

Estimating treatment effects in a staggered difference-in-differences framework

Partha Deb

Hunter College, CUNY and NBER

Edward C. Norton

University of Michigan and NBER

Jeffrey M. Wooldridge

Michigan State University

Jeffrey E. Zabel

Tufts University

October 1, 2026

Overview

- This session describes **flexdid**, a Stata command that implements a FLEXible linear model with covariates estimated by pooled OLS as an approach to estimating heterogeneous treatment effects in a staggered difference-in-differences design in the context of repeated cross-sectional data (Deb et al. 2026).

Overview

- Difference-in-differences (DID) study designs are important policy analysis tools.
- The workhorse specification is the two-way fixed effects (TWFE) model:

$$E(Y_{it} | R_g, P_t, x_{it}) = \tau Z_{it} + \sum_{g=1}^G \beta_g R_g + \sum_{t=2}^T \eta_t P_t + x_{it} \zeta$$

- Where Z_{it} is the treatment indicator, R_g are group indicators, P_t are time period indicators, and x_{it} are covariates.

Overview

- Difference-in-differences (DID) study designs are important policy analysis tools.
- The workhorse specification is the two-way fixed effects (TWFE) model:

$$E(Y_{it} | R_g, P_t, x_{it}) = \tau Z_{it} + \sum_{g=1}^G \beta_g R_g + \sum_{t=2}^T \eta_t P_t + x_{it} \zeta$$

- TWFE can produce biased estimates of the treatment effect, τ , when treatment is staggered over time and treatment effects are heterogeneous (e.g. [Goodman-Bacon \(2021\)](#)).

Overview

- Difference-in-differences (DID) study designs are important policy analysis tools.
- The workhorse specification is the two-way fixed effects (TWFE) model:

$$E(Y_{it} | R_g, P_t, x_{it}) = \tau Z_{it} + \sum_{g=1}^G \beta_g R_g + \sum_{t=2}^T \eta_t P_t + x_{it} \zeta$$

- TWFE can produce biased estimates of the treatment effect, τ , when treatment is staggered over time and treatment effects are heterogeneous (e.g., [Goodman-Bacon \(2021\)](#)).
- Several solutions have been proposed: [Borusyak et al. \(2024\)](#); [de Chaisemartin and D'Haultfoeuille \(2020\)](#); [Callaway and Sant'Anna \(2021\)](#).

Empirical Example: Medicaid Expansion

- Question: What is the effect of Medicaid expansion on working for low income/low education Americans?

Empirical Example: Medicaid Expansion

- Question: What is the effect of Medicaid expansions on working for low income/low education Americans?
- Before the Patient Protection and Affordable Care Act (ACA), millions of Americans lacked access to health insurance.
- One of the ACA's provisions was to expand Medicaid eligibility.

Empirical Example: Medicaid Expansion

- Question: What is the effect of Medicaid expansions on working for low income/low education Americans?
- Before the Patient Protection and Affordable Care Act (ACA), millions of Americans lacked access to health insurance.
- One of the ACA's provisions was to expand Medicaid eligibility.
- The expansion increased Medicaid coverage eligibility to include nearly all adults with income up to 138% of the federal poverty level.

Empirical Example: Medicaid Expansion

- Question: What is the effect of Medicaid expansions on working for low income / low education Americans?
- Before the Patient Protection and Affordable Care Act (ACA), millions of Americans lacked access to health insurance.
- One of the ACA's provisions was to expand Medicaid eligibility.
- The expansion increased Medicaid coverage eligibility to include nearly all adults with income up to 138% of the federal poverty level.
- Following a 2012 Supreme Court ruling, states could decide whether to implement the formerly required, expansion.

Empirical Example: Medicaid Expansion

- In this example, States are groups and Cohorts are groups of states that enacted Medicaid expansion in the same year

Cohort	Number of States	Percent	Cum.
0	10	19.61	19.61
2014	27	52.94	72.55
2015	3	5.88	78.43
2016	2	3.92	82.35
2019	2	3.92	86.27
2020	3	5.88	92.16
2021	2	3.92	96.08
2023	2	3.92	100
Total	51	100	

Why FLEX/flexdid?

■ Already in Stata:

- | | |
|-------------------------------------|--------------------|
| • Callaway and Sant'Anna: | csdid |
| • Sun & Abraham: | eventstudyinteract |
| • de Chaisemartin & D'Haultfœuille: | did_multiplegt |
| • Borusyak, Jaravel & Spiess: | did_imputation |
| • Gardner: | did2s |
| • Wooldridge: | jwddid |
| • Goodman-Bacon: | bacondecomp |
| • Cengiz et al.: | stackedid |

And

xthdidregress

Why FLEX/flexdid?

- Easy and transparent – one OLS regression.
- It delivers consistent estimates (given the appropriate assumptions).
- Allows for considerable functional form flexibility.
- Can be used with repeated cross sectional or panel data.
- Provides access to the OLS toolbox for inference and specification testing.

Why FLEX/flexdid?

- The standard approach is to estimate cohort by time treatment effects.
- flexdid allows one to estimate models with group by time treatment effects
 - a very important distinction if treatment is heterogeneous across groups within cohorts.
- For example, Alaska, Indiana, and Pennsylvania enacted Medicaid expansion in 2015 and it is likely that the impact on working for low income/low education households is quite different.

Why FLEX/flexdid?

- FLEX handles covariates in a very general (flexible) way
- Allows covariates (X) to be fully interact with
 - group/cohort-time treatment effects,
 - group/cohort indicators,
 - and time indicators.And hence allows for treatment heterogeneity/moderating effects.
- Or, as if often done, allow for only additive X.

Objectives

- We prove that asymptotically unbiased estimates of group-time heterogeneous treatment effects can be obtained in the repeated cross sectional setting by an imputation method.
- This imputation estimator extends [Borusyak et al. \(2024\)](#) and [Wooldridge \(2021\)](#)

Objectives

- We prove that asymptotically unbiased estimates of group-time heterogeneous treatment effects can be obtained in the repeated cross-section setting by an imputation method.
- This imputation estimator extends [Borusyak et al. \(2024\)](#) and [Wooldridge \(2021\)](#)
- We show that the estimates from this imputation estimator are identical to those from our FLEX model.
- Inference using FLEX is easy to carry out (no need to bootstrap standard errors).

Objectives

- We demonstrate the use of FLEX in the context of policy-relevant empirical examples.
- We illustrate the command and its post-estimation features using an empirical application that examines the effects of recent Medicaid expansion in the United States (among individuals of low socioeconomic status) on the probability of working.

Objectives

- We demonstrate the use of FLEX in the context of policy-relevant empirical examples.
- We illustrate the command and its post-estimation features using an empirical application that examines the effects of recent Medicaid expansion in the United States (among individuals of low socioeconomic status) on the probability of working.
- We compare the results from FLEX with other approaches including
 - Callaway and Sant'Anna (2021), **CS**
 - de Chaisemartin and D'Haultfœuille (2020), **dCdH**
 - Cengiz et al. (2019), **stacked DID**

Population Setting

- T time periods, indexed by $t = 1, 2, \dots, T$.
- Interventions (treatments) occur at the group level: $g = 1, 2, \dots, G$.
- A cohort, c , is a subset of groups that is first treated in period c : $1 < c \leq T$.
- Let $c(g)$ denote the cohort that group g belongs to.

Population Setting

- In our example, a group is a state that implements a policy (like ACA Medicaid expansion).
- FLEX allows for treatment effects at the group or cohort level.

Population Setting

- Let $\bar{g}$ groups receive treatment by time period $Q < T$.
- Treatment assignment is absorbing.
- No group is treated in $t=1$ and not all periods need to have new cohorts.
- Without loss of generality, let groups $g = 1, 2, \dots, \bar{g}$ index the treated groups ordered by their cohort membership.
- Never-treated groups $\bar{g} + 1, \dots, G$ are in cohort $c = \infty$.

Definitions

- Define potential outcomes by treatment group and time,

$$Y_t(g), g \in \{1, \dots, G\}, t \in \{1, 2, \dots, T\}.$$

- The ATETs of interest in staggered DID settings are the mean differences in the potential outcomes using $Y_t(\infty)$ as the reference outcome in a treated period t :

$$\tau_{gt} = E[Y_t(g) - Y_t(\infty) | R_g = 1], t = c(g), \dots, T; g = 1, \dots, \bar{g}$$

- The τ_{gt} 's can then be aggregated by group, calendar time, event time, or overall.

Assumption: Stable Unit Treatment Value Assumption (SUTVA)

$$Y_{it}(g_1, \dots, g_{i-1}, g_i, g_{i+1}, \dots) = Y_{it}(g_i), t = 1, \dots, T,$$

where g_i is the group assignment for unit i .

- For each unit i in the population, the potential outcome depends only on unit i 's assignment and not on the assignment of the other units.
- SUTVA ensures that we observe one potential outcome for each unit i corresponding to i 's treatment assignment.
- It also rules out spillover effects (though these can be specified in FLEX).

Assumption: No Bad Controls (NBC)

For a $1 \times K$ vector of covariates $\{\mathbf{X}_t(g) : t = 1, \dots, T; g = 1, \dots, G\}$, which may be time-varying, the covariates are the same across all potential treatment assignments: $\mathbf{X}_t(g) = \mathbf{X}_t(\infty) = \mathbf{X}_t$ for all $g \in \{1, \dots, G\}$ and $t = 1, \dots, T$.

- NBC implies that the covariates vary exogenously with treatment status.
- Covariates do not change with treatment assignment.
- Allowing the X's to time vary, means that we can include time-varying predictors of the outcome whose paths are not influenced by treatment.

Assumption: Conditional No Anticipation (CNA)

For groups $g \in 1, 2, \dots, \bar{g}$ and $t \in \{1, \dots, c(g) - 1\}$,

$$E[Y_t(g) | R_1, \dots, R_G, \mathbf{X}_t] = E[Y_t(\infty) | R_1, \dots, R_G, \mathbf{X}_t].$$

- Rules out anticipatory changes in the potential outcomes prior to the intervention.
- Rules out treatment assignment on the basis of pre-treatment outcomes.

Assumption: Conditional Parallel Trends (CPT)

- Let $R_g \in R_1, \dots, R_{\bar{g}}, R_{\bar{g}+1}, \dots, R_G$ denote a binary group-membership indicator.
- Let $P_t \in P_1, P_2, \dots, P_T$ denote binary time-period indicators.

For $t = 1, 2, \dots, T$,

$$E[Y_t(\infty) | R_1, \dots, R_G, \mathbf{X}_t] = \sum_G \beta_g R_g + \sum_G (R_g \cdot \mathbf{X}_t) \gamma_g + \mathbf{X}_t \pi_t + \eta_t.$$

The terms $\mathbf{X}_t \pi_t$ and $(R_g \cdot \mathbf{X}_t) \gamma_g$ allow relaxation of the usual parallel trends assumption.

Proposition: Flexible Linear Estimator with X (FLEX)

Using all of the data, consider the regression below that produces estimates $\tilde{\tau}_{gt}$

$Y_{i,t(i)}$ on $\mathbf{X}_{i,t(i)}$,

$$\begin{aligned} &R_{i,t(i),1}, \dots, R_{i,t(i),G}, R_{i,t(i),1} \cdot \mathbf{X}_{i,t(i)}, \dots, R_{i,t(i),G} \cdot \mathbf{X}_{i,t(i)}, \\ &P_{2,t(i)}, \dots, P_{T,t(i)}, P_{2,t(i)} \cdot \mathbf{X}_{i,t(i)}, \dots, P_{T,t(i)} \cdot \mathbf{X}_{i,t(i)} \\ &R_{i,t(i),q} \cdot P_{q,t(i)}, \dots, R_{i,t(i),q} \cdot P_{Q,t(i)}, \dots, R_{i,t(i),Q} \cdot P_{Q,t(i)}, \\ &R_{i,t(i),q} \cdot P_{q,t(i)} (\mathbf{X}_{i,t(i)} - \bar{\mathbf{X}}_{gt(i)}), \dots, R_{i,t(i),Q} \cdot P_{Q,t(i)} (\mathbf{X}_{i,t(i)} - \bar{\mathbf{X}}_{gt(i)}) \end{aligned}$$

where $\bar{\mathbf{X}}_{gt(i)}$ are the group-time means.

Proposition: Flexible Linear Estimator with X (FLEX)

Then

- (i) For all g and t , $\tilde{\beta}_j = \hat{\beta}_j$, $\tilde{\eta}_j = \hat{\eta}_j$, $\tilde{\gamma}_r = \hat{\gamma}_r$, and $\tilde{\pi}_r = \hat{\pi}_r$.
- (ii) For all $g \in \{g = 1, \dots, \bar{g}\}$ and $t \in \{c, \dots, T\}$,

$$\tilde{\tau}_{gt} = \hat{\tau}_{gt} \square$$

- There is a linear-in-parameters regression specification, FLEX, that when estimated by OLS is identical to the imputation estimator, $\hat{\tau}_{gt}$.

Proposition: Group-time Heterogeneity with lags and leads

In this specification, the τ_{gt} 's are the ATET's

$$\begin{aligned} E(Y_{it} | R_g, P_t, \mathbf{X}_{it}) = & \sum_{g=1}^{\bar{g}} \sum_{t=q}^T \tau_{gt} R_g P_t + \sum_{g=1}^{\bar{g}} \sum_{t=q}^T R_g P_t \cdot (\mathbf{X}_{it} - \bar{\mathbf{X}}_{gt}) \kappa_{gt} && \text{lags} \\ + \sum_{g=1}^{\bar{g}} \sum_{t=1}^{q-2} \tau_{gt} R_g P_t + \sum_{g=1}^{\bar{g}} \sum_{t=1}^{q-2} R_g P_t \cdot (\mathbf{X}_{it} - \bar{\mathbf{X}}_{gt}) \kappa_{gt} && \text{leads} \\ + \sum_{g=1}^G \beta_g R_g + \sum_{t=1}^T \eta_t P_t + \sum_{g=1}^G R_g \cdot \mathbf{X}_{it} \gamma_g + \sum_{t=1}^T P_t \cdot \mathbf{X}_{it} \pi_t + \mathbf{X}_{it} \zeta \end{aligned}$$

lags only model sets all pre-treatment effects = 0.

lags and leads model sets only the base-period pre-treatment effect = 0.

FLEX: Aggregating the Treatment Effects

- Overall ATET

$$\hat{\tau} = N^{-1} \sum_{g=1}^{\bar{g}} \sum_{t=c}^T N_{gt} \hat{\tau}_{gt} (R_g \cdot P_t)$$

where N_{gt} is the number of observations in the g^{th} treated group in the t^{th} time period and $N = \sum_{g=1}^{\bar{g}} \sum_{t=c}^T N_{gt}$.

- One can also aggregate the effects to the cohort, time, or event-time level.

The flexdid Command

■ Syntax

`flexdid ovar covarlist [if] [in] [weight], tx(txvar) group(groupvar) time(timevar) [options]`

- `ovar` outcome of interest.
- `covarlist` specifies the model's covariates and may contain factor variables.
- `txvar` a binary variable indicating the observations subject to treatment.
- `groupvar` categorical variable that indicates the group level at which treatment occurs.
- `timevar` integer-valued time variable.

The `tx()`, `group()`, and `time()` options are required.

The flexdid Command

- Options (some)
- **xnotinteractacted(xnotintvarlist):**
specifies covariates that are only included in the regression additively. xnotintvarlist must be distinct from covarlist.
- **verbose**
displays the output of the underlying OLS regression. Allows the user to see the table of flexdid regression estimates.
- **vce(vcetype)**
specifies the type of standard errors reported, which includes vce(cluster clustvar) that allows for intragroup correlation, and vce(robust) that allows for heteroskedasticity. vce(cluster groupvar) is the default.

The flexdid Command

- Post-estimation Commands
- `estat atet, required_option [options]`

where the required option can be overall, byget, byexposure, bycalendar, bycohort, or bygroup.

- **overall[(overall_list)]:**
estimates the overall ATET over all exposure time periods in **overall_list**. The list can contain negative integers, which indicate leads or pretreatment exposure time periods.
- Specifying overall without arguments estimates the overall ATET over all treated exposure time periods, i.e., from 0 to the last exposure time period observed in the data. This option does not produce a graph, i.e., it implies nograph.

The flexdid Command

- Post-estimation Commands: `required_option`.
- **byexposure[(exposure list)]**
estimates ATETs for each exposure time period in **exposure list**. The list can contain negative integers, which indicate leads or pretreatment periods appropriate for use in the lags-and-leads specification.
- Specifying **byexposure** without arguments estimates ATETs for all exposure time periods, i.e., from 0 in the lags-only specification and the earliest pre-treatment exposure period in the lags and-leads specification, to the latest exposure time observed in the data.

Empirical Example: Medicaid Expansion

- Question: What is the effect of Medicaid expansion on whether low income/low education Americans work?

Empirical Example: Medicaid Expansion

- We use individual-level data from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS).
- In the ASEC, the question about having worked during the year, our outcome of interest, is asked about the previous calendar year.
- Therefore, while we use data from the 2011-2024 surveys, our analysis period spans 2010-2023. This is a repeated cross sectional dataset.
- We limit our analyses to respondents between 25-59 years old with a high school diploma or less with annual incomes between 100%-300% of the federal poverty level.

Empirical Example: Medicaid Expansion

Mean Values for Model Variables		
Variable	Non-Expansion States	Expansion States
Worked=1	0.70	0.71
Age	41.97	41.77
Female=1	0.48	0.48
Minority=1	0.58	0.46
High School=1	0.70	0.71
Number of Obs	39,102	110,927

Empirical Example: Medicaid Expansion

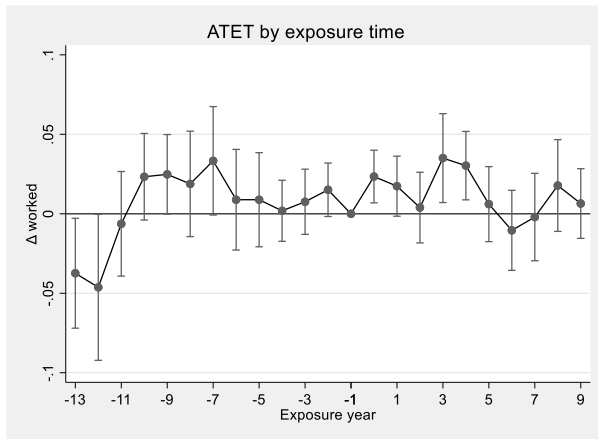
- flexdid worked, tx(expansion) group(statefips) time(year)
specification(lagsandleads)

Overall ATET				Number of obs = 150,029		
Outcome	Estimate	std. err.	t-stat	P> t	[95% conf. interval]	
Worked	0.01745	0.00707	2.47	0.017	0.00326	0.03164
Mean of Worked = 0.705						

- This is a $\frac{0.01745}{0.70488} \cdot 100\% = 1.745\%$ increase in working.

Lags and Leads Specification: ATET By Exposure Time

```
estat atet, byexposure graph(name(flex_llag_atet_et, replace) note("Data source: The Current Population Survey's Annual Social and Economic Supplement") title("The Effect of the ACA Expansions on Working"))
```



Model Comparison: CS vs FLEX: 2014 Cohort

exposure	CS		FLEX	
	Coef Est	Std Error	Coef Est	Std Error
-4	0.000545	0.009996	0.000545	0.010099
-3	0.006058	0.010074	0.006058	0.010178
-2	0.014518	0.00889	0.014518	0.008981
-1	0	0	0	0
0	0.023327	0.011657	0.023327	0.011777
1	-0.00127	0.011998	-0.00127	0.012122
2	0.038787	0.016692	0.038787	0.016865
3	0.033478	0.01159	0.033478	0.01171
...				
7	0.023319	0.014712	0.023319	0.014864
8	0.002885	0.010685	0.002885	0.010795
9	-0.00953	0.019866	-0.00953	0.020071
10	-0.00833	0.01728	-0.00833	0.017458

Model Comparison: CS and dCdH vs FLEX

Overall ATET				Number of obs = 150,029		
Model Specification	Estimate	std. err.	t-stat	P> t	[95% conf. interval]	
FLEX - group level	0.01745	0.00707	2.47	0.017	0.00326	0.03164
FLEX - cohort level	0.01692	0.00707	2.48	0.017	0.00319	0.03064
CS	0.01595	0.00687	2.32	0.02	0.00247	0.02949
dCdH	0.01692	0.01459	1.16	0.246	-0.01168	0.04551
Mean of worked = 0.705						

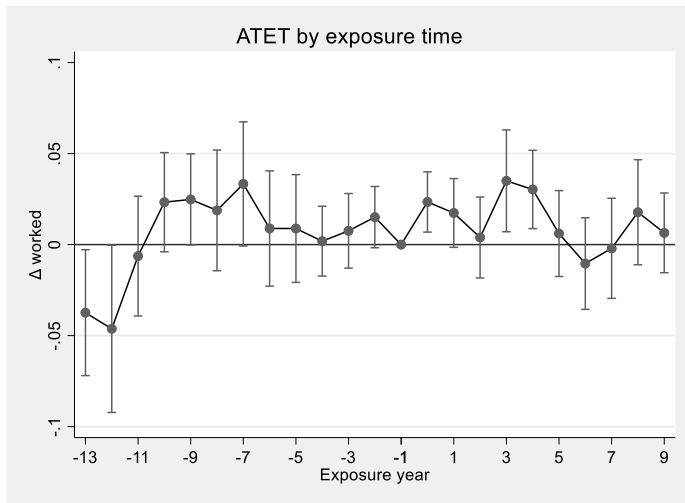
Model Comparison: CS and Stacked DID

Overall ATET		Number of obs: FLEX = 150,029, Stacked DID = 384,641				
Model Specification	Estimate	std. err.	t-stat	P> t	[95% conf. interval]	
FLEX - exposure level	0.01486	0.00631	2.36	0.018	0.00251	0.02722
Stacked DID	0.01643	0.00613	2.68	0.007	0.00441	0.02846
Mean of worked = 0.705						

Moderating Effects of Covariates using FLEX

Overall ATET				Number of obs = 150,029		
Subgroup	estimate	standar error	t-stat	P> t	[95% conf. interval]	
male	0.02267	0.01021	2.22	0.031	0.00217	0.04317
female	0.00628	0.01086	0.58	0.565	-0.01552	0.02809

Testing for Parallel Pre-trends



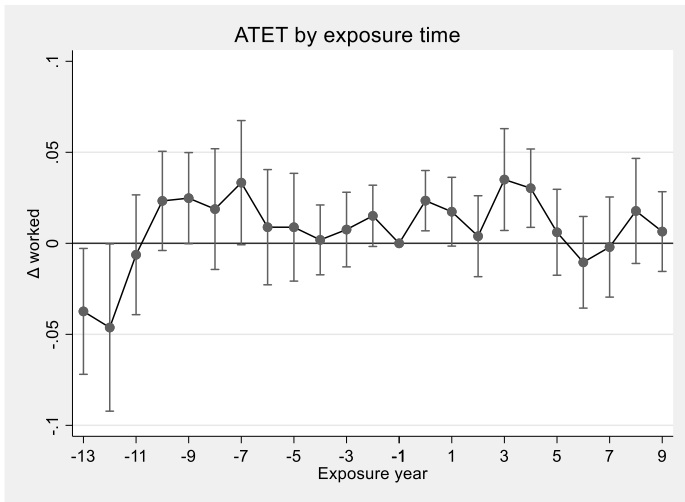
Testing for Parallel Pre-trends

- Test that the impacts of the leads are jointly 0 ($e = -2$ to -13)

$$\left. \begin{array}{l} H_0: \tau_{-2} = \tau_{-3} = \tau_{-13} = 0 \\ H_1: \text{not}(\tau_{-2} = \tau_{-3} = \tau_{-13} = 0) \end{array} \right\} \text{reject } H_0 \text{ if } F > F_{12,50,\alpha}$$

- Result: $F(10, 50) = 2.763$ Prob $> F = 0.0059$

Testing for Parallel Pre-trends



Testing for Parallel Pre-trends

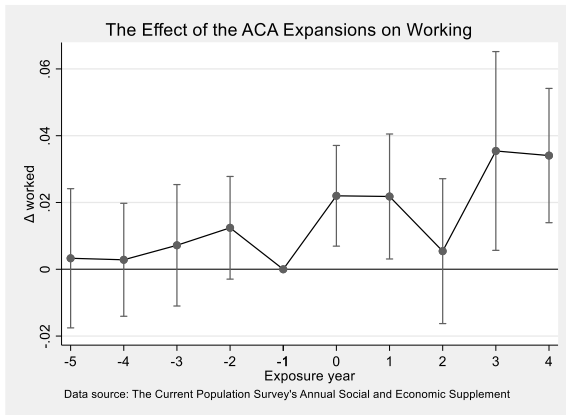
- Restrict the sample to exposure time = (-5,4)
- Test that the impacts of the leads are jointly 0 (e = -2 to -5)

$$\left. \begin{array}{l} H_0: \tau_{-2} =: \tau_{-3} =: \tau_{-5} = 0 \\ H_1: \text{not}(\tau_{-2} =: \tau_{-3} =: \tau_{-5} = 0) \end{array} \right\} \text{reject } H_0 \text{ if } F > F_{4,50,\alpha}$$

- Result: $F(4, 50) = 0.810$ Prob > F = 0.5248

FLEX: Exposure Year = (-5,4)

- estat atet, overall(0(1)4)
- ATET/p-value = 0.02317/0.003



END

Thank you!

Estimation by Imputation

- Under assumptions CNA and CPT, the parameters in CPT are identified using the untreated observations.

1. Use the untreated observations to estimate the parameters of the following regression by OLS:

$Y_{i,t(i)}$ on

Estimation by Imputation

- Under assumptions CNA and CPT, the parameters in CPT are identified using the untreated observations.

1. Use the untreated observations to estimate the parameters of the following regression by OLS:

$$Y_{i,t(i)} \text{ on } R_{i,t(i),1}, \dots, R_{i,t(i),G}, P_{2,t(i)}, \dots, P_{T,t(i)}, \mathbf{X}_{i,t(i)},$$

Estimation by Imputation

- Under assumptions CNA and CPT, the parameters in CPT are identified using the untreated observations.

1. Use the untreated observations to estimate the parameters of the following regression by OLS:

$$Y_{i,t(i)} \text{ on } R_{i,1}, \dots, R_{i,G}, P_{2,t(i)}, \dots, P_{T,t(i)}, \mathbf{X}_{i,t(i)}, \\ R_{i,1} \cdot \mathbf{X}_{i,t(i)}, \dots, R_{i,G} \cdot \mathbf{X}_{i,t(i)},$$

Estimation by Imputation

- Under assumptions CNA and CPT, the parameters in CPT are identified using the untreated observations.

1. Use the untreated observations to estimate the parameters of the following regression by OLS:

$$Y_{i,t(i)} \text{ on } R_{i,1}, \dots, R_{i,t(i),G}, P_{2,t(i)}, \dots, P_{T,t(i)}, \mathbf{X}_{i,t(i)}, \\ R_{i,t(i),1} \cdot \mathbf{X}_{i,t(i)}, \dots, R_{i,t(i),G} \cdot \mathbf{X}_{i,t(i)}, \\ P_{2,t(i)} \cdot \mathbf{X}_{i,t(i)}, \dots, P_{T,t(i)} \cdot \mathbf{X}_{i,t(i)}$$

Estimation by Imputation

2. For observation i at time t , impute $Y_{i,t}(\infty)$ as

$$\hat{Y}_{i,t(i)}(\infty) = \sum_{g=1}^G \hat{\beta}_g R_{i,t(i),g} + \sum_{g=1}^G (R_{i,t(i),g} \cdot \mathbf{x}_{i,t(i)}) \hat{\gamma}_g + \sum_{t=2}^T \hat{\eta}_t \mathbf{P}_{t(i)} + \sum_{g=1}^G (\mathbf{P}_{t(i)} \cdot \mathbf{x}_{i,t(i)}) \hat{\pi}_t + \mathbf{x}_{i,t(i)} \hat{\zeta}$$

3. For treatment group g , in period t , obtain

$$\begin{aligned}\hat{\tau}_{gt} &= N_{gt}^{-1} \sum_{t=1}^T R_{i,t(i),g} \cdot 1[t(i) = t] \cdot [Y_{i,t(i)}(g) - \hat{Y}_{i,t(i)}(\infty)] \\ &\equiv N_{gt}^{-1} \sum_{t=1}^T R_{i,t(i),g} \cdot 1[t(i) = t] \cdot T\hat{E}_{i,t(i)} \\ &= \bar{Y}_{gt} - N_{gt}^{-1} \sum_{t=1}^T R_{i,t(i),g} \cdot 1[t(i) = t] \cdot \hat{Y}_{i,t(i)}(\infty)\end{aligned}$$

where N_{gt} is the number of units in treatment group g at time t

Estimation by Imputation

- With a sufficient number of observations in each (g, t) cell, $\hat{\tau}_{gt}$ will have good statistical properties.
- The multi-step nature of the estimation makes inference challenging.
- [Borusyak et al. \(2024\)](#) use this imputation method in the panel data setting.
- Next, we prove that FLEX is identical to the imputation method for repeated cross-section study designs.