ECON220C Discussion Section 4

Staggered DID

Lapo Bini

Roadmap

1. Parallel Trend & No Anticipation
2. Staggered DID Framework
3. Exercise TWFE vs DID

Setup

- **Goal**: measure the impact of a policy intervention using differences-in-differences method.Assume there are only two periods, program participation only occurs between periods one and two.

- **Framework**: potential outcome. Remember from 220B $y_{it} = d_{it}y_{it}(1) + (1 - d_{it})y_{it}(0)$. Our target:

$$ATT \equiv E[y_{it}(1) - y_{it}(0)|d_{it} = 1]$$

- Before going full-mode on potential outcome + staggered did framework, let's introduce two key assumptions: **parallel trend** and **no anticipation effect**.

Let's Construct the Model

Model: $y_{it} = \alpha_i + \lambda_t + \beta d_{it} + u_{it} \quad \Longrightarrow$ Can we obtain this from Potential outcome notation?

1. $y_{it}(1) = E[y_{it}(1)] + u_{it}(1)$ and same for $y_{it}(0)$
2. Combine with $y_{it} = d_{it}y_{it}(1) + (1 - d_{it})y_{it}(0)$
3. We get our **individual** and **time** fixed effect model.

Identification β_{DID}

As before, the model is $y_{it} = \alpha_i + \lambda_t + \beta d_{it} + u_{it}$. The **identification strategy** for β is the following:

1. Take average $\forall i$ of y_{i1} and y_{i2} or first difference

2. Remove time and individual fixed effects using double-differencing

3. Parallel trends assumption

$\beta_{DID} = ATT$?

- First, let's rewrite $\beta_{DID} = E[\Delta y_{T2} - \Delta y_{C2}]$ in terms of **potential outcomes**:

- $\beta_{DID} - ATT =$

$\beta_{DID} = ATT$?

- **No Anticipation effect**: $E[y_{i1}(1)|d_{i2} = 1] = E[y_{i1}(0)|d_{i2} = 1]$, in plain english knowing that I would receive the treatment is not changing my behavior in the pre-treatment period.

- **New parallel trends** assumption:

$$E[y_{i2}(0) - y_{i1}(0)|d_{i2} = 1] = E[y_{i2}(0) - y_{i1}(0)|d_{i2} = 0]$$

What is telling us? Is the same as before?

Empirical Studies: DID Estimator and Dummies

Staggered ID - Setup

- **Setup**: multiple time periods, treatment could happen at any time but once treated, always treated. New notation: $y_{it}(g)$ where g is the period when unit i gets treated (if $t < g$ **not yet treated**), $y_{it}(\infty)$ means **never treated**.
- Example: 4 treatment groups and three time periods.

Staggered DID - Assumptions

1. **Parallel trends**:

$$E[y_{it}(\infty) - y_{it-1}(\infty)|G_i = g] = E[y_{jt}(\infty) - y_{jt-1}(\infty)|G_j = g']$$

   here $g, g' < \infty$ $\forall t$. If treatment hadn't happended, all adoption cohorts would have evolved in parallel at all periods. We are considering people in cohorts i and j, trated at g, g'.

2. **No anticipation effect** (same as before)

$$E[y_{it}(g)|G_i = g] = E[y_{it}(\infty)|G_i = g] \; \forall t < g$$

Staggered DID - Goal

- **ATT(g, t)**$= E[y_{it}(g) - y_{it}(\infty)|G_i = g]$ this is the object we aim to identify and estimate.

- **Claim**: given conditions above, ATT(g, t) is identified as (we will prove it for a simple case later)

$$ATT(g, t) = E[y_{it}(g) - y_{ig-1}(g)|G_i = g] - E[y_{it}(\infty) - y_{ig-1}(\infty)|G_i = \infty]$$

Exercise

Consider the **staggered DID** methodology for panel data in the potential outcomes framework. There are two periods $t = 1, 2$ and three potential treatment trajectories $\{d_1 = (1, 1); d_2 = (0, 1); d_3 = (0, 0)\}$ as listed in the table below:

d_i	t = 1	t=2
group 1	1	1
group 2	0	1
group 3	0	0

	t = 1	t=2
group 1	$\bar{Y}_{d_1,1} = 0.3$	$\bar{Y}_{d_1,2} = 1.2$
group 2	$\bar{Y}_{d_2,1} = 0$	$\bar{Y}_{d_2,2} = 0.6$
group 3	$\bar{Y}_{d_3,1} = 0.3$	$\bar{Y}_{d_3,2} = 0.3$

There are three potential outcomes $\{Y_{i1}(d_i), Y_{i2}(d_i)\}$ given the three different treatment status d_i. The sample is iid and each group has size $n_1 = n_2 = n_3 = 1/3$. The table on the right reports the average value of the outcome of interest for group d and time t.

Exercise - Q1

Let $ATT(d_2, t) = E[Y_{i2}(d_2) - Y_{i2}(d_3)|Di = d_2]$ be the average treatment effect on treated at time 2 for the group of individuals who chose d_2 as the treatment trajectory. Explain how you would identify $ATT(d_2, t)$.

Exercise - Q1

Answer

Exercise - Q2

Explain how would you estimate $ATT(d_2, t)$. Please first provide a formula for your proposed estimator and then give a numerical answer based on the information provided.

Exercise - Q2

Answer

Exercise - Q3

Suppose we estimate the following model $Y_{it}(D_i) = \alpha_i + \lambda_t + D_{it}\beta + e_i$ by the TWFE estimator which requires a two way demeaned version of the dummy variable of the treatment. Find $\tilde{D}_{it}$ and the TWFE estimate $\hat{\beta}_{TWFE}$. Please provide numerical answer.

Exercise - Q3

Answer

Exercise - Q4

Verify that $\hat{\beta}_{TWFE} = \hat{\beta}_{2,DID} - \frac{1}{2}\hat{\beta}_{1,DID}$ where $\hat{\beta}_{2,DID}$ is the estimate in Q2, and $\hat{\beta}_{1,DID} = (\bar{Y}_{d_1,2} - \bar{Y}_{d_1,1}) - (\bar{Y}_{d_3,2} - \bar{Y}_{d_3,1})$

Exercise - Q4

Answer