

Clase 22: Regresiones y Cuasi-Experimentos

Haciendo Economía I

Econ 2205

Ignacio Sarmiento-Barbieri

Universidad de los Andes

November 2, 2023

Agenda

① Anuncios

② Modelo de Resultados Potenciales Recap

③ Regresión y Cuasi Experimentos

Anuncios

- ▶ Las referencias para esta clase siguen siendo:
 - 1 Mastering Metrics de Angrist y Pischke (cap 1 y 2)
 - 2 Mostly Harmless Econometrics de Angrist y Pischke (cap 2)
 - 3 Causal Inference: The Mixtape de Cunningham (cap 4) (disponible online en su pagina web)
- ▶ Quiz sobre Resultados Potenciales → Martes 7 de Noviembre

Plan para hoy

1 Anuncios

2 Modelo de Resultados Potenciales Recap

3 Regresión y Cuasi Experimentos

Agenda

① Anuncios

② Modelo de Resultados Potenciales Recap

③ Regresión y Cuasi Experimentos

Regresión y Resultados Potenciales

- ▶ Asumimos que el efecto del tratamiento es constante ("homogéneo")

$$Y_{1i} - Y_{0i} = \tau \quad \forall i \quad (1)$$

- ▶ Entonces podemos escribir

$$Y_i = Y_{0i} + D_i (Y_{1i} - Y_{0i}) \quad (2)$$

- ▶ como

$$Y_i = \underbrace{\alpha}_{=E[Y_{0i}]} + D_i \underbrace{\tau}_{\frac{Y_{1i}-Y_{0i}}{Y_{0i}-E[Y_{0i}]}} + \underbrace{\eta_i}_{}$$
 (3)

Regresión y Resultados Potenciales

$$Y_i = \alpha + D_i \tau + \eta_i \quad (4)$$

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = \tau + \underbrace{E[\eta_i | D_i = 1] - E[\eta_i | D_i = 0]}_{\text{Sesgo de Selección}} \quad (5)$$

- ▶ Nuestra estimación del **efecto de tratamiento** (τ) es tan bueno como nuestra habilidad de "shut down" el *sesgo de selección*
- ▶ El *sesgo de selección* luce mucho como OVB, hay algo en el término de error η_i que está afectando Y_i y también está correlacionado con D_i (D_i es endógeno)

Agenda

1 Anuncios

2 Modelo de Resultados Potenciales Recap

3 Regresión y Cuasi Experimentos

Regresión y Cuasi Experimentos

- ▶ Cuando no podemos utilizar la asignación aleatoria, buscamos rutas alternativas hacia el conocimiento causal.
- ▶ La regresión puede tener un poder revelador de causalidad cuando un experimento no es posible.
- ▶ La regresión, compara sujetos de tratamiento y de control que tienen las mismas características observadas.
- ▶ La inferencia causal basada en regresiones descansa en el supuesto de que cuando las variables clave observadas se igualan en todos los grupos de tratamiento y de control, el sesgo de selección de las cosas que no podemos ver también se elimina (en su mayor parte).

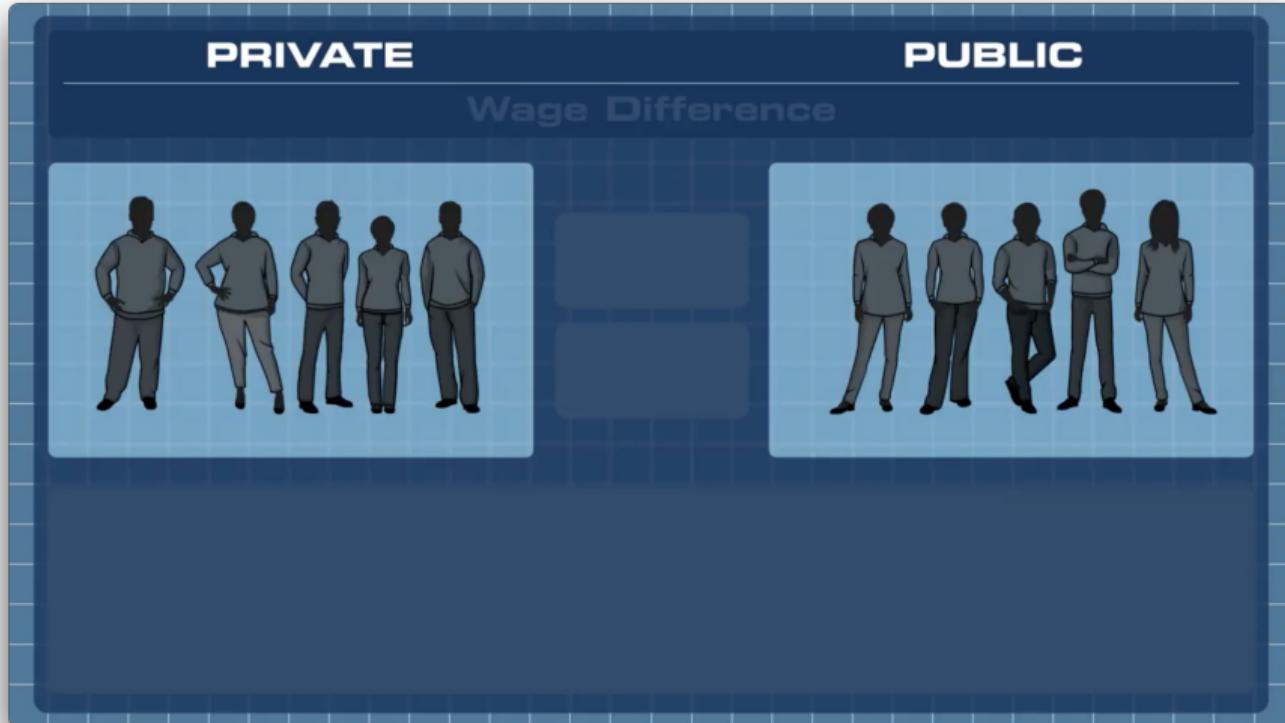
Retornos a Universidad Privada



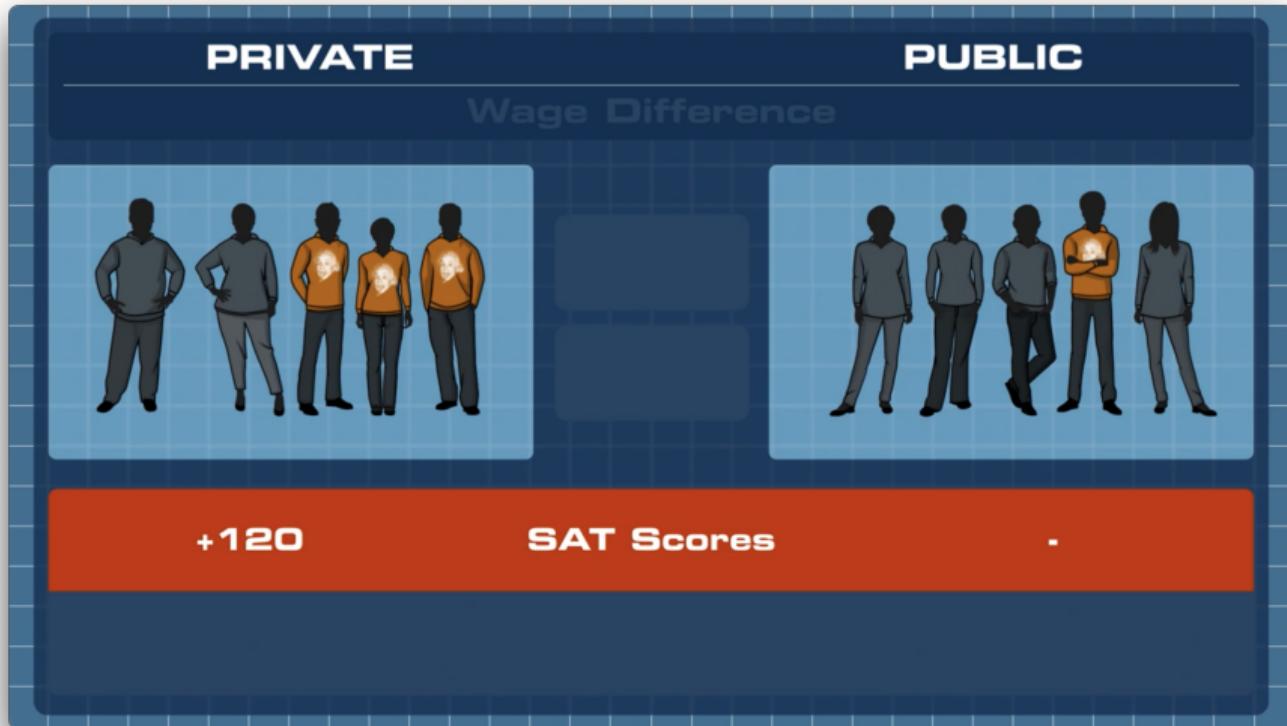
Retornos a Universidad Privada

Retornos promedios son de 14%!

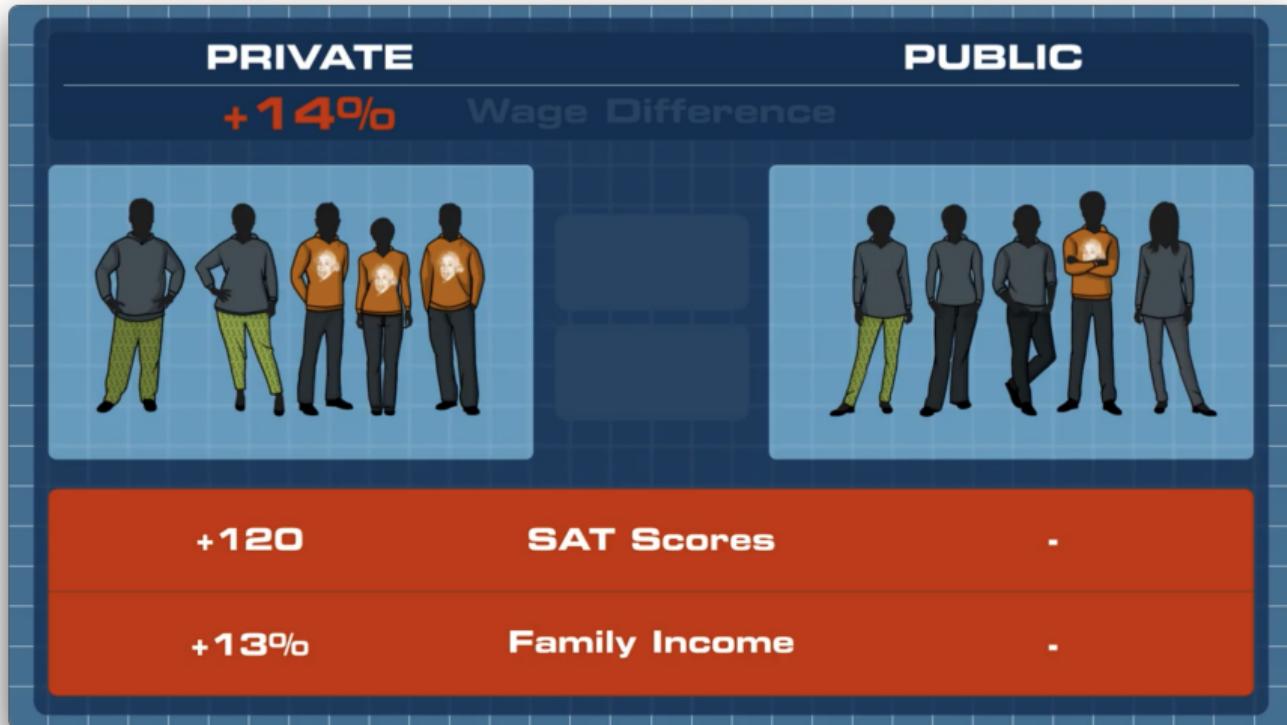
Retornos a Universidad Privada



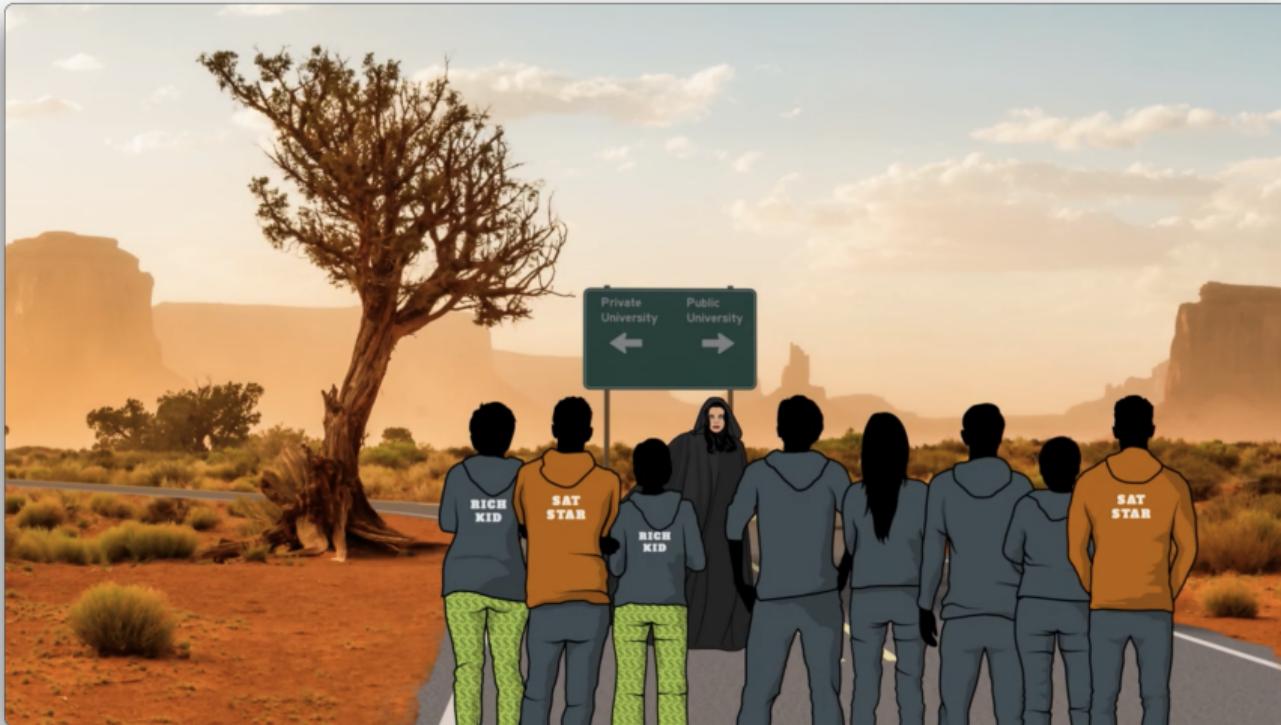
Retornos a Universidad Privada



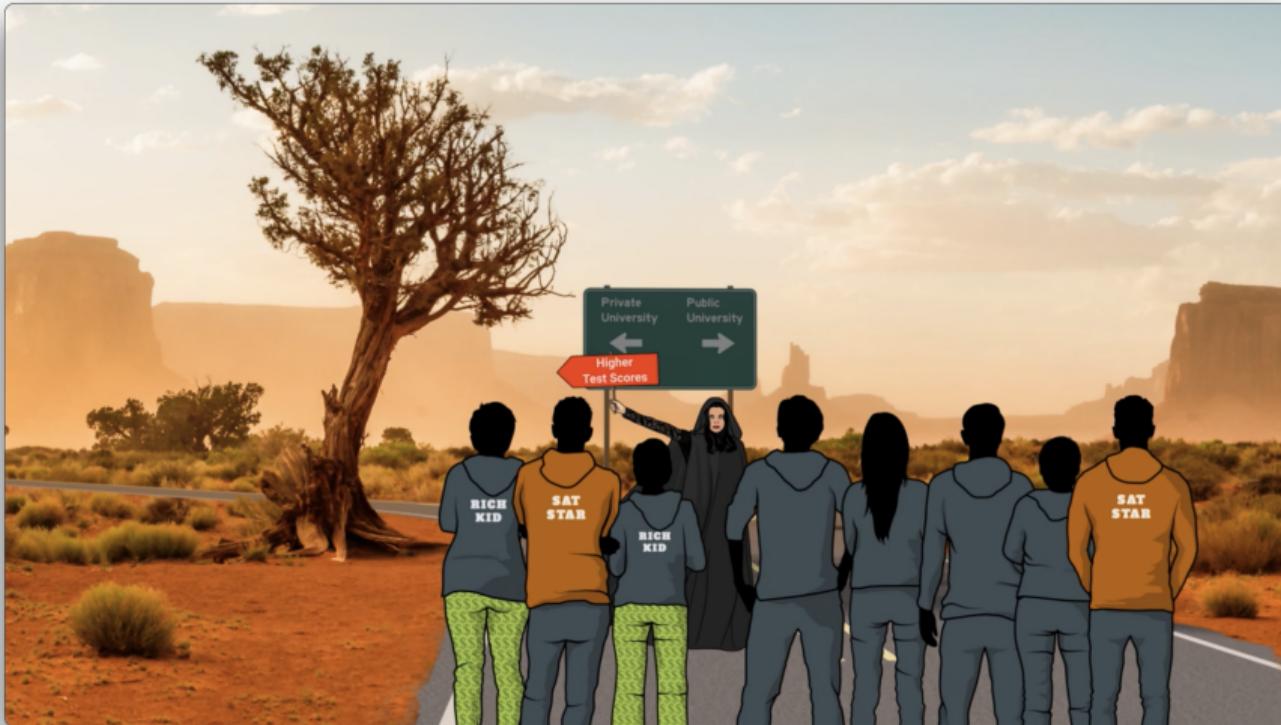
Retornos a Universidad Privada



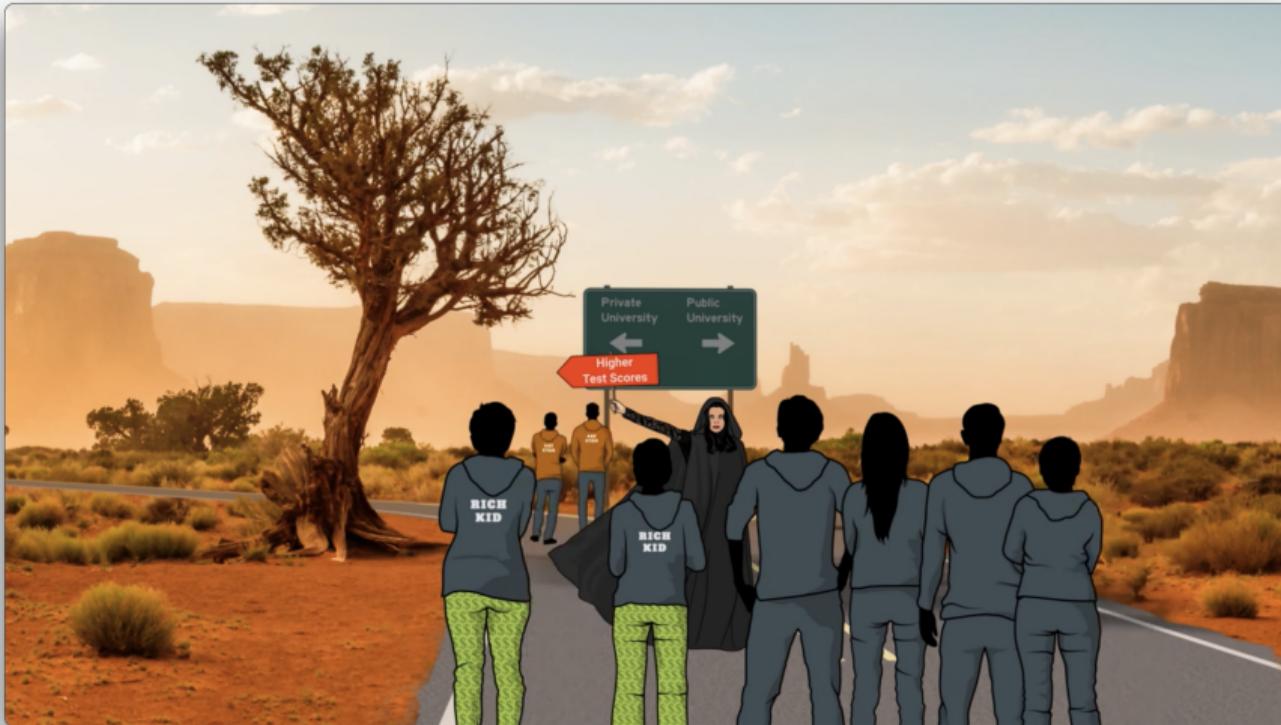
Retornos a Universidad Privada



Retornos a Universidad Privada



Retornos a Universidad Privada



Retornos a Universidad Privada



Retornos a Universidad Privada



Retornos a Universidad Privada



Retornos a Universidad Privada



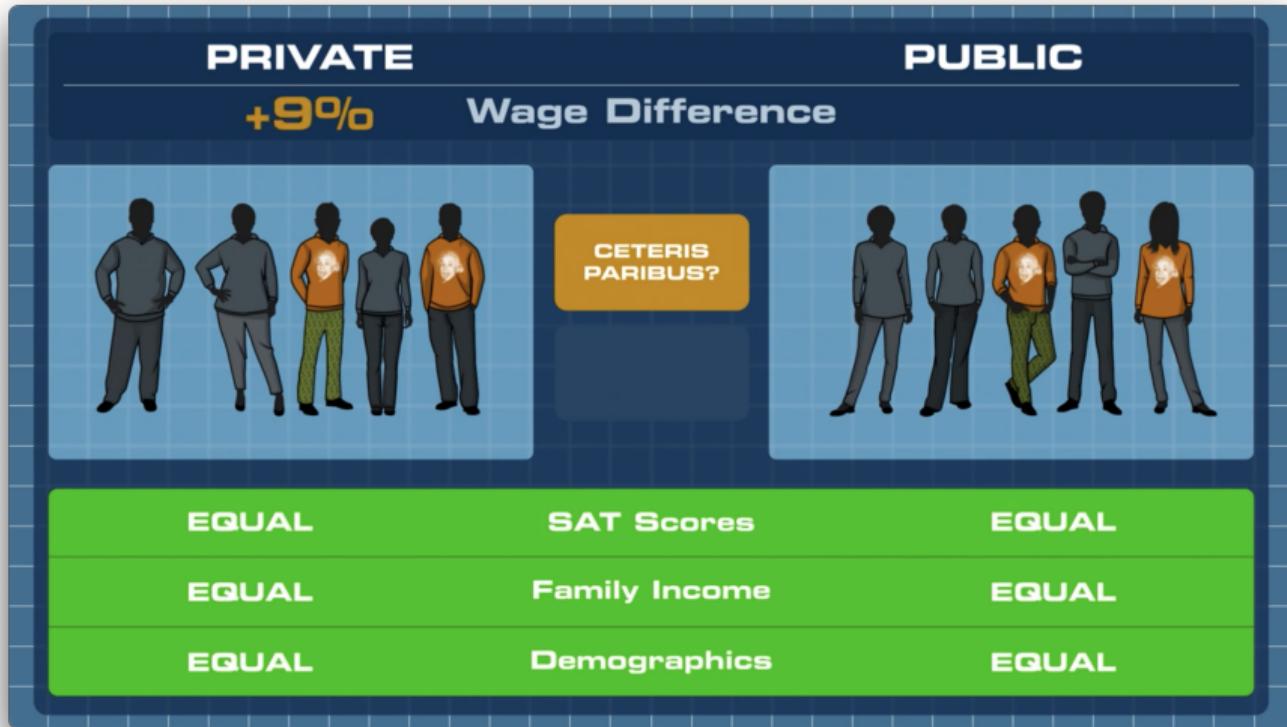
Retornos a Universidad Privada



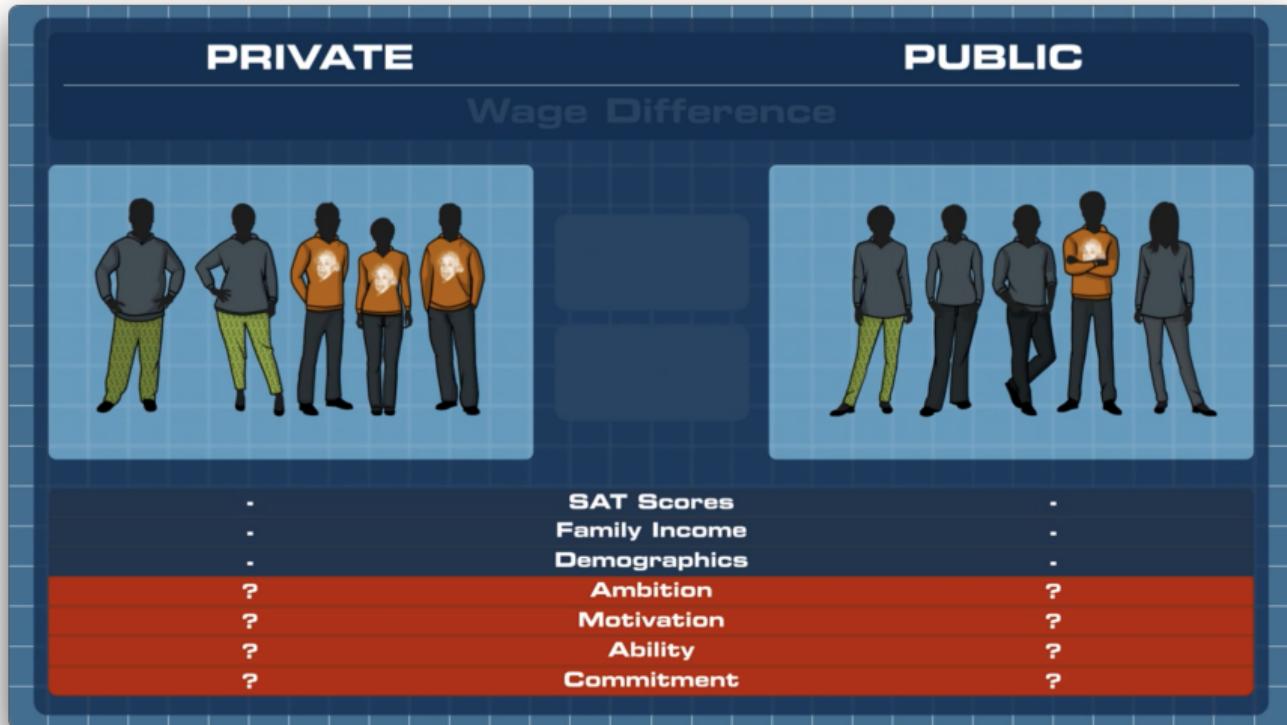
Retornos a Universidad Privada



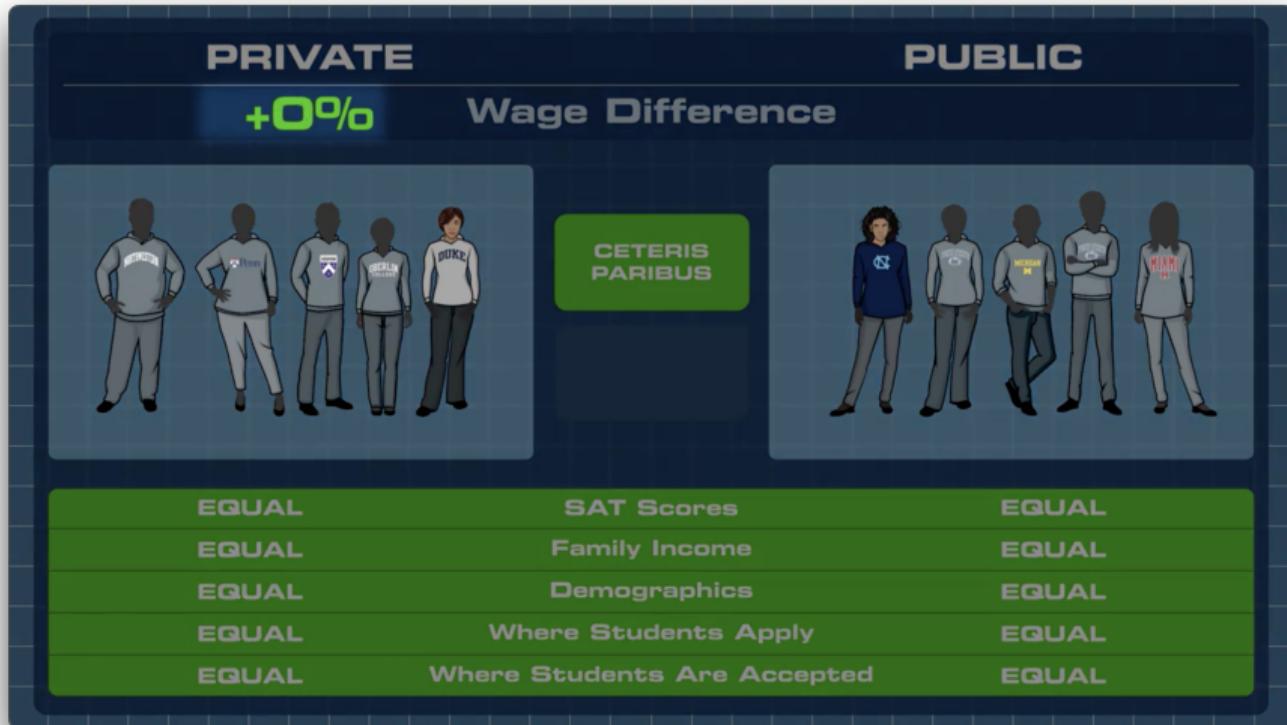
Retornos a Universidad Privada



Retornos a Universidad Privada



Retornos a Universidad Privada



Dale and Krueger

		Private			Public			1996 Earnings
Applicant Group	Student	Ivy	Leafy	Smart	All State	Ball State	Altered State	
∞	1		Reject	Admit		Admit		110,000
	A	2	Reject	Admit		Admit		100,000
		3	Reject	Admit		Admit		110,000
	B	4	Admit		Admit		Admit	60,000
		5	Admit		Admit		Admit	30,000
	C	6		Admit				115,000
		7		Admit				75,000
	D	8	Reject		Admit	Admit		90,000
		9	Reject		Admit	Admit		60,000

Notes: Students enroll at the college indicated in **bold**; enrollment decisions are also highlighted in grey.

Dale and Krueger

	No Selection Controls			Selection Controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private School	0.135 (0.055)	0.095 (0.052)	0.086 (0.034)	0.007 (0.038)	0.003 (0.039)	0.013 (0.025)
Own SAT score/100		0.048 (0.009)	0.016 (0.007)		0.033 (0.007)	0.001 (0.007)
Predicted log(Parental Income)			0.219 (0.022)			0.190 (0.023)
Female				-0.403 (0.018)		-0.395 (0.021)
Black				0.005 (0.041)		-0.040 (0.042)
Hispanic				0.062 (0.072)		0.032 (0.070)
Asian				0.170 (0.074)		0.145 (0.068)
Other/Missing Race				-0.074 (0.157)		-0.079 (0.156)
High School Top 10 Percent				0.095 (0.027)		0.082 (0.028)
High School Rank Missing				0.019 (0.033)		0.015 (0.037)
Athlete				0.123 (0.025)		0.115 (0.027)
Selection Controls	N	N	N	Y	Y	Y

Notes: Columns (1)-(3) include no selection controls. Columns (4)-(6) include a dummy for each group formed by matching students according to schools at which they were accepted or rejected. Each model is estimated using only observations with Barron's matches for which different students attended both private and public schools. The sample size is 5,583. Standard errors are shown in parentheses.

Dale and Krueger

	No Selection Controls			Selection Controls		
	(1)	(2)	(3)	(4)	(5)	(6)
School Avg. SAT Score/100	0.109 (0.026)	0.071 (0.025)	0.076 (0.016)	-0.021 (0.026)	-0.031 (0.026)	0.000 (0.018)
Own SAT score/100		0.049 (0.007)	0.018 (0.006)		0.037 (0.006)	0.009 (0.006)
Predicted log(Parental Income)			0.187 (0.024)			0.161 (0.025)
Female				-0.403 (0.015)		-0.396 (0.014)
Black				-0.023 (0.035)		-0.034 (0.035)
Hispanic				0.015 (0.052)		0.006 (0.053)
Asian				0.173 (0.036)		0.155 (0.037)
Other/Missing Race				-0.188 (0.119)		-0.193 (0.116)
High School Top 10 Percent				0.061 (0.018)		0.063 (0.019)
High School Rank Missing				0.001 (0.024)		-0.009 (0.022)
Athlete				0.102 (0.025)		0.094 (0.024)
Average SAT Score of Schools Applied To/100					0.138 (0.017)	0.116 (0.015) 0.089 (0.013)
Sent Two Application					0.082 (0.015)	0.075 (0.014) 0.063 (0.011)
Sent Three Applications					0.107 (0.017)	0.096 (0.014) 0.074 (0.011)

Revisión para el quiz

- ▶ ¿Qué es el sesgo de variable omitida?

- 1 Una muestra demasiado pequeña.
- 2 Sesgo de selección en un contexto de regresión.
- 3 Baja correlación entre variables de interés.
- 4 Otra forma de decir ceteris paribus.

Revisión para el quiz

- ▶ La asignación aleatoria es un arma ideal en la lucha contra el sesgo de selección porque
 - 1 Genera comparaciones de manzanas con manzanas.
 - 2 Genera un valor R cuadrado ideal.
 - 3 Es fácil de entender.
 - 4 Es fácil de implementar.

Revisión para el quiz

- ▶ La regresión intenta simular una asignación aleatoria mediante:
 - 1 Eliminar observaciones que no tengan efecto sobre la pregunta de investigación.
 - 2 Ignorar las similitudes entre las personas.
 - 3 Controlar los factores que importan.

Revisión para el quiz

- ▶ Supongamos que queremos estudiar los retornos a la educación. Supongamos que tiene acceso a datos sobre personas con diferentes niveles de escolaridad. ¿Cuál de ellos podría ser mejor para comparar?
 - 1 Mujeres que crecen en una misma ciudad.
 - 2 Gemelos que crecen en el mismo hogar.
 - 3 Personas del mismo código postal.
 - 4 Gemelos separados al nacer pero que viven en la misma ciudad.

Revisión para el quiz

- ▶ Supongamos que está interesado en los retornos de estudiar economía en lugar de historia del arte. ¿Cuál es la mejor manera de aprender sobre el efecto causal de estudiar economía sobre las ganancias futuras?
 - 1 Compare los ingresos de los estudiantes de economía con los de los estudiantes de historia del arte en Colombia.
 - 2 Compare los ingresos de los estudiantes de economía con los de los estudiantes de historia del arte con puntajes y ambiciones similares.
 - 3 Compare los ingresos de los estudiantes de economía con los de los estudiantes de historia del arte de la misma escuela.