Chapter 7. Difference in Differences

Joan Llull

Quantitative Statistical Methods II
Barcelona School of Economics

Difference in differences setup

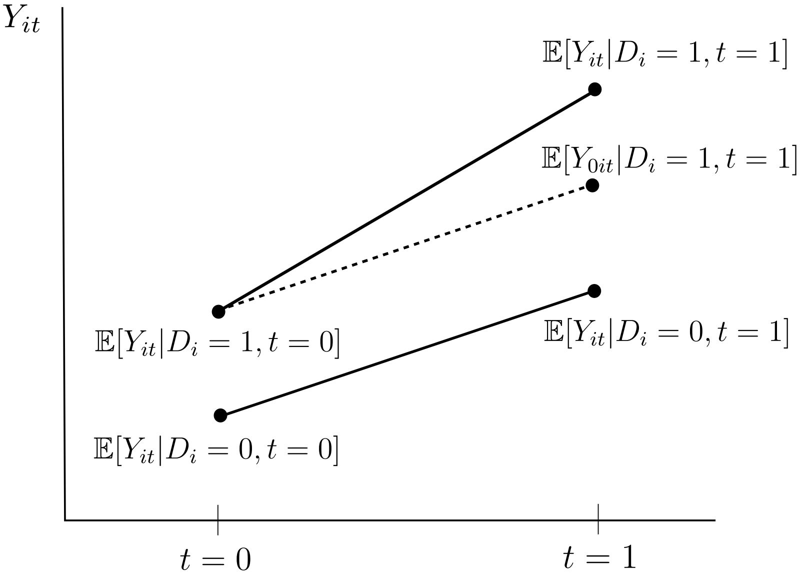
Randomized experiment $\Rightarrow$ simple comparison of the mean outcome in treatment and control groups ("difference" estimator), unbiased and consistent estimate of the ATE (Chapter 2).

The approach in this chapter: like in matching or sharp RD, adjusts somehow to **compensate confounders**.

Linking Chapter 6 to treatment effects approaches, we propose an alternative method to eliminate confounders that are **fixed over time** (like a fixed effect), using repeated observations.

Key assumption: common trend.

Difference in differences



Formal discussion

Figure suggests to use **trend observed for untreated** to predict the **counterfactual trend** for treated individuals in the absence of treatment:

$$
\begin{aligned}
\mathbb{E}[Y_{0it}|D_i = 1, t = 1] = {} & \underbrace{\mathbb{E}[Y_{it}|D_i = 0, t = 1]}_{\text{level for controls at } t=1} \\
& + \underbrace{\{\mathbb{E}[Y_{it}|D_i = 1, t = 0] - \mathbb{E}[Y_{it}|D_i = 0, t = 0]\}}_{\text{difference in levels at } t=0 \text{ difference}}.
\end{aligned}
$$

Fundamental DD assumption: common trend:

$$\mathbb{E}[Y_{0i1} - Y_{0i0}|D_i = 1] = \mathbb{E}[Y_{0i1} - Y_{0i0}|D_i = 0].$$

Hence, the difference in differences coefficient (which is an average treatment effect on the treated) is:

$$
\begin{aligned}
\beta = {} & \{\mathbb{E}[Y_{it}|D_i = 1, t = 1] - \mathbb{E}[Y_{it}|D_i = 1, t = 0]\} \\
& - \{\mathbb{E}[Y_{it}|D_i = 0, t = 1] - \mathbb{E}[Y_{it}|D_i = 0, t = 0]\}.
\end{aligned}
$$

Diff-in-diff and regression

The difference in differences coefficient can be obtained as the β coefficient in the following **regression**:

$$Y_{it} = \beta_0 + \beta_D D_i + \beta_T T_{it} + \beta D_i T_{it} + U_{it},$$

where $T_{it} = 1$ if individual i is treatment period $t = 1$, and $T_{it} = 0$ otherwise.

With similar arguments as in previous chapters:

- $\beta_0 = \mathbb{E}[Y_{it}|D_i = 0, t = 0]$,
- $\beta_0 + \beta_D = \mathbb{E}[Y_{it}|D_i = 1, t = 0]$,
- $\beta_0 + \beta_T = \mathbb{E}[Y_{it}|D_i = 0, t = 1]$,
- and β is the difference in differences coefficient.

Diff-in-diff and regression

This regression model can be expanded in several ways:

- **Further periods**: In such case, T_{it} is not a time dummy but, instead, a dummy that equals one in the post-treatment period. One could additionally include time effects, but the interaction term should be with the "post" dummy only.
- **Controls**: the regression allows for controls, X_{it} (works as in regression vs matching).
- **Panel data**: there is actually no need for panel data. However, in the repeated cross-section context, the researcher needs to sustain the assumption that the sample composition does not vary over time, which is satisfied by construction with panel data (also individual fixed effects)
- **Placebo analysis**: a regression that simulates the difference in differences analysis but for a point in time or group of individuals that resemble the treatment period or group but that was actually not treated.

Triple differences

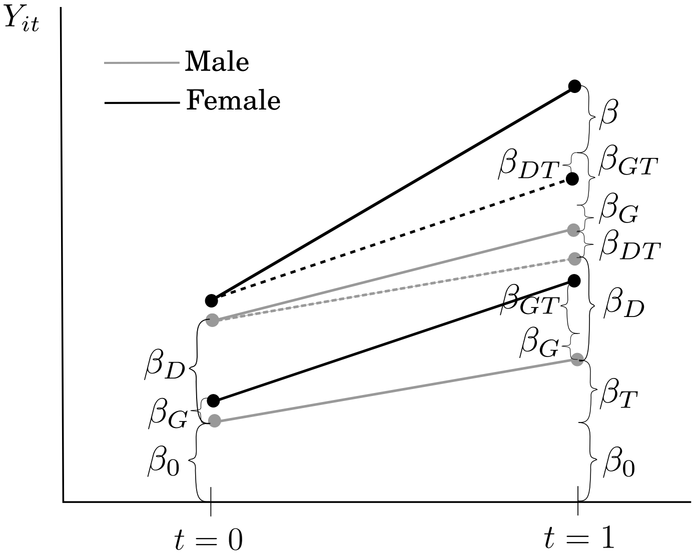
Triple difference: the difference in differences assumption does not hold, but the change in trends is assumed to be the same across sub-groups, some of which should be more affected than others.

Let G_i denote the (say sociodemographic) group to which individual i belongs. Then, the **triple-differences** model is:

$$\begin{aligned} Y_{it} = \beta_0 + \beta_D D_i + \beta_T T_{it} + \beta_G G_i + \beta_{GD} G_i D_i \\ + \beta_{GT} G_i T_{it} + \beta_{DT} D_i T_{it} + \beta G_i D_i T_{it} + U_{it}. \end{aligned}$$

Example: Maternity leave policies combined with a tax reform that affects young and old differently.

Difference in differences



Synthetic Control Methods

Synthetic control methods: use longitudinal data to build the weighted average of non-treated units that best reproduces the characteristics of the treated unit over time prior to the treatment.

Thus, we build an **artificial control** that has the best possible pre-trend possible, and then we compute the difference in differences estimate using such synthetic control group.