

Lecture 6

Difference in difference estimation

1. The setting

- Panel data or repeated cross sections.
- Idea: compare before / after change between the treatment and control groups.
- A constant treatment.
- Assume 2 time periods.
- Treatment takes place in period 2, if at all:

	period 1	period 2
Treatment group	0	1
Control group	0	0

2. The method in a nutshell

- Calculate the mean outcome
 - i. prior to treatment for the treated.
 - ii. prior to treatment for the controls.
 - iii. post treatment for the treated.
 - iv. post treatment for the controls.

NOTE: the controls and treated prior and post treatment may be different individuals *as long as eligibility can be determined*.

- Calculate the within-group differences.
- Calculate the difference of these.

3. Opening the black box

- A1: selection into treatment is independent of ε_{it} .
- Example of violation: Ashenfelter's dip.
- A2: (cross section data): the composition of treatment and control groups is stable over time.
- Example: heterogenous variation in labor market conditions.
- NOTE: *no* exclusion restrictions needed (IV / Lecture 7).
- Under A1 and A2, identifies that ATET.
- A1 and A2 are **very strong** assumptions.

4. Remedies

1. Adjusting for differential trends: find a period where a similar (macro) trend has taken place, and use that information to control for the different time effects.
2. Randomization inference test: use simulated treatments to construct the distribution of treatment effects. Then test the estimated treatment effect against this distribution.