STAT6061/STAT5008 – Causal Inference

Part 4-2. Difference-in-Differences (DiD) Method

An-Shun Tai

[1]*Department of Statistics*
National Cheng Kung University

[2]*Institute of Statistics and Data Science*
National Tsing Hua University

Before-and-after with control design

- In the presence of unmeasured confounding, a unit is often considered its own best control, which makes before–after comparisons intuitively appealing.

- However, a before–after comparison alone is not sufficient for causal inference, since outcome changes may reflect underlying time trends rather than treatment effects. (Why?)

- Alternatively, a treatment–control comparison at a single point in time (cross-sectional comparison) can yield unbiased estimates if all systematic differences between groups are properly adjusted for, but it remains biased in the presence of unmeasured confounding

- **Can we combine both comparisons?**
  That's the central idea behind the Difference-in-Differences (DID) approach.

- Design structure: Units from two or more groups are observed across multiple time periods, with treatment applied to some groups in some periods (e.g., repeated cross-sections or panel data).

Example: Minimum Wage and Employment

(Card and Krueger, 1994)

Research Question:
Did raising the minimum wage affect employment in the fast-food industry?

Design:
- **Treatment group**: New Jersey (raised minimum wage in 1992)
- **Control group**: Pennsylvania (no minimum wage change)

Time periods:
- Pre-treatment: before April 1992
- Post-treatment: after April 1992

Outcome measured:
Employment levels in fast-food restaurants.

	Before (1991)	**After (1992)**
New Jersey (Treatment)	Employment measured $Y_{T_0}\|A=1$	Employment measured $Y_{T_1}\|A=1$
Pennsylvania (Control)	Employment measured $Y_{T_0}\|A=0$	Employment measured $Y_{T_1}\|A=0$

Example: Minimum Wage and Employment (Conti.)

(Card and Krueger, 1994)

Strategy 1: Cross-sectional comparison between different units

$$\mathbb{E}\left(Y_{T_1} \middle| A = 1\right) - \mathbb{E}\left(Y_{T_1} \middle| A = 0\right) = \mathbb{E}\left(Y_{T_1}(1) \middle| A = 1\right) - \mathbb{E}\left(Y_{T_1}(0) \middle| A = 0\right)$$

- Require exchangeability.
- It is severely biased when unmeasured confounding is present.

	Before (1991)	**After (1992)**
New Jersey (Treatment)	Employment measured $Y_{T_0} \vert A = 1$	Employment measured $Y_{T_1} \vert A = 1$
Pennsylvania (Control)	Employment measured $Y_{T_0} \vert A = 0$	Employment measured $Y_{T_1} \vert A = 0$

Example: Minimum Wage and Employment (Conti.)

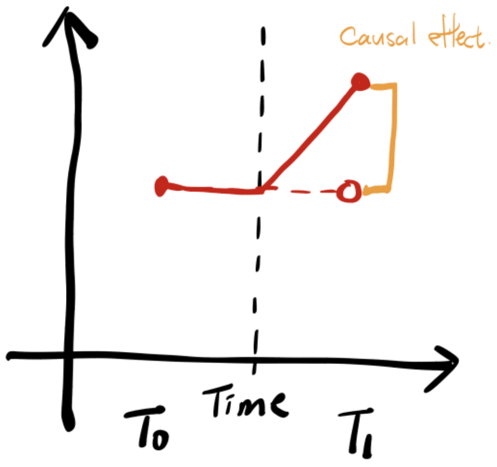
(Card and Krueger, 1994)

Strategy 2: Before-after comparison for the same unit:

$$\mathbb{E}\left(Y_{T_1} - Y_{T_0} | A = 1\right) = \mathbb{E}(Y_{T_1}(1) - Y_{T_0}(0) | A = 1)$$

- It is not the causal effect (ATT): $\mathbb{E}(Y_{T_1}(1) - Y_{T_1}(0) | A = 1)$

- An **extremely strong assumption** is required: $\mathbb{E}(Y_{T_1}(0) | A = 1) = \mathbb{E}(Y_{T_0}(0) | A = 1)$



1. Time Trend Bias (Secular Trend Bias)
Example: Measuring blood pressure before and after a health policy change, without accounting for general population trends or aging.

2. Regression to the Mean
Example: Selecting students with very low test scores for tutoring may show improvement post-tutoring even if the effect is due to natural score fluctuation.

3. History Effects
Example: A national campaign launched in the interim could affect behavior, not just the studied intervention.

4. Maturation Effects
Example: A child's improved reading ability may be due to natural development, not a specific intervention.

	Before (1991)	After (1992)
New Jersey (Treatment)	Employment measured $Y_{T_0} \| A = 1$	Employment measured $Y_{T_1} \| A = 1$
Pennsylvania (Control)	Employment measured $Y_{T_0} \| A = 0$	Employment measured $Y_{T_1} \| A = 0$

Example: Minimum Wage and Employment (Conti.)

(Card and Krueger, 1994)

Strategy 3: Combine cross-sectional and before–after comparisons

<u>The key concepts</u>

1. A cross-sectional comparison is vulnerable to between-group confounding, but it avoids temporal confounding and is less affected by external interventions.

2. A before–after comparison is free from between-group confounding (since it uses the same units), but it is susceptible to temporal confounding and external changes over time.

- As a result, compare before–after changes between groups by using

$$\mathbb{E}\left(Y_{T_1} - Y_{T_0} | A = 1\right) - \mathbb{E}\left(Y_{T_1} - Y_{T_0} | A = 0\right)$$

which is called Difference-in-Differences (DiD) method

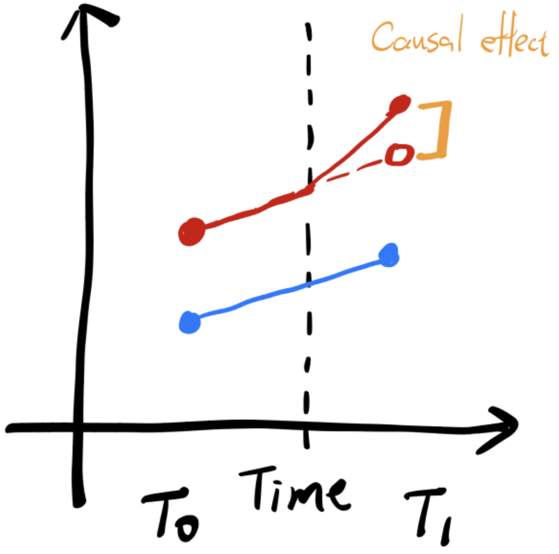
	Before (1991)	**After (1992)**
New Jersey (Treatment)	Employment measured $Y_{T_0} \| A = 1$	Employment measured $Y_{T_1} \| A = 1$
Pennsylvania (Control)	Employment measured $Y_{T_0} \| A = 0$	Employment measured $Y_{T_1} \| A = 0$

Assumption and identification for the DiD method

- **Parallel trends assumption**

- In the absence of treatment, the treatment and control groups are expected to exhibit parallel trends in outcomes over time. That is

$$\mathbb{E}\big(Y_{T_1}(0) - Y_{T_0}(0)|A = 1\big) = \mathbb{E}\big(Y_{T_1}(0) - Y_{T_0}(0)|A = 0\big)$$



- Also referred to as the additive equi-confounding assumption.

- **Fact: no causal effect of A on Y_{T_0}**

$$Y_{T_0}(1) = Y_{T_0}(0) = Y_{T_0}$$

- **Identification**

$\mathbb{E}\big(Y_{T_1} - Y_{T_0}|A = 1\big) - \mathbb{E}\big(Y_{T_1} - Y_{T_0}|A = 0\big)$

$= \mathbb{E}\big(Y_{T_1}(1) - Y_{T_0}(1)|A = 1\big) - \mathbb{E}\big(Y_{T_1}(0) - Y_{T_0}(0)|A = 0\big)$ **(Consistency assumption)**

$= \mathbb{E}\big(Y_{T_1}(1) - Y_{T_0}(1)|A = 1\big) - \mathbb{E}\big(Y_{T_1}(0) - Y_{T_0}(0)|A = 1\big)$ **(Parallel trends assumption)**

$= \mathbb{E}\big(Y_{T_1}(1) - Y_{T_0}|A = 1\big) - \mathbb{E}\big(Y_{T_1}(0) - Y_{T_0}|A = 1\big) = \mathbb{E}\big(Y_{T_1}(1) - Y_{T_1}(0)|A = 1\big) = \text{ATT}$

(no causal effect of A on Y_{T_0})

Alternative view for DiD

➢ **The DiD method**

$$\mathbb{E}\left(Y_{T_1} - Y_{T_0} | A = 1\right) - \mathbb{E}\left(Y_{T_1} - Y_{T_0} | A = 0\right)$$

represents the difference in before–after changes between the treatment and control groups.

- Assuming parallel trends, time-related bias in the treated group's before–after comparison can be adjusted for using the control group's trend.

➢ **Alternative view**

- (Constant Difference) Assumption:

$$\mathbb{E}\left(Y_{T_1}(0)\ | A = 1\right) - \mathbb{E}\left(Y_{T_1}(0)\ | A = 0\right) = \mathbb{E}\left(Y_{T_0}(0)\ | A = 1\right) - \mathbb{E}\left(Y_{T_0}(0)\ | A = 0\right)$$

- DiD formula can be rewritten as

$$\left[\mathbb{E}\left(Y_{T_1} | A = 1\right) - \mathbb{E}\left(Y_{T_1} | A = 0\right)\right] - \left[\mathbb{E}\left(Y_{T_0} | A = 1\right) - \mathbb{E}\left(Y_{T_0} | A = 0\right)\right]$$

represents the change in cross-sectional differences between pre- and post-treatment periods.

- Under the constant difference assumption, bias from covariate imbalance between groups can be removed by adjusting for pre-treatment differences.

References

Card, D., & Krueger, A. B. (1993). Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania. *The American Economic Review*, *84*(4), 772-793.