

Estimación

Diseño e implementación de experimentos en ciencias sociales
Departamento de Economía (UdelaR)

Experimentos aleatorizados de dos ramas

Efecto medio del tratamiento en la muestra (SATE)

- ▶ N : muestra
- ▶ Aleatorización completa de m en N
- ▶ $(Z_i = 1)$: tratados
- ▶ $(Z_i = 0)$: controles
- ▶ $SATE = E_{i \in N}[Y_i(1) - Y_i(0)]$
- ▶ $Y_i = Y_i(1) \times Z_1 + Y_i(0) \times (1 - Z_1)$

Efecto medio del tratamiento en la muestra (SATE)

- ▶ N : muestra
- ▶ Aleatorización completa de m en N
- ▶ $(Z_i = 1)$: tratados
- ▶ $(Z_i = 0)$: controles
- ▶ $SATE = E_{i \in N}[Y_i(1) - Y_i(0)]$
- ▶ $Y_i = Y_i(1) \times Z_1 + Y_i(0) \times (1 - Z_1)$

$$\hat{\tau} = \frac{\sum_1^m Y_i}{m} - \frac{\sum_{m+1}^N Y_i}{N - m}$$

Efecto medio del tratamiento en la muestra (SATE)

- ▶ N : muestra
- ▶ Aleatorización completa de m en N
- ▶ $(Z_i = 1)$: tratados
- ▶ $(Z_i = 0)$: controles
- ▶ $SATE = E_{i \in N}[Y_i(1) - Y_i(0)]$
- ▶ $Y_i = Y_i(1) \times Z_1 + Y_i(0) \times (1 - Z_1)$

$$\hat{\tau} = \frac{\sum_1^m Y_i}{m} - \frac{\sum_{m+1}^N Y_i}{N - m}$$

$$\widehat{se(\tau)} = \sqrt{\frac{V(Y_i(1))}{m} + \frac{V(Y_i(0))}{N - m}}$$

Ajuste por covariables

- ▶ Freedman (2008) critica la práctica de utilizar covariables en una regresión MCO para estimar el SATE con datos experimentales.

Ajuste por covariables

- ▶ Freedman (2008) critica la práctica de utilizar covariables en una regresión MCO para estimar el SATE con datos experimentales.
- ▶ Mientras que el estimador de diferencia de medias es insesgado para el efecto medio del tratamiento, el estimador MCO ajustado por covariables muestra un pequeño sesgo muestral (a veces denominado “sesgo Freedman”) que disminuye rápidamente a medida que aumenta el tamaño de la muestra.

Ajuste por covariables

- ▶ Freedman (2008) critica la práctica de utilizar covariables en una regresión MCO para estimar el SATE con datos experimentales.
- ▶ Mientras que el estimador de diferencia de medias es insesgado para el efecto medio del tratamiento, el estimador MCO ajustado por covariables muestra un pequeño sesgo muestral (a veces denominado “sesgo Freedman”) que disminuye rápidamente a medida que aumenta el tamaño de la muestra.
- ▶ Más preocupante es la crítica de que el ajuste de covariables puede incluso perjudicar la precisión.

Estimador de Lin

- ▶ Lin (2013) ofrece un estimador alternativo que garantiza ser al menos tan preciso como el estimador no ajustado.

Estimador de Lin

- ▶ Lin (2013) ofrece un estimador alternativo que garantiza ser al menos tan preciso como el estimador no ajustado.
- ▶ El problema se produce cuando:
 - ▶ la correlación de las covariables con el resultado es muy diferente en la condición de tratamiento que en la condición de control
 - ▶ los diseños están fuertemente desequilibrados en el sentido de tener grandes proporciones de unidades tratadas o no tratadas.

Estimador de Lin

- ▶ Lin (2013) ofrece un estimador alternativo que garantiza ser al menos tan preciso como el estimador no ajustado.
- ▶ El problema se produce cuando:
 - ▶ la correlación de las covariables con el resultado es muy diferente en la condición de tratamiento que en la condición de control
 - ▶ los diseños están fuertemente desequilibrados en el sentido de tener grandes proporciones de unidades tratadas o no tratadas.
- ▶ El estimador de Lin ajusta por covariables en cada brazo del experimento por separado.

Estimador de Lin

- ▶ Lin (2013) ofrece un estimador alternativo que garantiza ser al menos tan preciso como el estimador no ajustado.
- ▶ El problema se produce cuando:
 - ▶ la correlación de las covariables con el resultado es muy diferente en la condición de tratamiento que en la condición de control
 - ▶ los diseños están fuertemente desequilibrados en el sentido de tener grandes proporciones de unidades tratadas o no tratadas.
- ▶ El estimador de Lin ajusta por covariables en cada brazo del experimento por separado.
- ▶ Equivale a la inclusión de un conjunto completo de interacciones del tratamiento con cada covariable.

Estimador de Lin

- ▶ Lin (2013) ofrece un estimador alternativo que garantiza ser al menos tan preciso como el estimador no ajustado.
- ▶ El problema se produce cuando:
 - ▶ la correlación de las covariables con el resultado es muy diferente en la condición de tratamiento que en la condición de control
 - ▶ los diseños están fuertemente desequilibrados en el sentido de tener grandes proporciones de unidades tratadas o no tratadas.
- ▶ El estimador de Lin ajusta por covariables en cada brazo del experimento por separado.
- ▶ Equivale a la inclusión de un conjunto completo de interacciones del tratamiento con cada covariable.
- ▶ El preprocesamiento de los datos mediante quitarle la media a las covariables convierte el coeficiente del regresor del tratamiento en una estimación del SATE global.

Estimador de Lin

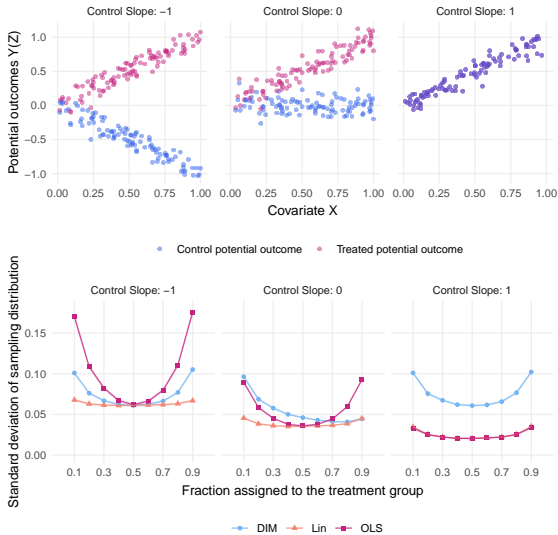
- ▶ Lin (2013) ofrece un estimador alternativo que garantiza ser al menos tan preciso como el estimador no ajustado.
- ▶ El problema se produce cuando:
 - ▶ la correlación de las covariables con el resultado es muy diferente en la condición de tratamiento que en la condición de control
 - ▶ los diseños están fuertemente desequilibrados en el sentido de tener grandes proporciones de unidades tratadas o no tratadas.
- ▶ El estimador de Lin ajusta por covariables en cada brazo del experimento por separado.
- ▶ Equivale a la inclusión de un conjunto completo de interacciones del tratamiento con cada covariable.
- ▶ El preprocesamiento de los datos mediante quitarle la media a las covariables convierte el coeficiente del regresor del tratamiento en una estimación del SATE global.
- ▶ El estimador “lm_lin” del paquete “estimatr” implementa este preprocesamiento en un solo paso.

Estimador de Lin

$$Y_i = \tau Z_i + \beta X_i + \epsilon_i$$

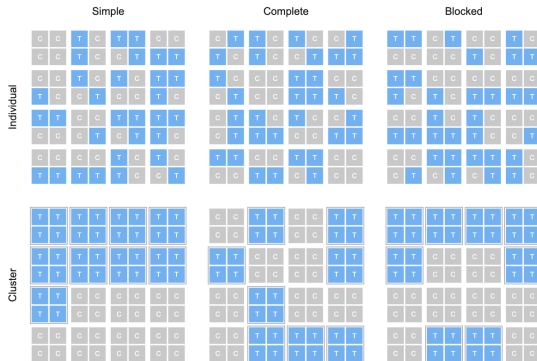
$$Y_i = \tau Z_i + \beta X_i^c + \gamma X_i^c Z_i + \epsilon_i$$

Estimador de Lin (vs OLS y DIM)

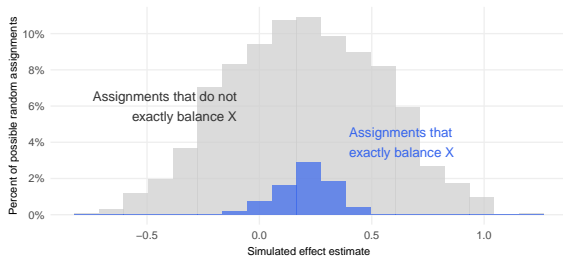


Experiments blocked

Estrategia



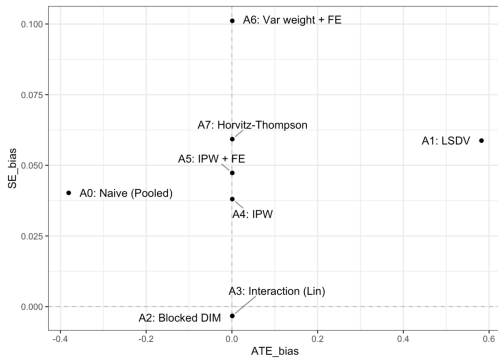
Asignaciones



Supone un conjunto de bloques j

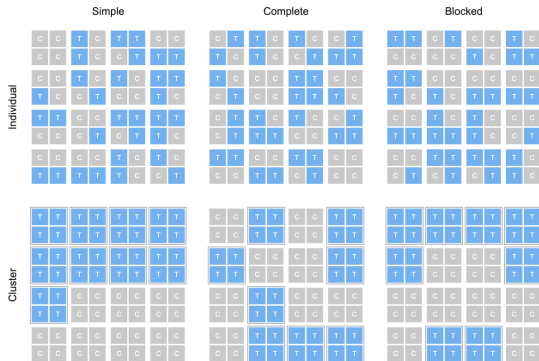
$$\hat{\tau} = \sum_{j=1}^J \frac{N_j}{N} \hat{\tau}_j$$

Problema de controlar por efectos fijos (bloque)



Experimentos de asignación por conglomerados

Estrategia



$$ICC = \frac{\sigma_{between}^2}{\sigma_{between}^2 + \sigma_{within}^2}$$

