

# Medición de los efectos de la educación superior en Colombia: Dónde estamos

Ben Domingue

Institute of Behavioral Sciences, University of Colorado Boulder  
ICFES  
ben.domingue@gmail.com

ICFES: Noviembre 2013

# Propósito de la charla

Proveer una visión general sobre:

- ▶ Qué decisiones hemos tomado en los últimos dos años acerca de la operacionalización del valor agregado (especificaciones de los datos, selección del resultado, etc.)
- ▶ Ilustrar los resultados

## Distintas cohortes y tamaño

Los modelos de valor agregado frecuentemente recaen en grandes bases de datos para obtener estimaciones que sean estables y precisas.

- ▶ Para maximizar el tamaño de muestra, usamos resultados de 3 rondas de SABER PRO (2011-2, 2012-1, 2012-2).
- ▶ Para asegurarnos que no tenemos estimadores basados en muestras muy pequeñas, solamente estimamos efectos para aquellos grupos del grupo de referencia dentro de una institución (IBR) unidades de al menos 35 estudiantes.

Después de imponer nuestras restricciones a los datos, tenemos 64,000 estudiantes en 18 grupos de referencia.

## Grupos de Referencia/IBR

- ▶ Tratamos de enfocar nuestro trabajo de tal manera que nos enfoquemos en las comparaciones más relevantes para los diseñadores de políticas y futuros estudiantes. En particular, procuramos comparar diferencias en efectividad para instituciones dentro de grupos de referencia.
- ▶ Para hacerlo, estimamos modelos separadamente dentro de cada grupo de referencia. Por lo tanto, los resultados no son comparables entre grupos de referencia.

# Resultados SABER PRO

De los cuatro resultados de SABER PRO decidimos concentrarnos en

- ▶ Razonamiento cuantitativo
- ▶ Lectura crítica

Dado que estas medidas son importantes para todos los estudiantes y son estimadas en forma confiable por SABER PRO.

# Resultados SABER PRO

Actualmente no estamos trabajando en:

- ▶ Inglés—fuertemente influenciado por factores externos a la escuela.
- ▶ Escritura— difícil de medir y relacionada débilmente con los puntajes de SABER 11.

## Puntajes previos de SABER 11

La selección de puntajes previos es exigente dado que:

- ▶ Los puntajes SABER 11 tienen una correlación relativamente débil con los puntajes SABER PRO. Correlaciones entre puntajes PRO y 11 son inferiores a 0.5. En contraste, correlaciones entre grados en Lectura y Matemáticas en EEUU son típicamente superiores a 0.75.
- ▶ SABER 11 fue un instrumento relativamente ruidoso. Su precisión era típicamente alrededor de 0.5 mientras que en EEUU las evaluaciones más parecidas tenían precisiones superiores a 0.85.

El error de medición en los puntajes previos y la relación entre previos y puntajes de resultado son dos factores importantes para hacer análisis de VA.

# Puntajes previos de SABER 11

Dados estos desafíos, decidimos trabajar con un conjunto de cuatro puntajes previos para los dos resultados SABER PRO:

- ▶ Lenguaje
- ▶ Matemáticas
- ▶ Química
- ▶ Ciencias sociales



# Modelos

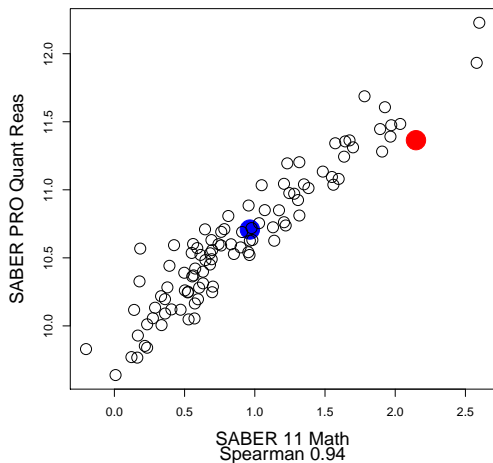
Nos centramos en tres modelos.

- ▶ Línea de base VAM: Cuatro puntajes previos SABER 11.
- ▶ INSE VAM: Línea de base VAM + media IBR INSE (recordar que INSE es un indicador socio-económico).
- ▶ SB11 VAM: Línea de base VAM + IBR media SB11 (matemáticas para razonamiento cuantitativo, lenguaje para lectura crítica).

# Resultados

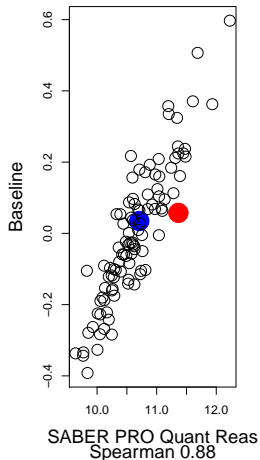
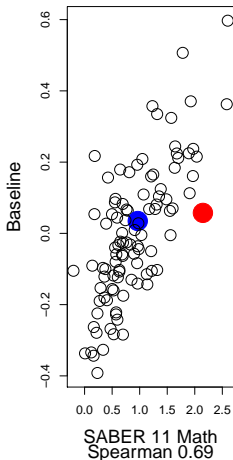
- ▶ Comparemos las estimaciones VA de nuestros tres modelos con medidas de status (media SABER 11 y puntajes PRO) y entre ellas.
- ▶ Nos concentraremos en ingeniería (15,000 estudiantes de 101 instituciones) en razonamiento cuantitativo.

# SABER 11 versus PRO



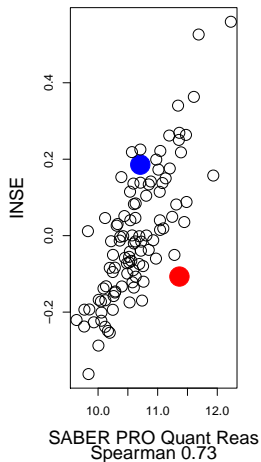
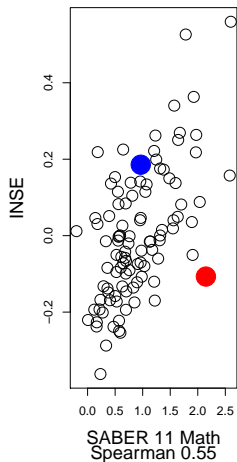
- ▶ Instituciones con alto puntaje medio en SABER 11 (en matemáticas) tienen puntajes medios PRO más altos (en RC).
- ▶ Nos concentraremos en lo que ocurre en las instituciones azules y rojas en las diapositivas siguientes.

## Resultados: status versus línea de base



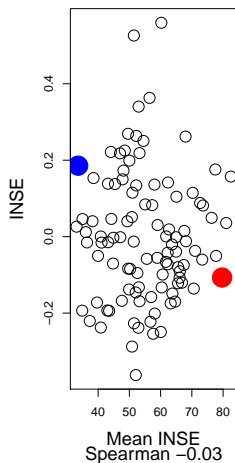
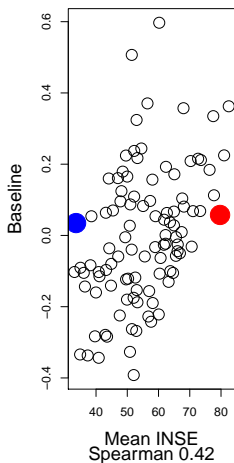
- ▶ Menor correlación entre estimación línea de base VA y media SABER 11, frente a PRO. Sin embargo, ambas correlaciones son altas.
- ▶ A pesar de diferentes puntos de inicio y terminación, las dos escuelas son muy similares en su medición de línea de base VA.

# Resultados: status versus VA-INSE



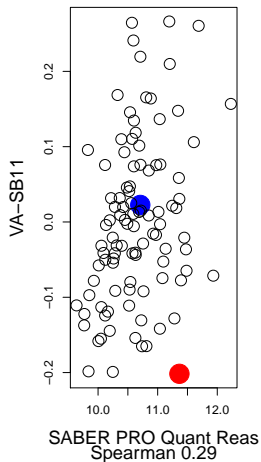
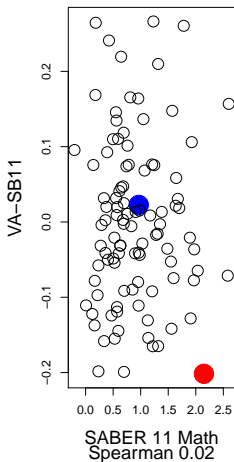
- ▶ Menores correlaciones entre estimaciones INSE y medidas status que estimaciones línea de base y medidas estatus.
- ▶ Dos instituciones tienen muy diferentes VA-INSE estimaciones. ¿Por qué?

# VA versus INSE



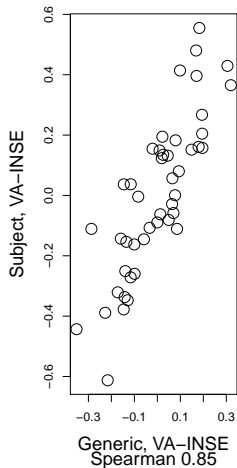
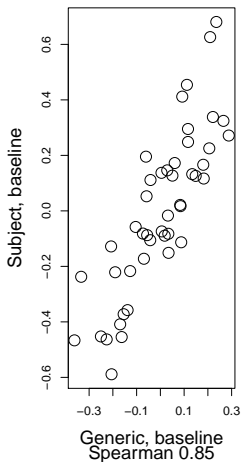
- ▶ Escuela azul es muy bajo en INSE.
- ▶ Nótese que las correlaciones son un poco menores que nuestras otras medidas de estatus.

## Resultados: status versus VA-SB11



- ▶ Estimaciones VA no están correlacionadas con puntajes SABER 11 (por diseño) y su relación con puntajes SABER PRO es muy baja.
- ▶ El rango de estimaciones de eficacia también ha disminuido (con este modelo, la efectividad de las escuelas varía menos).

## Resultado específico de un área (Derecho)



- ▶ Correlaciones entre modelos VA fueron 0.89 y 0.95 entre genéricos y específicos de un área (estimaciones son ligeramente más sensibles a la elección del resultado que a la elección del modelo).
- ▶ También tenemos evidencia que hay mayor variabilidad en efectividad de escuelas cuando usamos resultados específicos de un área.



## Consideraciones pendientes

- ▶ Manejar la atrición
- ▶ Medidas adicionales de áreas específicas
- ▶ Modelamiento multidimensional de los resultados SABER PRO en forma conjunta
- ▶ Manejar heteroscedasticidad

## Consideraciones pendientes

- ▶ Manejar la atrición
- ▶ Medidas adicionales de áreas específicas
- ▶ Modelamiento multidimensional de los resultados SABER PRO en forma conjunta
- ▶ Manejar heteroscedasticidad

¡Estamos atentos a discutir este proyecto conforme avanzamos!