7. Difference-in-Differences

Empirical Evaluation of Economic Policy

Iris Kesternich[1,2] Alexander Wintzéus[2]

Universität Hamburg[1]

KU Leuven[2]

February 9, 2024

Overview

1. Panel data methods
   - Error-components model
   - Estimators
2. Difference-in-Differences designs
   - Binary designs
   - Complex designs
3. Application

Overview

Error-components model

- Traditionally, the most commonly used model to study relationships in panel data settings
- **Two-way error-components model:**

$$y_{it} = \mathbf{x}_{it}'\beta + \varepsilon_{it} \quad i \in \{0, \ldots, N\},\ t \in \{0, \ldots, T\} \qquad (1)$$

with

$$\varepsilon_{it} = \alpha_i + \gamma_t + \eta_{it}. \qquad (2)$$

Here,

- α_i: **unobserved unit-specific effect** (e.g., innate ability)
- γ_t: **unobserved period-specific effect** (e.g., national policies)
- η_{it}: remaining disturbance

Two-way fixed effects model

Plugging (2) into equation (1) gives us the **Two-way fixed effects model:**

$$y_{it} = \alpha_i + \gamma_t + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_{it} \tag{3}$$

where

- α_i: unit-specific intercept
- γ_t: period-specific intercept

Note: Effects of *unit-specific time-invariant* characteristics **and** *period-specific unit-invariant* characteristics will be absorbed by α_i and γ_t, resp.

Estimating fixed effects models

Recall, α_i and γ_t are **unobserved**, but primary interest is $\boldsymbol{\beta}$

Three possible ways to go about estimation:

- Within-estimator
- Least Squares Dummy Variables (LSDV) estimator
- First-difference estimator (will not be discussed)

Within-estimator – OWFE estimator

If $\gamma_t = 0$ for all t (no period-specific effects), (3) reduces to a **One-way fixed effects model:**

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_{it} \tag{4}$$

We will first consider the within-estimator in this simplified setting

Within-estimator – OWFE estimator

Consider transforming all variables – **including** individual fixed effects α_i – as follows (taking the dependent variable as example):

$$\tilde{y}_{it} = y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it}$$

Clearly, given this transformation, the individual-specific effects α_i will be **eliminated** from the regression:

$$\begin{aligned}\tilde{\alpha}_i &= \alpha_i - \frac{1}{T}\sum_{t=1}^{T} \alpha_{it} \\ &= 0.\end{aligned}$$

Within-estimator – OWFE estimator

The **within-estimator** is then simply defined as the *pooled* OLS estimator of the **transformed** model:

$$\tilde{y}_{it} = \tilde{\mathbf{x}}_{it}' \boldsymbol{\beta} + \tilde{\eta}_{it}$$

In the one-way fixed effects model, the within estimator is often referred to as the **One-way fixed effects (OWFE) estimator**

Within-estimator – TWFE estimator

Consider transforming all variables – **including** individual and time fixed effects α_i and γ_t – as follows (taking the dependent variable as example):

$$\check{y}_{it} = y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it} - \frac{1}{N}\sum_{i=1}^{N} y_{it} + \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} y_{it}$$

Clearly, given this transformation, **both** the individual-specific effects α_i and period-specific effects γ_t will be **eliminated** from the regression:

$$\check{\alpha}_i = \check{\gamma}_t = 0.$$

Within-estimator – TWFE estimator

Again, the **within-estimator** in this setting is defined as the *pooled* OLS estimator of the **transformed** model:

$$\check{y}_{it} = \check{\mathbf{x}}'_{it}\boldsymbol{\beta} + \check{\eta}_{it}$$

In the one-way fixed effects model, the within estimator is often referred to as the **Two-way fixed effects (TWFE) estimator**

Note: The **TWFE estimator** is historically thé most commonly used estimator in Difference-in-Differences settings

LSDV estimator

It can be shown that the TWFE estimator is *equivalent* to a particular **Least Squares Dummy Variables (LSDV) estimator**:

$$y_{it} = \sum_{j=1}^{N} \alpha_j 1\{i = j\} + \sum_{t'=1}^{T} \gamma_{t'} 1\{t = t'\} + \mathbf{x}'_{it}\beta + \eta_{it} \qquad (5)$$

This can easily be implemented in `Stata` using `regress`:

```
regress y i.unit_id i.period_id x
```

Overview

Simple DiD

Often in economics, having randomized treatments is **not possible**

- RCTs may be unfeasible due to practical or ethical reasons
- RCTs may lack *external validity*

Hence, researchers rely on **natural experiments** to estimate treatment effects

- Induced by policy changes (e.g., changes in U.S. state's minimum wage laws)
- Assignment to treatment is generally **not randomized**
- Simply comparing control and treated units leads to biased estimates for treatment effects

Simple DiD

Consider two locations $g \in \{s, c\}$ and two time periods $t \in \{0, 1\}$:

- $Y_{g,t}(0)$: Potential outcome at location g in t *without* treatment
  $\rightarrow$ employment level with low minimum wages

- $Y_{g,t}(1)$: Potential outcome at location g in t *with* treatment
  $\rightarrow$ employment level with high minimum wages

In $t = 0$, both locations g are untreated so that $Y_{g,0} = Y_{g,0}(0)$.
However, in $t = 1$:

- For the treated group s: $Y_{s,1} = Y_{s,1}(1)$
- For the control group c: $Y_{c,1} = Y_{c,1}(0)$

Simple DiD

We would like to estimate $E[Y_{s,1}(1) - Y_{s,1}(0)]$
$\rightarrow$ average effect of increasing minimum wages in location s at period $t = 1$ (Card & Krueger, 1994)

To overcome that treatment assignment to g may not be random, we could use the following **simple DiD estimator**:

$$\textbf{DiD} = (Y_{s,1} - Y_{s,0}) - (Y_{c,1} - Y_{c,0})$$

Under the assumption that in the absence of treatment, both locations g would have experienced the **same** average outcome **evolution**, the **simple DiD estimator** is unbiased:

$$E[\textbf{DiD}] = E[Y_{s,1}(1) - Y_{s,1}(0)].$$

Dynamic potential outcomes

To make the exposition slightly more general, we extend the standard potential outcomes framework

- Assume a panel of G groups observed for T periods
- Assume treatment is assigned at the (g, t) level and is binary $D_{g,t} \in \{0, 1\}$
- Assume **SUTVA** holds: Potential outcomes of group g only depend on treatments received by group g

Let $(d_1, \ldots, d_T) \in \{0, 1\}^T$ be a particular sequence of treatments for group g in all periods t, then

$$Y_{g,t}(d_1, \ldots, d_T)$$

will denote the associated potential outcomes

Assumptions

To identify treatment effects, we will rely on **two** key assumptions:

- **No anticipation:** For all groups g and all possible treatment values $(d_1, \ldots, d_T) \in \{0,1\}^T$

$$Y_{g,t}(d_1, \ldots, d_T) = Y_{g,t}(d_1, \ldots, d_t) \qquad \text{(NA)}$$

- **Parallel trends:** For all time periods $t \geq 2$

$$E\left[Y_{g,t}(\mathbf{0}_t) - Y_{g,t-1}(\mathbf{0}_{t-1})\right] \qquad \text{(PT)}$$

does not vary across groups g.

Assumptions

Sometimes, an additional assumption is made:

- **No dynamic effects:** For all groups g and all possible treatment values $(d_1, \ldots, d_T) \in \{0, 1\}^T$

$$Y_{g,t}(d_1, \ldots, d_t) = Y_{g,t}(d_t) \tag{ND}$$

In this case, **parallel trends** reduces to: For all time periods $t \geq 2$

$$E\left[Y_{g,t}(0) - Y_{g,t-1}(0)\right] \tag{PT}$$

does not vary across groups g.

Overview

Definition

Assume treatment is binary and that all groups g are treated at the same time – i.e., **no variation in treatment timing**

- Let $T_g \in \{0, 1\}$ be an indicator for treatment groups
- Assume there is at least one treated group and one control group
- Let $F \geq 2$ denote the time period at which all treatment groups become treated

Formally, this design is summarized as follows: For all groups g

$$D_{g,t} = 1\{t \geq F\} T_g. \tag{D1}$$

TWFE is simple DiD

In design (D1), the TWFE estimator is a simple DiD estimator

- Let G_0 and G_1 denote the number of control and treatment groups, resp.
- Let T_0 and T_1 denote the number of control and treatment periods, resp.

It can be shown that

$$\beta_{TWFE} = \left(\frac{1}{G_1 T_1} \sum_{g: T_g = 1, t \geq F}^{T} Y_{g,t} - \frac{1}{G_1 T_0} \sum_{g: T_g = 1, t < F}^{T} Y_{g,t} \right) \quad (6)$$

$$- \left(\frac{1}{G_0 T_1} \sum_{g: T_g = 0, t \geq F}^{T} Y_{g,t} - \frac{1}{G_0 T_0} \sum_{g: T_g = 0, t < F}^{T} Y_{g,t} \right) \quad (7)$$

TWFE is unbiased for ATT

A natural target parameter in design (D1) is the **Average Treatment Effect on the Treated (ATT)**:

$$\mathbf{ATT} = \frac{1}{G_1 T_1} \sum_{(g,t):D_{g,t}=1} E[Y_{g,t}(\mathbf{0}_{F-1}, \mathbf{1}_{t-F+1}) - Y_{g,t}(\mathbf{0}_t)].$$

If assumptions (NA) and (PT) hold, then the TWFE estimator is **unbiased** for the **ATT**:

$$E[\beta_{TWFE}] = \mathbf{ATT}.$$

Dynamic effects and event-study designs

To estimate dynamic treatment effects in design (D1), researchers have often relied on two-way fixed effects **event-study** regressions:

$$
\begin{aligned}
Y_{g,t} = \alpha_0 + \alpha_1 T_g + & \sum_{t'=1, t' \neq F-1}^{T} \gamma_t 1\{t = t'\} \\
& + \sum_{\ell=-F+2, \ell \neq 0}^{T-F+1} \beta^{\ell} 1\{t = F - 1 + \ell\} T_g + \varepsilon_{g,t}. \qquad \text{(ES1)}
\end{aligned}
$$

where $1\{t = F - 1 + \ell\} T_g$ are *relative-time* indicators equal to one **if** at period t group g has been treated for ℓ periods.

TWFE ES is simple DiD

The TWFE estimator in event-study setup (ES1) is again a simple DiD estimator:

$$\beta^{\ell}_{TWFE} = \frac{1}{G_1}\sum_{g:\, T_g=1}^{T}(Y_{g,F-1+\ell} - Y_{g,F-1}) - \frac{1}{G_0}\sum_{g:\, T_g=0}^{T}(Y_{g,F-1+\ell} - Y_{g,F-1}).$$

For $\ell \leq -1$, β^{ℓ}_{TWFE} is often referred to as a **pre-trend** or **placebo estimator**: Can be used to formally test (NA) and (PT)

TWFE ES is unbiased for ATT_ℓ

In design (D1), two-way fixed effects event-studies allow us to consider dynamic treatment effects – $\textbf{ATT}_\ell$:

$$\textbf{ATT}_\ell = \frac{1}{G_1} \sum_{g:\, T_g = 1} E[Y_{g,F-1+\ell}(\mathbf{0}_{F-1}, \mathbf{1}_\ell) - Y_{g,F-1+\ell}(\mathbf{0}_{F-1})].$$

Again, if assumptions (NA) and (PT) hold, then for all periods $\ell > 0$ the TWFE ES estimator is **unbiased** for the $\textbf{ATT}_\ell$:

$$E[\beta^\ell_{TWFE}] = \textbf{ATT}_\ell.$$

TWFE ES allows to test NA and PT

Furthermore, if assumptions (NA) and (PT) hold, then for all periods $\ell < 0$:

$$E[\beta^{\ell}_{TWFE}] = 0. \tag{8}$$

Hence, the **placebo estimators** can be used to formally test the null of (NA) and (PT) **jointly** being satisfied

Overview

Staggered adoption

Assume treatment is binary **but** groups g may get treated at different time periods – i.e., **staggered treatment adoption**

- Let F_g denote the time period at which group g becomes treated
- Assume that for untreated groups $F_g > T$
- Assume that all groups not yet treated at $t = 1$, do not all receive treatment in the same time period

Formally, in a staggered adoption design: For all groups g

$$D_{g,t} = 1\{t \geq F_g\}. \tag{D2}$$

Failure of TWFE in staggered designs

If there is variation in treatment timing across groups g and if treatment effect may vary across time t, then

- β_{TWFE} may be biased for the **ATT**
- β^{ℓ}_{TWFE} may be biased for the $\mathbf{ATT}_{\ell}$

Why? TWFE effectively estimates a **weighted sum** of treatment effects across all treated (g, t) cells with weights that may become **negative** if treatment effects vary across t

Robust estimators

Most of recent lit. using TWFE (ES) regressions are embedded in more complex designs than the simple binary treatment, no variation in treatment timing design (D1) . . .

In a staggered adoption design like (D2), estimators **robust** to heterogeneous effects across time **exist** – see, e.g., Callaway and Sant'Anna (2021) and Sun and Abraham (2021)

For an overview of (robust) estimators in more complex and general designs, see de Chaisemartin and D'Haultfoeuille (2023)

Overview

Bailey and Goodman-Bacon (2015)

What is the impact of providing access to primary care on longer-term health?

- Use the rollout of Community Health Centers (CHCs)
- CHCs can help lower **mortality among elderly** by providing accessible preventive care

Exploit the staggered adoption of CHCs across U.S. counties:

- *Our empirical strategy uses variation in when and where CHC programs were established to quantify their effects on mortality rates*
- Since CHCs are started in different counties in different time periods, effects are estimated in event-time – i.e., relative to initial rollout

Bailey and Goodman-Bacon (2015)

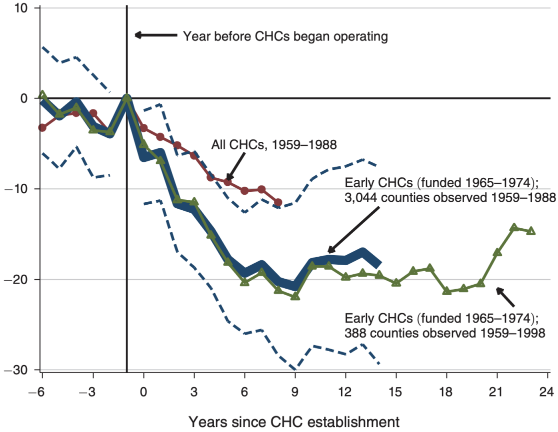
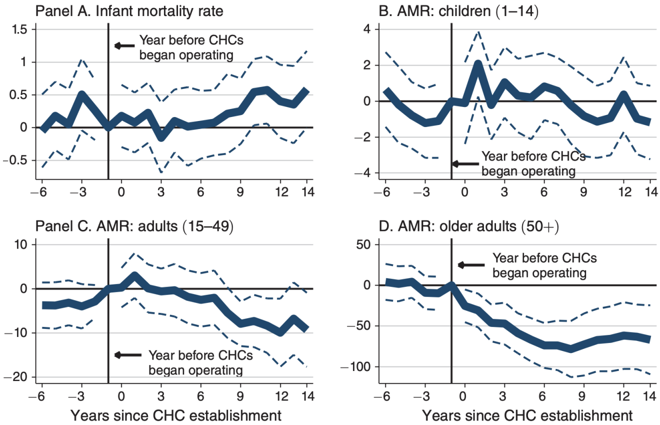


Figure 5. The Relationship between Community Health Centers and Mortality Rates

Bailey and Goodman-Bacon (2015)



References I

Bailey, M. J., & Goodman-Bacon, A. (2015). The war on poverty's experiment in public medicine: Community health centers and the mortality of older americans. *American Economic Review*, *105*(3), 1067-1104.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, *225*(2), 200-230.

Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fastfood industry in New Jersey and Pennsylvania. *American Economic Review*, *84*(4), 772-793.

de Chaisemartin, C., & D'Haultfoeuille, X. (2023). *Difference-in-differences for simple and complex natural experiments*. Available at SSRN 4487202.

References II

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, *225*(2), 175-199.