Causal Inference with Panel Data

Kee Hyun Park   Xiaonan Wang

University of Maryland

GVPT Methods Workshop

March 8, 2021

Potential outcomes framework [5,15]

Suppose I have COVID-19. Would going to the hospital improve my health?

Defining *causal effect* of "go hospital"(D) on my "health"(Y):

- 2 potential outcomes: $Y_i(D = \text{"go"})$ or $Y_i(D = \text{"stay"})$
- Individual causal effect *defined* by the difference in potential outcomes:

$$Y_i(D = \text{"go"}) - Y_i(D = \text{"stay"})$$

Fundamental problem [13]:

- We only observe one of the two potential outcomes
  $\rightarrow$ can't "go" or "stay" at the same time. The other remains as counterfactual
- Cannot infer individual causal effects

Instead, we aim to recover the **average treatment effect** with multiple units

Average Treatment Effect [15]

Assume that:

1. Treatment assigned independently from outcome and pre-treatment attributes (ignorability, or *"unconfoundedness"*; *"exogeneity"*; *"selection on observables"*; *"omitted variables"*)
2. Each unit has some chance to receive a treatment (overlap; *common support*)
3. Units receive same version of treatment (Stable Unit Treatment Value Assumption)
4. No spillover or carry-over effects across units and times
5. Treatment effect is linear

Under above assumptions, aim to recover following *causal estimands*:

- **A**verage **T**reatment **E**ffect:

$$\mathrm{E}[Y_i(\text{go}) - Y_i(\text{stay})]$$

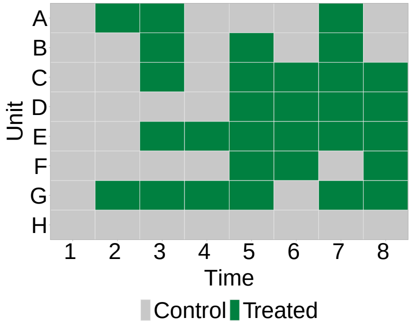
- **A**verage **T**reatment Effect on the **T**reated:

$$\mathrm{E}[Y_i(\text{go}) - Y_i(\text{stay}) \mid \underbrace{D_i = \text{“go”}}_{\text{treated group}}]$$

Observational “panel” or “TSCS” data [7]

- Panel: repeated cross-section data (different units sampled over time)
- Time-series-cross-section: “same” units *repeatedly* observed across time periods
  $\rightarrow$ Balanced panel
- Today’s methods generally apply to both cases

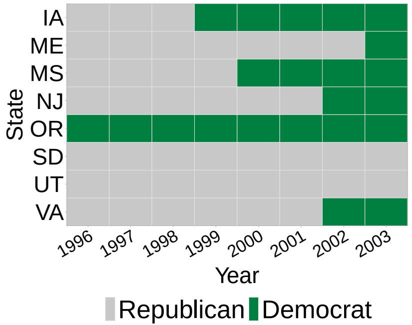
Mapping panel data onto grids



- Y-axis indicates unit (N)
- X-axis indicates time (T)
- Green cells are treated units
- Grey cells are control units

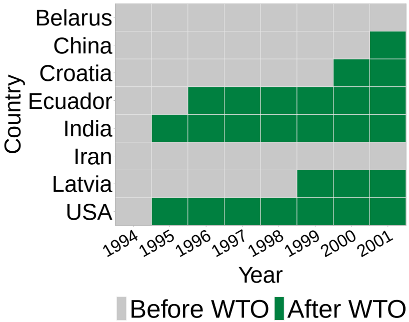
- Treatment is binary (either get treated or remain as control)
- Aim to recover **average treatment effect** or **average treatment effect for the treated** under various settings

Examples



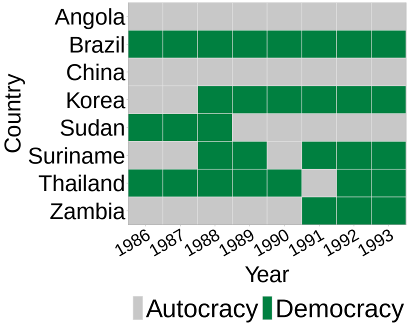
- Does a state governor's party ID affect policy performances?
  (Dynes & Holbein 2019) [11]

Examples



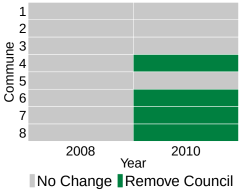
- Does WTO accession increase trade?
  (Rose 2004; Tomz, Goldstein & Rivers 2007)
  [18,20]

Examples

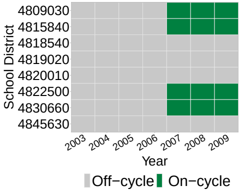


- Does a country's regime type matter for economic growth?
  (Acemoğlu et al. 2019) [1]

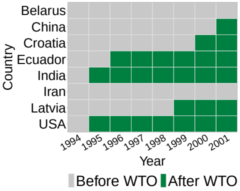
Panel/TSCS data with heterogeneous treatment paths



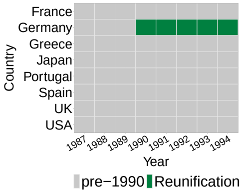
(a) Simple DiD [17]



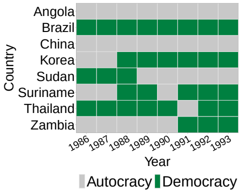
(b) Multi-period DiD [6]



(c) Staggered Adoption

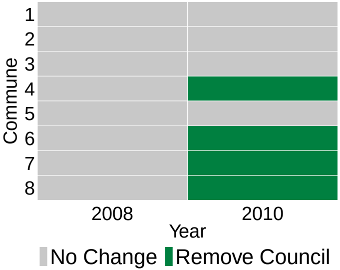


(d) Synthetic Control [3,4]



(e) Treatment Reversal [1]

Difference-in-Differences design



Malesky et al. (2014, *APSR* [17])

- Question: Does decentralization lead to better public service delivery?
- Treatment: Removal of elected councils in Vietnamese districts
- 30 outcome variables (public transport, TV broadcasting, post office, health care$\cdots$)
- Two periods (2008, 2010), treatment assigned on April 25, 2009
- Two groups (communes that removed vs. maintained councils)

Identification under a DiD design

Assumptions:

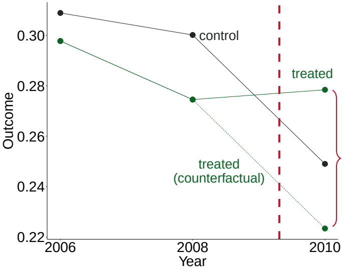
- Treatment is exogenous $\rightsquigarrow$ not random (treatment carefully assigned)
- Same treatment for each unit $\rightsquigarrow$ “Removal of councils”
- No spill-over or carry-over effect
- Linear treatment effect

Additionally, we assume the parallel trends assumption

“Parallel Trends” assumption

Assume that:

- Pre-treatment trends of the DV is same for treated & control groups



average treatment effect for the treated

$$\mathbb{E}[\underbrace{Y_{i1}(0)}_{\text{post-treat DV}} - \underbrace{Y_{i0}(0)}_{\text{pre-treat DV}} \mid \underbrace{D_i = 1]}_{\text{treated}} = \mathbb{E}[\underbrace{Y_{i1}(0)}_{\text{post-treat DV}} - \underbrace{Y_{i0}(0)}_{\text{pre-treat DV}} \mid \underbrace{D_i = 0]}_{\text{control}}$$

Estimation

The DiD estimator:

$$\begin{aligned}&(\text{E[treated units' post-treat outcomes]} - \text{E[treated units' pre-treat outcomes]})\\ &- (\text{E[control units' post-treat outcomes]} - \text{E[control units' pre-treat outcomes]})\end{aligned}$$

Under the parallel trends assumption, we compute this DiD estimator via:

$$Y_{it} \sim \alpha + \theta D_i + \gamma T_t + \beta(D_i \times T_t)$$

In R, equivalent to:

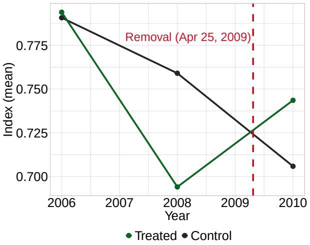
```
lm(y ~ treat + post + treat*post, data = data)
```

- `treat` (D_i) $\in \{0,1\}$ (treatment or control group)
- `post` (T_t) $\in \{0,1\}$ (pre- or post- treatment period)
- control variables can be added

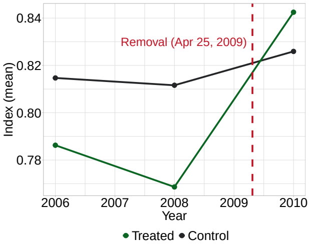
Coefficient for the treat*post (β) numerically equivalent to the DiD estimator (ATT)

Check the parallel trends assumption

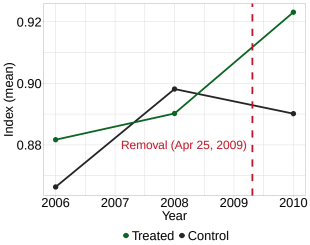
- Plot DV trends for both treated/control group:



(a) Public Transport



(b) TV Broadcasting



(c) Post Office

- What to do when the parallel trends seem implausible?
  $\Rightarrow$ Use information from multiple "pre-treatment" period (ongoing works) [21]

Summary

Difference-in-Differences:

- Special case of panel/TSCS data settings
- Importance of the parallel trends assumption
- Estimation is straightforward when all assumptions are plausible
- Linear treatment effect

Statistical software on Difference-in-Differences design

- In R:
  - `panelView` (by Xu)
  - `did` (by Callaway & Sant'Anna)
  - `DIDdesign` (by Egami & Yamauchi)
- In STATA:
  - `DID_MULTIPLEGT` (by de Chaisemartin & D'Haultfœuille)
- caution! all under development (read documentation carefully)

Two-way Fixed Effects: Setup [14]

From two units and two time periods to

- multiple units
- multiple periods
- varying treatment adoptions (staggered adoption; treatment status on and off)

Two-way Fixed Effects (2FE) Model

- Two problems
  - treatment effects heterogeneity
  - unboserved time-varying confounders
- Practical solutions

Two-way Fixed Effects: Setup [14]

$$Y_{it} = \alpha_i + \gamma_t + \beta(D_i \times T_t) + \epsilon_{it}$$

- α_i is time-invariant unit specific fixed effects (e.g. culture)
- γ_t is unit-invariant time specific fixed effects (e.g. pandemic)
- $D_i = 1$ if the unit is ever treated
- $T_t = 1$ if it is the post-treatment time periods
- β is the estimate of interests
- Covariates can be easily added

Two-way Fixed Effects: Setup [14]

2FE model and DID

- **2FE is not equivalent to DID with multiple units and time periods**
- 2FE is equivalent to DID only with two units and two time periods
- The connection between 2FE and DID will be illustrated with an example later

Two-way Fixed Effects: Setup [9]

Interpret 2FE using potential outcome framework

- Demean: compare Y_{it} with the average of other observations from the same group
- Estimated counterfactuals in 2FE model: $M_i + M_t - M_o$
- M_i: mean of other observations from the same group
- M_t: mean of other observations from the same time period
- M_o: mean of other observations from neither the same group nor the same time period
- **There is no observation that is both in the same group and the same time period**

Two-way Fixed Effects: Setup [9]

Causal assumptions

1. No unobserved time-varying confounders (parallel trend)
2. Past outcomes do not affect current treatment status (exogeneity)
3. **Linearity assumption (functional form)**
   - Even when 1 and 2 are satisfied, violation of 3 results in inconsistent estimation of β

Note

- Past outcomes are allowed to affect current outcomes: need to cluster standard errors
- Past treatments are allowed to affect currect outcomes: add lags and leads
- The leads are often used to test for parallel trend
- Trade-off between fixed effects and lagged dependent variables: assumption 2

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

Consequences

- Inconsistent estimation: β_{2fe} could be negative even when ATE for each unit is positive

Why?

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

2FE model: $Y_{it} = \alpha_i + \gamma_t + \beta D_i \times T_t + \epsilon_{it}$

Decompose β_{2fe}

- β_{2fe} is a weighted average of treatment effects of each treated unit in each time period
- The weights can be negative

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

Data generating process:

```
# Treatment assignments
DT <- list(c(0,0,1),c(0,1,1))
# Unit and time specific treatment effects
beta <- list(c(0,0,1),c(0,1,4))
# time invariant unit specific effects
alpha <- list(1.7,2.6)
# unit invariant time specific effects
gamma <- c(0,0,0)
# Data generating process for unit 1 and unit 2
Y_1 <- alpha[[1]] + gamma + beta[[1]]*DT[[1]]
Y_2 <- alpha[[2]] + gamma + beta[[2]]*DT[[2]]
```

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

```
##            id time DT   Y
## 1   Maryland    1  0 1.7
## 2   Maryland    2  0 1.7
## 3   Maryland    3  1 2.7
## 4 California    1  0 2.6
## 5 California    2  1 3.6
## 6 California    3  1 6.6
```

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

Estimate β_{fe} using two-way fixed effects model

```
# 2FE model
m <- lm(data = d_example, Y ~ DT + as.factor(id) + as.factor(time))
beta_fe <- coef(m)[2]
beta_fe

##   DT
## -0.5
```

Why is β_{fe} negative when $\beta_{MD,t=3} = 1$, $\beta_{CA,t=2} = 1$, and $\beta_{CA,t=3} = 4$ are all positive?

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

```
# calculate the weights
m_weight <- lm(data=d_example,DT ~ as.factor(id) + as.factor(time))
d_example <- d_example %>% mutate(W = m_weight$residuals)
d_example

##           id time DT   Y          W
## 1   Maryland    1  0 1.7  0.1666667
## 2   Maryland    2  0 1.7 -0.3333333
## 3   Maryland    3  1 2.7  0.1666667
## 4 California    1  0 2.6 -0.1666667
## 5 California    2  1 3.6  0.3333333
## 6 California    3  1 6.6 -0.1666667
```

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

$$\omega_{MD,t=3} = \frac{1}{6}, \omega_{CA,t=2} = \frac{1}{3}, \omega_{CA,t=3} = -\frac{1}{6}$$

After normalization, $\omega_{MD,t=3} = \frac{1}{2}, \omega_{CA,t=2} = 1, \omega_{CA,t=3} = -\frac{1}{2}$

$$\beta_{fe} = \frac{1}{2} \times 1 + 1 \times 1 - \frac{1}{2} \times 4 = -0.5$$

Why β_{fe} is negative: **heterogenous treatment effects and negative weights**

Two-way Fixed Effects: Problems [9]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

The connection between β_{fe} and DID estimator

```
##   time California Maryland
## 1    1          0        0
## 2    2          1        0
## 3    3          1        1
```

Two DID estimators: $\beta_{fe} = \dfrac{DID_1 + DID_2}{2}$

$$DID_1 = (Y_{MD,t=3} - Y_{MD,t=2}) - (Y_{CA,t=3} - Y_{CA,t=2})$$

$$DID_2 = (Y_{CA,t=2} - Y_{CA,t=1}) - (Y_{MD,t=2} - Y_{MD,t=1})$$

Two-way Fixed Effects: Problems [9,8,16]

Problem: treatment effects heterogeneity across units or over time

An example: two units and three time periods with staggered treatment adoption

The connection between β_{fe} and DID estimator

$$DID_1 = (2.7 - 1.7) - (6.6 - 3.6) = -2$$

$$DID_2 = (3.6 - 2.6) - (1.7 - 1.7) = 1$$

$$\beta_{fe} = \frac{-2 + 1}{2} = -0.5$$

Why DID_1 is negative: **treated observations (CA in time 2 and 3) are used as controls**

Two-way Fixed Effects: Problems [9,8,16]

Problem: treatment effects heterogeneity across units or over time

Summary

- 2FE estimator is a weighted average of treatment effects of each treated unit in each time period
- The weights can be negative
- β_{fe} cannot be consistenly estimated if treatment effects are heterogenous
- Units adopting the treatment earlier are more likely to receive some negative weights

Two-way Fixed Effects: Problems [9,8,16]

Problem: treatment effects heterogeneity across units or over time

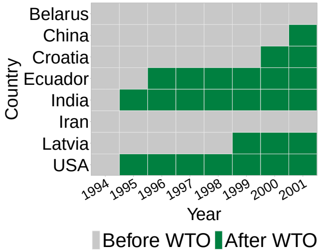
Solutions

- First, compare the treated only with untreated observations (e.g. define $DID_1 = 0$)
- Second, calculate unit and time specific treatment effects
- Third, aggregate individual effects in a way without negative weights

Two-way Fixed Effects: Problems [9,8,16]

Problem: treatment effects heterogeneity across units or over time

Solutions: matching based approach



Estimate $\beta_{Croatia,2000}$

- Find control units from Belarus, China, and Iran
- Estimate $DID_{2000,pre2000}$

Estimate $\beta_{Croatia,2001}$

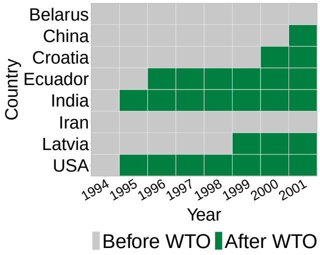
- Find control units from Belarus, Iran
- Estimate $DID_{2001,pre2000}$

Aggregate $DID_{2000,pre2000}$ and $DID_{2001,pre2000}$

Two-way Fixed Effects: Problems [5,16,19]

Problem: treatment effects heterogeneity across units or over time

Solutions: model based approach (R package: *fect*)



- Use all untreated observations (the gray ones), and fit the following 2FE model

$$Y_{it} = \alpha_i + \gamma_t + \epsilon_{it}$$

- Use the estimated $\hat{\alpha}$ and $\hat{\gamma}$ to imputate counterfactuals for treated observations (the green ones): $Y(\hat{0})_{it} = \hat{\alpha}_i + \hat{\gamma}_t$
- Calculate treatment effects for unit i at time period t: $\beta_{it} = Y_{it}^{observed} - Y(\hat{0})_{it}$
- Aggregate β_{it}

Two-way Fixed Effects: Problems [5,16,19]

Problem: treatment effects heterogeneity across units or over time

Solutions: summary

- Only use untreated observations as control units
- Calculate unit and time specific treatment effects, and then aggregate
- Flexibility in choosing control units
- Flexibility in estimating counterfactuals
- Flexibility in aggregating results

Two-way Fixed Effects: Problems [2,3,4,10]

Problem: unobserved time-varying confounders (violation of parallel trend)

Solutions

- Test for (no) pre-trend, allowing heterogenous effects (otherwise it is problematic)
- Directly model unobserved time-varying confounders
- Both need long pre-treatment periods

Two-way Fixed Effects: Problems [2,3,4,10]

Problem: unobserved time-varying confounders (violation of parallel trend)

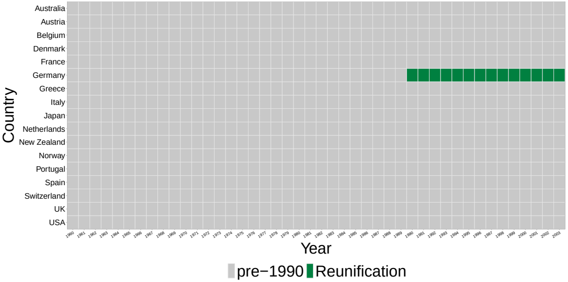
Solutions: test for (no) pre-trend, allowing heterogenous effects

- The trend between treated and control units before the adoption of treatment
- If parallel pre-treatment, more likely to be parallel post-treatment
- $\beta_{it} = Y_{it}^{observed} - \hat{Y_{it}}(0)$
- Plot β_{it} from $t = -m, -m+1, ..., -2, -1, 0, 1, 2, ..., m-1, m$
- Whether the effects in the pre-treatment periods exhibit upward or downward trend
- Visualize
- Formally test:
  - jointly zero
  - equivalent test (flip the null hypothesis)
  - placebo test (move treatment adoption backward)

Two-way Fixed Effects: Problems [2,3,4,10]

Problem: unobserved time-varying confounders (violation of parallel trend)

Solutions: model unobserved time-varying confounders (synthetic control)



- $Y_{Germany} = 0.2Y_{Italy} + 0.3Y_{France} + 0.5Y_{Austria}$
- No control alone follows a parallel trend with Germany
- But a weighted combination of Italy, France, and Austria can be seen as counterfactual outcomes to Germany
- Require a relatively long pre-treatment period

Two-way Fixed Effects: Problems

Problem: unobserved time-varying confounders (violation of parallel trend)

Solutions: model unobserved time-varying confounders (synthetic control)

$$Y_{Germany} = 0.2Y_{Italy} + 0.3Y_{France} + 0.5Y_{Austria}$$

How to understand the weights

- The outcome trajectories of some countries are more similar than others to Germany
  - DID: equal weights are assigned to each control unit
  - countries with very different trends are assigned 0 as weight
- Germany is correlated with other countries in some way
  - estimate the correlations using pre-treatment observations
  - assume that the correlations hold in post-treatment periods
  - use estimated correlations to impute counterfactuals for Germany at each post-treatment period

R package: *Synth*

Two-way Fixed Effects: Problems

Problem: unobserved time-varying confounders (violation of parallel trend)

Solutions: model unobserved time-varying confounders (synthetic control)

With multiple treated units

- Repeat the estimation for each treated unit and aggregate the effects
- Use R package *fect* with interactive fixed effects model (a factor model)

Note

- Parallel trend is not guaranteed, unless it mimics the true data generating process
- But this is easy to check (pre-trend)

Two-way Fixed Effects: Summary

Two problems

- Treatment effects heterogeneity
- unobserved time-varying confounders (violation of parallel trend)

Solutions

Other notes

- Number of N and T
- Statistical inference (standard errors): randomized inference or bootstrap

fect: Fixed Effects Counterfactual Estimators

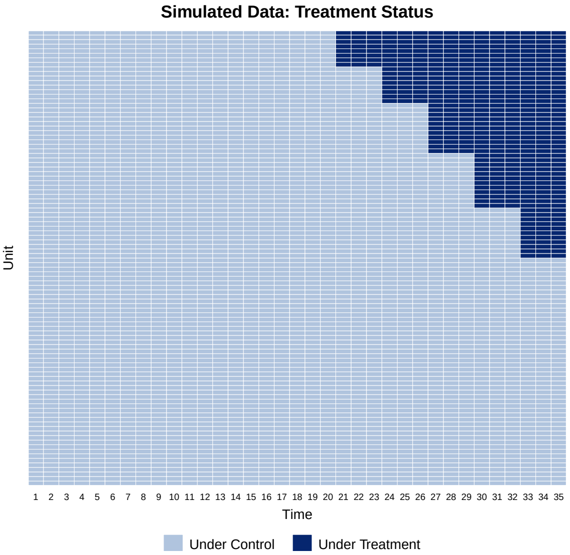
Installation and user instruction: http://yiqingxu.org/software/fect/fect.html

Reference: Licheng Liu, Ye Wang, Yiqing Xu (2019). “A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data.” Available at SSRN: https://papers.ssrn.com/abstract=3555463.

fect: Fixed Effects Counterfactual Estimators

```r
# load required packages
library(fect)
library(panelView)
# load the data
data(fect)
# view the data
panelView(Y ~ D, data = simdata1, index = c("id","time"),
          by.timing = TRUE,
          axis.lab = "time", xlab = "Time", ylab ="Unit",
          show.id = c(1:100),
          background = "white",
          main = "Simulated Data: Treatment Status")
```

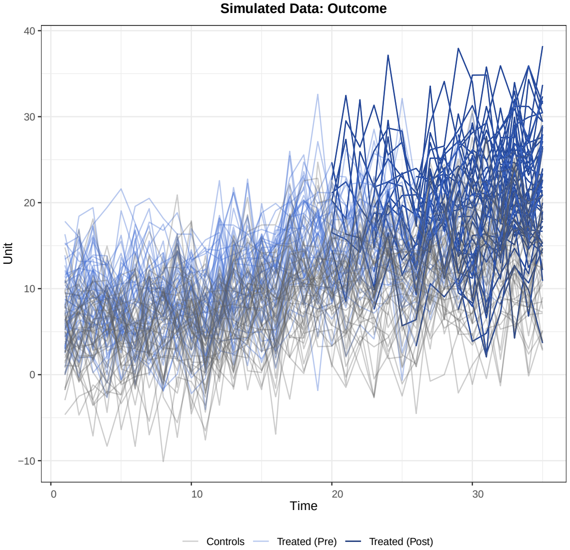
fect: Fixed Effects Counterfactual Estimators



fect: Fixed Effects Counterfactual Estimators

```
# view the outcome
panelView(Y ~ D, data = simdata1, index = c("id","time"),
  axis.lab = "time", xlab = "Time", ylab = "Unit", show.id = c(1:100),
  theme.bw = TRUE, type = "outcome", main = "Simulated Data: Outcome")
```

fect: Fixed Effects Counterfactual Estimators



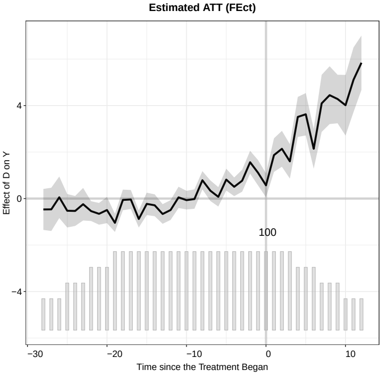
fect: Fixed Effects Counterfactual Estimators

```
# 2FE counterfactual estimator
out.fect <- fect(Y ~ D + X1 + X2,
                 data = simdata1, index = c("id","time"),
                 force = "two-way", se = TRUE,  nboots = 200)
```

```
# ATE: equally weight each observation
out.fect$est.avg

##       ATT.avg      S.E. CI.lower CI.upper p.value
## [1,] 3.489078 0.3885647 2.727505  4.25065       0

# ATE: equally weight each unit
out.fect$est.avg.unit

##      ATT.avg.unit      S.E. CI.lower CI.upper      p.value
## [1,]     2.928551 0.3678433 2.207591 3.649511 1.776357e-15
```

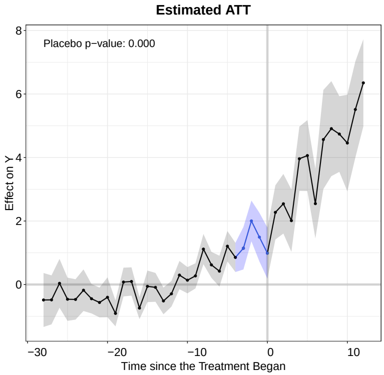
fect: Fixed Effects Counterfactual Estimators

```r
# plot the estimated ATT
plot(out.fect, main = "Estimated ATT (FEct)", ylab = "Effect of D on Y",
  cex.main = 0.8, cex.lab = 0.8, cex.axis = 0.8)
```



fect: Fixed Effects Counterfactual Estimators

```r
# a placebo test
out.fect.p <- fect(Y ~ D + X1 + X2, data = simdata1,
                   index = c("id", "time"), force = "two-way", se = TRUE, nboots = 200,
                   placeboTest = TRUE, placebo.period = c(-3, 0))
plot(out.fect.p, cex.text = 0.8)
```



fect: Fixed Effects Counterfactual Estimators

Not parallel, what can we do

- Interactive fixed effects model
  - A factor model
  - Helpful only when the data follows a factor structure
  - Need a cross validation process to choose the number of factors

```
out.ife <- fect(Y ~ D + X1 + X2, data = simdata1, index = c("id","time"),
                force = "two-way", method = "ife",
                CV = TRUE, r = c(0, 5),
                se = TRUE, nboots = 200,)
```

- Match each treated unit with similar control units from the never treated

Thank you!

Slides and other materials available from:

`https://github.com/gsa-gvpt/gvpt-methods`

Follow-up Questions:

khpark@umd.edu

or

xwang21@terpmail.umd.edu

References I

[1] Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. 2019. "Democracy Does Cause Growth." *Journal of Political Economy* 127(1): 47–100. [link]

[2] Abadie, Alberto. 2020. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Forthcoming in Journal of Economic Literature*. [link]

[3] Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490): 493–505. [link]

[4] Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59(2): 495–510. [link]

[5] Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

[6] Anzia, Sarah F. 2012. "The Election Timing Effect: Evidence from a Policy Intervention in Texas." *Quarterly Journal of Political Science* 7(3): 209–48. [link]

[7] Beck, Nathaniel. 2001. "Time-Series-Cross-Sectional Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4: 271–93. [link]

[8] Callaway, Brantly, and Pedro H. C. Sant'Anna. 2020. "Difference-in-Differences with Multiple Time Periods." *Forthcoming in Journal of Econometrics*. [link]

References II

[9] de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators With Heterogeneous Treatment Effects." *American Economic Review* 110(9): 2964–96. [link]

[10] Doudchenko, Nikolay, and Guido W. Imbens. 2017. "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis." Working paper. [link]

[11] Dynes, Adam, and John B. Holbein. 2019. "Noisy Retrospection: The Effect of Party Control on Policy Outcomes." *American Political Science Review* 114(1): 237–57. [link]

[12] Egami, Naoki, and Soichiro Yamauchi. 2021. "Using Multiple Pre-treatment Periods to Improve Difference-in-Differences and Staggered Adoption Designs." Working paper. [link]

[13] Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396): 945–60. [link]

[14] Imai, Kosuke, and In Song Kim. 2020. "On the Use of Linear Fixed Effects Regression Models for Causal Inference." *Forthcoming in Political Analysis*. [link]

[15] Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

[16] Liu, Licheng, Ye Wang, and Yiqing Xu. 2020. "A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data." Working paper. [link]

References III

[17] Malesky, Edmund J., Cuong Viet Nguyen, and Anh Tran. 2014. “The Impact of Recentralization on Public Services: A Difference-in-Differences Analysis of the Abolition of Elected Councils in Vietnam.” *American Political Science Review* 108(1): 144–68. [link]

[18] Rose, Andrew K. 2004. “Do We Really Know That the WTO Increases Trade?” *American Economic Review* 94(1): 98–114. [link]

[19] Sun, Liyang, and Sarah Abraham. 2020. “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.” *Forthcoming in Journal of Econometrics*. [link]

[20] Tomz, Michael, Judith L. Goldstein, and Douglas Rivers. 2007. “Do We Really Know That the WTO Increases Trade? Comment.” *American Economic Review* 97(5): 2005–18. [link]