

# Inferencia, Causalidad y Políticas Públicas

## ECO-60116

Week 02: Experimentos Aleatorios Controlados (RCTs)

Eduard F. Martinez Gonzalez, Ph.D.

Departamento de Economía, Universidad Icesi

September 4, 2025

# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- Diseño Experimental
- Estimación del Efecto
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Recap Week-01

- **Efecto individual:**  $\tau_i = Y_i(1) - Y_i(0)$ .
- **Problema fundamental de la inferencia causal:** para cada unidad  $i$  sólo se observa uno de los dos resultados potenciales.

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

# Recap Week-01

- **Efecto individual:**  $\tau_i = Y_i(1) - Y_i(0)$ .
- **Problema fundamental de la inferencia causal:** para cada unidad  $i$  sólo se observa uno de los dos resultados potenciales.

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

- **Diferencia de medias observada:**

$$\underbrace{\mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0]}_{\text{Diferencia observada}} = \underbrace{\mathbb{E}[Y(1) - Y(0)]}_{\text{Efecto causal (ATE)}} + \underbrace{\left( \mathbb{E}[Y(0) \mid D = 1] - \mathbb{E}[Y(0) \mid D = 0] \right)}_{\text{Sesgo de selección}}$$

# Recap Week-01

- **Efecto individual:**  $\tau_i = Y_i(1) - Y_i(0)$ .
- **Problema fundamental de la inferencia causal:** para cada unidad  $i$  sólo se observa uno de los dos resultados potenciales.

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

- **Diferencia de medias observada:**

$$\underbrace{\mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0]}_{\text{Diferencia observada}} = \underbrace{\mathbb{E}[Y(1) - Y(0)]}_{\text{Efecto causal (ATE)}} + \underbrace{\left( \mathbb{E}[Y(0) \mid D = 1] - \mathbb{E}[Y(0) \mid D = 0] \right)}_{\text{Sesgo de selección}}$$

- **Condición clave para interpretación causal:** se requiere *variación exógena* en  $D_i$  (diseño/identificación), no basta con estimar una regresión.

$$\mathbb{E}[\varepsilon \mid \mathbf{X}] = 0$$

# Recap Week-01

- **Efecto individual:**  $\tau_i = Y_i(1) - Y_i(0)$ .
- **Problema fundamental de la inferencia causal:** para cada unidad  $i$  sólo se observa uno de los dos resultados potenciales.

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

- **Diferencia de medias observada:**

$$\underbrace{\mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0]}_{\text{Diferencia observada}} = \underbrace{\mathbb{E}[Y(1) - Y(0)]}_{\text{Efecto causal (ATE)}} + \underbrace{\left( \mathbb{E}[Y(0) \mid D = 1] - \mathbb{E}[Y(0) \mid D = 0] \right)}_{\text{Sesgo de selección}}$$

- **Condición clave para interpretación causal:** se requiere *variación exógena* en  $D_i$  (diseño/identificación), no basta con estimar una regresión.

$$\mathbb{E}[\varepsilon \mid \mathbf{X}] = 0$$

- **Bajo exogeneidad:** el estimador de Mínimos Cuadrados Ordinarios (OLS) es insesgado.

$$\mathbb{E}[\hat{\beta} \mid \mathbf{X}] = \beta$$

# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- Diseño Experimental
- Estimación del Efecto
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Intuición del RCT

Un RCT bien diseñado se considera el “*Gold standard*” en evaluación de impacto, pues ofrece alta credibilidad interna. El RCT permite construir un **contrafactual creíble**.



# Intuición del RCT

Un RCT bien diseñado se considera el “*Gold standard*” en evaluación de impacto, pues ofrece alta credibilidad interna. El RCT permite construir un **contrafactual creíble**.

**Idea central:** la asignación aleatoria del tratamiento garantiza independencia entre  $D_i$  y los resultados potenciales.

$$D_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\}$$

# Intuición del RCT

Un RCT bien diseñado se considera el “*Gold standard*” en evaluación de impacto, pues ofrece alta credibilidad interna. El RCT permite construir un **contrafactual creíble**.

**Idea central:** la asignación aleatoria del tratamiento garantiza independencia entre  $D_i$  y los resultados potenciales.

$$D_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\}$$

**Consecuencia:** en valor esperado, los grupos de tratamiento y control son comparables en todos los aspectos.

$$\mathbb{E}[Y(0) \mid D = 1] = \mathbb{E}[Y(0) \mid D = 0]$$

# Intuición del RCT

Un RCT bien diseñado se considera el “*Gold standard*” en evaluación de impacto, pues ofrece alta credibilidad interna. El RCT permite construir un **contrafactual creíble**.

**Idea central:** la asignación aleatoria del tratamiento garantiza independencia entre  $D_i$  y los resultados potenciales.

$$D_i \perp\!\!\!\perp \{Y_i(1), Y_i(0)\}$$

**Consecuencia:** en valor esperado, los grupos de tratamiento y control son comparables en todos los aspectos.

$$\mathbb{E}[Y(0) \mid D = 1] = \mathbb{E}[Y(0) \mid D = 0]$$

**Estimación simple (muestra grande):**

$$\widehat{ATE} = \bar{Y}_{\text{trat}} - \bar{Y}_{\text{ctrl}}$$

# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- **Diseño Experimental**
- Estimación del Efecto
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Unidad de Aleatorización

## Unidad de aleatorización

- Puede ser:
  - ▶ **Individual:** hogares, estudiantes, pacientes.
  - ▶ **Por grupos o clusters:** escuelas, veredas, municipios, hospitales.

# Unidad de Aleatorización

## Unidad de aleatorización

- Puede ser:
  - ▶ **Individual:** hogares, estudiantes, pacientes.
  - ▶ **Por grupos o clusters:** escuelas, veredas, municipios, hospitales.
- La elección depende de la naturaleza de la intervención y del riesgo de **spillovers**:
  - ▶ Ej.: un programa educativo impartido a nivel de escuela → aleatorización a nivel de escuela.
  - ▶ Previene contaminación: los controles no reciben efectos indirectos (ej. interacción entre alumnos).

# Unidad de Aleatorización

## Unidad de aleatorización

- Puede ser:
  - ▶ **Individual:** hogares, estudiantes, pacientes.
  - ▶ **Por grupos o clusters:** escuelas, veredas, municipios, hospitales.
- La elección depende de la naturaleza de la intervención y del riesgo de **spillovers**:
  - ▶ Ej.: un programa educativo impartido a nivel de escuela → aleatorización a nivel de escuela.
  - ▶ Previene contaminación: los controles no reciben efectos indirectos (ej. interacción entre alumnos).
- **Costos y logística:** aleatorizar por cluster suele ser más barato y operativo, pero requiere muestras mayores para mantener poder estadístico.

# Unidad de Aleatorización

## Unidad de aleatorización

- Puede ser:
  - ▶ **Individual:** hogares, estudiantes, pacientes.
  - ▶ **Por grupos o clusters:** escuelas, veredas, municipios, hospitales.
- La elección depende de la naturaleza de la intervención y del riesgo de **spillovers**:
  - ▶ Ej.: un programa educativo impartido a nivel de escuela → aleatorización a nivel de escuela.
  - ▶ Previene contaminación: los controles no reciben efectos indirectos (ej. interacción entre alumnos).
- **Costos y logística:** aleatorizar por cluster suele ser más barato y operativo, pero requiere muestras mayores para mantener poder estadístico.
- **Implicación estadística:**
  - ▶ Los resultados de unidades en un mismo cluster están correlacionados →  $\rho$  (intracluster correlation coefficient).
  - ▶ Esto aumenta la varianza de los estimadores: se debe ajustar con errores estándar cluster-robust y calcular el **factor de diseño**



# Estratificación

## Estratificación o blocking

- Consiste en dividir la muestra en **estratos homogéneos** usando variables predeterminadas (sexo, edad, escuela, valor en línea base de  $Y$ ).
- Dentro de cada estrato, se realiza la aleatorización **separadamente**.

# Estratificación

## Estratificación o blocking

- Consiste en dividir la muestra en **estratos homogéneos** usando variables predeterminadas (sexo, edad, escuela, valor en línea base de  $Y$ ).
- Dentro de cada estrato, se realiza la aleatorización **separadamente**.
- **Beneficios:**
  - ▶ Mejora la **precisión**: controla la varianza, balancea covariables clave.
  - ▶ Asegura **representatividad** en subgrupos relevantes (p.ej., que mujeres y hombres estén bien distribuidos).
  - ▶ Reduce el riesgo de “mala suerte” en la asignación.

# Estratificación

## Estratificación o blocking

- Consiste en dividir la muestra en **estratos homogéneos** usando variables predeterminadas (sexo, edad, escuela, valor en línea base de  $Y$ ).
- Dentro de cada estrato, se realiza la aleatorización **separadamente**.
- **Beneficios:**
  - ▶ Mejora la **precisión**: controla la varianza, balancea covariables clave.
  - ▶ Asegura **representatividad** en subgrupos relevantes (p.ej., que mujeres y hombres estén bien distribuidos).
  - ▶ Reduce el riesgo de “mala suerte” en la asignación.
- **Requisitos:**
  - ▶ Tamaños mínimos adecuados en cada estrato → evitar celdas muy pequeñas.
  - ▶ Variables de estratificación deben definirse antes del tratamiento y no ser afectadas por él.

# Estratificación

## Estratificación o blocking

- Consiste en dividir la muestra en **estratos homogéneos** usando variables predeterminadas (sexo, edad, escuela, valor en línea base de  $Y$ ).
- Dentro de cada estrato, se realiza la aleatorización **separadamente**.
- **Beneficios:**
  - ▶ Mejora la **precisión**: controla la varianza, balancea covariables clave.
  - ▶ Asegura **representatividad** en subgrupos relevantes (p.ej., que mujeres y hombres estén bien distribuidos).
  - ▶ Reduce el riesgo de “mala suerte” en la asignación.
- **Requisitos:**
  - ▶ Tamaños mínimos adecuados en cada estrato → evitar celdas muy pequeñas.
  - ▶ Variables de estratificación deben definirse antes del tratamiento y no ser afectadas por él.
- **Regla práctica:** si usas estratificación, debes **controlar por los estratos** en la estimación (regresión/ANOVA), incluso si el balance es perfecto.

# Chequeos de balance

## Bajo aleatorización válida:

$$X \perp D \quad \Rightarrow \quad \mathbb{E}[X \mid D = 1] = \mathbb{E}[X \mid D = 0]$$

donde  $X$  son características predeterminadas (línea base) y  $D$  indica tratamiento.

# Chequeos de balance

## Bajo aleatorización válida:

$$X \perp D \Rightarrow \mathbb{E}[X \mid D = 1] = \mathbb{E}[X \mid D = 0]$$

donde  $X$  son características predeterminadas (línea base) y  $D$  indica tratamiento.

## ¿Qué se hace en la práctica?

- **Tabla de balance:** reportar medias por grupo, diferencias, errores estándar y p-valores para cada covariable relevante (edad, educación, línea base del outcome, etc.).
- **Si hubo estratificación:** incluir dummies de estrato en la estimación para reflejar el diseño.

# Chequeos de balance

## Bajo aleatorización válida:

$$X \perp D \Rightarrow \mathbb{E}[X \mid D = 1] = \mathbb{E}[X \mid D = 0]$$

donde  $X$  son características predeterminadas (línea base) y  $D$  indica tratamiento.

## ¿Qué se hace en la práctica?

- **Tabla de balance:** reportar medias por grupo, diferencias, errores estándar y p-valores para cada covariable relevante (edad, educación, línea base del outcome, etc.).
- **Si hubo estratificación:** incluir dummies de estrato en la estimación para reflejar el diseño.

## Enfoques complementarios

- **Prueba conjunta:** estimar una regresión  $D_i = \alpha + X_i' \gamma + u_i$  y aplicar un test  $F$  de significancia global.

**Mensaje:** el balance empírico brinda evidencia de que la asignación aleatoria funcionó como se esperaba.

# Poder estadístico y tamaño muestral

## ¿Qué es el poder estadístico?

- Probabilidad de rechazar  $H_0$  cuando el efecto realmente existe  $(1 - \beta)$ .
- Depende de: tamaño muestral  $n$ , varianza del outcome  $\sigma^2$ , nivel de significancia  $\alpha$ , y magnitud del efecto mínimo detectable  $\Delta$ .

$$n \approx \frac{2\sigma^2 (z_{1-\alpha/2} + z_{1-\beta})^2}{\Delta^2}$$



# Poder estadístico y tamaño muestral

## ¿Qué es el poder estadístico?

- Probabilidad de rechazar  $H_0$  cuando el efecto realmente existe  $(1 - \beta)$ .
- Depende de: tamaño muestral  $n$ , varianza del outcome  $\sigma^2$ , nivel de significancia  $\alpha$ , y magnitud del efecto mínimo detectable  $\Delta$ .

$$n \approx \frac{2\sigma^2 (z_{1-\alpha/2} + z_{1-\beta})^2}{\Delta^2}$$

## Cluster RCT: pérdida de eficiencia

- Observaciones en un cluster correlacionadas  $\Rightarrow$  mayor varianza.
- Factor de diseño (DEFF):

$$DEFF \approx 1 + (\bar{m} - 1)\rho$$

donde  $\bar{m}$  = tamaño promedio de cluster,  $\rho$  = ICC (Intraclass Correlation Coefficient). Siendo la muestra efectiva:  $n_{\text{ef}} = n/DEFF$ .

**Implicación:** en RCTs por clusters se requieren más unidades para alcanzar el mismo poder que en un RCT individual.

# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- Diseño Experimental
- **Estimación del Efecto**
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Estimación básica en un RCT

**Idea central:** la aleatorización asegura que  $D_i$  es exógeno, por lo que el estimador de diferencia de medias es insesgado.

**Diferencia de medias:**

$$\hat{\tau} = \bar{Y}_{D=1} - \bar{Y}_{D=0}$$

**Regresión equivalente:**

$$Y_i = \alpha + \beta D_i + \varepsilon_i$$

donde:

$$\mathbb{E}[Y_i \mid D_i = 1] = \alpha + \beta, \quad \mathbb{E}[Y_i \mid D_i = 0] = \alpha$$

**Interpretación:**

$$\hat{\beta} = \hat{\tau} \Rightarrow \text{Efecto promedio del tratamiento (ATE/ITT)}$$

**Clave:** si la aleatorización funcionó, no hay riesgo de sesgo por variable omitida.

# ¿Por qué incluir controles en un RCT?

**El estimador ya es insesgado, pero los controles pueden:**

- Aumentar la **precisión** de la estimación al reducir  $\text{Var}(\varepsilon_i)$ .
- Explicar parte de la variabilidad en  $Y_i \rightarrow$  mayor  $R^2$ .
- Ganar **poder estadístico**: el estimador  $\hat{\beta}$  es más preciso.

# ¿Por qué incluir controles en un RCT?

**El estimador ya es insesgado, pero los controles pueden:**

- Aumentar la **precisión** de la estimación al reducir  $\text{Var}(\varepsilon_i)$ .
- Explicar parte de la variabilidad en  $Y_i \rightarrow$  mayor  $R^2$ .
- Ganar **poder estadístico**: el estimador  $\hat{\beta}$  es más preciso.

**Regresión con controles:**

$$Y_i = \alpha + \beta D_i + \gamma' X_i + \varepsilon_i$$

**Controles típicos:**

- Variables predeterminadas y no afectadas por el tratamiento.
- El valor de  $Y_i$  en la línea base es uno de los más poderosos.

## ¿Por qué incluir controles en un RCT?

**El estimador ya es insesgado, pero los controles pueden:**

- Aumentar la **precisión** de la estimación al reducir  $\text{Var}(\varepsilon_i)$ .
- Explicar parte de la variabilidad en  $Y_i \rightarrow$  mayor  $R^2$ .
- Ganar **poder estadístico**: el estimador  $\hat{\beta}$  es más preciso.

**Regresión con controles:**

$$Y_i = \alpha + \beta D_i + \gamma' X_i + \varepsilon_i$$

**Controles típicos:**

- Variables predeterminadas y no afectadas por el tratamiento.
- El valor de  $Y_i$  en la línea base es uno de los más poderosos.

**Mensaje:** el efecto estimado será similar sin controles, pero **menos significativo** al tener mayor varianza residual.

# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- Diseño Experimental
- Estimación del Efecto
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Consideraciones Éticas en los RCTs

- En ciencias sociales, un RCT puede decidir quién recibe un beneficio esencial (ej. alimentación, salud, educación). **¿Estamos jugando con la vida de las personas?**



# Consideraciones Éticas en los RCTs

- En ciencias sociales, un RCT puede decidir quién recibe un beneficio esencial (ej. alimentación, salud, educación). **¿Estamos jugando con la vida de las personas?**
- **Condición fundamental:** sobresuscripción
  - ▶ La demanda excede la capacidad de la intervención.
  - ▶ No se puede reducir el número de beneficiarios que era posible atender con el presupuesto/capacidad disponible.

# Consideraciones Éticas en los RCTs

- En ciencias sociales, un RCT puede decidir quién recibe un beneficio esencial (ej. alimentación, salud, educación). **¿Estamos jugando con la vida de las personas?**
- **Condición fundamental:** sobresuscripción
  - ▶ La demanda excede la capacidad de la intervención.
  - ▶ No se puede reducir el número de beneficiarios que era posible atender con el presupuesto/capacidad disponible.
- La asignación aleatoria se aplica **dentro de los elegibles**, lo que:
  - ▶ Garantiza igualdad de oportunidades.
  - ▶ Puede reducir la discrecionalidad (evitar que los más conectados siempre ganen).

# Consideraciones Éticas en los RCTs

- En ciencias sociales, un RCT puede decidir quién recibe un beneficio esencial (ej. alimentación, salud, educación). **¿Estamos jugando con la vida de las personas?**
- **Condición fundamental:** sobresuscripción
  - ▶ La demanda excede la capacidad de la intervención.
  - ▶ No se puede reducir el número de beneficiarios que era posible atender con el presupuesto/capacidad disponible.
- La asignación aleatoria se aplica **dentro de los elegibles**, lo que:
  - ▶ Garantiza igualdad de oportunidades.
  - ▶ Puede reducir la discrecionalidad (evitar que los más conectados siempre ganen).
- **Desafío práctico:** Todos deben inscribirse como beneficiarios potenciales antes del sorteo → manejar expectativas y asegurar transparencia.

# Amenazas a la validez interna (I)

## ① Attrition (pérdida de muestra):

- ▶ Cuando las tasas de respuesta difieren entre tratamiento y control, la comparabilidad inicial se pierde.
- ▶ Ej.: encuesta de seguimiento con 90% en tratados y 70% en controles.
- ▶ *Mitigación:*
  - ★ Maximizar esfuerzos de seguimiento (visitas adicionales, incentivos).
  - ★ Verificar balance en línea base de quienes permanecen.
  - ★ Usar métodos de corrección (bounds de Lee, imputaciones).

# Amenazas a la validez interna (I)

## ① Attrition (pérdida de muestra):

- ▶ Cuando las tasas de respuesta difieren entre tratamiento y control, la comparabilidad inicial se pierde.
- ▶ Ej.: encuesta de seguimiento con 90% en tratados y 70% en controles.
- ▶ *Mitigación:*
  - ★ Maximizar esfuerzos de seguimiento (visitas adicionales, incentivos).
  - ★ Verificar balance en línea base de quienes permanecen.
  - ★ Usar métodos de corrección (bounds de Lee, imputaciones).

## ② Spillovers / Externalidades:

- ▶ Ocurren cuando el tratamiento afecta indirectamente a no tratados.
- ▶ *Tipos:*
  - ① Externalidades biológicas (ej. vacunación, desparasitación).
  - ② Interacción social (redes de empleo, educación).
  - ③ Equilibrio general (mercados laborales, precios).
- ▶ *Mitigación:*
  - ★ Aleatorización por clusters o uso de buffers geográficos.
  - ★ Medición explícita de la exposición indirecta.

# Amenazas a la validez interna (II)

## ③ Efecto John Henry:

- ▶ El grupo de control se esfuerza más al sentirse en competencia con los tratados.
- ▶ Ej.: estudiantes en control aumentan su rendimiento para “demostrar” que no son peores.
- ▶ *Consecuencia*: sesgo hacia abajo en la estimación del efecto.
- ▶ *Mitigación*: anonimizar la asignación cuando sea posible; usar clusters que reduzcan competencia directa.

# Amenazas a la validez interna (II)

## ③ Efecto John Henry:

- ▶ El grupo de control se esfuerza más al sentirse en competencia con los tratados.
- ▶ Ej.: estudiantes en control aumentan su rendimiento para “demostrar” que no son peores.
- ▶ *Consecuencia*: sesgo hacia abajo en la estimación del efecto.
- ▶ *Mitigación*: anonimizar la asignación cuando sea posible; usar clusters que reduzcan competencia directa.

## ④ Efecto Hawthorne:

- ▶ Los individuos cambian su comportamiento solo por saberse observados.
- ▶ Ej.: trabajadores aumentan productividad durante el experimento, independientemente del tratamiento.
- ▶ *Consecuencia*: los resultados no reflejan un escenario “normal”, reduciendo validez externa e interna.
- ▶ *Mitigación*: estandarizar la observación para tratados y controles; usar períodos de adaptación (“run-in”).

# Otras consideraciones en los RCTs (I)

- **Noncompliance (incumplimiento):**

- ▶ No todos los asignados reciben el tratamiento (no-take-up, crossovers).
- ▶ En este caso, la estimación principal es el **ITT** (Intention-to-Treat).
- ▶ Importante discutir las implicaciones: ¿quiénes cumplen y quiénes no?, ¿qué significa para la validez de la intervención?



# Otras consideraciones en los RCTs (I)

- **Noncompliance (incumplimiento):**

- ▶ No todos los asignados reciben el tratamiento (no-take-up, crossovers).
- ▶ En este caso, la estimación principal es el **ITT** (Intention-to-Treat).
- ▶ Importante discutir las implicaciones: ¿quiénes cumplen y quiénes no?, ¿qué significa para la validez de la intervención?

- **Heterogeneidad ex ante:**

- ▶ Incluso con aleatorización, puede haber interés en estimar efectos por subgrupos definidos antes del tratamiento (ej.: hombres vs. mujeres, hogares rurales vs. urbanos).
- ▶ Relevante para la política pública: permite ver **para quién funciona** y si hay equidad en los efectos.
- ▶ Recomendación: definir subgrupos de antemano y evitar análisis ad hoc.

# Otras consideraciones en los RCTs (II)

- **Múltiples outcomes:**

- ▶ Cuando se miden muchos resultados, aumenta el riesgo de encontrar efectos espurios (*data mining*).
- ▶ **Buenas prácticas:**
  - ① Pre-registro del plan de análisis.
  - ② Construcción de índices compuestos (ej.: educación, salud).
  - ③ Ajustes por multiplicidad (Bonferroni, FDR).

# Otras consideraciones en los RCTs (II)

- **Múltiples outcomes:**

- ▶ Cuando se miden muchos resultados, aumenta el riesgo de encontrar efectos espurios (*data mining*).
- ▶ **Buenas prácticas:**
  - ① Pre-registro del plan de análisis.
  - ② Construcción de índices compuestos (ej.: educación, salud).
  - ③ Ajustes por multiplicidad (Bonferroni, FDR).

- **Validez externa:**

- ▶ La muestra del RCT puede ser muy específica o poco representativa (ej.: ONGs ejemplares, contextos piloto).
- ▶ La implementación a pequeña escala puede diferir de lo que ocurriría en políticas nacionales (problemas de **scaling up**).
- ▶ Recomendación: siempre discutir límites de generalización y complementar con evidencia observacional.

# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- Diseño Experimental
- Estimación del Efecto
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Este Paper

**Pregunta central:** ¿Pueden los incentivos financieros ligados a asistencia, monitoreada con tecnología, reducir el ausentismo docente y mejorar el aprendizaje?

## Contexto:

- En India, cerca del 24% de los docentes estaban ausentes en horas de clase.
- Escuelas no formales en Rajasthan, operadas por la ONG *Seva Mandir*.
- Docentes: *para-teachers*, con salarios bajos, contratos flexibles y alta rotación.

## Contribución del paper:

- Ensayo controlado aleatorizado (RCT) de 30 meses (2003–2006).
- Uso de cámaras digitales con sello de hora como herramienta de monitoreo.
- Estimación experimental + modelo estructural dinámico de oferta laboral.

# Diseño Experimental

## Muestra y asignación:

- 120 escuelas no formales seleccionadas: 60 tratamiento, 60 control.
- Aleatorización balanceada en características iniciales (asistencia y desempeño estudiantil).

## Tratamiento:

- Docentes con cámaras digitales: fotos al inicio y fin de la jornada, con sello de fecha/hora.
- Pago no lineal: salario base + bono (Rs. 50) por cada día trabajado por encima de 20 al mes.
- Multas por inasistencia o falsificación de registros.

## Grupo control:

- Salario fijo de Rs. 1,000/mes, sin monitoreo adicional.

**Periodo de estudio:** 2003–2006, con visitas mensuales no anunciadas a todas las escuelas.

## Fuentes principales:

- **Asistencia docente:**

- ▶ Visitas aleatorias mensuales (tratamiento y control).
- ▶ Fotos con sello de hora (solo en tratamiento).

- **Actividad en aula:** uso de tablero, interacción con estudiantes, número de niños presentes.

- **Resultados de aprendizaje:**

- ▶ Prueba inicial (línea base).
- ▶ Mid-test (abril 2004).
- ▶ Post-test (septiembre 2004).
- ▶ Resultados en Hindi y Matemáticas, medidos en desviaciones estándar.

- **Verificación de integridad:**

- ▶ Balance inicial en asistencia y puntajes entre grupos.
- ▶ Seguimiento completo durante los 30 meses.

# Resultados

## Impactos principales:

### ● Asistencia docente:

- ▶ Ausentismo cayó de 42% a 21% en el grupo tratamiento.
- ▶ Efecto persistente incluso 4 años después del inicio.

### ● Resultados de aprendizaje:

- ▶ +0.17 desviaciones estándar en puntajes de los estudiantes.
- ▶ +10 pp de probabilidad de transición a escuelas formales.

### ● Elasticidad de oferta laboral:

- ▶ Modelo estructural: 0.20–0.30 respecto al bono.
- ▶ Evidencia de fuerte respuesta de docentes a incentivos monetarios.

### ● Costo-efectividad:

- ▶ Reducción de costos unitarios de educación.
- ▶ Incentivos simples + monitoreo confiable → gran impacto en capital humano.



# Roadmap

## 1 Experimentos Aleatorios Controlados (RCTs)

- Intuición
- Diseño Experimental
- Estimación del Efecto
- Limitaciones y/o Amenazas a la Validez

## 2 Incentives Work: Getting Teachers to Come to School

- Este Paper
- Diseño Experimental
- Datos
- Resultados

## 3 Hands-on: replicar el RCT en R

# Replication Package

**Repositorio original:** Disponible en ICPSR: [doi.org/10.3886/E112523V1](https://doi.org/10.3886/E112523V1)

**Repositorio en R (GitHub):** Descarga directa del paquete de replicación (código y datos en R): [replication\\_package.zip](#)

**Instrucciones:** Antes de salir al break:

- Descarguen y descompriman el paquete.
- Ejecuten el script inicial para cargar librerías:

**Script en R:**

- *require(pacman)*
- *p\_load(tidyverse, rio, fixest, broom)*

Cuando regresemos de la pausa, las librerías ya estarán cargadas y listas para usar.