

Midiendo los efectos de instituciones de Post-secundaria de Colombia en el aprendizaje de los estudiantes

Ben Domingue
con agradecimientos a Ed Wiley

Instituto de Ciencias del Comportamiento
Universidad de Colorado Boulder
ben.domingue@gmail.com

ICFES: Noviembre 2012

¿Por qué medir los efectos en el aprendizaje?

- El aprendizaje de los estudiantes es la función principal de la educación post–secundaria.
- ¿Podemos medir diferencias institucionales en el aprendizaje estudiantil?
- En caso afirmativo, esto sería útil para una variedad de actores (estudiantes, padres, gobierno, empresarios, etc.)

¿Cómo podemos medir si las instituciones influyen en el aprendizaje estudiantil?

¿Por qué medir los efectos en el aprendizaje?

- El aprendizaje de los estudiantes es la función principal de la educación post–secundaria.
- ¿Podemos medir diferencias institucionales en el aprendizaje estudiantil?
- En caso afirmativo, esto sería útil para una variedad de actores (estudiantes, padres, gobierno, empresarios, etc.)

¿Cómo podemos medir si las instituciones influyen en el aprendizaje estudiantil?

- Idealmente, queremos una una medida del desempeño de los estudiantes antes y después de recibir educación post–secundaria.
- Una medida *estandarizadas* de la habilidad al finalizar la educación post–secundaria es más o menos inexistente en la mayoría de los países.

Colombia es única

- Colombia tiene la oportunidad única de que los mismos estudiantes presentan SABER 11o. y SABER PRO.
- Estar en esta posición única, hace de esta investigación especialmente atractiva.

Colombia es única

- Colombia tiene la oportunidad única de que los mismos estudiantes presentan SABER 11o. y SABER PRO.
- Estar en esta posición única, hace de esta investigación especialmente atractiva.

En Colombia, ya se hizo un primer estudio sobre la efectividad de la educación post-secundaria. Sirve como buen punto de partida para discutir este proyecto.

Hallazgos de Saavedra & Saavedra

Importante: Interés sobre el efecto de la educación post secundaria, no el efecto de instituciones individuales de educación post secundaria.

- Saavedra & Saavedra usaron datos de la prueba de valoración de destrezas de graduados (ACER's MC Test) para alrededor de 2000 estudiantes que estaban terminando la universidad (emparejados con estudiantes que entran a la universidad) de una muestra de 17 instituciones (representativa de las 100 primeras universidades del país) para estimar “cuánto valor agregan las universidades a los estudiantes en pensamiento crítico, solución de problemas y destrezas de comunicación”.

Hallazgos de Saavedra & Saavedra

Importante: Interés sobre el efecto de la educación post secundaria, no el efecto de instituciones individuales de educación post secundaria.

- Saavedra & Saavedra usaron datos de la prueba de valoración de destrezas de graduados (ACER's MC Test) para alrededor de 2000 estudiantes que estaban terminando la universidad (emparejados con estudiantes que entran a la universidad) de una muestra de 17 instituciones (representativa de las 100 primeras universidades del país) para estimar “cuánto valor agregan las universidades a los estudiantes en pensamiento crítico, solución de problemas y destrezas de comunicación”.
- En la prueba en su conjunto, terminar la universidad está asociado con un aumento de aproximadamente 0.5 DE en la medida GSA.
- Esto es menor que el efecto de 1 DE encontrado para universidades de Estados Unidos (Klein et al., 2007).

Hallazgos Saavedra & Saavedra

Preguntas adicionales:

- ¿Son las instituciones privadas y públicas efectivas en forma diferente?
- ¿Las medidas de calidad (selectividad, % de profesores con PhD) están asociadas a la efectividad?

Hallazgos Saavedra & Saavedra

Preguntas adicionales:

- ¿Son las instituciones privadas y públicas efectivas en forma diferente?
- ¿Las medidas de calidad (selectividad, % de profesores con PhD) están asociadas a la efectividad?

Encuentran:

- Las instituciones privadas fueron más efectivas que las instituciones públicas (0.31 DE superior).
- Las medidas típicas de calidad (como selectividad % de profesores con PhD) no son especialmente predictivas de la efectividad. (nota: esto coincide con estudios sobre la efectividad de los maestros americanos donde las cualidades de los maestros por las que típicamente reciben dinero adicional, experiencia, educación adicional (por ejemplo grado de maestría en currículo) no se traduce en efectividad adicional.

Hallazgos Saavedra & Saavedra

También encuentran que las universidades son efectivas de manera diferencial, motivación empírica para este trabajo.

Limitaciones de Saavedra & Saavedra

- Saavedra & Saavedra usan datos de corte transversal (estudiantes que entraron a la Universidad fueron contrastados con estudiantes que dejaban la universidad). Esto puede generar un efecto sesgado por diversas razones.
- Principalmente, porque las cohortes de estudiantes pueden ser diferentes.

Limitaciones de Saavedra & Saavedra

- Saavedra & Saavedra usan datos de corte transversal (estudiantes que entraron a la Universidad fueron contrastados con estudiantes que dejaban la universidad). Esto puede generar un efecto sesgado por diversas razones.
- Principalmente, porque las cohortes de estudiantes pueden ser diferentes.
- Recuerden que en la ideal “torre de marfil”, medimos los cambios en logro para *cada estudiante que ha ingresado a la universidad*, vinculando el puntaje de cada estudiante al momento de entrada con el de la salida de la universidad.
- Los datos longitudinales disponibles no son perfectos, pero son potencialmente superiores que los utilizados en análisis de corte transversal.
- El nuevo estudio examinará la efectividad de las universidades usando datos longitudinales, con énfasis en estudiantes que tomaron SABER PRO en 2011.

Puntajes

Puesto que los puntajes en las pruebas están en el centro de este proyecto, mirémoslos de cerca.

- El puntaje de salida de un estudiante de universidad (*resultados*), será el puntaje del estudiante en SABER PRO 2011.
- El puntaje de entrada de un estudiante a la universidad (*inicial*) será el puntaje del estudiante cuando tomó SABER 11o.

Puntajes

Este estudio se basa en la vinculación de los puntajes de SABER PRO y SABER 11o. También utiliza datos de la encuesta administrada a los estudiantes que aplican SABER PRO. Aplicación SABER PRO—2011

- Lectura, Escritura, Razonamiento cuantitativo e Inglés
- Puntajes con desviación estándar de 1.

SABER 11—Cualquier estudiante que participó en SABER PRO 2011 y que presentó las pruebas de:

- Matemáticas, Lenguaje e Inglés.
- Puntajes estandarizados dentro de cada aplicación.

Puntajes

Este estudio se basa en la vinculación de los puntajes de SABER PRO y SABER 11o. También utiliza datos de la encuesta administrada a los estudiantes que aplican SABER PRO. Aplicación SABER PRO—2011

- Lectura, Escritura, Razonamiento cuantitativo e Inglés
- Puntajes con desviación estándar de 1.

SABER 11—Cualquier estudiante que participó en SABER PRO 2011 y que presentó las pruebas de:

- Matemáticas, Lenguaje e Inglés.
- Puntajes estandarizados dentro de cada aplicación.

Dada la multitud de pruebas, ¿cómo podemos decidir cuál resultado SABER 11o. debemos considerar como el pre-test del puntaje SABER PRO?

Correlación entre puntajes

Una manera de tomar esta decisión es usar correlaciones entre puntajes.

| | | SABER PRO | | | |
|----------|----------|-------------|-------------|-------------|-------------|
| | | Lectura | Escrit | Raz. Cuant. | Inglés |
| SABER 11 | Matem. | 0.38 | 0.20 | 0.48 | 0.43 |
| | Lenguaje | 0.56 | 0.32 | 0.45 | 0.49 |
| | Inglés | 0.46 | 0.29 | 0.44 | 0.69 |

Correlación entre puntajes

Una manera de tomar esta decisión es usar correlaciones entre puntajes.

| | | SABER PRO | | | |
|----------|----------|-------------|-------------|-------------|-------------|
| | | Lectura | Escrit | Raz. Cuant. | Inglés |
| SABER 11 | Matem. | 0.38 | 0.20 | 0.48 | 0.43 |
| | Lenguaje | 0.56 | 0.32 | 0.45 | 0.49 |
| | Inglés | 0.46 | 0.29 | 0.44 | 0.69 |

Todas las pruebas tienen correlaciones altas, excepto Lenguaje–Escritura.

- La prueba de Inglés no fue siempre aplicada en SABER 11o.
- Este estudio se concentra en la secuencia Lenguaje–Escritura, debido a que la correlación es más alta que la de Matemáticas–Razonamiento Cuantitativo.

¿Qué necesitamos?

- Recordemos nuestro objetivo: medir *correctamente* la efectividad universitaria usando datos longitudinales.
- Que necesitamos: una medida de la variación en la efectividad entre instituciones, que elimine las diferencias con respecto al desempeño inicial de quienes ingresan y a las características de los estudiantes.
- Los modelos estadísticos que intentan medir la efectividad de las instituciones, como lo hemos descrito aquí, se conocen como modelos de *Valor Agregado*.

Modelos de Valor Agregado (VA)

Un modelo de VA representa una aproximación estadística que procura aislar el efecto de las instituciones del efecto generado por otros factores que contribuye al aprendizaje (por ejemplo, riqueza de los padres, la motivación de los estudiantes, etc.). Aunque hay muchas maneras para especificar el modelo de VA todas tienen dos cosas en común:

Modelos de Valor Agregado (VA)

Un modelo de VA representa una aproximación estadística que procura aislar el efecto de las instituciones del efecto generado por otros factores que contribuye al aprendizaje (por ejemplo, riqueza de los padres, la motivación de los estudiantes, etc.). Aunque hay muchas maneras para especificar el modelo de VA todas tienen dos cosas en común:

- Todos los modelos de VA controlan por el nivel de logro de los estudiantes al ingreso. Esto se hace para asegurar que a la institución no se le considere responsable injustamente por diferencias preexistentes entre los estudiantes. El énfasis del modelo de VA es en el logro condicional o *crecimiento* del logro.

Modelos de Valor Agregado (VA)

Un modelo de VA representa una aproximación estadística que procura aislar el efecto de las instituciones del efecto generado por otros factores que contribuye al aprendizaje (por ejemplo, riqueza de los padres, la motivación de los estudiantes, etc.). Aunque hay muchas maneras para especificar el modelo de VA todas tienen dos cosas en común:

- Todos los modelos de VA controlan por el nivel de logro de los estudiantes al ingreso. Esto se hace para asegurar que a la institución no se le considere responsable injustamente por diferencias preexistentes entre los estudiantes. El énfasis del modelo de VA es en el logro condicional o *crecimiento* del logro.
- Ningún Modelo de VA puede revelar por qué algunas instituciones son más efectivas que otras. Comprender las diferencias entre prácticas educativas, que causan que los estudiantes aprendan más en determinados ambientes, requiere de investigación adicional.

El razonamiento detrás del trabajo de VA

- Nuestra meta es utilizar un modelo estadístico para (1) medir el “aprendizaje” de estudiantes individuales, luego (2) valorar si los estudiantes en ciertas universidades aprenden más que en otras.
- Podríamos simplemente tomar el puntaje de cada estudiante en SABER PRO y restarle el puntaje SABER 11o., y después promediar la diferencia para cada universidad.
- Sin embargo, esto sería insuficiente para explicar las diferencias entre tipos de estudiantes que se matriculan a las instituciones. Por ejemplo, algunas instituciones aceptan estudiantes con altos puntajes en SABER 11 o. y algunas instituciones matriculan estudiantes con bajos puntajes en SABER 11o.

El razonamiento detrás del trabajo de VA

- Adicionalmente, las instituciones tienen mezclas diferentes de características demográficas de los estudiantes. Veremos, por ejemplo, que la mediana de los estudiantes en algunas instituciones se ubicará en el primer estrato económico cuando en otras la mediana está en el quinto.
- ¿Cómo podemos hacer comparaciones justas y sensatas entre grupos tan diferentes?

El razonamiento detrás del trabajo de VA

- Adicionalmente, las instituciones tienen mezclas diferentes de características demográficas de los estudiantes. Veremos, por ejemplo, que la mediana de los estudiantes en algunas instituciones se ubicará en el primer estrato económico cuando en otras la mediana está en el quinto.
- ¿Cómo podemos hacer comparaciones justas y sensatas entre grupos tan diferentes?
- Aquí es donde entran los modelos de Valor Agregado. Construimos un modelo de Valor Agregado en forma tal que procure ajustar por las diferencias en los grupos de estudiantes que asisten a las distintas universidades: diferencias en puntajes SABER 11, y/o diferencias en características demográficas de los estudiantes, como el estatus económico o la educación de la madre.

Preguntas al construir el modelo de VA

- ¿Qué hacer si tenemos puntajes de SABER 9o. y SABER 11?o.
¿Debemos usarlos ambos?
- ¿Debemos ajustar por diferencias en la población estudiantil usando cada una de las medidas demográficas de los estudiantes que están disponibles? ¿Solamente un subconjunto?
¿Algunas variables son más importantes que otras?
- ¿Cómo manejar los datos faltantes? ¿Debemos excluir estudiantes que no tienen los datos completos?
- La mayoría estudiantes completó SABER PRO en los semestres 9 y 10 después de matricularse. Sin embargo, algunos estudiantes tomaron la prueba en el primero, segundo, o tercer semestre. ¿Debemos tener en cuenta estas diferencias en el modelo de Valor Agregado?

Muchas preguntas

- El principal desafío al hacer el análisis de valor agregado, es que cada decisión tomada a lo largo del estudio puede derivarse en cambios en la efectividad estimada de algunas universidades. En la mayoría de los casos, no hay una única combinación obvia de decisiones aceptadas por todos los estadísticos y diseñadores de política.
- En consecuencia, desarrollar un modelo de VA requiere de una gran cantidad de previsión, investigación y consideración de las alternativas.

Muchas preguntas

- El principal desafío al hacer el análisis de valor agregado, es que cada decisión tomada a lo largo del estudio puede derivarse en cambios en la efectividad estimada de algunas universidades. En la mayoría de los casos, no hay una única combinación obvia de decisiones aceptadas por todos los estadísticos y diseñadores de política.
- En consecuencia, desarrollar un modelo de VA requiere de una gran cantidad de previsión, investigación y consideración de las alternativas.

En este punto, podemos tener una discusión técnica para seleccionar los modelos de VA. Sin embargo, es más importante demostrar el tipo de resultados que se pueden obtener de un modelo de VA.

Por lo tanto, para el resto de la presentación se utilizará solo modelo como punto de partida de las discusiones acerca de lo que los modelos de VA pueden (y no pueden) hacer.

Demográficas

Puesto que controlar por diferencias demográficas será importante para hacer comparaciones razonables, es importante conocer las variables demográficas que vamos a utilizar.

trabaja Indicador de si el estudiante trabaja.

estrato Indicador socioeconómico

semestre_cursa semestre de universidad que cursa en estudiantes en el cual el presentó SABER PRO.

educa_madre nivel educativo de la madre.

year11 año en el cual el estudiante presentó SABER 11o.

Demográficas

Puesto que controlar por diferencias demográficas será importante para hacer comparaciones razonables, es importante conocer las variables demográficas que vamos a utilizar.

trabaja Indicador de si el estudiante trabaja.

estrato Indicador socioeconómico

semestre_cursa semestre de universidad que cursa en estudiantes en el cual el presentó SABER PRO.

educa_madre nivel educativo de la madre.

year11 año en el cual el estudiante presentó SABER 11o.

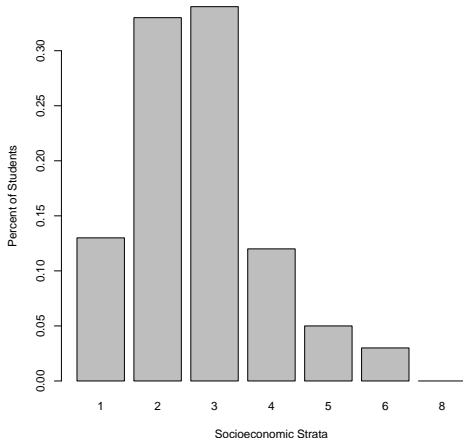
La inclusión de variables demográficas permite que los puntajes esperados de SABER PRO varíen de acuerdo a éstas características. Por ejemplo, podríamos esperar un menor incremento en el conocimiento de los estudiantes que trabajaban mientras asistían a la universidad.

Niveles de estrato

1–6 niveles.

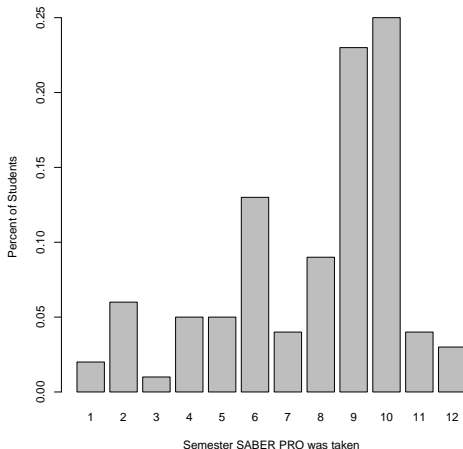
8 Estudiantes rurales sin clasificación.

Distribución del estrato



- La mayoría de los estudiantes se encuentran en los estratos más bajos 1 a 3.
- Grupo muy pequeño de estudiantes sin información (nivel 8).

Distribución de semestre_cursa

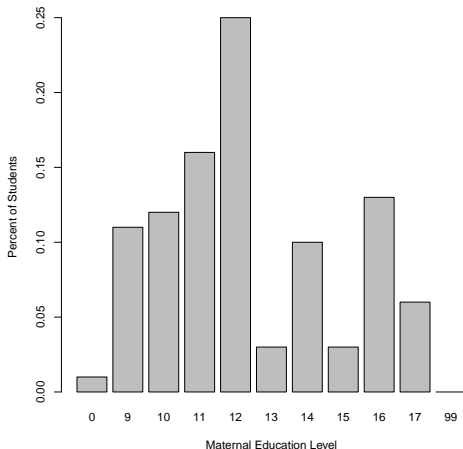


- En el análisis descrito, diferentes tipos de institución son analizados conjuntamente.

Niveles de educa_madre

- 0 Ninguno
- 9 Primaria incompleta
- 10 Primaria completa
- 11 Secundaria incompleta
- 12 Secundaria completa
- 13 Educación técnica sin grado
- 14 Educación técnica con grado
- 15 Educación profesional sin grado
- 16 Educación profesional con grado
- 17 Postgrado
- 99 Desconocido

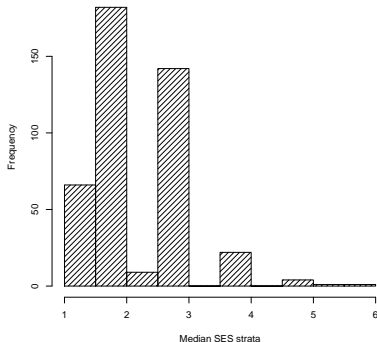
Distribución de educa_madre



- Mayores valores en 12, 14 y 16 corresponden a secundaria completa, institución comercial y universidad, respectivamente.

Demográficas a nivel institucional

Es importante recordar que no solo los estudiantes varían en términos de sus características demográficas, también las instituciones tienen diferentes tipos de estudiantes.

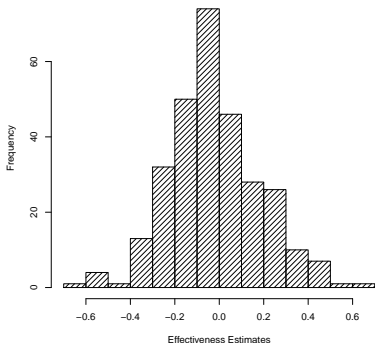


Efecto institución

Gran pregunta: ¿qué tanto varían las estimaciones de efectividad institucional?

Variabilidad en la Efectividad Estimada

Las instituciones varían en los efectos que tienen sobre el aprendizaje.

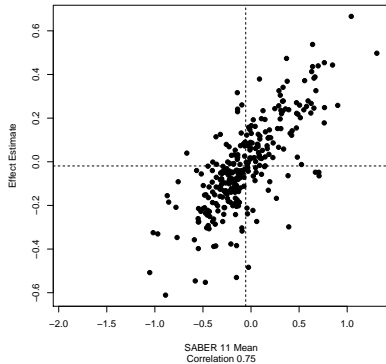


- La “Efectividad estimada” se miden con relación al *promedio* de las instituciones.
- En consecuencia, un estimador negativo no significa que una institución *eliminó* lo que sus estudiantes aprendieron al presentar SABER 11o; simplemente refleja que la institución fue menos efectiva que la institución promedio.

La efectividad en contexto

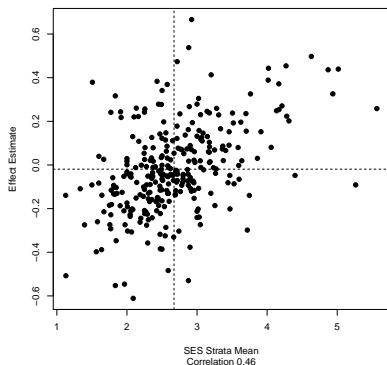
Ahora que sabemos que la efectividad institucional varía, miremos detenidamente las propiedades que tienen los efectos, especialmente cuando se relacionan con otras características de las instituciones.

Crecimiento versus SABER 11



- Como referencia, los promedios a nivel institucional entre SABER PRO y SABER 11o. tienen una correlación de 0.79.
- Propósito: “ponerlos en igualdad de condiciones” al romper la correlación entre el puntaje de los estudiantes que entran y sus puntajes de salida.
- Esta correlación es más fuerte de la observada comúnmente entre las escuelas primarias americanas.

Crecimiento versus Estrato económico



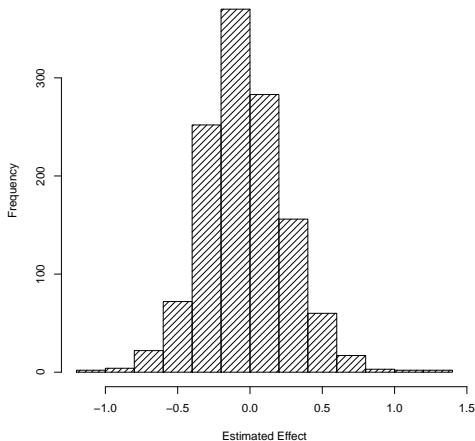
- Como se espera, vemos una correlación positiva entre el estrato socioeconómico del estudiante promedio la institución y el estimador de efectividad.
- A lo largo del espectro de condiciones socioeconómicas, observamos una variación sustancial en la efectividad entre instituciones que comparten niveles similares de composición socioeconómica.

VA con diferentes unidades

De la misma manera como los modelos de VA se usan para estimar la efectividad de las instituciones de educación superior, también pueden ser usados para estimar la efectividad de otros aspectos de la educación terciaria, tales como:

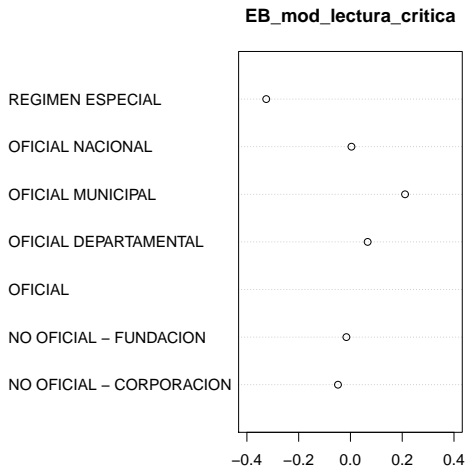
- Curso de estudio: 1441 programas de estudio (por ejemplo, diferentes disciplinas de ingeniería, historia, matemáticas)
- Clasificación institucional
- Acreditación

Curso de estudio



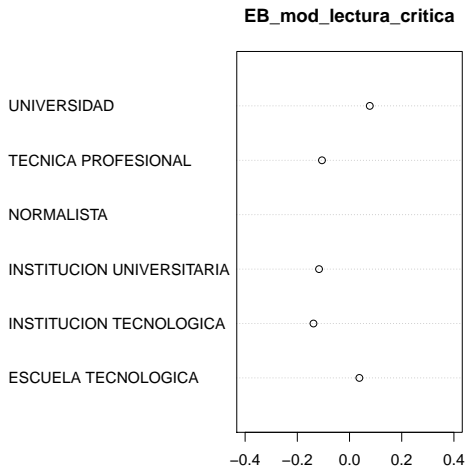
- Variación sustancial por programa de estudio (Desviación estándar de cerca de 0.3 comparado con 0.2 para las instituciones).

Público/Privado



- Sin estimador para institución “oficial” debido a que solo hay una institución de este tipo.

Carácter académico



- Como la diapositiva anterior, solamente hay una Institución “normalista”.

Acreditación

Las estimaciones muestran que las instituciones acreditadas fueron 0.3 unidades más efectivas que las no acreditadas.

- Relativo al estudiante general (Desviación estándar de 1 en esta prueba), es de un tamaño importante.

Temas para considerar

- La proliferación de resultados que se generarían si tratamos de analizar todos los aspectos de SABER PRO separadamente entre múltiples agregaciones (universidad, área de estudio, et.) serían muy difícil de interpretar. Posiblemente, es beneficioso considerar un compuesto.
- Teniendo en cuenta el rango de semestres durante los cuales los estudiantes completan sus estudios antes de presentar SABER PRO, los análisis pueden dividirse entre instituciones tradicional y vocacionales/comerciales.
- El modelo podría ser expandido para incluir variables adicionales, donde probablemente SABER 11o. sería la más importante.

Temas para considerar

- Tasas de deserción diferentes para diferentes universidades. Más específicamente, diferentes tipos de estudiantes desertan de diferentes universidades. Esto puede potencialmente sesgar las estimaciones.
- Cualquier estimador estadístico contiene incertidumbre. Los estimadores de efectividad pueden tener “intervalos de confianza” que sugieran la variabilidad contenida en esos estimadores. Esto es especialmente importante cuando se discuten los estimadores de manera individual.
- El error de medición en la prueba de SABER 11o. puede conducir a estimadores sesgados. Hay métodos que pueden corregir esto y que vale la pena explorar.

Ideas para cerrar

- Colombia tiene en una posición global única para hacer este tipo de investigación.
- Un uso potencial de los estimadores de efectividad sería ofrecer algo así como una medida de “efecto por dólar” para todas las instituciones del país. A pesar de las limitaciones, éste sería un importante indicador para los actores involucrados.

Gracias!

Contácteme si tiene comentarios o preguntas.

ben.domingue@gmail.com