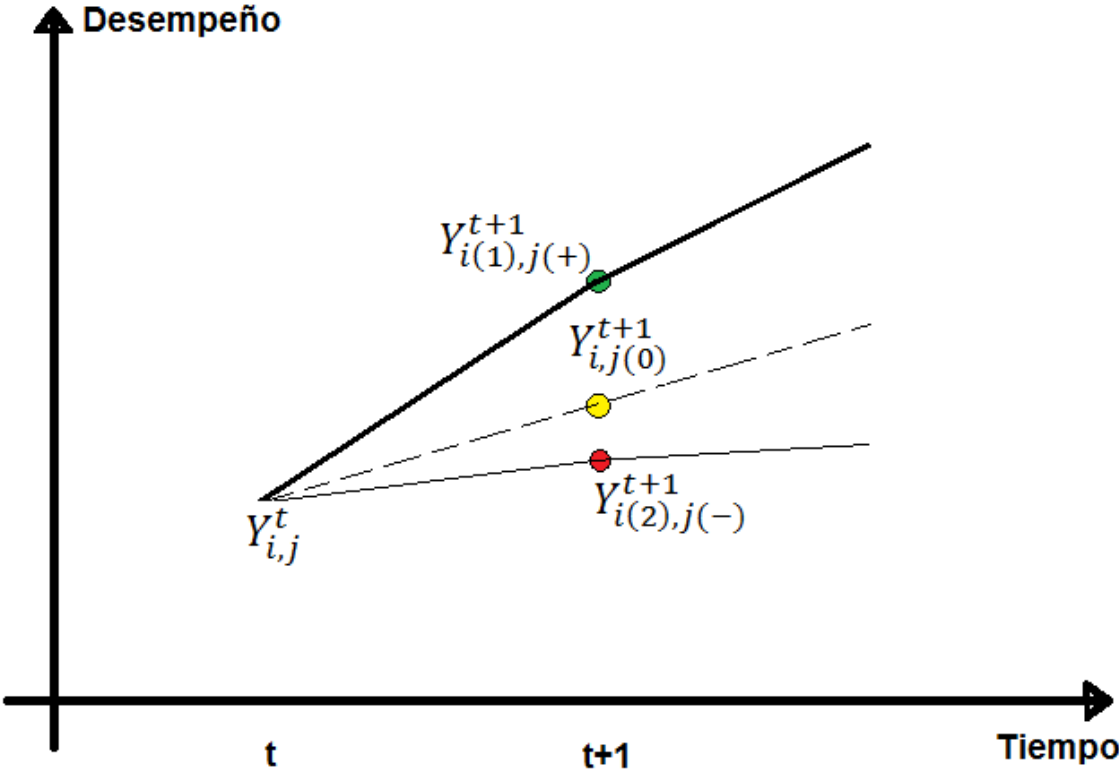
Tabla 6.3 Ranking de universidades por Valor Agregado en Razonamiento Cuantitativo

Grupo Ingenierías. Mejores veinte clasificados.

	Cross- classification (VIII)	HLM 2 niveles (IV)	Diferencia con cross	Modelo Nulo (I)	Diferencia con cross
UNAL Bogota	0.927	0.916	0.011	1.774	-0.847
UDEA	0.666	0.680	-0.014	1.113	-0.447
UNIANDES	0.636	0.594	0.042	1.471	-0.835
UNAL Medellín	0.555	0.559	-0.005	1.076	-0.521
Esc Col INGENIERIA	0.447	0.433	0.013	0.675	-0.228
UIS	0.443	0.453	-0.010	0.972	-0.529
UNIVALLE	0.435	0.451	-0.016	0.826	-0.392
JAVERIANA	0.421	0.379	0.042	0.957	-0.536
UNAL Manizales	0.419	0.415	0.005	0.811	-0.391
DISTRITAL	0.418	0.429	-0.012	0.810	-0.393
U. LLANOS	0.410	0.435	-0.024	0.726	-0.315
Fund U. NORTE	0.392	0.375	0.017	0.948	-0.557
SURCOLOMBIANA	0.382	0.402	-0.020	0.887	-0.506
U. ATLANTICO	0.355	0.361	-0.006	0.641	-0.286
PEDAGOGICA	0.346	0.341	0.005	0.794	-0.447
PUJ Cali	0.328	0.338	-0.010	0.680	-0.352
ICESI	0.325	0.328	-0.003	0.836	-0.512
UNICAUCA	0.285	0.299	-0.014	0.553	-0.268
EAFIT	0.264	0.231	0.033	0.875	-0.611
S. ARBOLEDA	0.262	0.241	0.021	0.479	-0.217

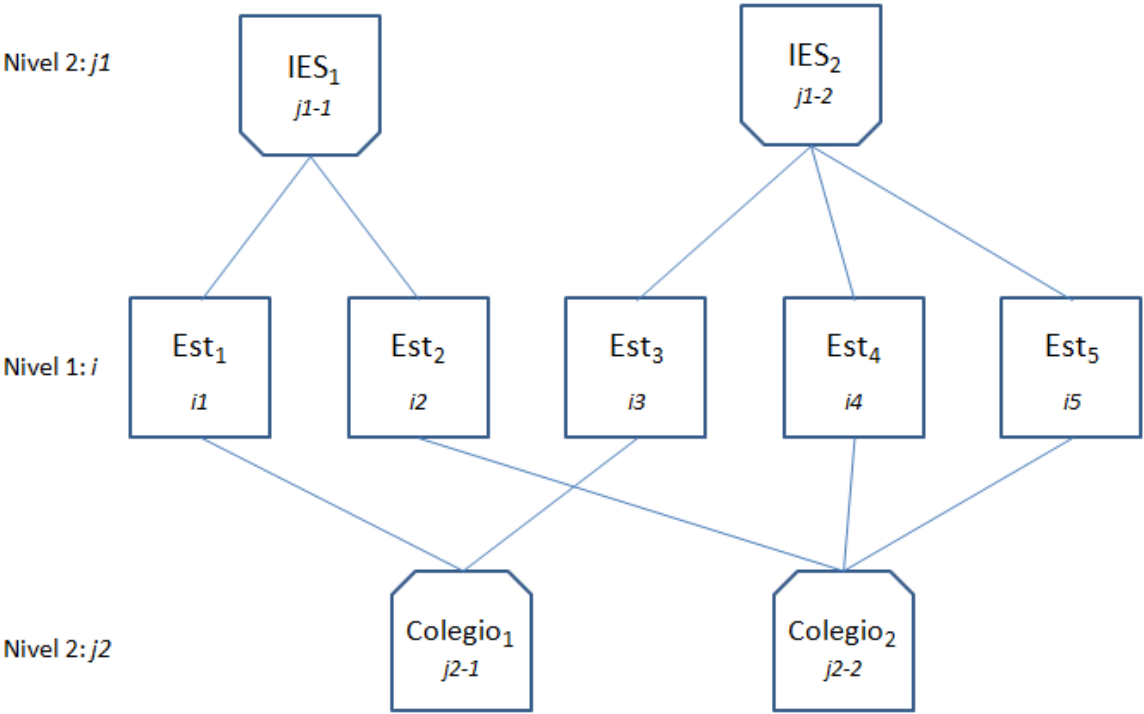
Fuente: cálculos propios con las bases públicas por el ICFES.

Figura 3.1 Metodología de estimación valor agregado – modelo de crecimiento del estudiante.



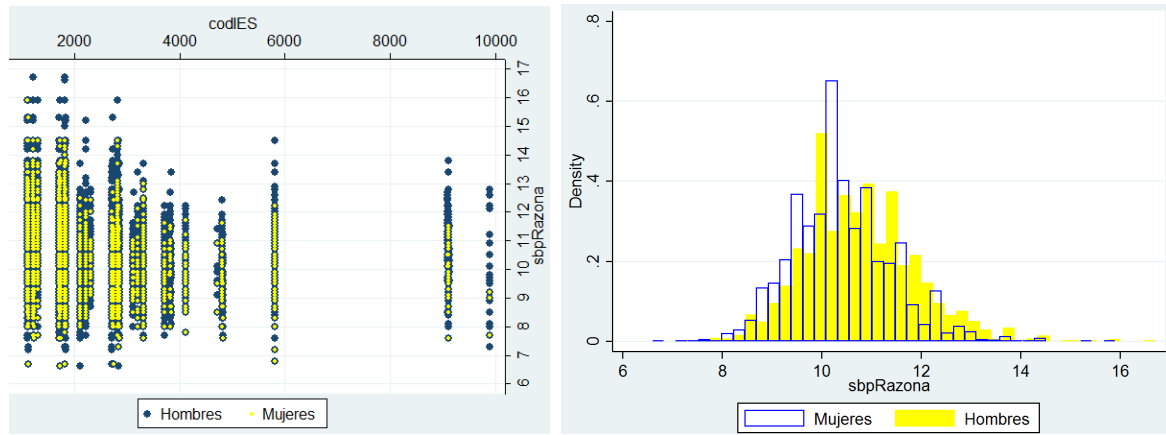
Fuente: adaptación del autor basada en Raudenbush (2004).

Figura 4.1 Modelo de clasificación cruzada en el nivel 2 (IES y Colegio) con estudiantes en el nivel 1.



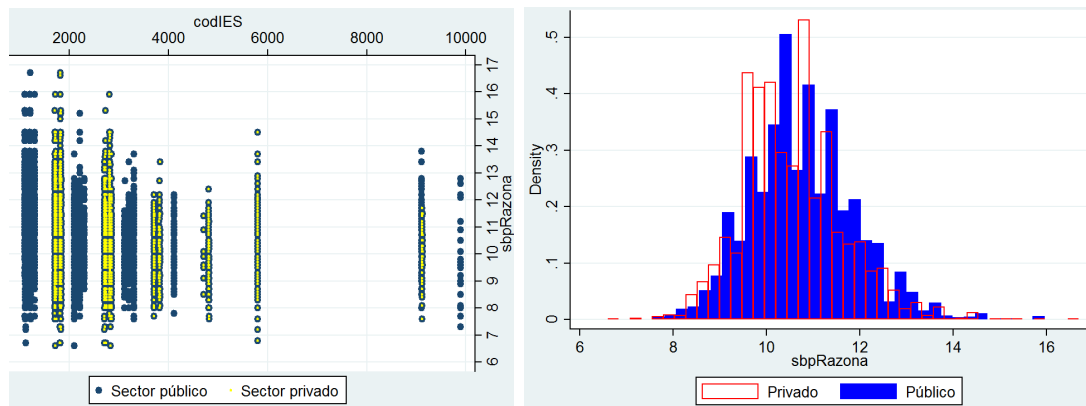
Fuente: adaptación del autor basada en Goldstein (2011).

Figura 5.1 Desempeños Razonamiento Cuantitativo Ingenierías Hombres-Mujeres.



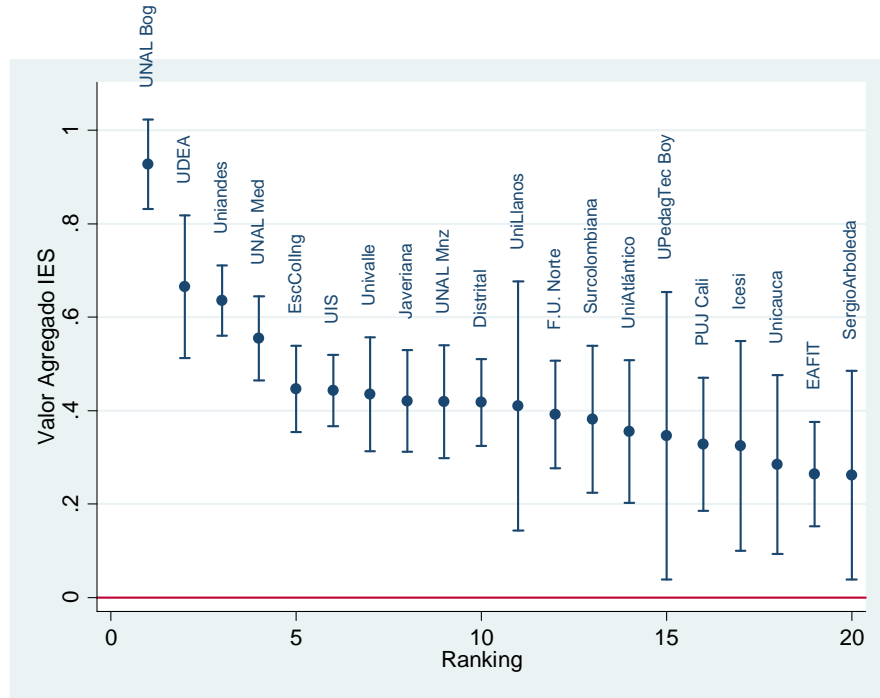
Fuente: cálculos propios con las bases públicas por el ICFES.

Figura 5.2 Desempeños Razonamiento Cuantitativo Ingenierías Público-Privado.



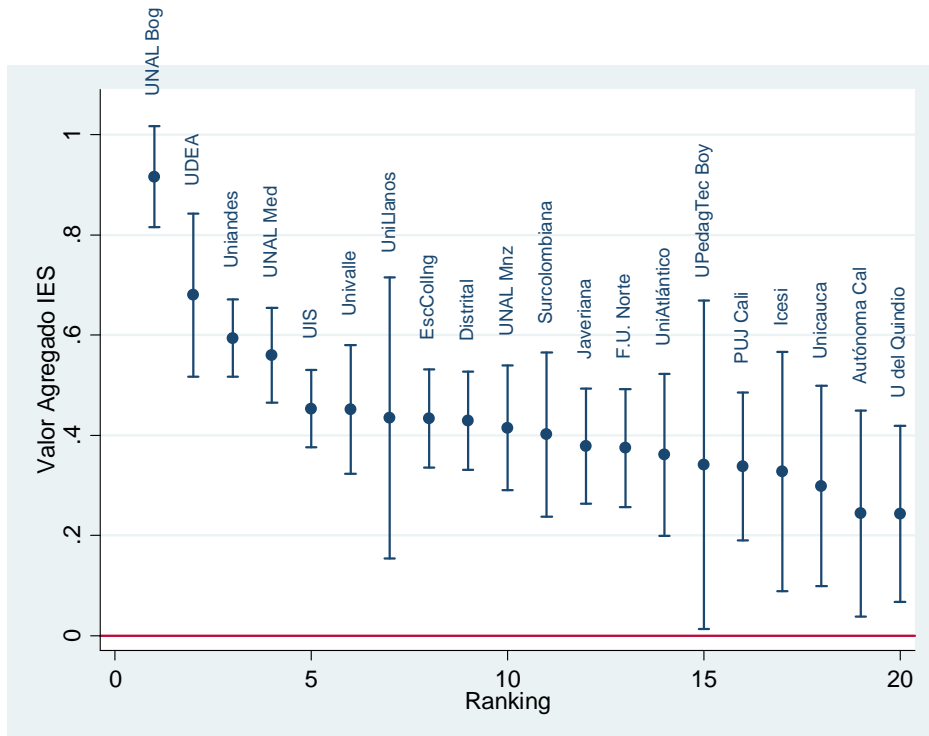
Fuente: cálculos propios con las bases públicas por el ICFES.

Figura 6.1 Mejores veinte, modelo clasificación cruzada, control Saber 11–Ingreso.



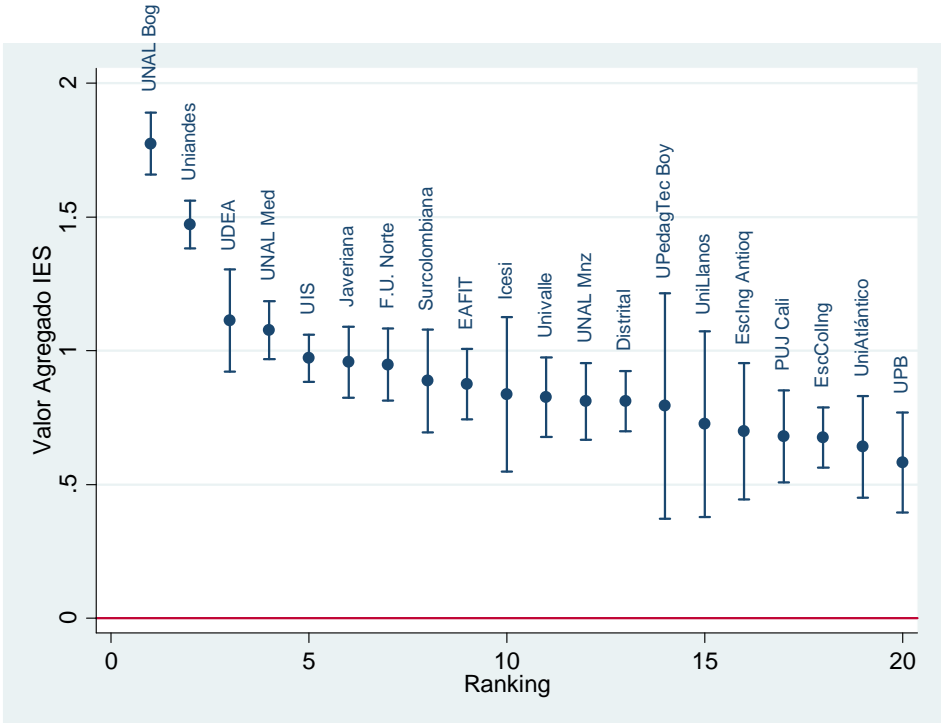
Fuente: cálculos propios con las bases públicas por el ICFES.

Figura 6.2 Mejores veinte, modelo dos niveles, control Saber 11–Ingreso.



Fuente: cálculos propios con las bases públicas por el ICFES.

Figura 6.3 Mejores veinte IES, modelo nulo.



Fuente: cálculos propios con las bases públicas por el ICFES.