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: $y_{it} = \alpha_i + \lambda_t + \beta d_{it} + u_{it} \implies$ 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.

$$y_{it}(0) = E[y_{it}(0)] + u_{it}(0)$$
$$y_{it}(1) = E[y_{it}(1)] + u_{it}(1)$$
$$y_{it} = x_{it}\, y_{it}(1) + (1 - x_{it})\, y_{it}(0)$$

$$y_{it} = x_{it}\left(E[y_{it}(1)] + u_{it}(1)\right) + (1 - x_{it})\left(E[y_{it}(1)] + u_{it}(1)\right)$$
$$= \underbrace{E[y_{it}(0)]} + x_{it}\left(\underline{E[y_{it}(1) - y_{it}(0)]}\right) + u_{it}$$
$$= \alpha_i + \lambda_t + x_{it}\beta + u_{it}$$

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 (WITHIN TREADED (T) OR CONTROL GROUP (C))

   $\Delta y_{T2} = y_{T2} - y_{T1} = (\lambda_2 - \lambda_1) + x_{T2}\beta + \Delta u_{T2}$

   $\Delta y_{C1} = y_{C2} - y_{C1} = (\lambda_2 - \lambda_1) + \Delta u_{C2}$

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

   $\Delta y_{T2} - \Delta_{C2} = (\lambda_2 - \lambda_1) + x_{T2}\beta + \Delta u_{T2} - (\lambda_2 - \lambda_1) - \Delta u_{C2}$

   $\Delta y_{T2} - \Delta_{C2} = \beta + \underline{\Delta u_{T2} - \Delta_{C2}}$ WE WANT TO GET RID OF THIS

3. Parallel trends assumption

IF : $E[\Delta u_{T2} - \Delta u_{C2}] = 0$, THAT IS $E[u_{T2} - u_{T1}] = E[u_{C2} - u_{C1}]$

THEN $\beta_{DID} = E[\Delta y_{T2}] - E[\Delta y_{C2}] \Rightarrow \hat{\beta}_{DID} = \left(\frac{1}{N_T}\sum_{i \in T}(y_{i2} - y_{i1})\right) - \left(\frac{1}{N_C}\sum_{j \in C}(y_{j2} - y_{j1})\right)$

$\beta_{DID} = ATT$?

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

$$\beta_{DID} = E[\Delta y_{T2}] - E[\Delta y_{C2}] = E[y_{T2} - y_{T1}] - E[y_{C2} - y_{C1}]$$

$$= E[y_{i2}(1) - y_{i1}(1) \mid x_{i2} = 1] - E[y_{i2}(0) - y_{i1}(0) \mid x_{i2} = 0]$$

- $\beta_{DID} - \underline{ATT} =$

$$\left(E[y_{i2}(1) - y_{i1}(1) \mid x_{i2} = 1] - E[y_{i2}(0) - y_{i1}(0) \mid x_{i2} = 0] \right)$$

$$- E[y_{i2}(1) - y_{i2}(0) \mid x_{it} = 1]$$

$$= \Big(\underbrace{E[y_{i2}(0) \mid d_{it} = 1]}_{\text{NOT IDENTIFIED}} - E[y_{i1}(1) \mid x_{i2} = 1] \Big) - E[y_{i2}(0) - y_{i1}(0) \mid x_{i2} = 0]$$

$\beta_{DID} = ATT$?

LET'S USE THIS

1. • **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.

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

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

2. • **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?

BY 1 + 2 WE GET $\beta_{DID} - ATT = 0$

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

Empirical Studies: DID Estimator and Dummies

$\hat{\beta}_{DID} = \left(\frac{1}{N_T}\sum_{i\in T}(y_{i2} - y_{i1})\right) - \left(\frac{1}{N_C}\sum_{j\in C}(y_{j2} - y_{j1})\right)$ CAN BE OBTAINED BY RUNNING

$$y_{it} = \beta_0 + \mathbb{1}\{\text{TREATED}\}_{it} \cdot \beta_1 + \mathbb{1}\{t=2\}_{it} \cdot \beta_2 + \mathbb{1}\{\text{TREATED} \cap t=2\}_{it} \cdot \beta_3$$

$$\therefore \hat{\beta}_{DID} = \Big[\underbrace{(\hat{\beta}_3 + \hat{\beta}_2 + \hat{\beta}_1 + \hat{\beta}_0)}_{\bar{y}_{T2}} - \underbrace{(\hat{\beta}_0 + \hat{\beta}_1)}_{\bar{y}_{T1}}\Big] - \Big[\underbrace{(\hat{\beta}_2 + \hat{\beta}_0)}_{\bar{y}_{C2}} - \underbrace{\hat{\beta}_0}_{\bar{y}_{C1}}\Big]$$

$\longrightarrow$ EQUIVALENT TO TWFE ESTIMATOR

$$\Delta y_{i2} = \underbrace{\beta_0}_{\text{what is This}} + \underbrace{\Delta x_{i2}}_{}\beta + u_{it}$$

$\Delta x_{i2} = x_{i2} \rightarrow$ CLEAR?

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.

	TREATMENT DECISION			OBSERVED OUTCOME		
	$t=1$	$t=2$	$t=3$	$t=1$	$t=2$	$t=3$
GROUP 1	1	1	1	$y_{i1}(1)$	$y_{i2}(1)$	$y_{i3}(1)$
GROUP 2	0	1	1	$y_{i1}(2)$	$y_{i2}(2)$	$y_{i3}(2)$
GROUP 3	0	0	1	$y_{i1}(3)$	$y_{i2}(3)$	$y_{i3}(3)$
GROUP 4	0	0	0	$y_{i1}(\infty)$	$y_{i2}(\infty)$	$y_{i3}(\infty)$

$\Longrightarrow$ NEVER TREATED, CONTROL GROUP

Staggered DID - Assumptions

PEOPLE TAKING TREATMENT AT TIME g AND g'

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)

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

KNOWING THAT I GO INTO TREATMENT DOES NOT CHANGE MY BEHAVIOR BEFORE

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)

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

WHEN I TAKE TREATMENT

EFFECT OF TREATMENT TODAY WITH $t > g$

WHEN I WAS THE SAME AS CONTROL GROUP

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	0
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

$$ATT(d_2, t) = E[Y_{i2}(d_2) - \overbrace{Y_{i2}(d_3)}^{\text{NOT OBSERVED}} \mid d_i = D_2]$$

$$= E[Y_{i2}(d_2) \mid d_i = d_2] - E[Y_{i2}(d_3) \mid d_i = D_2]$$

$$= E[Y_{i2}(d_2) \mid d_i = d_2] - E[Y_{i2}(d_3) \mid d_i = D_2]$$

$$+ E[Y_{i2}(d_3) \mid d_i = d_3] - E[Y_{i2}(d_3) \mid d_i = d_3]$$

ASSUME PARALLEL TRENDS

$$E[Y_{i2}(d_3) - Y_{i1}(d_3) \mid d_i = d_2] = E[Y_{i2}(d_3) - Y_{i1}(d_3) \mid d_i = d_3]$$

EQUAL TO

$$E[Y_{i2}(d_3) \mid d_i = d_2] - E[Y_{i2}(d_3) \mid d_i = d_3] = E[Y_{i1}(d_3) \mid d_i = d_2] - E[Y_{i1}(d_3) \mid d_i = d_3]$$

Exercise - Q1

Answer

LHS OF SELECTION BIAS (WITH MINUS SIGN)

$$ATT(d_2,t) = E[Y_{i2}(d_2) \mid d_i = d_2] - E[Y_{i2}(d_3) \mid d_i = D_2]$$
$$+ E[Y_{i2}(d_3) \mid d_i = d_3] - E[Y_{i2}(d_3) \mid d_i = d_3]$$

$$= E[Y_{i2}(d_2) \mid d_i = d_2] \underline{- E[Y_{i1}(d_3) \mid d_i = d_2]} + E[Y_{i1}(d_3) \mid d_i = d_3]$$
$$- E[Y_{i2}(d_3) \mid d_i = d_3]$$

WE SUBSTITUTE RHS

NOW WE IMPOSE NO ANTICIPATION EFFECT

$$\underline{E[Y_{i1}(d_3) \mid d_i = d_2]} = E[Y_{i1}(d_2) \mid d_i = d_2]$$

AND WE SUBSTITUTE IN ATT TO GET

$$ATT(d_2,t) = \{E[Y_{i2}(d_2) \mid d_i = d_2] - E[Y_{i1}(d_2) \mid d_i = d_2]\} - \{E[Y_{i2}(d_3) - Y_{i1}(d_3) \mid d_i = d_3]\}$$

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

PUT AN HAT :

$$ATT(d_2, t) = \{E[Y_{i2}(d_2) \mid d_i = d_2] - E[Y_{i1}(d_2) \mid d_i = d_2]\} - \{E[Y_{i2}(d_3) - Y_{i1}(d_3) \mid d_i = d_3]\}$$

$$= \left\{ \frac{1}{N_T} \sum (y_{T2} - y_{T1}) \right\} - \left\{ \frac{1}{N_C} \sum (y_{C2} - y_{C1}) \right\}$$

$$= \{ \bar{Y}_{d_2,2} - \bar{Y}_{d_2,1} \} - \{ \bar{Y}_{d_3,2} - \bar{Y}_{d_3,1} \}$$

$$= (0.6 - 0.0) - (0.3 - 0.3)$$

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.

$\tilde{D}_{it} = D_{it} - \bar{D}_i - \bar{D}_t - \bar{D}$ => FIRST OF ALL : WHY ?

BY FWL THEOREM : given $y_{it} = x_{it}\beta + w_{it}\gamma + u_{it}$

IF INTERESTED IN BETA , THE FOLLOWING ARE EQUIVALENT :

(1) $y_{it} \sim x_{it} + w_{it}$

(2) $\tilde{y}_{it} \sim \tilde{x}_{it}$ RESIDUALIZED WRT w_{it}

(3) $y_{it} \sim \tilde{x}_{it}$ => FWL + ANGRIST (2009)

Exercise - Q3 OUR MODEL $y_{it} = \alpha_i + \lambda_t + D_{it}\beta + u_{it}$

Answer $\tilde{x}_{it}$ IS $\tilde{D}_{it}$ IN OUR CASE

$\tilde{\omega}_{it}$ ARE α_i AND λ_t. HOW DO WE REMOVE THEM? AUX. REG: $D_{it} = \alpha_i^D + \lambda_t^D + u_{it}^D$

1. $\bar{D}_i := \frac{1}{T}\sum_t (\alpha_i^D + \lambda_t^D + u_{it}^D) = \alpha_i^D + \bar{\lambda}_t^D + \frac{1}{T}\sum_t u_{it}^D$

2. $D_{it} - \bar{D}_i = (\alpha_i^D + \lambda_t^D + u_{it}^D) - (\alpha_i^D + \bar{\lambda}_t^D + \frac{1}{T}\sum_t u_{it}^D)$

$$= \lambda_t^D - \bar{\lambda}_t^D + u_{it}^D - \frac{1}{T}\sum u_{it}^D$$

3. $\bar{D}_t := \frac{1}{N}\sum_i (\alpha_i^D + \lambda_t^D + u_{it}^D) = \bar{\alpha}_i^D + \lambda_t^D + \frac{1}{N}\sum_i u_{it}^D$

4. $D_{it} - \bar{D}_i - \bar{D}_t = (\lambda_t^D - \bar{\lambda}_t^D + u_{it}^D - \frac{1}{T}\sum u_{it}^D) - (\bar{\alpha}_i^D + \lambda_t^D + \frac{1}{N}\sum_i u_{it}^D)$

$$= -\bar{\alpha}_i^D - \bar{\lambda}_t^D + u_{it}^D - \frac{1}{T}\sum_t u_{it}^D - \frac{1}{N}\sum_i u_{it}^D$$

Exercise - Q3

Answer

5. $$\bar{D} = \frac{1}{NT}\sum_i \sum_t (\alpha_i^D + \lambda_t^D + u_{it}^D)$$

$$= \left(\frac{1}{NT}\sum_i T\alpha_i^D\right) + \left(\frac{1}{NT}\sum_t N\lambda_t^D\right) + \frac{1}{NT}\sum_i\sum_t u_{it}^D$$

$$= \left(\frac{1}{N}\sum \alpha_i^D\right) + \left(\frac{1}{T}\sum \lambda_t^D\right) + \frac{1}{NT}\sum_i\sum_t u_{it}^D = \bar{\alpha}_i^D + \bar{\lambda}_t^D + \frac{1}{NT}\sum_i\sum_t u_{it}^d$$

6. $$D_{it} - \bar{D}_i - \bar{D}_t + \bar{D} = -\bar{\alpha}_i^D - \bar{\lambda}_t^D + u_{it}^D - \frac{1}{T}\sum_t u_{it}^D - \frac{1}{N}\sum_i u_{it}^D$$

$$+ \bar{\alpha}_i^D + \bar{\lambda}_t^D + \frac{1}{NT}\sum_i\sum_t u_{it}^d$$

$$= u_{it} - \frac{1}{N}\sum_i u_{it}^D - \frac{1}{T}\sum_t u_{it}^d + \frac{1}{NT}\sum\sum u_{it}^D := \tilde{D}_{it}$$

Exercise - Q3

Answer

$$\begin{array}{c|cc|l} & \bar{D}_{t=1} & \bar{D}_{t=2} & \\ \hline d_1 & 1 & 1 & 1 = \bar{D}_{i=1} \\ d_2 & 0 & 1 & 1/5 = \bar{D}_{i=2} \\ d_3 & 0 & 0 & 0 = \bar{D}_{i=3} \\ \hline & 1/3 & 2/3 & \bar{D} = 1/2 \end{array}$$

$$\tilde{D}_{it} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 1/2 & 1/2 \\ 0 & 0 \end{vmatrix} - \begin{vmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \\ 1/3 & 2/3 \end{vmatrix} + \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{vmatrix}$$

$$\tilde{D}_{it} = \begin{vmatrix} 1/6 & -1/2 \\ -1/3 & 1/3 \\ 1/6 & -1/6 \end{vmatrix}$$

Exercise - Q3

PROBLEM : WE DON'T HAVE INDIVIDUAL OBSERVATION BUT GROUP VALUES

Answer

$$\hat{\beta}_{TWFE} = \frac{\widehat{Cov}(Y_{it}, \tilde{D}_{it})}{\widehat{Var}(\tilde{D}_{it})} = \frac{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{2}\tilde{D}_{it}Y_{it}}{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{2}(\tilde{D}_{it}^{2})}$$

$$= \frac{\frac{1}{N}\sum_{t=1}^{2}\sum_{i=1}^{N}\tilde{D}_{it}Y_{it}}{\frac{1}{N}\sum_{t=1}^{2}\sum_{i=1}^{N}(\tilde{D}_{it})^{2}} = \frac{\sum_{g=1}^{3}\sum_{t=1}^{2}\frac{m_g}{m_g}\frac{1}{N}\sum_{i\in g}\tilde{D}_{it}Y_{it}}{\sum_{g=1}^{3}\sum_{t=1}^{2}\frac{m_g}{m_g}\frac{1}{m}\sum_{i\in g}(\tilde{D}_{it})^{2}}$$

$$= \frac{\sum_{g=1}^{3}\sum_{t=1}^{2}\frac{m_g}{N}\tilde{D}_{gt}\left(\frac{1}{m_g}\sum_{i\in g}Y_{it}\right)}{\sum_{g=1}^{3}\sum_{t=1}^{2}\frac{m_g}{m}\frac{1}{m_g}\sum_{i\in g}(\tilde{D}_{it})^{2}} = \frac{\sum_{g=1}^{3}\sum_{t=1}^{2}w_g\tilde{D}_{gt}\bar{Y}_{gt}}{\sum_{g=1}^{3}\sum_{t=1}^{2}w_g\frac{1}{m_g}m_g\tilde{D}_{gt}^{2}}$$

Exercise - Q3

Answer

$$\hat{\beta}_{TWFE} = \frac{\sum_{g=1}^{3} \sum_{t=1}^{2} \omega_g \tilde{D}_{gt} \bar{Y}_{gt}}{\boxed{\sum_{g=1}^{3} \sum_{t=1}^{2} \omega_g \tilde{D}_{gt}^2} := K}$$

$$= \frac{1}{K} \sum_{g=1}^{3} \sum_{t=1}^{2} \omega_g \tilde{D}_{gt} \bar{Y}_{gt}$$

$$= \sum_{g=1}^{3} \sum_{t=1}^{2} \frac{\omega_g \tilde{D}_{gt}}{K} \bar{Y}_{gt} = \sum_{g=1}^{3} \sum_{t=1}^{2} \alpha_{gt} \cdot \bar{Y}_{gt}$$

$$= 0.15$$

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 $\hat{\beta}_{1,DID} = (\bar{Y}_{d_1,2} - \bar{Y}_{d_1,1}) - (\bar{Y}_{d_3,2} - \bar{Y}_{d_3,1})$

$$\hat{\beta}_{1\,DID} = (1.2 - 0.3) - (0.3 - 0.3) = 0.9$$

$$\hat{\beta}_{2\,DID} - \frac{1}{2}\hat{\beta}_{1\,DID} = 0.6 - \frac{1}{2}0.9 = 0.15 \checkmark$$

Exercise - Q4

Answer