TA Session 7

Bruno Amat

November 6, 2024

OUTLINE

Difference-in-Difference

The Evaluation Problem

In this research design we want to assess the impact of some policy under a random group of agents. Therefore, we need to define the following aspects

- Treatment and Control Groups
- Treatment

Treatment and Control Groups

To guarantee that our estimates is consistent we need that the treatment group and the control group are random. That is, the fact that one individual is treated is not related to the treatment itself.

We have different ways to address this selection problem

- RCT
- Natural Experiment
- Matching Methods
- Instrumental Methods
- Discontinuity Methods

TREATMENT

Suppose I assign random students to have access to their notes during the exam.

Therefore, some students have access to their notes and others not.

Hence, I can quantify how the access to their notes impact on their grade.

- **Control Group**: Students that does not have access to their notes
- **Treatment Group**: Students that have access to their notes

Treatment

However, the treatment might have homogeneous or heterogeneous response.

- **Homogeneous:** The treatment group have access to their notes since the beginning of the exam
- **Heterogeneous:** Some students in the treatment group have access to their notes since the beginning of the Exam, others only after 1 hour of the exam
- **Heterogeneous:** Some students in the treatment group have access to their notes the others have access to theirs and the TA's notes.

General Setup

Let d denote the treatment indicator. Moreover, let y_{it}^1 and y_{it}^0 be the outcome for the treated and the non-treated groups respectively,

$$y_{it}^1 = \beta + \alpha_i + u_{it}$$
$$y_{it}^0 = \beta + u_{it}$$

Hence, the observable outcome is characterized by:

$$y_{it} = d_{it} y_{it}^1 + (1 - d_{it}) y_{it}^0$$

Common Treatment Parameters

- **Average Treatment Effect:** Outcome if the individuals were assigned random to treatment. $\alpha^{ATE} = E[\alpha]$
- **Average Treatment on the Treated:** Average impact for the individuals who were assigned to treatment. $\alpha^{ATT} = E[\alpha|d_{it} = 1]$
- **Average Treatment on the Non Treated:** Average impact for the individuals who were not assigned to treatment. $\alpha^{ATNT} = E[\alpha|d_{it} = 0]$

The Social Experiment Approach: Randomization

Let's head back to our example of the access to the notes during the exam. Note that in that case I am defining the treatment group randomly independent of the grade. Hence, we satisfy two hypothesis

$$E[u_i|d_{it} = 1] = E[u_i|d_{it} = 0] = E[u_i]$$
$$E[\alpha_i|d_{it} = 1] = E[\alpha_i|d_{it} = 0] = E[\alpha_i]$$

Hence, the OLS estimator will be the ATE.

Difference in Difference

To perform the Diff-in-Diff we need 4 groups of Data

- Before Treatment - Treatment Group
- Before Treatment - Control Group
- After Treatment - Treatment Group
- After Treatment - Control Group

Using this information we can quantify the difference between the control group and the treatment group after the treatment was imposed.

Difference in Difference

Hence, we define our treatment variable as:

$$d_{it} \begin{cases} 1 \text{ if } t > t^* \text{ and } i \text{ in the treatment group} \\ 0 \text{ otherwise} \end{cases}$$

where t^* is the period which the treatment occurred

Difference in Difference

Hence, we estimate the following method

$$y_{it} = \beta + \alpha_i d_{it} + \phi_i + \theta_t + \epsilon_{it}$$

where ϕ_i is the individual fixed effect and θ_t is the time fixed effect

Difference in Difference

We can compute the Diff-in-Diff estimator as

$$\alpha^{DID} = [\bar{y}_{t1}^{1} - \bar{y}y_{t0}^{1}] - [\bar{y}_{t1}^{0} - \bar{y}_{t0}^{0}]$$

We need the Parallel Trends Assumption

MATCHING

The matching constructs the correct sample counterpart for the missing information on the treated outcomes.

Note that to perform the matching we need another variable X, to perform the matching and not use the outcome y

Matching Algorithms:

- **Nearest-Neighbor Matching**: assigns weight 1 to the closest non-treated observation and 0 to all others
- **Kernel Matching**: Defines a neighborhood for each treated observation and constructs a counterfactual based on all control observations

Matching - Algorithms

Nearest Neighbor:

$$C(P_i) = \min_j |P_i - P_j|$$

Kernel:

$$\tilde{w}_{ij} = \frac{K\left(\frac{P_j - P_i}{h}\right)}{\sum_{k \in C} K\left(\frac{P_k - P_i}{h}\right)}$$

Matching Diff-in-Diff

$$\hat{\alpha}^{MDID} = \sum_{i \in T} \left[(y_{it1} - y_{it0}) - \sum_{j \in C} w_{ij}((y_{jt1} - y_{jt0}) \right] w_i$$