Two-way fixed effects estimation - Part 2: Solutions

Hans H. Sievertsen (h.h.sievertsen@bristol.ac.uk)

21. juni 2022

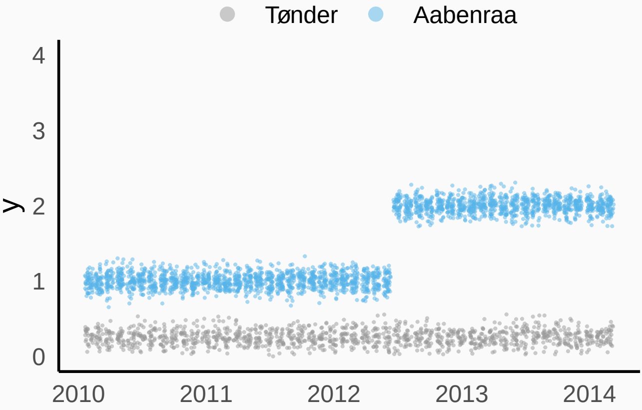
Plan for today

1. Recap: Two-way fixed effects $\neq$ difference-in-differences.
2. Apply solutions in Stata on simulated and real example.

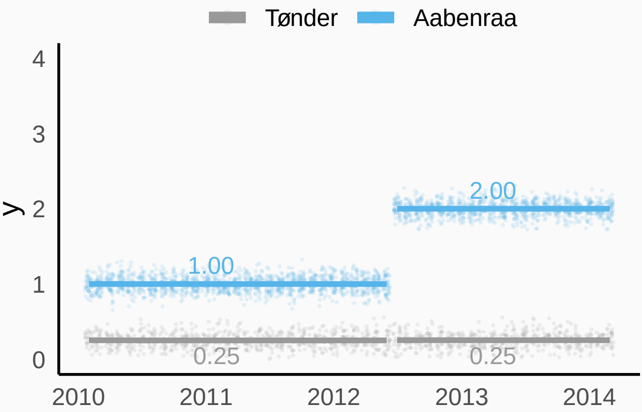
- I will share do files, slides, and zoom recording.

Recap: Difference-in-Differences - Chart

- Aabenraa Kommune is treated with an intervention in June 2012 (D=1).
- Tønder Kommune is never treated (D=0)
- Outcome of interest y
- Data for individuals in Sønderborg (N=50 individuals) and Tønder (N=25 individuals) for the period January 2010 to December 2011.



Recap: Difference-in-Differences - Means



- Difference Aabenraa: 2.00-1.00=1.00
- Difference Tønder: 0.25-0.25=0.00
- Difference-in-Differences=1.00-0.00=1.00

Recap: Difference-in-Differences - Regression

We use OLS to estimate

$$y = \beta_0 + \beta_1 treated + \beta_2 after + \beta_{DiD} after \times treated + u$$

- *after* 1 if June 2012 or later, 0 otherwise.
- *treated* 1 if Aabenraa, 0 otherwise.

```
feols(y~treated+after+afterXtreated,data=analysisdata)
```

```
## OLS estimation, Dep. Var.: y
## Observations: 3,750
## Standard-errors: IID
##               Estimate Std. Error   t value  Pr(>|t|)
## (Intercept)   0.252057   0.003756  67.09981 < 2.2e-16 ***
## treated       0.748393   0.004601 162.66940 < 2.2e-16 ***
## after         0.002826   0.005796   0.48761   0.62585
## afterXtreated 0.997073   0.007099 140.45191 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.101092   Adj. R2: 0.978482
```

Recap: Difference-in-Differences - TWFE

We use OLS to estimate the Two-Way Fixed Effects (TWFE) model

$y = \alpha + \beta_{TWFE} D + \tau' t + \mu' m + u$

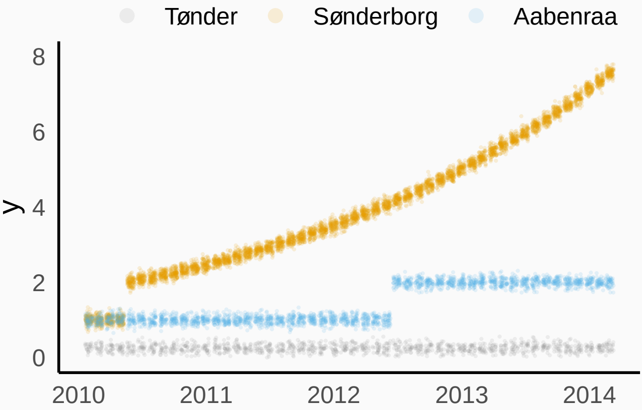
- t a vector of month dummies
- m a vector of municipality dummies
- D =1 if treated, 0 otherwise

```
feols(y~D|t+m,data=analysisdata)
```

```
## OLS estimation, Dep. Var.: y
## Observations: 3,750
## Fixed-effects: t: 50,  m: 2
## Standard-errors: Clustered (t)
##   Estimate Std. Error t value  Pr(>|t|)
## D 0.997073   0.006866 145.214 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.100425     Adj. R2: 0.97849
##                  Within R2: 0.842178
```

Recap: More data - More problems

- Sønderborg Kommune is treated with the same intervention in May 2010 and onwards.



Recap: TWFE with more groups

We again use OLS to estimate the Two-Way Fixed Effects (TWFE) model

$y = \alpha + \beta_{TWFE} D + \tau' t + \mu' m + u$

- t a vector of month dummies
- m a vector of municipality dummies
- D =1 if in Sønderborg May 2010 or later, 1 if in Aabenraa in June 2012 or later, 0 otherwise

```
feols(y~D|t+m,data=analysisdata_update)
```

```
## OLS estimation, Dep. Var.: y
## Observations: 8,750
## Fixed-effects: t: 50,  m: 3
## Standard-errors: Clustered (t)
##    Estimate Std. Error  t value Pr(>|t|)
## D -0.354829    0.38736 -0.91602  0.36414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.764277     Adj. R2: 0.856435
##                  Within R2: 0.012088
```

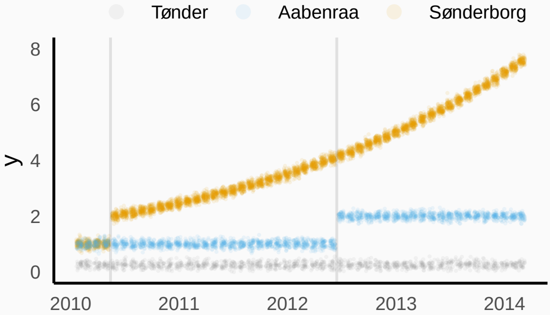
Recap: How does β_{TWFE} relate to β_{DiD}?

What is β_{TWFE} actually capturing?

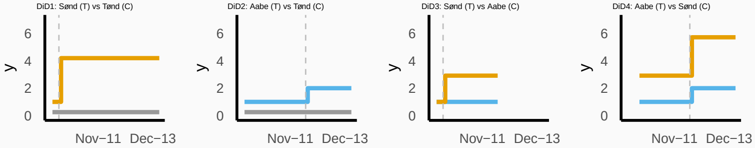
- *Goodman-Bacon (GB): "Difference-in-Differences with Variation in Treatment Timing" (2021, JoE)*
  - **Decomposition of $\hat{\beta}_{TWFE}$ in weighted $\hat{\beta}_{DiD}$s**
  - Applicable to staggered adoption designs.
- *Chaisemartin & D'Haultfoeuille (CD): "Two-way fixed effects estimators with heterogeneous treatment effects" (2020, AER)*
  - **Decomposition of $\hat{\beta}_{TWFE}$ in weighted TEs across (t, m) cells.**
  - Applicable to 2-way (e.g., group & time) fixed effects approaches.

Recap: The Goodman-Bacon Decomposition

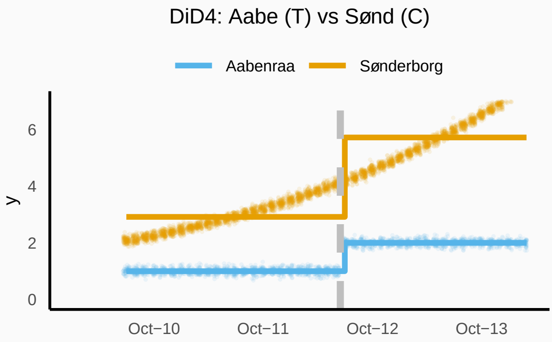
- Goodman Bacon: we can decompose



- into four 2X2 DiDs:



Recap: DiD4 the Bad Guy!



- Difference Aabenraa (Treated): 2.00-1.00=1.00
- Difference Sønderborg (Always Treated: Our Control!): 5.72-2.65=3.07
- Difference-in-Differences: 1.00-3.07=-2.00

We use Sønderborg as a control group, because it doesn't change treatment status. However, because of dynamic treatment effects (for Sønderborg), Sønderborg is a poor control because the number of periods it has been treated changes over time!

- Conclusion: Trend for Sønderborg is not a good counterfactual for Aabenraa!

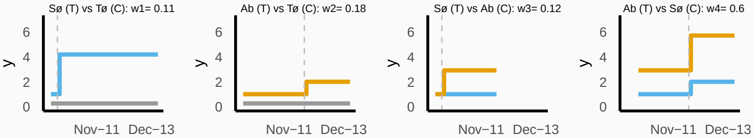
Recap: The Goodman-Bacon β_{TWFE} decomposition

β_{TWFE} ***is the weighted average across these 4 DiDs***:

$$\hat{\beta}_{TWFE} = w_1\hat{\beta}_{DiD1} + w_2\hat{\beta}_{DiD2} + w_3\hat{\beta}_{DiD3} + w_4\hat{\beta}_{DiD4}$$

- Key insight: $w \neq$ population shares, but also **depends on when a group gets treated**.
- See Theorem 1 in GB for the general expression of (3) (equation 10a in his paper) and the definition of the weights.

Recap: Goodman-Bacon Weights in our Example



Recap: Decomposing the TWFE

$$\hat{\beta}_{TWFE} = 0.106 \times 3.195 + 0.175 \times 0.997 + 0.115 \times 1.917 + 0.604 \times \text{-}1.802 = \text{-}0.35$$

- Weight Sønderborg:0.11+0.12=0.23. $N_{\text{Sø}} = 100$
- Weight Aabenraa:0.18+0.60=0.78. $N_{\text{Ab}} = 50$
- Aabenraa gets a larger weight because it is treated more in the middle.
- Not necessarily a problem if you know that it is not population weighted!

Recap: The CD decomposition

Decompose β_{TWFE} in weighted TEs across $(m,t) : D_{m,t} = 1$ cells

- A m, t cell is time (one of the fixed effects) period times area (the other fixed effect) treated unit.
- β_{TWFE} is then given by the weighted average across all these cells.

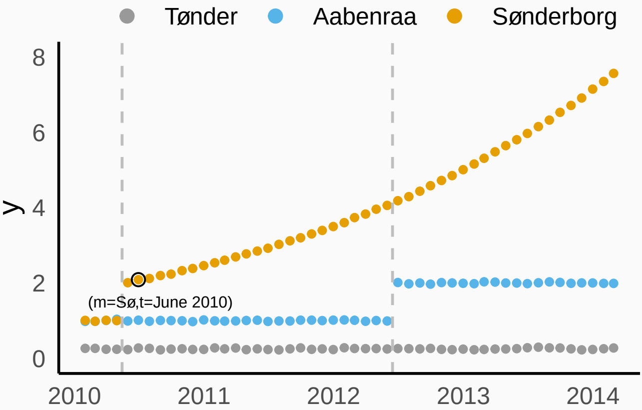
$$\beta_{TWFE} = w_{S\emptyset nderborg, May2010} TE_{S,5:2010} + w_{S,6:2010} TE_{S,6:2010} \cdots$$

$+$

- (This is more "general" than GB)
- Cells that are treated when no one else is treated => large weight
- A treated cell for a group that is rarely treated => large weight!

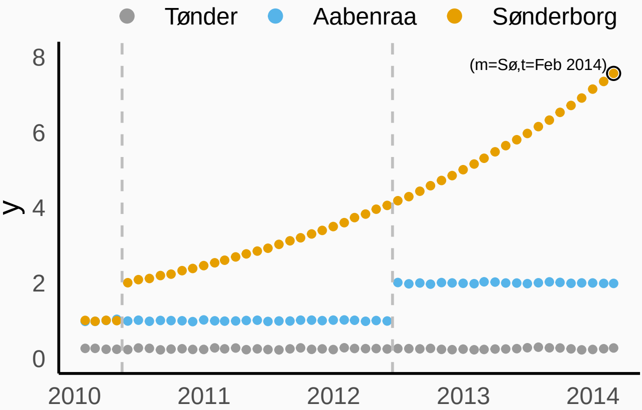
Recap: The CD decomposition

Decompose β_{TWFE} in weighted TEs across $(m, t) : D_{m,t} = 1$ cells

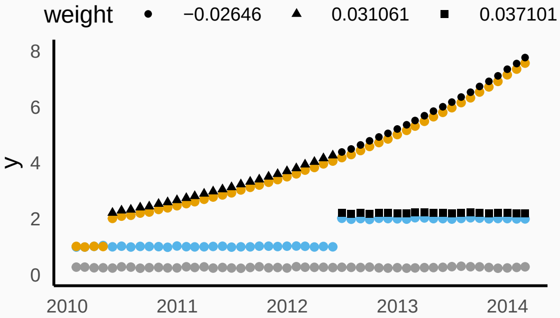


Recap: The CD decomposition

Decompose β_{TWFE} in weighted TEs across $(m, t) : D_{m,t} = 1$ cells



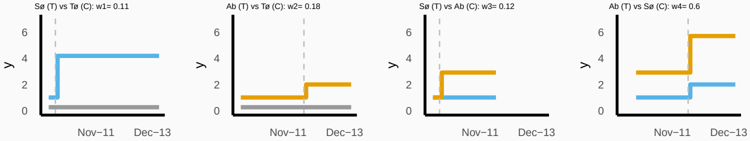
Recap: The CD weights in our example



- Sønderborg cells get negative weight after $t = 30$ because they are from a group that is mostly treated and at a time that that is mostly treated.
- Note that in contrast to GB, we cannot empirically decompose β_{TWFE} because we don't know the TEs in the m, t cells!

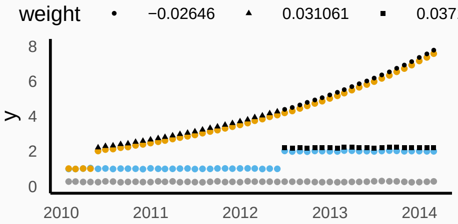
Recap: Linking GB and CD

GB weights



- Weight Sønderborg:0.11+0.12=0.23.
- Weight Aabenraa:0.18+0.60+0.78.

CD weights



- Weight Sønderborg:$-0.026 \times 21 + 0.031 \times 25$ =0.23.
- Weight Aabenraa:0.037×21=0.78.

What to do?

So you have a TWFE analysis. Ask:

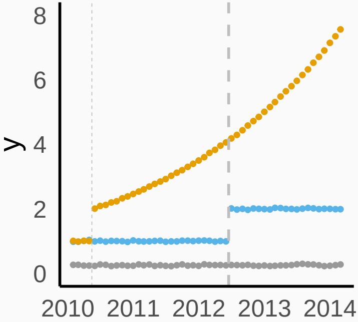
1. Are treatment effects homogeneous?
2. Are treatment effects dynamic?
3. Do you have negative weights?

Solutions: A Roadmap:

1. Ruling out dynamic and heterogeneous treatments effects a priori is often a strong assumption. So I typically assume that could be the case.
2. Calculate weights (GB and/or CD)
3. Use estimator that handles the weight issue and (potentially) allows for dynamic effects.

Solution 1: CD's estimator

- Chaisemartin & D'Haultfoeuille (CD): (2020, AER)
- Not allowing dynamic effects!
- Compares adjacent periods for switching in and out cells.
- Switching in (joiners) relies on // assumption on untreated outcomes.
- Switching out (leavers) relies on // assumption on treated outcomes.



Allowing for dynamic effects

- Goal: Estimate event-study design with OLS for effects at l periods to treatment.

$$y_{g,t} = \alpha_g + \tau_t + \sum \beta_{TWFE,l} 1\{l == 1\} + e_{g,t}$$

- This approach also suffers from the issues listed above,

$$E[\hat{\beta}_{TWFE,l}] = \sum w_{g,l} TE_{g,l} + \sum \sum w_{g,l'} TE_{g,l'}$$

1. First sum might have negative weights (very similar as before)!
2. And also contamination from other periods $l' \neq l$ treatment effect.

- Decide on how to aggregate effects across groups!

(see Sun and Abraham (2021) and Chaisemartin's "Advances in Difference-in-Differences in Econometrics" talk in December 2022).

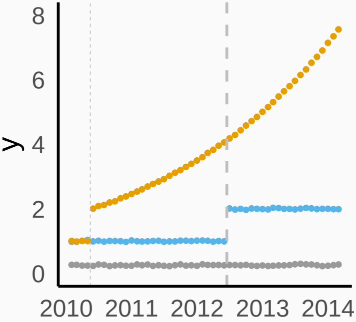
Solution 2: Callaway and Sant'Anna (2021, JE) estimator

(Sun & Abraham (2021, JE) is very much of the same spirit)

- Create groups, g, that start treatment at the same time, c.
- To get the effect of having been treated for l periods:

    1. Compute difference in y for group c between period $c+l$ and period $c-1$.
    2. Compute average difference in y for never treated groups between period $c+l$ and period $c-l$.
    3. Compute difference between 1 and 2.

- Key assumption: // based on never having been treated!

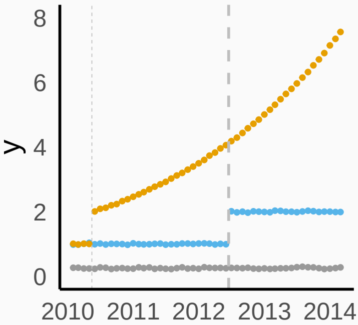


Solution 3: Borusyak, Jaravel, and Spiess (BJS) (2022)

- As CS, but to get the effect of having been treated for l periods:

  1. Compute difference in y for group c between period $c + l$ and average across all periods $t = 0$ and $c - 1$.
  2. Compute average difference in y for never treated groups between period $c + l$ and average across all periods $t = 0$ and $c - 1$.
  3. Compute difference between 1 and 2.

- More efficient than CS because it uses more data.
- But more sensitive to violation of common trends assumption!
- Bias-variance trade-off



Let's try them in Stata on our simulated exampel and on a real example

A real example

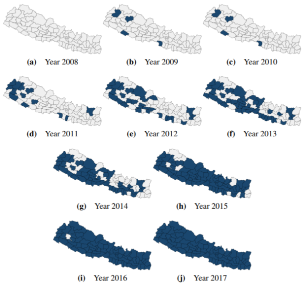


Figure 1: CHX cord application roll-out across districts over time (adopted CHX=blue).

Valente, Sievertsen, & Puri (2021)

- Staggered rollout of chlorhexidine gel (CHX) treatment across districts in Nepal.
- Estimate effect of CHX on mortality using district and month of fixed effects.