

EVIDENCIA PRELIMINAR PARA CAMBIOS EN LA PSU-M

Nancy Lacourly

Universidad de Chile
Santiago, 5/12/2017

Puntos a abordar

- Antecedentes y recomendaciones entregadas por el informe de Pearson Education
- Estudios realizados
- Conclusiones

Antecedentes y Recomendaciones

- Informe Pearson (2013)

- Déficit en capacidad predictiva
- Déficit en equidad.

- Recomendaciones de Pearson

- Estudiar nuevo marco para desarrollo de todas las PSU
- Focalizar las pruebas en contenidos hasta 2° medio
- Adecuar especialmente la prueba de Matemática (PSU-M) a la población que la rinde (actual grado de dificultad inadecuado).

Tres preguntas

1. ¿Es posible reducir contenidos en la PSU-M a fin de corregir el inadecuado grado de dificultad reportado en el informe Pearson, sin comprometer significativamente la confiabilidad de la medición?
2. ¿Es posible acotar los contenidos curriculares evaluados en la PSU-M a aquellos de 1° y 2° medio sin perjudicar significativamente su actual capacidad predictiva?
3. ¿Contribuiría la PSU-M en contenidos curriculares de 1° y 2° medio, a reducir la brecha de rendimiento que se observa entre los postulantes de colegios particulares pagados y municipales, en especial los que asisten a la FMTP?

Datos y indicadores

- Base de datos del SUA del rendimiento universitario primer año (2013).
- Bases de datos PSU-M del DEMRE (admisión 2013).
- Muestra: Alumnos Promoción 2012
- Puntajes brutos por nivel escolar:
 - Items Básicos (1° y 2° medio)
 - Items Avanzados (Mayor que 2° medio)
 - Todos los ítems
- Puntajes brutos por eje temático:
 - Números
 - Álgebra
 - Geometría
 - Datos y azar

Estudios realizados

1. Evaluación del impacto de la reducción de contenidos en el grado de dificultad y confiabilidad de la prueba
2. Estudio de reducción de contenidos y validez predictiva
3. Estudio de reducción de contenidos y brecha de rendimiento

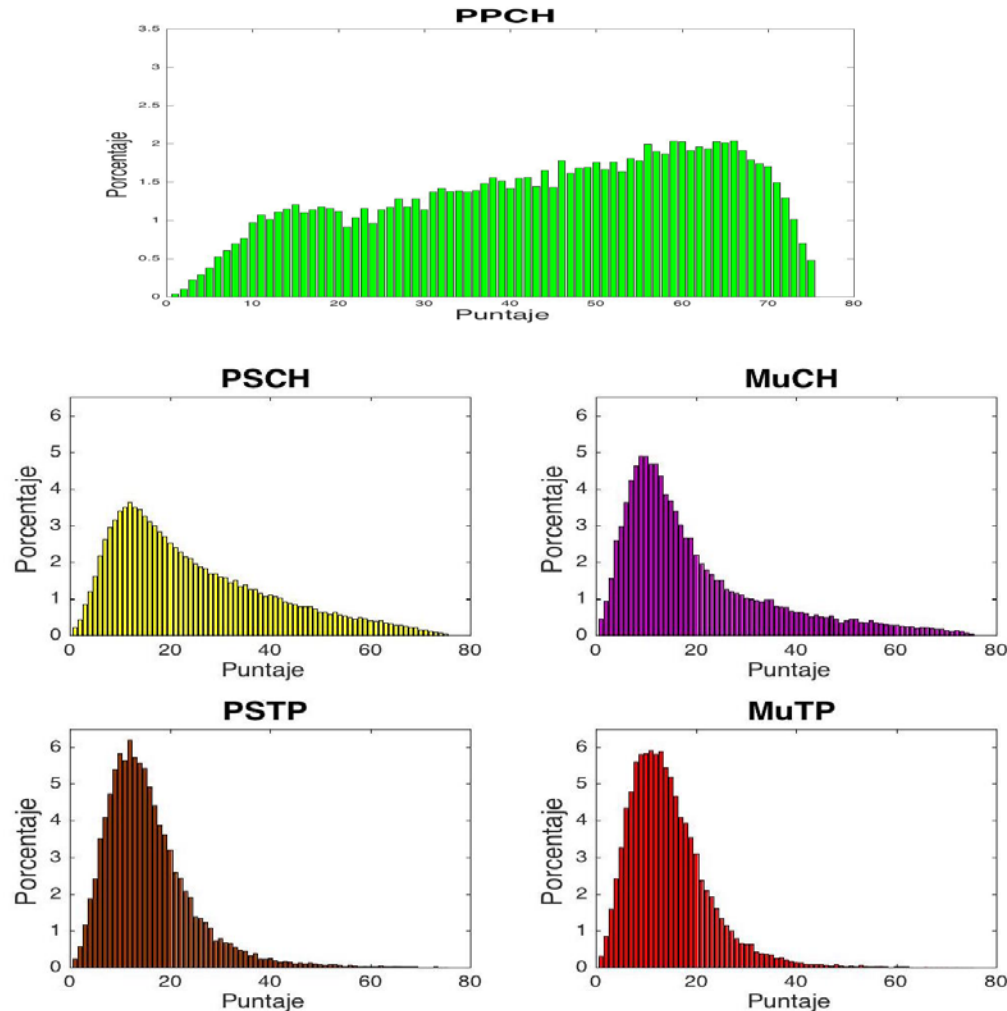
Evaluación del impacto de la reducción de contenidos en el grado de dificultad y confiabilidad

	Estadísticos	Ítems básicos	Ítems avanzados	Total
Números	Dificultad Promedio	0.55 (0.05)	0.30 (0.22)	0.41 (0.21)
	Alpha Cronbach	0.67	0.62	0.79
	Número ítems	5	6	11
Álgebra	Dificultad Promedio	0.32 (0.15)	0.26 (0.15)	0.28 (0.15)
	Alpha Cronbach	0.91	0.89	0.93
	Número ítems	30	18	48
Geometría	Dificultad Promedio	0.26 (0.08)	0.21 (0.12)	0.23 (0.11)
	Alpha Cronbach	0.76	0.86	0.9
	Número ítems	8	14	22
Datos y azar	Dificultad Promedio	0.42 (0.21)	0.31 (0.22)	0.36 (0.21)
	Alpha Cronbach	0.52	0.62	0.75
	Número ítems	5	6	11
Total	Dificultad Promedio	0.36 (0.17)	0.25 (0.16)	0.30 (0.18)
	Alpha Cronbach	0.91	0.94	0.96
	Número ítems	30	44	74

Indicadores de Teoría Clásica
(N=170.000)

Algebra y Geometría son más difíciles, tanto ítems básicos como avanzados

Distribución del grado de dificultad por tipo de colegio

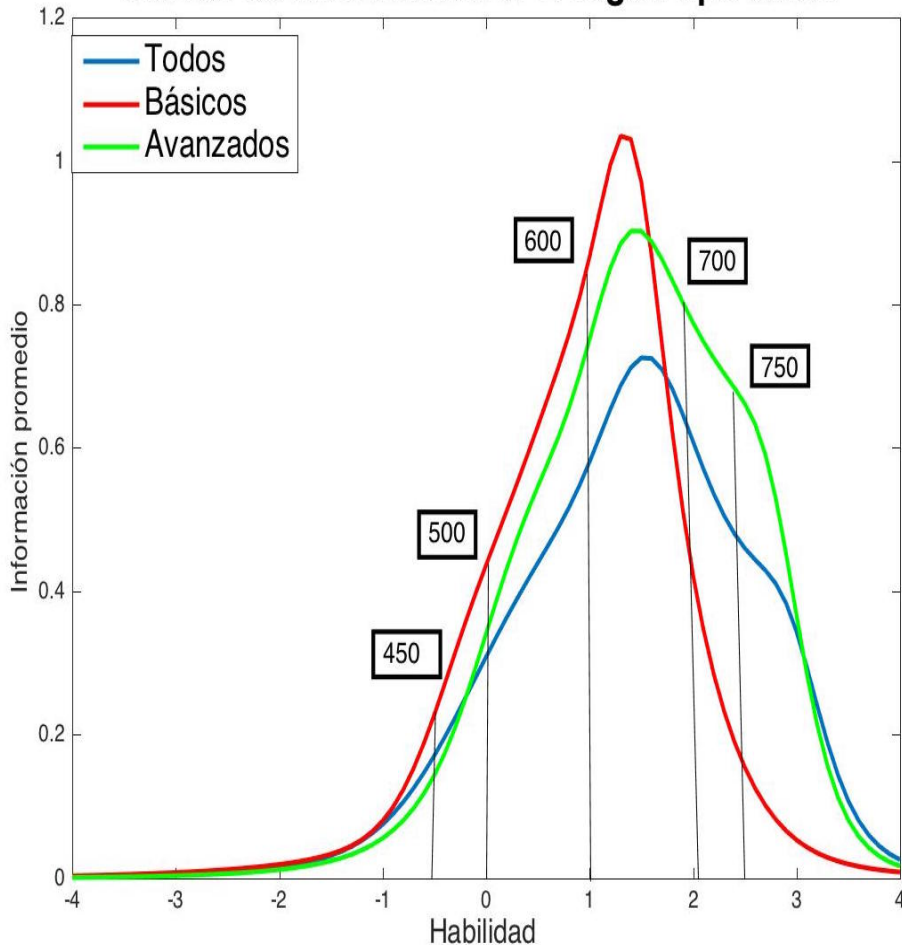


Desempeño heterogéneo entre tipos de colegio.

A excepción de los PPCH, para el resto de los alumnos, la PSU-M es muy difícil, en especial para los TP.

Resultados IRT 3 PL por tipo de ítems

Curvas de información 3PL según tipo ítems



- Las curvas difieren según el tipo de ítems.
- La prueba alcanza su mayor información entre las habilidades 1 y 2.5 (600 y 750).
- Poca información en los puntos de corte importante para la selección en muchas universidades (450 y 500).

Conclusiones

- Los ítems de contenidos avanzados como los de contenidos básicos resultan muy difíciles para la población que la rinde, en especial para el grupo proveniente de la educación técnica profesional.
- La prueba entrega poca información en algunos puntos clave a la izquierda de la distribución.
- Solución: dos pruebas de matemática.

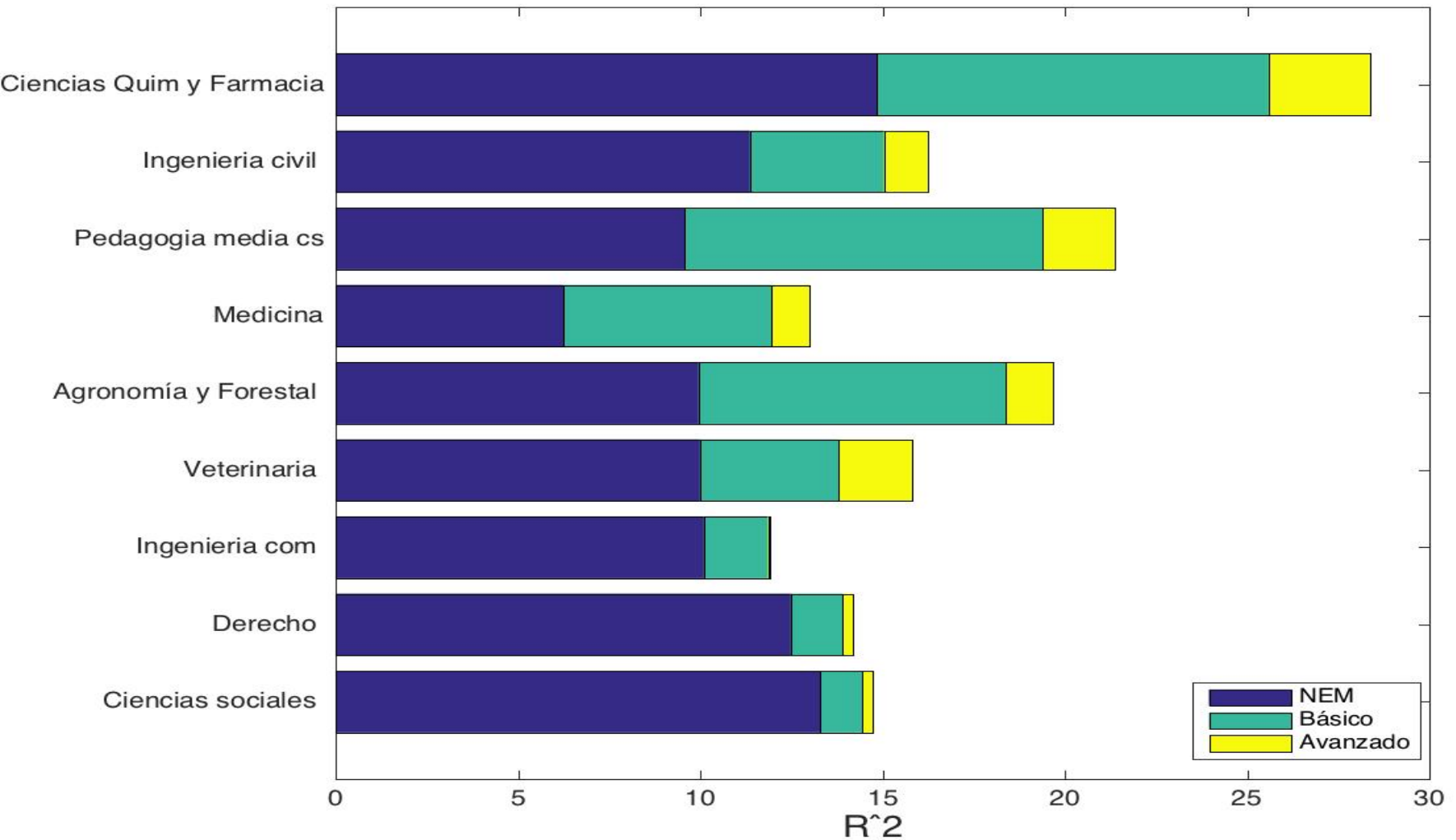
Estudio de reducción de contenidos y validez predictiva

¿Disminuye la capacidad predictiva al usar solo ítems básicos?

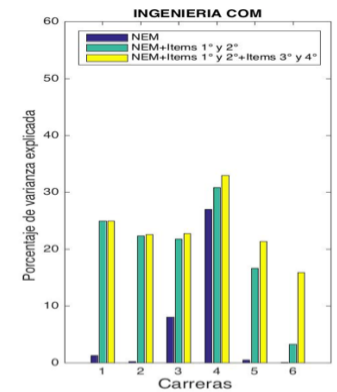
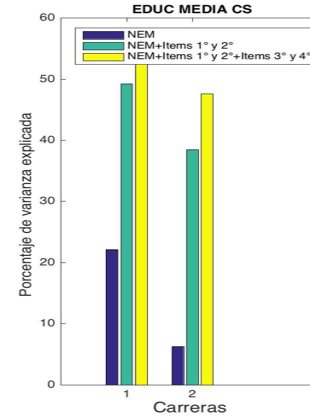
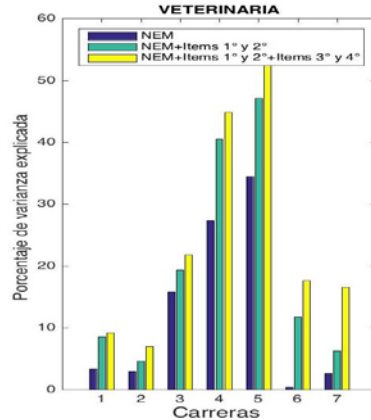
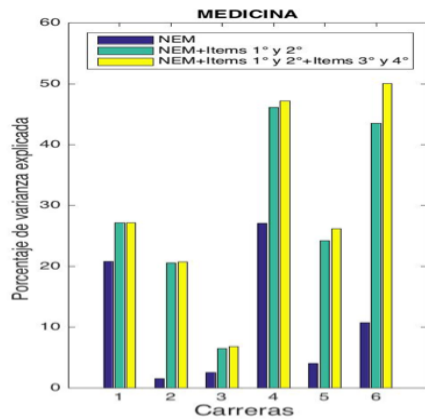
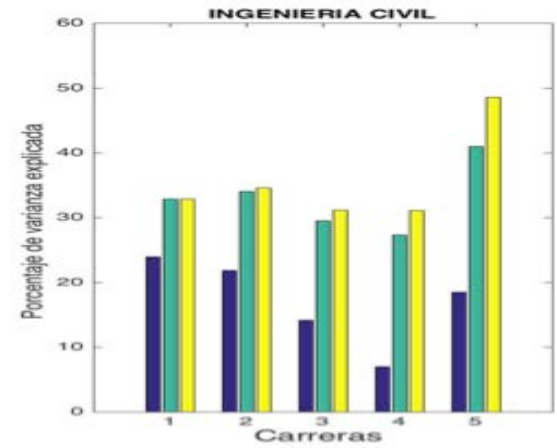
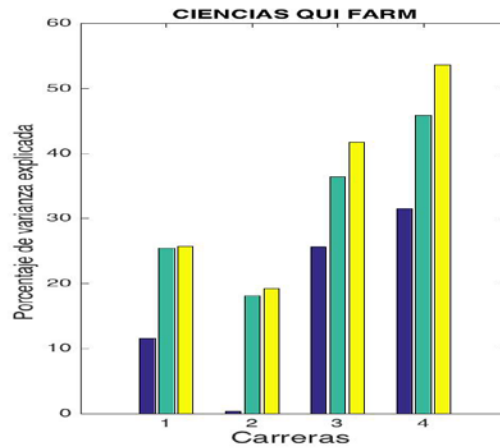
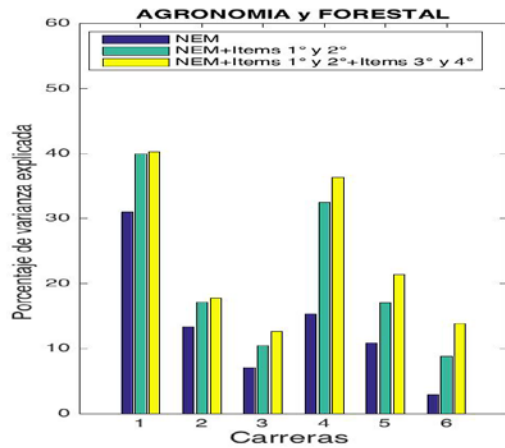
Se usaron correlaciones, regresiones y incremento del R^2 de diferentes modelos de regresión (N=47.000).

- Variable criterio: Promedio ponderado (ramos aprobados y reprobados) del primer año.
- Predictores: NEM, puntajes estimados por subconjuntos de ítems de nivel básico y nivel avanzado.
- Análisis por carrera con más de 30 alumnos (600), área (29) y universidad (33).

Porcentajes R^2 por área



Carreras dentro de su área



Conclusiones

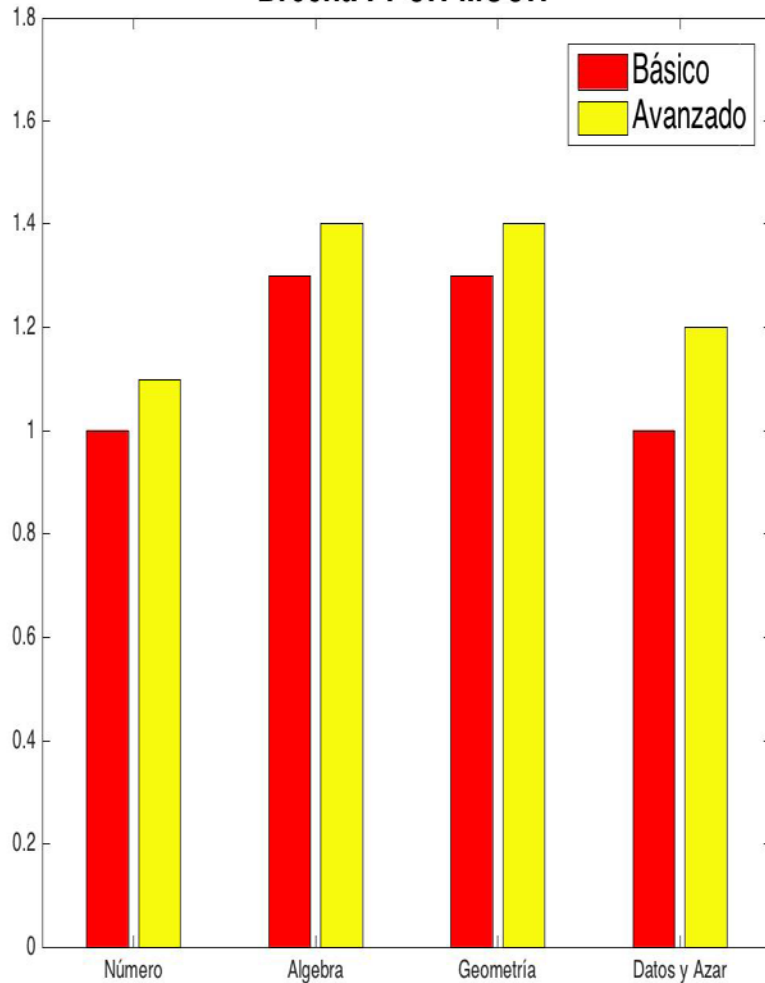
- Hay una gran variabilidad en la capacidad predictiva tanto de NEM como PSU-M entre áreas y entre carreras de una misma área.
- En algunas carreras de la misma área, los ítems avanzados hacen un aporte significativo a la predicción por sobre la NEM y los ítems básicos; en otras no.

Estudio de reducción de contenidos y brecha de rendimiento

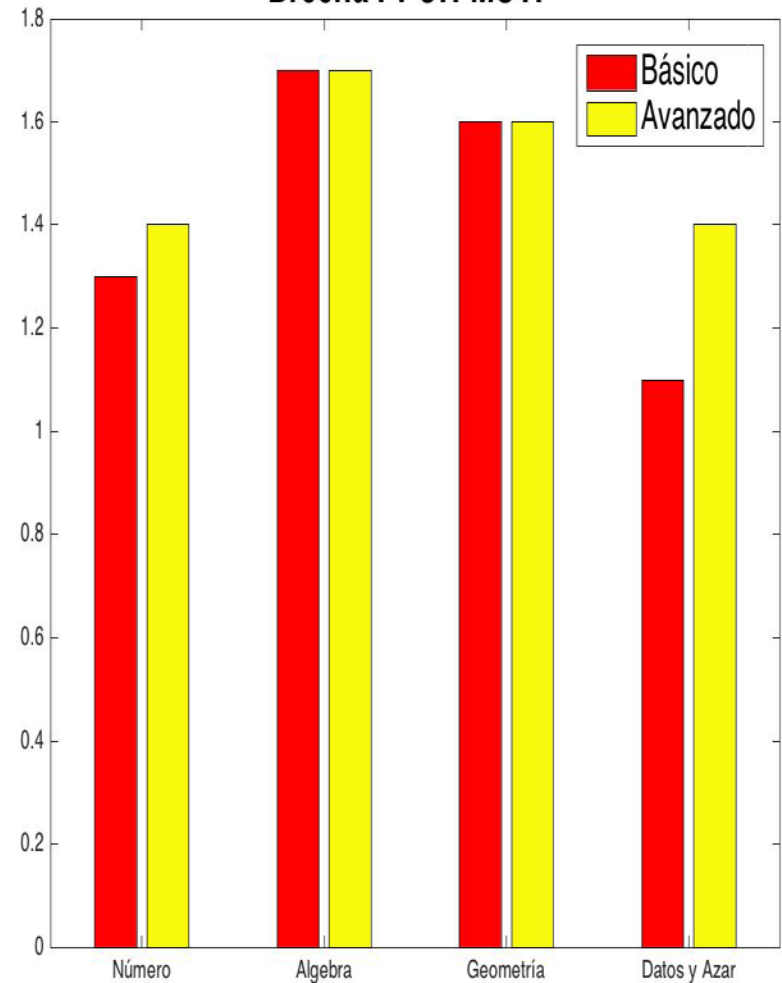
- Se busca responder si focalizar la PSU-M en preguntas de contenidos básicos podría contribuir o no a la reducción de la brecha de rendimiento que se observa entre los postulantes provenientes de establecimientos particulares pagados y municipales, en especial los que asisten a la EMTP.
- La brecha estandarizada es la diferencia de los rendimientos promedio de dos grupos dividida por la desviación estándar (de los dos grupos juntos).
- Tomamos como grupo referencia a los alumnos de colegios PBCH

Brechas estandarizadas

Brecha PPCH-MUCH



Brecha PPCH-MUTP



Conclusiones

- Las brechas mostradas en los gráficos son de gran magnitud y consistentemente favorecen a los alumnos de colegios particulares pagados. Se encuentra brechas similares para los alumnos PS.
- Las mayores brechas corresponden a los alumnos de la educación EMTP.
- En términos de la magnitud de las brechas, las de menor tamaño, para todos los grupos, corresponden a los ejes temáticos de Números y Datos y Azar de nivel básico.
- Las brechas correspondientes a los ejes temáticos de Álgebra y Geometría de nivel básico son prácticamente tan altas como las de nivel avanzado para los grupos provenientes de la EMTP.

Conclusiones generales

Grado de dificultad:

- Hay ítems cuyo grado de dificultad escapa de los rangos esperables en este tipo de prueba. No es de extrañar puesto que la prueba se enfoca en el currículo científico humanista, que no es cubierto en todos los establecimientos educacionales del país. En este sentido, hay evidencia contundente que el grado de dificultad de la PSU-M es mayor para los alumnos de la EMTP.
- Privilegiar ítems de 1er y 2do medio no resuelve completamente el problema del inadecuado grado de dificultad. Hay que estudiar incluir ítems de 7mo y 8vo básicos.

Confiabilidad:

- Acotar contenidos hasta 2do Medio en la PSU-M no

Conclusiones generales

Validez predictiva:

- Para algunas carreras los contenidos de los dos últimos años de enseñanza media hacen un aporte adicional a la predicción.
- Pero para muchas carreras el uso de contenidos avanzados no contribuye a mejorar significativamente la predicción.

Brechas:

- La magnitud de las brechas varía según el tipo de ítems.
- Debería privilegiarse ítems que contribuyen a la predicción y presenten las menores brechas de rendimiento.

Limitaciones

- Los puntajes de ítems básicos y avanzados no corresponden a pruebas separadas y los puntajes obtenidos por los postulantes pueden verse afectados por el hecho de que deben prepararse para rendir todos los contenidos. Por ende, es razonable suponer que la preparación del alumno puede, en algún grado, verse influida por la extensión de los contenidos examinados y ello podría afectar el rendimiento. Este es un aspecto que debe ser examinado a través de estudios empíricos.
- En el año 2012 existía penalización por respuesta errada. Hay que replicar estudios con promoción más reciente.
- Para introducir cambios en la PSU-M se deben realizar estudios complementarios a nivel de las carreras (cuantitativos y cualitativos) que permitan determinar qué conocimientos matemáticos requieren dominar los alumnos. Ello a fin de