ECON3389 Econometric Methods

Module 5
Treatment Effects

Alberto Cappello

Department of Economics, Boston College

Spring 2025

Differences in Differences: Two Time Periods

- Consider a *natural* experiment where an exogenous policy change (called a treatment) affects one group more than another.
- Let y denote the outcome and d denote the treatment:
  - $d = 1$ if treated and $d = 0$ if not treated.
- **Method 1: Treatment-control comparison (at a point in time)**
  - Treatment effect: $\bar{y}_{d=1} - \bar{y}_{d=0}$.
  - **Problem:** Misleading if the treated and untreated groups differ in characteristics.
  - Example: Policy targeted towards poor people.
- **Method 2: Before-after comparison over time for treated only**
  - Treatment effect: $\bar{y}_{\text{treated, post}} - \bar{y}_{\text{treated, pre}}$.
  - **Problem:** Misleading if other factors also affect the treated group over time.
- **Differences-in-Differences (DID): Combines Methods 1 and 2**
  - Uses change over time for the untreated to control for non-treatment changes over time.
  - Assumes both groups have the same time trend.

Differences in Differences Formula

- Introduce time before (**pre**) and after (**post**) the policy comes into effect:
  - $t = 0$: Time period before the policy.
  - $t = 1$: Time period after the policy.
- The difference-in-difference (DID) estimate of the effect of treatment:
  - $\text{DID} = \Delta \bar{y}_{\text{treated}} - \Delta \bar{y}_{\text{untreated}}$
  - $= (\bar{y}_{d=1,\text{post}} - \bar{y}_{d=1,\text{pre}}) - (\bar{y}_{d=0,\text{post}} - \bar{y}_{d=0,\text{pre}})$
- Alternatively:
  - $\text{DID} = (\bar{y}_{d=1,\text{post}} - \bar{y}_{d=0,\text{post}}) - (\bar{y}_{d=1,\text{pre}} - \bar{y}_{d=0,\text{pre}})$
  - The post-period difference in the two groups minus the pre-period difference.
- **Estimation:**
  - Compute the four separate means and calculate the differences.

Regression Computation

- The same difference-in-difference (DID) estimate can be obtained as the coefficient of $t \cdot d$ in the OLS regression:

$$y_i = \beta_1 + \beta_2 t_i + \beta_3 d_i + \beta_4 (t_i \cdot d_i) + u_i$$

- Where:
  - $t_i = 1$ in the post-period and $t_i = 0$ in the pre-period.
  - $d_i = 1$ if treated and $d_i = 0$ if not treated.
  - $t_i \cdot d_i = 1$ if treated and in the post-period; $= 0$ otherwise.
- **Proof:** The model implies the following outcomes:

	Treated ($d = 1$)	Not Treated ($d = 0$)	Difference Over Treatment
Pre ($t = 0$)	$\beta_1 + \beta_3$	β_1	β_3
Post ($t = 1$)	$\beta_1 + \beta_2 + \beta_3 + \beta_4$	$\beta_1 + \beta_2$	$\beta_3 + \beta_4$
Change Over Time	$\beta_2 + \beta_4$	β_2	DID $= \beta_4$

Differences in Differences Regression Computation

- Suppose we have data on each individual, not just the means.
- The OLS regression model is:

$$y_i = \beta_1 + \beta_2 t_i + \beta_3 d_i + \beta_4 (t_i \cdot d_i) + u_i$$

- This is often rewritten as:

$$y_i = \beta_1 + \beta_2 \text{Post}_i + \beta_3 \text{Treat}_i + \beta_4 (\text{Post}_i \cdot \text{Treat}_i) + u_i$$

- The difference-in-differences (DID) estimate is β_4.
- **Advantages of OLS regression:**
  1. A t-test of $H_0 : \beta_4 = 0$ tests the statistical significance of the treatment.
  2. We can add control variables as additional regressors.
  3. Robust standard errors for $\hat{\beta}_4$ can be computed.

Introduction

- DID is a method for causal inference.
- Useful for analyzing exogenous policy impacts affecting groups differently.
- Often used with repeated cross-sectional data over time or subgroup comparisons.
- Relies on the **parallel trends assumption**:
    - In absence of treatment, trends for treated and untreated groups are equal.

Outline

Differences in Differences: Two Time Periods

- Natural experiment: Exogenous policy affects one group more than another.
- Notation:
  - y: Outcome, d: Treatment indicator ($d = 1$ treated, $d = 0$ untreated).
- Methods:
  1. Treatment-control comparison (misleading if groups differ in characteristics).
  2. Before-after comparison for treated only (misleading if other factors change).
  3. DID combines these methods, using untreated to control for non-treatment changes.

Differences in Differences Formula

- Time periods: $t = 0$ (pre) and $t = 1$ (post).
- DID estimate: $\Delta\bar{y}_{treated} - \Delta\bar{y}_{untreated}$
- Equivalent forms:
  - $\text{DID} = (\bar{y}_{d=1,\text{post}} - \bar{y}_{d=1,\text{pre}}) - (\bar{y}_{d=0,\text{post}} - \bar{y}_{d=0,\text{pre}})$
  - $\text{DID} = (\bar{y}_{d=1,\text{post}} - \bar{y}_{d=0,\text{post}}) - (\bar{y}_{d=1,\text{pre}} - \bar{y}_{d=0,\text{pre}})$
- Compute by averaging group means.

Example: Access to Health Care and Health Outcomes

- Data: 1,071 South African children aged 1-4 in 54 communities.
- 1993: 26 communities had clinics; 1998: all had clinics.
- Outcome: Weight-for-age z-score (waz).
- Treatment: $d = 1$ if access to clinic.
- Time: $t = 0$ (1993), $t = 1$ (1998).

Results: Manual Computation

	High Treated	Low Treated
Before (1993)	-0.545	-0.414
After (1998)	0.321	-0.069
Change	0.867	0.345
DID	0.521	

- DID estimate: $0.867 - 0.345 = 0.521$.
- Substantial effect: A third of a standard deviation increase in *waz*.

Parallel Trends Assumption

- Causal interpretation requires parallel trends:
  - Change over time in outcome (without treatment) is the same for treated and untreated groups.
- Mathematical form: $E[Y_{\text{post}}(0)|d = 1] - E[Y_{\text{pre}}(0)|d = 1] = E[Y_{\text{post}}(0)|d = 0] - E[Y_{\text{pre}}(0)|d = 0]$.

References

- Cameron, A. Colin (2022). *Analysis of Economics Data: An Introduction to Econometrics*.
- Angrist, Joshua D., and Pischke, Jörn-Steffen (2015). *Mastering Metrics*.
- Cunningham, Scott (2021). *Causal Inference: The MixTape*.
- Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data*.

Introduction to Treatment Effects

Treatment effects refer to the impact of a policy, intervention, or treatment on an outcome of interest. In causal inference, we are concerned with estimating how an individual's outcome would change due to exposure to treatment.

Key concepts:

- **Average Treatment Effect (ATE)**: The average effect of the treatment across the entire population.
- **Average Treatment Effect on the Treated (ATT)**: The average effect on individuals who actually receive the treatment.

Example: Minimum Wage and Employment

Study: A study by Card and Krueger (1994) investigated the impact of an increase in the minimum wage on employment in the fast food industry in New Jersey.

Research Question:

- Did the increase in the minimum wage reduce employment in fast food restaurants in New Jersey compared to Pennsylvania (a control group)?

Approach: The authors used a Difference-in-Differences (DiD) approach to estimate the treatment effect of the minimum wage increase.

Difference-in-Differences (DiD) Setup

Two groups: Treatment group and control group.
Two time periods: Before and after the treatment.

- Let Y_{it} be the outcome for unit i at time t.
- Let T_i be a binary indicator for whether unit i is in the treatment group.
- Let $Post_t$ be a binary indicator for whether time t is after the treatment period.

DiD Model Specification

The DiD model can be specified as:

$$Y_{it} = \alpha + \beta T_i \cdot Post_t + \gamma_i + \lambda_t + \epsilon_{it}$$

where:

- $T_i \cdot Post_t$ is the interaction term that captures the treatment effect.
- γ_i is unit-specific fixed effects (controls for time-invariant individual heterogeneity).
- λ_t is time-specific fixed effects (controls for common shocks across time).
- ϵ_{it} is the error term.

Identifying the Treatment Effect

The DiD estimator is given by the coefficient $\hat{\beta}$ on the interaction term $T_i \cdot Post_t$, which represents the causal effect of the treatment.

$$\hat{\beta} = \text{(Change in Treatment Group)} - \text{(Change in Control Group)}$$

This difference accounts for any pre-existing trends that affect both groups.

Assumptions for DiD Identification

To correctly identify the treatment effect, we need to make the following assumptions:

- **Parallel Trends Assumption**: In the absence of the treatment, the treatment and control groups would have followed parallel paths over time.
- **No Anticipation of the Treatment**: Individuals should not change their behavior before the treatment because they anticipate it.

Graphical Interpretation of DiD

- A plot of the outcomes over time for both the treatment and control groups helps visualize the parallel trends assumption.
- The DiD estimate is the difference in the post-treatment period, adjusted for pre-treatment trends.

Deriving the DiD Estimator

The DiD estimator is derived as follows:

$$\hat{\beta} = \frac{1}{N_{\text{treated}}} \sum_{i \in \text{treated}} \left(Y_{it}^{\text{post}} - Y_{it}^{\text{pre}} \right) - \frac{1}{N_{\text{control}}} \sum_{i \in \text{control}} \left(Y_{it}^{\text{post}} - Y_{it}^{\text{pre}} \right)$$

The difference in changes between the two groups gives the treatment effect after adjusting for baseline differences and time trends.

Potential Bias in DiD

Bias in DiD estimates can arise from:

- **Violation of Parallel Trends Assumption**: If the pre-treatment trends are not parallel, the DiD estimate will be biased.
- **Spillover Effects**: If the control group is affected by the treatment, the estimate may be biased.

Testing the Parallel Trends Assumption

We can test for parallel trends by:

- Checking the pre-treatment trends visually in a plot.
- Adding leads of the treatment indicator in the model:

$$Y_{it} = \alpha + \sum_k \beta_k \cdot T_i \cdot \text{Post}_{t-k} + \gamma_i + \lambda_t + \epsilon_{it}$$

where β_k tests for pre-treatment trends.

Two-way Fixed Effects in DiD

To control for both unit and time heterogeneity, we use **two-way fixed effects**:

$$Y_{it} = \alpha + \beta T_i \cdot Post_t + \gamma_i + \lambda_t + \epsilon_{it}$$

where:

- γ_i is the unit-specific fixed effect.
- λ_t is the time-specific fixed effect.

This model controls for any unobserved factors that vary across units or time.

Extensions of DiD

Multiple Groups and Time Periods:

- DiD can be extended to handle multiple treatment groups or multiple time periods.
- The model becomes:

$$Y_{it} = \alpha + \beta T_{i,t} + \gamma_i + \lambda_t + \epsilon_{it}$$

where $T_{i,t}$ now indicates multiple treatments over time.

Application to Card and Krueger (1994)

DiD Applied:

- Card and Krueger (1994) used a DiD approach to estimate the effect of an increase in minimum wage on fast food employment.
- By comparing changes in employment in New Jersey (treatment group) and Pennsylvania (control group), they found no evidence of a negative impact.

Conclusion and Limitations of DiD

Conclusion:

- DiD is a powerful method for estimating causal treatment effects.
- The key assumptions (parallel trends) must hold for unbiased estimation.
  Limitations:
  - Violations of parallel trends lead to biased estimates.
  - There may be spillover effects between groups.