

Desarrollo Económico

Andrés Barrios Fernández

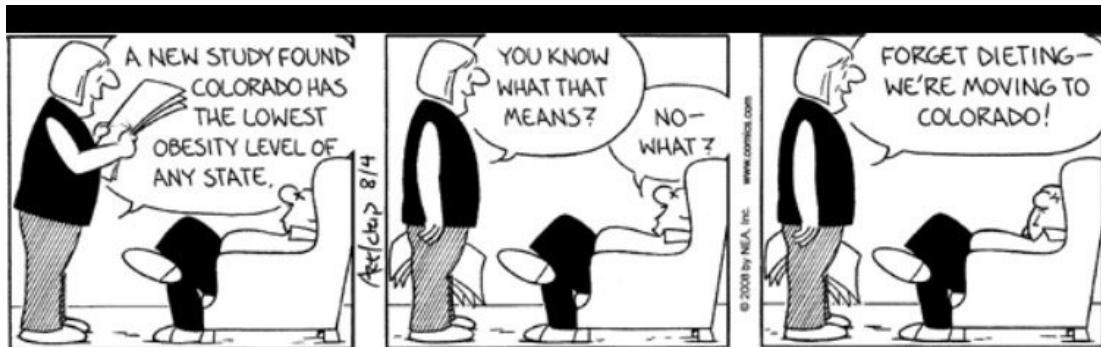
abarriosf@uandes.cl

Universidad de los Andes

August 26, 2026

Preguntas Causales

¿Cómo determinar causalidad?



© NEA, Inc.

Preguntas Causales

Escenarios Contrafactuales



Preguntas Causales

Escenarios Contrafactuales

Preguntas Causales

Escenarios Contrafactuales

Preguntas Causales

Sesgo de Selección

Para cada individuo, existen dos *outcomes* potenciales:

- Sueldo en caso de asistir a universidad privada (y_{0i}) y sueldo en caso de asistir a universidad pública (y_{1i}).
- Estamos interesados en la diferencia $y_{0i} - y_{1i}$.
- Sólo observamos un potencial *outcome*.

¿Por qué no comparar entonces a egresados de universidades privadas ($D_i = 1$) y públicas ($D_i = 0$)?

$$\begin{aligned} E[y_i | D_i = 1] - E[y_i | D_i = 0] &= E[y_{1i} | D_i = 1] - E[y_{0i} | D_i = 0] \\ &= \underbrace{E[y_{1i} | D_i = 1] - E[y_{0i} | D_i = 1]}_{\text{ATT}} + \underbrace{E[y_{0i} | D_i = 1] - E[y_{0i} | D_i = 0]}_{\text{Sesgo de selección}} \end{aligned}$$

Preguntas Causales

¿Cómo eliminar el sesgo de selección?

figure05.png

Preguntas Causales

¿Cómo eliminar el sesgo de selección?

$$\begin{aligned} E[y_i|D_i = 1] - E[y_i|D_i = 0] &= E[y_{1i}|D_i = 1] - E[y_{0i}|D_i = 0] \\ &= \underbrace{E[y_{1i}|D_i = 1] - E[y_{0i}|D_i = 1]}_{\text{ATT}} + \underbrace{E[y_{0i}|D_i = 1] - E[y_{0i}|D_i = 0]}_{\text{Sesgo de selección}} \end{aligned}$$

Aprovechar fuentes de variación exógena y usar los “*big five*”:

- 1 Randomized controlled trials (RCT)
- 2 Linear regression
- 3 Differences-in-differences (DID)
- 4 Regression discontinuity design (RDD)
- 5 Instrumental variables (IV)



Josh Angrist (MIT) & Steve Pischke (LSE)

Preguntas Causales

Práctica

- ¿Cuál es el efecto de un año adicional de escolaridad en empleabilidad y salarios?
- ¿Cuál es el efecto de un aumento del salario mínimo en empleo?
- ¿Cuál es el efecto de tener un hijo/a en los ingresos de los padres?
- ¿Cuál es el efecto del barrio en qué vivimos en nuestras trayectorias educativas y laborales?

Pregunta	Outcome	Tratamiento	Outcomes contrafactuales	Desafíos de identificación
			$Y_{0i}:$	
			$Y_{1i}:$	

Preguntas Causales

¿Qué hemos aprendido?

- No es trivial responder preguntas causales: sólo observamos la realización de un *outcome* potencial.
- Una comparación simple de individuos típicamente nos conducirá a respuestas sesgadas:
 - Los *outcomes* potenciales de individuos siguiendo distintas trayectorias típicamente son distintos.
- Solución: encontrar fuentes de variación exógena que nos permitan usar alguno de los “big five”:
 - RCT
 - Regresión lineal
 - DiD
 - RDD
 - IV
- Fuentes para profundizar los contenidos de la clase de hoy:
 - Mastering Metrics. Joshua Angrist & Jorn-Steffen Pischke.
 - Mostly Harmless Econometrics. Joshua Angrist & Jorn-Steffen Pischke.
 - <https://www.masteringmetrics.com/online-metrics-resources/>.

Desarrollo Económico

Andrés Barrios Fernández

abarriosf@uandes.cl

Universidad de los Andes

August 26, 2026

Preguntas Causales

Sesgo de seleccion y herramientas para resolverlo

Sesgo de selección:

$$\begin{aligned} E[y_i|D_i = 1] - E[y_i|D_i = 0] &= E[y_{1i}|D_i = 1] - E[y_{0i}|D_i = 0] \\ &= \underbrace{E[y_{1i}|D_i = 1] - E[y_{0i}|D_i = 1]}_{\text{ATT}} + \underbrace{E[y_{0i}|D_i = 1] - E[y_{0i}|D_i = 0]}_{\text{Sesgo de selección}} \end{aligned}$$

¿Cómo podemos solucionarlo?

- RCT
- Regresión lineal
- DiD
- RDD
- IV

Preguntas Causales

Sesgo de selección y herramientas para resolverlo

Volvamos al ejemplo que estabamos usando la clase anterior: universidad privada vs pública.

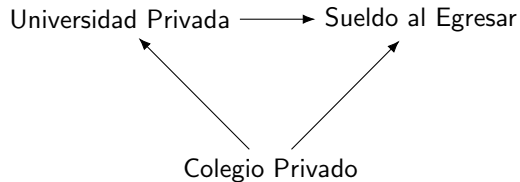


Universidad Privada —→ Sueldo al Egresar

Preguntas Causales

Sesgo de selección y herramientas para resolverlo

Volvamos al ejemplo que estabamos usando la clase anterior: universidad privada vs pública.



Preguntas Causales

Sesgo de selección y herramientas para resolverlo

¿Cómo nos ayudaría un RCT en este caso?



Asignar la universidad de manera aleatoria rompe la relación entre el *confounder* y el tratamiento:

Universidad Privada → Sueldo al Egresar

Colegio Privado ↗

$$\begin{aligned} E[y_i|D_i = 1] - E[y_i|D_i = 0] &= E[y_{1i}|D_i = 1] - E[y_{0i}|D_i = 0] \\ &= \underbrace{E[y_{1i}|D_i = 1] - E[y_{0i}|D_i = 1]}_{\text{ATT}} + \underbrace{E[y_{0i}|D_i = 1] - E[y_{0i}|D_i = 0]}_{\text{Sesgo de selección}} \end{aligned}$$

Preguntas Causales

Sesgo de selección y herramientas para resolverlo

¿Por qué no usar siempre RCTs entonces?



I called you all here because
I need you to stop helping people.
It's really messing up my impact assessment.

Algunas desventajas de los RCTs:

- Consideraciones éticas (e.g., vacunas contra COVID-19).
- Viabilidad: hay variables que no podemos manipular (e.g., sexo).
- Costo: la implementación puede ser muy cara.

¿Qué alternativas tenemos?

- Explotar variación exógena generada por experimentos naturales.

Preguntas Causales

Sesgo de selección y herramientas para resolverlo

¿Por qué no usar siempre RCTs entonces?



I called you all here because
I need you to stop helping people.
It's really messing up my impact assessment.

Algunas desventajas de los RCTs:

- Consideraciones éticas (e.g., vacunas contra COVID-19).
- Viabilidad: hay variables que no podemos manipular (e.g., sexo).
- Costo: la implementación puede ser muy cara.

¿Qué alternativas tenemos?

- Explotar variación exógena generada por **experimentos naturales**.

Preguntas Causales

Regresión Lineal

Controlar por **todos** los potenciales *confounders* en una regresión lineal elimina el sesgo de selección:



$$y_i = \beta_0 + \beta_1 D_i + \underbrace{\sum_{k=2}^K \beta_k x_{ki}}_{\text{Confounders}} + \varepsilon_i$$

Universidad Privada $\longrightarrow$ Sueldo al Egresar

Colegio Privado $\nearrow$

$$\begin{aligned} E[y_i | D_i = 1] - E[y_i | D_i = 0] &= E[y_{1i} | D_i = 1] - E[y_{0i} | D_i = 0] \\ &= \underbrace{E[y_{1i} | D_i = 1] - E[y_{0i} | D_i = 1]}_{\text{Desarrollo Económico}} + \underbrace{E[y_{0i} | D_i = 1] - E[y_{0i} | D_i = 0]}_{\text{Sesgo de Selección}} \end{aligned}$$

Preguntas Causales

Regresión Líneal

Table: Efecto de Asistir a una Universidad Privada en Ingresos

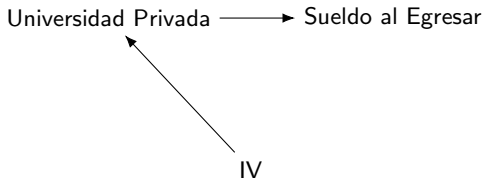
	Ln(Earnings) (1)	Ln(Earning) (2)	Ln(Earnings) (3)	Ln(Earnings) (4)
Universidad privada = 1	0.288*** (0.110)	0.194** (0.095)	0.201** (0.098)	0.201** (0.096)
Colegio privado = 1		0.095** (0.045)	0.073** (0.035)	0.073** (0.033)
NEM (100)			0.060*** (0.021)	0.060*** (0.020)
Mujer = 1				-0.143*** (0.052)
Constante	19.008*** (2.351)	18.939*** (2.247)	15.607*** (2.189)	16.021*** (2.234)
Observaciones	25634	25634	25634	25634

Preguntas Causales

Variables Instrumentales

→ Variables Instrumentales (IV):

- Relevancia: $IV \rightarrow$ universidad privada.
- Exclusion restriction: No afecta al outcome de interés de manera directa o indirecta.



Usar IV nos ayuda a solucionar el sesgo de selección, pero tiene costos:

- No es fácil encontrar un IV válido: conocimiento del contexto e instituciones es fundamental.
- Podemos identificar el efecto promedio local del tratamiento (LATE), no el efecto promedio (ATT):
 - Efecto en el subconjunto de individuos que responde al instrumento.
 - ¿Qué tan representativo es este subconjunto respecto a la población general?
 - ¿Qué tan grande es este subconjunto?

Preguntas Causales

Diferencias-en-diferencias

Imaginemos que en una ciudad hay dos universidades: Universidad los Alpes (UA) y Universidad del Himalaya (UH).

- Originalmente las dos universidades eran públicas.
- En el año 2000, la Universidad de los Alpes se hizo privada.

Este setting nos permitirían usar DiD para estimar el efecto de asistir a una universidad privada:

- Diferencia 1: Δ en ingreso entre estudiantes que asistieron a la UA cuando era pública y privada.
- Diferencia 2: Δ en ingreso entre estudiantes que asistieron a la UA y a la UH.
- Esto nos permite controlar por evolución temporal de salarios (ciclo) y por diferencias en el nivel de salarios entre universidades.

Diferencias-en-diferencias

El gráfico muestra la evolución del sueldo mensual al quinto año de egreso para dos tipos de universidades: pública y privada. El eje vertical representa el sueldo mensual y el eje horizontal el año. Una línea vertical roja discontinua marca el año 2000. Para la UA pública, el sueldo aumenta de manera constante y moderada. Para la UA privada, el sueldo también aumenta, pero a una tasa mucho mayor, especialmente después del año 2000, lo que resulta en una brecha salarial creciente entre ambas.

- Trayectorias paralelas: los estudiantes de UA y UH hubiesen seguido trayectorias paralelas en la ausencia del cambio.
- No hay cambios entre tratados y controles: no hay cambios en la composición de los estudiantes asociados a la privatización de UA.
- Ausencia de externalidades: lo que ocurre con los estudiantes de la UA no afecta lo que ocurre con los estudiantes de la UH.

Preguntas Causales

Regresiones Discontinuas (RDD)

Imaginemos que estamos interesados en estudiar el efecto de estudiar ingeniería comercial en la U. de los Andes (UA) en vez de en la U. de Chile (UCh). Podemos aprovechar la variación generada por el sistema de admisión universitario.

En el proceso de admisión 2022, los puntajes de corte fueron:

- Universidad de Chile: 698.2
- Universidad de los Andes: 682.1

Podemos acceder a datos de postulaciones e identificar a estudiantes en el margen de ir a la UCh/UA:

- $\text{Puntaje} \geq 698.2 \rightarrow$ estudiante admitido en la UCh.
- $682.1 \leq \text{puntaje} < 698.2 \rightarrow$ estudiante admitido en la UA.

Idea de RDD: Comparar a estudiantes muy parecidos y que por alguna razón aleatoria obtuvieron un puntaje ligeramente por encima/debajo del puntaje de corte.

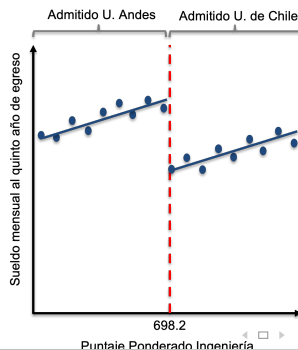
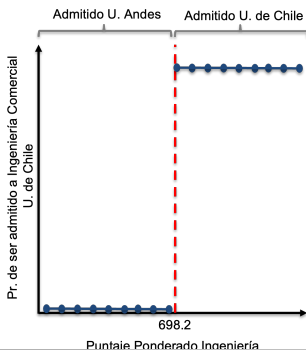
Preguntas Causales

Regresiones Discontinuas (RDD)

Un RDD nos permite identificar el efecto causal de manera local si:

- No existe manipulación de la variable de asignación.
- No hay discontinuidades en *confounders* en el *cutoff*.

$$y_i = \beta_0 + \beta_1 A_i + \beta_2 \text{Puntaje}_i + \beta_3 A_i \times \text{Puntaje}_i + \varepsilon_i$$



Preguntas Causales

¿Qué aprendimos esta semana?

Las preguntas causales son particularmente interesantes, pero responderlas no es trivial.

- Sólo observamos un *outcome* potencial.
- Tratados y controles típicamente difieren en múltiples otras dimensiones relevantes (*confounders*).
- Para responder estas preguntas necesitamos encontrar una fuente de variación exógena en take-up del tratamiento.
- Los RCT generan esta variación por diseño, pero no siempre son viables.
- En la ausencia de un RCT, podemos explotar experimentos naturales y usar las siguientes estrategias:
 - Regresión lineal.
 - Variables instrumentales.
 - Diferencias-en-diferencias.
 - Regresiones Discontinuas.

Recuerden! Mastering Metrics and Mostly Harmless Econometrics son excelentes fuentes para repasar estos conceptos.

Desarrollo Económico

Andrés Barrios Fernández

abarriosf@uandes.cl

Universidad de los Andes

August 26, 2026