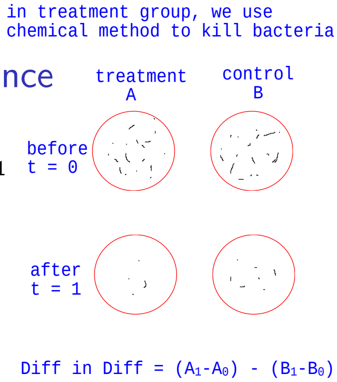
Difference in Difference

Chungsang Tom Lam[1]

[1]Department of Economics
Clemson University

April 9, 2020



Idea

- Consider a situation we have the outcome variable Y_{it} for each individual $i = 1, 2, 3....N$ and $t = 1, 2$.
- There is a policy change or a treatment (D) in period 2 which affects only some individuals.
- The difference in difference estimator is the sample analog of:

$$\begin{gathered}(E[Y_{i2}|D=1] - E[Y_{i1}|D=1]) \\ -(E[Y_{i2}|D=0] - E[Y_{i1}|D=0])\end{gathered}$$

- The $D = 0$ individuals are in the control group and only the Y in the treatment group in period 2 is affected by the treatment.

Regression Model

y	i	time	D	P	D*P
	1	1	1	0	0
	1	2	1	0	0
	1	3	1	1	1
	1	4	1	1	1
	2	1	1	0	0
	2	2	1	0	0
	2	3	1	1	1
	2	4	1	1	1
	3	1	0	0	0
	3	2	0	0	0
	5	3	0	1	0
	3	4	0	1	0
	4	1	0	0	0
	4	2	0	0	0
	4	3	0	1	0
	4	4	0	1	0

$$Y_{it} = \alpha + \theta D_i + \delta P_t + \gamma D_i * P_t + \epsilon_{it}$$

where

$$P_t = \begin{cases} 1 & \text{if } t \geq t^* \\ 0 & \text{o.w} \end{cases}$$

$$P_t = I(t \geq t^*)$$

How this works?

D = 1 = P

$$E[Y_{i2}|D = 1] = \alpha + \theta + \delta + \gamma$$

D = 1 P = 0

$$E[Y_{i1}|D = 1] = \alpha + \theta$$

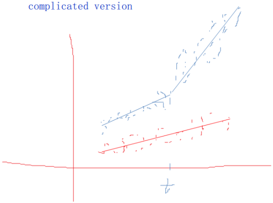
D = 0 P = 1

$$E[Y_{i2}|D = 0] = \alpha + \delta$$

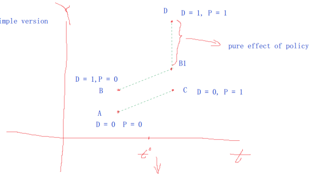
D = 0 P = 0

$$E[Y_{i1}|D = 0] = \alpha$$

complicated version



simple version



after this time point t, policy take effect.

(P = 1)

start at point A, after policy take effect(but A not in treatment group), A moves to C, due to non-policy effect
so, for point B, it should move to B1, however, it move to D
the distance between B1 and D is the pure effect of the policy

Regression Model

- Hence,

$$
\begin{aligned}
&(E[Y_{i2}|D=1]-E[Y_{i1}|D=1])-(E[Y_{i2}|D=0]-E[Y_{i1}|D=0])\\
&=(\alpha+\theta+\delta+\gamma)-(\alpha+\theta)-((\alpha+\delta)-\alpha)\\
&=\gamma
\end{aligned}
$$

- Therefore the coefficient estimator of γ gives you the Difference in difference estimator.
- In this approach we can add other regressors to the regression

$$Y_{it}=\alpha+\beta X_{it}+\theta D_i+\delta P_t+\gamma D_i * P_t+\epsilon_{it} \tag{3}$$

- Time trend or Time fixed effects?