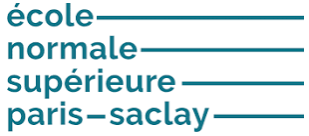


Public Policy Evaluation
Lecture 2: Difference-in-differences

Yao Thibaut Kpegli

Licence 3 in Economics, ENS Paris-Saclay
2023–2024

March 2024

Outline

Context

- Unobserved heterogeneity: unobserved characteristics that explain both participation in the program and the outcome variable
- Heterogeneity is fixed and additive
- Repeated observations are available on both treated and untreated

Context

Example: Treatment aimed at promoting children's growth

- First method: comparison after treatment between treated children and untreated children.
- neglects the possibility that differences may exist between children in the absence of treatment

Context

Example: treatment aimed at promoting children's growth

- Second method: comparison of treated children before and after the treatment.
- neglects the possibility that the outcome variable (in this case, height) may vary over time independently of treatment.

Context

- treated vs untreated and before vs after provide biased estimate of the impact
- DiD: mix the two methods

Context

Common trend assumption: in the absence of treatment, the difference between the two groups would be constant ("fixed") over time

Context

Two dates: t=0 and t=1

- At t=0, the program does not exist.
- At t=1, the program is in place and has produced its effects.

Potential outcomes

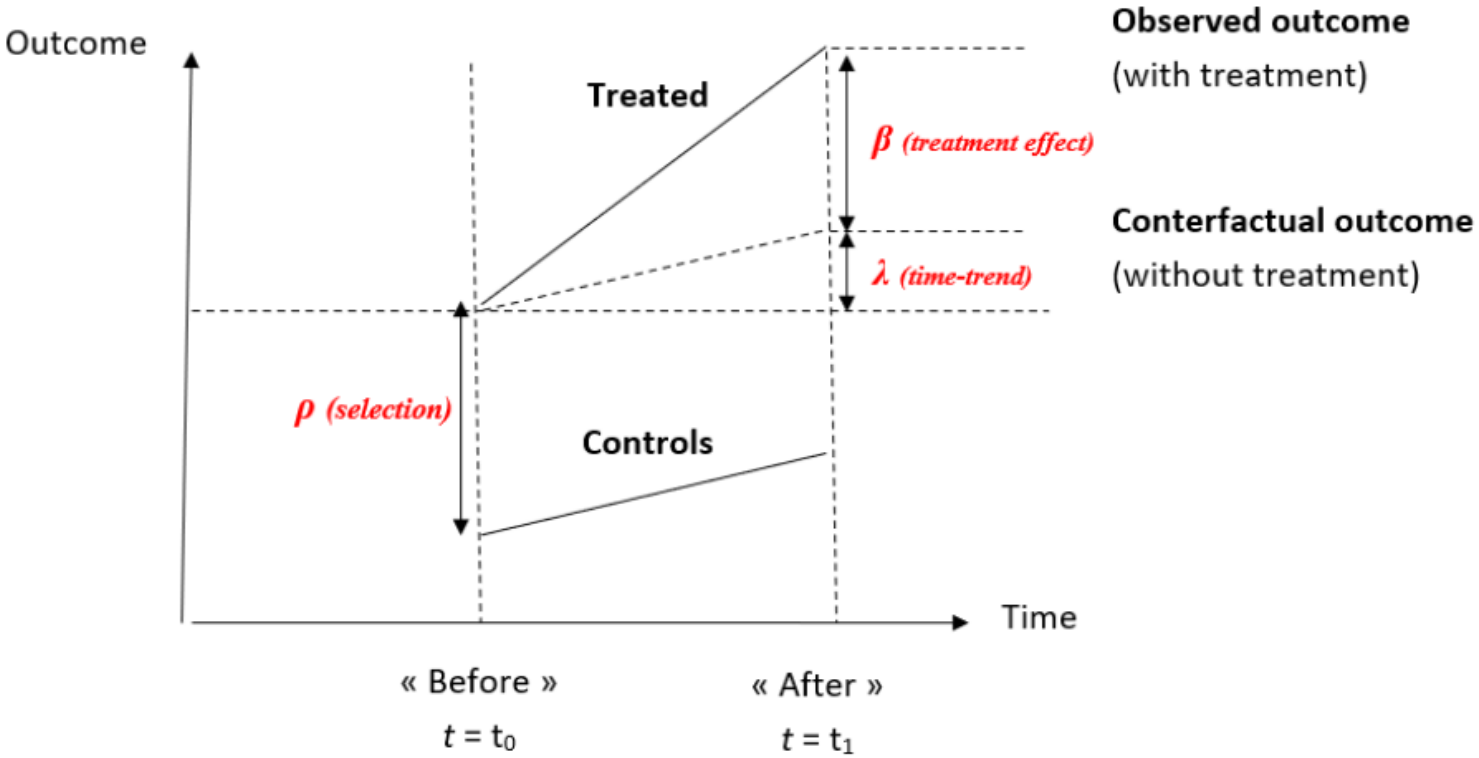
- Y_{1it}: value of the outcome variable for treated individual i at time t
- Y_{0it}: value of the outcome variable for untreated individual i at time t

Treatment Effect

$$DiD = E[Y_{1i1} - Y_{1i0}|T_i = 1] - E[Y_{0i1} - Y_{0i0}|T_i = 0]$$

Outline

DiD in regression



- ρ: selection bias
- λ: trend biais

DiD in regression

	Untreated	Treated
Before	α	$\alpha + \rho$
After	$\alpha + \lambda$	$\alpha + \lambda + \rho + \beta$

- DiD is a comparison between 4 cell-levels means
- Only one cell is treated: Treatment $\times$ Post-treatment

DiD in regression

$$\boldsymbol{Y}_{it} = \alpha + \lambda t + \rho \boldsymbol{T}_i + \beta \boldsymbol{T}_i \times \boldsymbol{t} + e_{it}$$

- $E[Y_{i0}|T_i = 0] = \alpha$
- $E[Y_{i1}|T_i = 0] = \alpha + \lambda$
- $E[Y_{i0}|T_i = 1] = \alpha + \rho$
- $E[Y_{i1}|T_i = 1] = \alpha + \lambda + \rho + \beta$

$$\begin{array}{lcccc} \text{DiD} & = & \Big(E[Y_{i1}|T_i = 1] - E[Y_{i0}|T_i = 1]\Big) & - & \Big(E[Y_{i1}|T_i = 0] - E[Y_{i0}|T_i = 0]\Big) \\ & = & (\alpha + \lambda + \rho + \beta - \alpha - \rho) & - & (\alpha + \lambda - \alpha) \\ & = & (\lambda + \beta) & - & (\lambda) \\ & = & & \beta & \end{array}$$

Controls

- Increase precision
- Common trend assumption may hold after conditioning on controls
- With interaction, it allows to estimate heterogeneous treatment effects

More than two dates: panel fixed effects regression

$$\boldsymbol{Y}_{it} = \alpha + \mu_i + \lambda_t + \beta \boldsymbol{T}_i + \boldsymbol{X}_{it}\gamma + e_{it}$$

- μ_i: individual fixed effect
- λ_t: time fixed effects

Outline

Application

The impact of building secondary schools on the primary education level of children - ecampus

Thank you for your attention !