

Difference in Differences

Erick F Molina

Summer, 2024

ITAM

- Difference-in-Differences (DiD) is a quasi-experimental identification strategy to estimate causal effects.
- It uses natural variation in the implementation of treatment across different groups over time.

John Snow and the Cholera Hypothesis

- John Snow used a natural experiment to show that cholera was transmitted through contaminated water.
- He compared cholera mortality rates between two water companies: Lambeth and Southwark and Vauxhall.
- Lambeth moved its water intake upstream, avoiding contamination, while Southwark and Vauxhall did not.

Results from Snow's Study

- In 1849, Southwark and Vauxhall had 135 cholera cases per 10,000 households, while Lambeth had 85.
- In 1854, Southwark and Vauxhall had 147 cases per 10,000 households, while Lambeth had only 19.
- The estimated difference-in-differences (DiD) shows 78 fewer cholera deaths per 10,000 households in Lambeth compared to Southwark and Vauxhall.

Estimating the Causal Effect

- The goal is to estimate the causal effect of clean water (D) on cholera deaths (Y).
- The basic DiD formula is:

$$\delta^{DD} = (Y_{\text{post, treated}} - Y_{\text{pre, treated}}) - (Y_{\text{post, control}} - Y_{\text{pre, control}}) \quad (1)$$

- This removes time fixed effects and time-invariant differences across groups.

Intuitively:

- $Y_{\text{post, treated}} - Y_{\text{pre, treated}} = \text{treatment effect} + \text{time trend}$
- $Y_{\text{post, control}} - Y_{\text{pre, control}} = \text{time trend}$
- $\delta^{DD} = \text{treatment effect}$

Estimating the causal effect

$$\begin{aligned}\delta^{DD} &= (Y_{\text{post, treated}} - Y_{\text{pre, treated}}) - (Y_{\text{post, control}} - Y_{\text{pre, control}}) \\ &= (Y_{\text{post, treated}} - Y_{\text{post, control}}) - (Y_{\text{pre, treated}} - Y_{\text{pre, control}}) \quad (2)\end{aligned}$$

Intuitively:

- $Y_{\text{post, treated}} - Y_{\text{post, control}} = \text{treatment effect} + \text{selection bias}$
- $Y_{\text{pre, treated}} - Y_{\text{pre, control}} = \text{selection bias}$
- $\delta^{DD} = \text{treatment effect}$

DiD estimation

	Treated	Control
Pre	$Y_{\text{pre, treated}}$	$Y_{\text{pre, control}}$
Post	$Y_{\text{post, treated}}$	$Y_{\text{post, control}}$

- Intuitively, DiD is a comparison across the four cells.

- If δ is the true program impact, then:

$$\delta = \mathbb{E}[Y_i(1) | T_i = 1, t = \tau] - \mathbb{E}[Y_i(0) | T_i = 1, t = \tau] \quad (3)$$

- We assume δ does not depend on the time period, nor on individuals' characteristics.
- In the absence of the program, individual i 's outcome in period t is given by:

$$\mathbb{E}[Y_i | T_i = 0, t = \tau] = \gamma_i + \lambda_\tau \quad (4)$$

- Selection bias is related to individuals' fixed characteristics γ_i .
- The time trend (λ_τ) is the same for the treated and control groups: the **parallel trends assumption**.

Then:

$$\begin{aligned}\mathbb{E}[Y_{\text{pre, control}}] &= \mathbb{E}[Y_i(0) | T_i = 0, t = \textit{pre}] = \mathbb{E}[\gamma_i | T_i = 0] + \mathbb{E}[\lambda_\tau | t = \textit{pre}] \\ \mathbb{E}[Y_{\text{post, control}}] &= \mathbb{E}[Y_i(0) | T_i = 0, t = \textit{post}] = \mathbb{E}[\gamma_i | T_i = 0] + \mathbb{E}[\lambda_\tau | t = \textit{post}] \\ \mathbb{E}[Y_{\text{pre, treated}}] &= \mathbb{E}[Y_i(0) | T_i = 1, t = \textit{pre}] = \mathbb{E}[\gamma_i | T_i = 1] + \mathbb{E}[\lambda_\tau | t = \textit{pre}] \\ \mathbb{E}[Y_{\text{post, treated}}] &= \mathbb{E}[Y_i(1) | T_i = 1, t = \textit{pre}] = \delta + \mathbb{E}[\gamma_i | T_i = 1] + \mathbb{E}[\lambda_\tau | t = \textit{post}]\end{aligned}$$

Comparisons

Treated vs. control:

$$\begin{aligned}\mathbb{E}[Y_{\text{post, treated}}] - \mathbb{E}[Y_{\text{post, control}}] &= \delta + \mathbb{E}[\gamma_i | T_i = 1] + \mathbb{E}[\lambda_\tau | t = \textit{post}] - \\ &\quad \mathbb{E}[\gamma_i | T_i = 0] - \mathbb{E}[\lambda_\tau | t = \textit{post}] \\ &= \delta + \underbrace{\mathbb{E}[\gamma_i | T_i = 1] - \mathbb{E}[\gamma_i | T_i = 0]}_{\text{selection bias}}\end{aligned}$$

Post vs. pre:

$$\begin{aligned}\mathbb{E}[Y_{\text{post, treated}}] - \mathbb{E}[Y_{\text{pre, treated}}] &= \delta + \mathbb{E}[\gamma_i | T_i = 1] + \mathbb{E}[\lambda_\tau | t = \textit{post}] - \\ &\quad \mathbb{E}[Y_{i0} | T_i = 1, t = \textit{pre}] \\ &= \mathbb{E}[\gamma_i | T_i = 1] - \mathbb{E}[\lambda_\tau | t = \textit{pre}] \\ &= \delta + \underbrace{\mathbb{E}[\lambda_\tau | t = \textit{post}] - \mathbb{E}[\lambda_\tau | t = \textit{pre}]}_{\text{time trend}}\end{aligned}$$

Comparisons

$$\begin{aligned}\delta^{DD} &= (Y_{\text{post, treated}} - Y_{\text{pre, treated}}) - (Y_{\text{post, control}} - Y_{\text{pre, control}}) \\ &= (\delta + \mathbb{E}[\gamma_i | T_i = 1] + \mathbb{E}[\lambda_\tau | t = \text{post}] - \mathbb{E}[\gamma_i | T_i = 1] - \mathbb{E}[\lambda_\tau | t = \text{pre}]) - \\ &\quad (\mathbb{E}[\gamma_i | T_i = 0] + \mathbb{E}[\lambda_\tau | t = \text{post}] - \mathbb{E}[\gamma_i | T_i = 0] - \mathbb{E}[\lambda_\tau | t = \text{pre}]) \\ &= (\delta + \mathbb{E}[\lambda_\tau | t = \text{post}] - \mathbb{E}[\lambda_\tau | t = \text{pre}]) - \\ &\quad (\mathbb{E}[\lambda_\tau | t = \text{post}] - \mathbb{E}[\lambda_\tau | t = \text{pre}]) \\ &= \delta\end{aligned}$$

- DiD does not rely on homogeneous treatment effects.
- Under homogeneous treatment effects, the estimate corresponds to the average treatment effect on the treated (ATT).

DiD as a regression

We will estimate the following regression:

$$Y_{it} = \alpha + \beta T_i + \zeta Post_t + \delta (T_i * Post_t) + \varepsilon_{it}$$

where:

- $Post_t$ is a dummy equal to 1 if $t = 2$
- δ is the treatment effect.
- $\alpha = E[\gamma_i | T_i = 0] + \lambda_1$: pre-program mean for the control group.
- $\beta = E[\gamma_i | T_i = 1] - E[\gamma_i | T_i = 0]$: selection bias.
- $\zeta = \lambda_2 - \lambda_1$: time trend.

Two-Way Fixed Effects (TWFE)

- With more than two time periods, statistical power increases.

$$Y_{i,t} = \alpha + \eta_i + \nu_t + \gamma D_{i,t} + \varepsilon_{i,t}$$

- η_i are unit fixed effects.
- ν_t are time fixed effects.

Event studies

- The idea is to include dummies for each post-treatment period.

$$Y_{i,t} = \alpha + \eta_i + \nu_t + \gamma_1 T1_{i,t} + \gamma_2 T2_{i,t} + \gamma_3 T3_{i,t} + \dots + \varepsilon_{i,t}$$

When treatment intensity is continuous:

$$Y_{i,t} = \alpha + \beta Intensity_i + \zeta Post_t + \delta (Intensity_i * Post_t) + \varepsilon_{i,t}$$

The parallel trends assumption

The parallel trends assumption

- The key assumption is that outcomes in the treatment and control groups would follow the same time trend in the absence of treatment.
 - This does not imply they must have the same mean outcome level.
- However, the assumption cannot be directly tested.
 - Without it, we cannot identify the ATT.

The parallel trends assumption

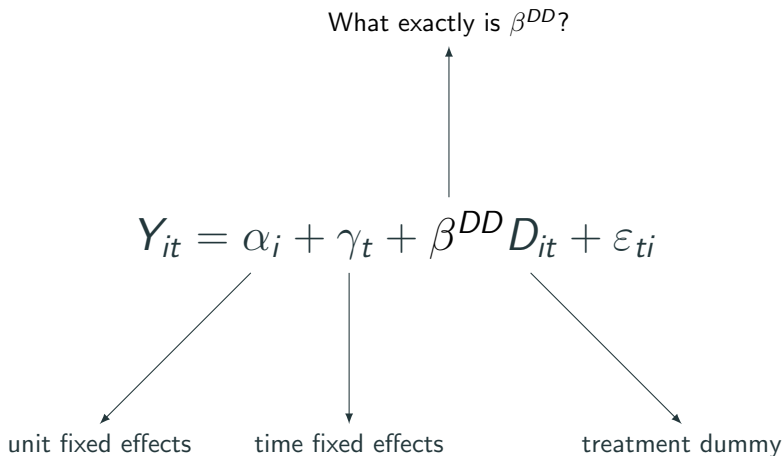
How can we defend this assumption?

1. Plot outcomes over time!
2. Falsification tests (look further into the past or into the future).
3. Control directly for time trends.

- The regression we will estimate is:

$$Y_{its} = \gamma_s + \lambda_t + \sum_{\tau=-1}^{-q} \gamma_{\tau} T_{s\tau} + \sum_{\tau=0}^m \delta_{\tau} T_{s\tau} + x_{ist} + \varepsilon_{ist}$$

- Treatment occurs at period zero.



Frisch-Waugh (1933): TWFE is equivalent to a univariate regression:

$$\tilde{Y}_{it} = \tilde{D}_{it} + \zeta_{ti}$$

where:

$$\tilde{Y}_{it} = Y_{it} - \bar{Y}_i - \left(\bar{Y}_t - \bar{\bar{Y}} \right)$$

and:

$$\tilde{D}_{it} = D_{it} - \bar{D}_i - \left(\bar{D}_t - \bar{\bar{D}} \right)$$

Standard errors

- Bertrand, Duflo, and Mullainathan (2004) point out that conventional standard errors often underestimate the standard deviation of the estimators. In other words, standard errors are biased downward.
- To address this, they propose:
 - Using bootstrap standard errors.
 - Using clustered standard errors at the group level.
 - Aggregating the data at the group level.