

# Semana 4. Experimentos Sociales Controlados (RCT)

Equipo Econometría Avanzada

Universidad de los Andes

2 de septiembre de 2022



# Contenido

- 1 Contexto y pregunta de investigación
- 2 Consecuencias de la aleatoriedad de  $D$
- 3 Supuestos de identificación
- 4 Estimación

# Contexto

Nores, Bernal y Barnett (2019) utilizan datos de niños y niñas de comunidades vulnerables de la intervención aeioTU.

- Intervención en **centros de cuidado** basada en un currículo estructurado alineado con la filosofía Reggio Emilia.
- La **filosofía Reggio Emilia** enfatiza en la co-construcción del aprendizaje de los niños involucrando tanto a los niños, los maestros y a la familia.
- Se **capacitó** a las maestras antes y durante la intervención.
- Dado que había sobre-inscripción, **los cupos fueron asignados en un sorteo público**.

# Pregunta de investigación

¿Cuál es el efecto de la **asignación** de un cupo en un centro de cuidado ( $D_i$ ) sobre el desarrollo cognitivo ( $Y_i$ ) ?

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

- ¿Cuál es nuestro parámetro de interés?

# Consecuencias de la aleatoriedad de $D$

¿Qué pasa cuando la  $D$  es aleatoria?

- 1 Es plausible el supuesto de  $Y_{0i}, Y_{1i} \perp D_i$ : se “soluciona” el sesgo de selección.
- 2 En promedio, la única diferencia entre  $D = 1$  y  $D = 0$  es la asignación del cupo en el centro de cuidado.
- 3  $E[Y_{0i}|D_i = 0]$  es plausiblemente una buena aproximación de  $E[Y_{0i}|D_i = 1]$ .

Sin embargo, esto solo sucede si se cumplen ciertas condiciones.

# Supuestos de identificación

- ❶ **Correcta aleatorización:** la asignación al grupo de tratamiento es efectivamente aleatoria.
  - ❷ **No hay sesgo por la aleatorización:** la aleatorización únicamente determina si un individuo participa o no en el programa, pero no afecta nada más.
    - ▶ Efectos placebo, Hawthorne y Jhon Henry.
  - ❸ **SUTVA (Stable Unit Treatment Value Assumption):** La asignación del tratamiento no afecta al grupo de control (no hay externalidades ni contagio).
- ¿A cuáles podemos aproximarnos de forma teórica?
  - ¿A cuáles podemos aproximarnos de forma empírica?

# Balance Muestral

Son una aproximación empírica a qué tan bien hecha está la aleatorización y a qué tanto se parecen  $D = 1$  a  $D = 0$ . Las más comunes son:

- 1 Diferencia de medias con las características observadas en línea de base (o  $\chi^2$  de Pearson en caso de ser categóricas).
- 2 Análisis de regresión: ver si algo predice la asignación al tratamiento

$$D_i = \delta_0 + \delta_1 X_{1i} + \delta_2 X_{2i} + \dots + \delta_K X_{Ki} + \epsilon_i$$

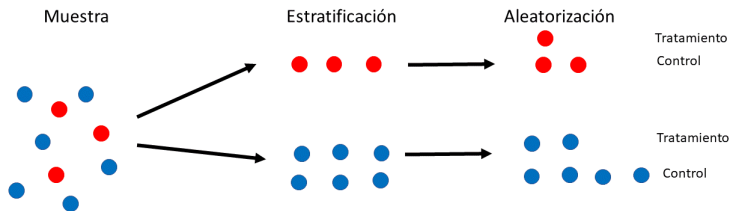
**Nota:** Si la estimación es por MCO y la aleatorización individual se **deben** usar errores estándar robustos de White.

- ¿Vale la pena siempre incluir controles?
- ¿A qué variables se les debe revisar el balance muestral?
- Si no hay diferencias en las variables observadas, entonces, ¿sabemos con certeza que no hay sesgo?
- Si encontramos al menos una variable que difiere entre grupos, entonces, ¿el experimento se invalida?

# Estratificación

$$Y_i = \beta_0 + \tau D_i + X_i\gamma + u_i$$

- ¿Por qué puede ser deseable estratificar la aleatorización?
- ¿Hay que incluir controles adicionales si la aleatorización es estratificada?





## ¿Qué estima una regresión con controles?

- Definamos:

$$\sigma_D^2(X_i) = \mathbb{E}[(D_i - \mathbb{E}[D_i|X_i])^2|X_i]$$

$$\tau(x) = \mathbb{E}[Y_i|D_i = 1, X = x] - \mathbb{E}[Y_i|D_i = 0, X = x]$$

Note que, bajo aleatorización,  $\tau(x)$  puede interpretarse como un efecto de tratamiento promedio condicional en las covariables.

- Se puede demostrar (inténtelo) que

$$\tau = \frac{\mathbb{E}[\sigma_D^2(X_i)\tau(X_i)]}{\mathbb{E}[\sigma_D^2(X_i)]}$$

- Esto es, el parámetro poblacional de la regresión múltiple que acompaña a la variable de tratamiento corresponde a un *variance-weighted treatment effect*.
- Esta aproximación es óptima bajo ciertas **condiciones y criterios**.

# Conclusión

¿Cómo se conduce un análisis por RCT entonces?

- ❶ Verifique si es un diseño individual o por clusters
  - ▶ Eso tendrá implicaciones sobre la estimación de los errores estándar
- ❷ ¿Son los grupos de tratamiento y control comparables?  
Analice si es razonable que se cumplan los supuestos
  - ▶ Verifique que la aleatorización se hizo correctamente.
  - ▶ Verifique si la atrición se relaciona con el tratamiento.
- ❸ Estime el efecto de  $D$  en una regresión preferiblemente multivariada.
  - ▶ Verifique el efecto diferenciado en sus grupos de interés.

¡Gracias!