Difference-in-Differences

ECON 31720 Applied Microeconometrics

Francesco Ruggieri

The University of Chicago

December 2, 2020

1. Difference-in-Differences with Two Time Periods
   - Example: Participation to a Program (Wing and Cook 2013)
2. Difference-in-Differences with Multiple Time Periods
   - Monte Carlo Simulation: Linear Regression Implementation
3. Changes-in-Changes (Athey and Imbens 2006)
4. Summary

Setup

- i and $t \in \{1, 2\}$ indicate **units** and **time periods**, respectively
- $Y_{it} \in \mathbb{R}$ is a scalar **outcome** of interest
- $G_i \in \{0, 1\}$ is a time-invariant **binary treatment group**
- $D_{it} \equiv G_i \mathbb{I}[t = 2]$ is a **binary treatment** available to units in $G_i = 1$ in period $t = 2$
- D_{it} and Y_{it} are linked by **potential outcomes** $Y_{it}(0), Y_{it}(1)$

Identification

- Assume **common trends** in **untreated potential outcomes** across treatment groups:

$$\mathbb{E}\left[Y_{i2}(0) - Y_{i1}(0) | G_i = 0\right] = \mathbb{E}\left[Y_{i2}(0) - Y_{i1}(0) | G_i = 1\right]$$

- The **average change** in **untreated potential outcomes** is **group-invariant**

- Thus, the **average untreated potential outcome** among **treated units** in $t = 2$ is

$$\begin{aligned}\mathbb{E}\left[Y_{i2}(0) | G_i = 1\right] &= \mathbb{E}\left[Y_{i1}(0) | G_i = 1\right] + \mathbb{E}\left[Y_{i2}(0) - Y_{i1}(0) | G_i = 0\right] \\ &= \mathbb{E}\left[Y_{i1} | G_i = 1\right] + \mathbb{E}\left[Y_{i2} - Y_{i1} | G_i = 0\right]\end{aligned}$$

where the second equality follows from the fact that **all** units are **untreated** in $t = 1$

Identification

- The **Average Treatment Effect on the Treated** (ATT) can be identified as

$$\begin{aligned}
\mathrm{ATT} &\equiv \mathbb{E}\left[Y_{i2}(1) - Y_{i2}(0) | G_i = 1\right] \\
&= \mathbb{E}\left[Y_{i2}(1) | G_i = 1\right] - \mathbb{E}\left[Y_{i2}(0) | G_i = 1\right] && \text{(linearity of } \mathbb{E}[\cdot]) \\
&= \mathbb{E}\left[Y_{i2} | G_i = 1\right] - \mathbb{E}\left[Y_{i2}(0) | G_i = 1\right] && (D_{it} \equiv G_i \mathbb{I}[t = 2]) \\
&= \mathbb{E}\left[Y_{i2} | G_i = 1\right] - \left(\mathbb{E}\left[Y_{i1} | G_i = 1\right] + \mathbb{E}\left[Y_{i2} - Y_{i1} | G_i = 0\right]\right) && \text{(common trends)} \\
&= \mathbb{E}\left[Y_{i2} - Y_{i1} | G_i = 1\right] - \mathbb{E}\left[Y_{i2} - Y_{i1} | G_i = 0\right]
\end{aligned}$$

- The **Average Treatment Effect on the Untreated** (ATU) cannot be identified because

$$\begin{aligned}
\mathrm{ATU} &\equiv \mathbb{E}\left[Y_{i2}(1) - Y_{i2}(0) | G_i = 0\right] \\
&= \mathbb{E}\left[Y_{i2}(1) | G_i = 0\right] - \mathbb{E}\left[Y_{i2}(0) | G_i = 0\right] && \text{(linearity of } \mathbb{E}[\cdot]) \\
&= \mathbb{E}\left[Y_{i2}(1) | G_i = 0\right] - \mathbb{E}\left[Y_{i2} | G_i = 0\right] && (D_{it} \equiv G_i \mathbb{I}[t = 2])
\end{aligned}$$

  and treated potential outcomes are **never observed** among untreated units

Setup

- As above, i and $t \in \{1, 2\}$ indicate **units** and **time periods**, respectively
- $X_i \in \mathcal{X} \subseteq \mathbb{R}$ is a **time-invariant, predetermined and observable** random variable
- $D_{it} \in \{0, 1\}$ denotes **participation to a program** that is only available in period $t = 2$
  - As in the standard case, $D_{i1} = 0$ for all i
  - Program participation in period $t = 2$ is determined as $D_{i2} \equiv \mathbb{I}[X_i \geq \overline{x}]$, with $\overline{x}$ **known**
  - E.g. a program that applies retroactively to individuals above an age cutoff at a given date
- $Y_{it} \in \mathbb{R}$ is a scalar **outcome** of interest

Identification

- D_{it} and Y_{it} are linked by **potential outcomes** $Y_{it}(0), Y_{it}(1)$
- $\mathbb{E}\left[Y_{it}(d)|X_i = x\right]$ is **continuous** for all $x \in \mathcal{X}$ and $d = 0, 1$
- Assume that the **average change** in **untreated potential outcomes** is **constant**:

$$\mathbb{E}\left[Y_{i2}(0) - Y_{i1}(0)|X_i = x\right] = \alpha \in \mathbb{R} \quad \forall x \in \mathcal{X}$$

- Goal: determine the **largest set** of X for which one can **point identify**

$$\mathbb{E}\left[Y_{i2}(1) - Y_{i2}(0)|X_i = x\right]$$

i.e., the (conditional) average treatment effect

Identification

- Consider any $x < \bar{x}$. Then $D_{i1} = 0$ and

$$\alpha = \mathbb{E}\left[Y_{i2}(0) - Y_{i1}(0) | X_i = x\right] = \mathbb{E}\left[Y_{i2} - Y_{i1} | X_i = x\right] \quad \forall x < \bar{x}$$

  which implies that α **is point identified**

- Consider any $x \geq \bar{x}$. Then $D_{i2} = 1$ and

$$\mathbb{E}\left[Y_{i2}(1) | X_i = x\right] = \mathbb{E}\left[Y_{i2} | X_i = x\right] \quad \forall x \geq \bar{x}$$

- In addition, for any $x \geq \bar{x}$,

$$\begin{aligned}
\mathbb{E}\left[Y_{i2}(0) | X_i = x\right] &= \mathbb{E}\left[Y_{i2}(0) | X_i = x\right] + \mathbb{E}\left[Y_{i1}(0) - Y_{i1}(0) | X_i = x\right] \\
&= \mathbb{E}\left[Y_{i2}(0) - Y_{i1}(0) | X_i = x\right] + \mathbb{E}\left[Y_{i1}(0) | X_i = x\right] \\
&= \alpha + \mathbb{E}\left[Y_{i1}(0) | X_i = x\right] \qquad (\alpha \text{ point identified}) \\
&= \alpha + \mathbb{E}\left[Y_{i1} | X_i = x\right] \qquad (D_{i1} = 0 \ \forall i)
\end{aligned}$$

Identification

- The **target parameter** can be **point identified** for any $x \geq \overline{x}$:

$$\begin{aligned}\mathbb{E}[Y_{i2}(1) - Y_{i2}(0)|X_i = x] &= \mathbb{E}[Y_{i2}(1)|X_i = x] - \mathbb{E}[Y_{i2}(0)|X_i = x] \quad \text{(linearity of } \mathbb{E}[\cdot]) \\ &= \mathbb{E}[Y_{i2}|X_i = x] - (\alpha + \mathbb{E}[Y_{i1}|X_i = x]) \\ &= \mathbb{E}[Y_{i2} - Y_{i1}|X_i = x] - \alpha \quad \text{(linearity of } \mathbb{E}[\cdot])\end{aligned}$$

- The target parameter **cannot** be point identified for $x < \overline{x}$ because

$$\begin{aligned}\mathbb{E}[Y_{i2}(1) - Y_{i2}(0)|X_i = x] &= \mathbb{E}[Y_{i2}(1)|X_i = x] - \mathbb{E}[Y_{i2}(0)|X_i = x] \quad \text{(linearity of } \mathbb{E}[\cdot]) \\ &= \mathbb{E}[Y_{i2}(1)|X_i = x] - \mathbb{E}[Y_{i2}|X_i = x] \quad (D_{i2} = 0 \text{ for } x < \overline{x})\end{aligned}$$

and treated potential outcomes are **never observed** among untreated units

Setup

- i and $t \in \{1, \ldots, t_0, t^*, \ldots, \overline{t}\}$ indicate **units** and **time periods**, respectively
- $Y_{it} \in \mathbb{R}$ is a scalar **outcome** of interest
- $G_i \in \{0, 1\}$ is a time-invariant **binary treatment group**
- $D_{it} \equiv G_i \mathbb{I}[t \geq t^*]$ is a **binary treatment** available to units in $G_i = 1$ in periods $t \geq t^*$
    - $\{1, \ldots, t_0\}$ is the set of **pre-periods** and $\{t^*, \ldots, \overline{t}\}$ is the set of **post-periods**
- D_{it} and Y_{it} are linked by **potential outcomes** $Y_{it}(0), Y_{it}(1)$

Identification

- Assume **common trends** in **untreated potential outcomes** across treatment groups:

$$\mathbb{E}\left[Y_{is}(0) - Y_{ir}(0) | G_i = 0\right] = \mathbb{E}\left[Y_{is}(0) - Y_{ir}(0) | G_i = 1\right]$$

for any $r \in \{1, \ldots, t_0\}$ and any $s \in \{t^*, \ldots, \overline{t}\}$

- All **average changes** in **untreated potential outcomes** are **group-invariant**

- Thus, the **average untreated potential outcome** among **treated units** in $t = s$ is

$$\begin{aligned} \mathbb{E}\left[Y_{is}(0) | G_i = 1\right] &= \mathbb{E}\left[Y_{ir}(0) | G_i = 1\right] + \mathbb{E}\left[Y_{is}(0) - Y_{ir}(0) | G_i = 0\right] \\ &= \mathbb{E}\left[Y_{ir} | G_i = 1\right] + \mathbb{E}\left[Y_{is} - Y_{ir} | G_i = 0\right] \end{aligned}$$

where the second equality follows from the fact that **all** units are **untreated** in $t = r$

Identification

- **Period-specific ATTs** can be identified as

$$\begin{aligned} \text{ATT}_s &\equiv \mathbb{E}\left[Y_{is}(1) - Y_{is}(0) | G_i = 1\right] \\ &= \mathbb{E}\left[Y_{is}(1) | G_i = 1\right] - \mathbb{E}\left[Y_{is}(0) | G_i = 1\right] \qquad \text{(linearity of } \mathbb{E}\left[\cdot\right]\text{)} \\ &= \mathbb{E}\left[Y_{is} | G_i = 1\right] - \mathbb{E}\left[Y_{is}(0) | G_i = 1\right] \qquad (D_{it} \equiv G_i \mathbb{I}\left[t \geq t^*\right]) \\ &= \mathbb{E}\left[Y_{is} | G_i = 1\right] - (\mathbb{E}\left[Y_{ir} | G_i = 1\right] + \mathbb{E}\left[Y_{is} - Y_{ir} | G_i = 0\right]) \qquad \text{(common trends)} \\ &= \mathbb{E}\left[Y_{is} - Y_{ir} | G_i = 1\right] - \mathbb{E}\left[Y_{is} - Y_{ir} | G_i = 0\right] \end{aligned}$$

- **Period-specific ATUs** cannot be identified because

$$\begin{aligned} \text{ATU}_s &\equiv \mathbb{E}\left[Y_{is}(1) - Y_{is}(0) | G_i = 0\right] \\ &= \mathbb{E}\left[Y_{is}(1) | G_i = 0\right] - \mathbb{E}\left[Y_{is}(0) | G_i = 0\right] \qquad \text{(linearity of } \mathbb{E}\left[\cdot\right]\text{)} \\ &= \mathbb{E}\left[Y_{is}(1) | G_i = 0\right] - \mathbb{E}\left[Y_{is} | G_i = 0\right] \qquad (D_{it} \equiv G_i \mathbb{I}\left[t \geq t^*\right]) \end{aligned}$$

and treated potential outcomes are **never observed** among untreated units

Linear Regression Implementation

Period-specific ATTs can be equivalently computed with linear regression:

- **Common trends** implies **additive separability of unit and time effects** in $\mathbb{E}\left[Y_{it}(0)|G_i\right]$:

$$\mathbb{E}\left[Y_{it}(0)|G_i = g\right] = \mathbb{I}\left[G_i = g\right] + \beta_t = \alpha_i + \beta_t \quad \text{for } g = 0, 1$$

- Under common trends, the conditional mean of the **observed outcome** is

$$\mathbb{E}\left[Y_{it}|G_i = g\right] = \alpha_i + \beta_t + \sum_{j \geq t^*} \mathbb{I}\left[G_i = 1, t = j\right] \mathrm{ATT}_j$$

- This is the **linear regression** implementation of a difference-in-differences design
    - **Not fully saturated**, but $\{\mathrm{ATT}_j\}_{j \geq t^*}$ are **exactly** (not approximately) point identified

Linear Regression Implementation

Let us compare three common regression specifications:

1. **Two-way fixed effects** regression with **post-period interactions**

$$Y_{it} = \alpha_i + \beta_t + \sum_{j \geq t^*} \gamma_j G_i \mathbb{I}\left[t = j\right] + U_{it}$$

2. **Two-way fixed effects** regression with **a single post-period interaction**

$$Y_{it} = \alpha_i + \beta_t + \gamma G_i \mathbb{I}\left[t \geq t^*\right] + U_{it}$$

3. **Two-way fixed effects** regression with **some pre- and post-period interactions**

$$Y_{it} = \alpha_i + \beta_t + \sum_{j \in \mathcal{J}} \gamma_j G_i \mathbb{I}\left[t = j\right] + U_{it} \quad \text{where } \mathcal{J} = \left\{t^* - \bar{l}, \ldots, t^*, \ldots, t^* + \overline{m}\right\}$$

1. Difference-in-Differences with Two Time Periods
   - Example: Participation to a Program (Wing and Cook 2013)
2. Difference-in-Differences with Multiple Time Periods
   - Monte Carlo Simulation: Linear Regression Implementation
3. Changes-in-Changes (Athey and Imbens 2006)
4. Summary

Data Generating Process

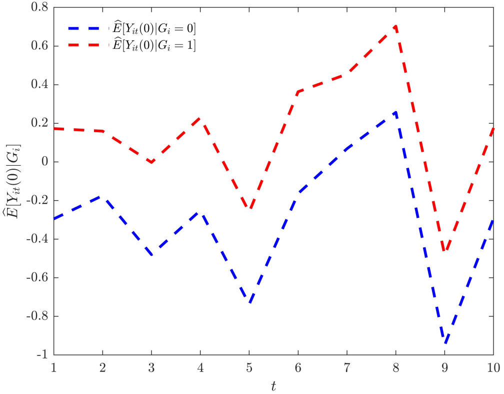
$$Y_{it}(0) = A_i + B_t + U_{it}$$
$$Y_{it}(1) - Y_{it}(0) = \sin(t)\left(A_i + 0.3G_i\right) + V_{it}$$

- Time periods indexed by $t \in \{1, \dots, 10\}$
- $\mathbb{P}\left(G_i = 1\right) = 0.3$
- $A_i | G_i = g \sim \mathcal{N}\left(-0.2 + 0.5g, (1 + 0.3g)^2\right)$
- $B_t \sim \mathcal{N}\left(0, 0.09\right)$, and independent of all other variables
- $U_{it} \sim \mathcal{N}\left(0, 1\right)$, and independent of all other variables
- $V_{it} \sim \mathcal{N}\left(0, 0.04\right)$, and independent of all other variables
- The binary treatment is defined as $D_{it} \equiv G_i \mathbb{I}\left[t \geq 6\right]$

Common Trends

Common trends holds because

$$\mathbb{E}\left[Y_{it}(0) - Y_{i1}(0) | G_i = g\right] = B_t - B_1 + \mathbb{E}\left[U_i | G_i = g\right] = B_t - B_1 \quad \text{for } g = 0, 1$$



Monte Carlo Simulation

- Perform a **Monte Carlo simulation** to compare difference-in-differences specifications
- **Period-specific ATTs** can be estimated as $\{\gamma_j\}_{j=6}^{10}$ in

$$Y_{it} = \alpha_i + \beta_t + \sum_{j=6}^{10} \gamma_j G_i \mathbb{I}\left[t = j\right] + R_{it}$$

Parameter	Mean Estimate
γ_6	-0.166
γ_7	0.394
γ_8	0.596
γ_9	0.249
γ_{10}	-0.327

Notes: This table reports mean OLS estimates of $\{\gamma_j\}_{j=6}^{10}$ across 1000 Monte Carlo simulations.

Monte Carlo Simulation

- The TWFE regression with **one post-period interaction** identifies $\overline{\widehat{\gamma}} = \frac{1}{5}\sum_{j=6}^{10}\widehat{\mathrm{ATT}}_j = 0.148$

- Consider the two-way fixed effects specification with **some pre- and post-period interactions**:

$$Y_{it} = \alpha_i + \beta_t + \sum_{j=4}^{7} \gamma_j G_i \mathbb{I}\left[t=j\right] + R_{it}$$

Parameter	Mean Estimate
γ_4	-0.085
γ_5	-0.086
γ_6	-0.252
γ_7	0.308

Notes: This table reports mean OLS estimates of $\{\gamma_j\}_{j=4}^{7}$ across 1000 Monte Carlo simulations.

- In this case, mean estimates are **significantly different** from the **estimated ATTs**

Motivation

- **Difference-in-differences** is subject to a **nonlinearity critique**
  - E.g. if common trends holds for Y, common trends cannot hold for $\log(Y)$, and viceversa
  - There may be valid economic reasons why this critique is not particularly salient
- **Changes-in-changes** (CiC) is immune to this nonlinearity critique
- In a standard difference-in-differences design, **common trends** implies that

$$\mathbb{E}[Y_{is}(0)|G_i = 1] = \mathbb{E}[Y_{ir}|G_i = 1] + \mathbb{E}[Y_{is} - Y_{ir}|G_i = 0]$$

for any $r \in \{1, \ldots, t_0\}$ and any $s \in \{t^*, \ldots, \bar{t}\}$

- CiC argument: identify the **marginal distributions** of $\boldsymbol{Y(0)}$ among **treated units** in **post-periods** by assuming **rank invariance** of the marginal distributions of $Y(0)$ over time

Setup

For simplicity, consider the following framework:

- i and $t \in \{1, 2\}$ indicate **units** and **time periods**, respectively
- $Y_{it} \in \mathbb{R}$ is a scalar and continuously distributed **outcome** of interest
- $G_i \in \{0, 1\}$ is a time-invariant **binary treatment group**
- $D_{it} \equiv G_i \mathbb{I}[t = 2]$ is a **binary treatment** available to units in $G_i = 1$ in period $t = 2$
- D_{it} and Y_{it} are linked by **potential outcomes** $Y_{it}(0), Y_{it}(1)$
- $U_i \in \mathbb{R}$ is a **time-invariant** scalar **latent variable**

Identification

- The **difference-in-differences** model assumes the **additive single index** structure

$$Y_{it}(0) = h_t\left(U_i\right) = \phi\left(U_i + \delta t\right) = U_i + \delta t$$

where $\phi(\cdot)$ is the **identity function**

- The **changes-in-changes** model assumes the **additive single index** structure

$$Y_{it}(0) = h_t\left(U_i\right) = \phi\left(U_i + \delta t\right)$$

where $\phi(\cdot)$ is a generic **strictly increasing function**

- The h_t functions are **unknown** and strictly increasing

- Further assume the marginal distributions of $Y(0)$ are **rank invariant over time**:

$$F_{Y_1(0)}\left(Y_{i1}(0)\right) = F_{Y_2(0)}\left(Y_{i2}(0)\right) = U_i$$

Identification

- The **marginal distribution** of the **untreated potential outcome** in the **pre-period** is

$$\begin{aligned} F_{Y_1(0)|G}(y|g) &\equiv \mathbb{P}\left(Y_{i1}(0) \leq y | G_i = g\right) \\ &= \mathbb{P}\left(h_1(U_i) \leq y | G_i = g\right) \qquad \left(Y_{i1}(0) = h_1\left(U_i\right)\right) \\ &= \mathbb{P}\left(U_i \leq h_1^{-1}(y) | G_i = g\right) \qquad \left(h_1 \text{ strictly increasing}\right) \\ &= \mathbb{P}\left(h_2(U_i) \leq h_2\left(h_1^{-1}(y)\right) | G_i = g\right) \qquad \left(h_2 \text{ strictly increasing}\right) \\ &= \mathbb{P}\left(Y_{i2}(0) \leq h_2\left(h_1^{-1}(y)\right) | G_i = g\right) \qquad \left(Y_{i2}(0) = h_2\left(U_i\right)\right) \\ &\equiv F_{Y_2(0)|G}\left(h_2\left(h_1^{-1}(y)\right) | g\right) \end{aligned}$$

- Because $D_{i1} = 0$ for all i, $h_2\left(h_1^{-1}(y)\right)$ can be **point identified** as

$$\phi(y) \equiv h_2\left(h_1^{-1}(y)\right) = F_{Y_2(0)|G}^{-1}\left(F_{Y_1(0)|G}(y|0)\,|0\right) = F_{Y_2|G}^{-1}\left(F_{Y_1|G}(y|0)\,|0\right)$$

Identification

- Thus, the **marginal distribution** of $\boldsymbol{Y(0)}$ among **treated units** in the **post-period** is

$$F_{Y_2(0)|G}\left(\phi\left(y\right)|1\right) = F_{Y_1(0)|G}\left(y|1\right) = F_{Y_1|G}\left(y|1\right)$$

or, equivalently,

$$F_{Y_2(0)|G}\left(y|1\right) = F_{Y_1(0)|G}\left(\phi^{-1}\left(y\right)\middle|1\right) = F_{Y_1|G}\left(\phi^{-1}\left(y\right)\middle|1\right)$$

- As usual, the **marginal distribution** of $\boldsymbol{Y(1)}$ among **treated units** in the **post-period** is

$$F_{Y_2(1)|G}\left(y|1\right) \equiv \mathbb{P}\left(Y_{i2}(1) \leq y|G_i = 1\right) = \mathbb{P}\left(Y_{i2} \leq y|G_i = 1\right) \equiv F_{Y_2|G}\left(y|1\right)$$

where point identification follows from the fact that $D_{it} \equiv G_i\mathbb{I}\left[t = 2\right]$

Identification

- Any **target parameter** that is a function of $F_{Y_2(0)|G}(y|1)$ and $F_{Y_2(1)|G}(y|1)$ is **identified**

- For instance, the **Average Treatment Effect on the Treated** can be point identified as

$$\begin{aligned}
\text{ATT} &\equiv \mathbb{E}\left[Y_{i2}(1) - Y_{i2}(0)|G_i = 1\right] \\
&= \mathbb{E}\left[Y_{i2}(1)|G_i = 1\right] - \mathbb{E}\left[Y_{i2}(0)|G_i = 1\right] \qquad (\text{linearity of } \mathbb{E}[\cdot]) \\
&= \mathbb{E}\left[Y_{i2}|G_i = 1\right] - \mathbb{E}\left[Y_{i2}(0)|G_i = 1\right] \qquad (D_{it} \equiv G_i \mathbb{I}[t = 2]) \\
&= \mathbb{E}\left[Y_{i2}|G_i = 1\right] - \mathbb{E}_{F_{Y_1|G}}\left[\phi^{-1}(Y_{i1})\,|G_i = 1\right]
\end{aligned}$$

where the last equality uses the point identified distribution of $Y_2(0)|G = 1$

Summary

- Difference-in-differences with **two time periods** identifies the ATT, not the $\mathrm{ATU}/\mathrm{ATE}$
- Difference-in-differences with **multiple time periods** identifies period-specific ATTs
- The changes-in-changes model does **not** hinge on **common trends** but assumes **rank invariance** of the distributions of untreated potential outcomes **over time**