Problem Set 06
Difference in Differences (DiD)

TAs: Aida Hatami, Sam Fathinejad, Fatemeh Salehi, Reza Sahour

May 8, 2025

Question 1

a) Staggered DID: 2) This question is based on Carey, Miller, and Wherry (2020) *"The Impact of Insurance Expansions on the Already Insured: The Affordable Care Act and Medicare"*, American Economic Journal: Applied Economics. The whole question is written assuming you will be coding in Stata or R.

a) Staggered DID:

0. Install relevant packages
For R, you will need the following packages: `did`, `dplyr`, `fixest`, `bacondecomp`, `here`, `HonestDiD`, and `haven`.
For Stata, you will need `csdid`, `drdid`, `reghdfe`, `honestdid`, and `ddtiming`. Note that `ddtiming` is not on SSC, but can be installed with:

```
net install ddtiming,
from(https://raw.githubusercontent.com/tgoldring/ddtiming/master)
```

Similarly, for `honestdid`:

```
net install honestdid,
from("https://raw.githubusercontent.com/mcaceresb/stata-honestdid/main")
```

`replace` honestdid$_plugin_check$

1. Download "ehec_data.dta"

The provided dataset `ehec_data.dta` contains a state-level panel dataset on health insurance coverage and Medicaid expansion. The variable `dins` shows the share of low-income childless adults with health insurance in the state. The variable `yexp2` gives the year that a state expanded Medicaid coverage under the Affordable Care Act, and is missing if the state never expanded. The variable `year` gives the year of the observation and the variable `stfips` is a state identifier. (The variable `W` is the sum of person-weights for the state in the ACS; for simplicity, we will treat all states equally and ignore the weights, although if you'd like an additional challenge feel free to re-do everything incorporating the population weights!)

2. Estimate the ATT(g,t) using Callaway and Sant'Anna's estimator

Use the `attgt` function in the `did` package (R) or the `csdid` function in the `csdid` package (Stata) to estimate the group-time specific $\text{ATT}_{g,t}$ for the outcome `dins`. In R, I recommend using the control group option `"notyettreated"`, which uses as a comparison group all units who are not-yet-treated at a given period (including never-treated units). In Stata, use the option `notyet`. (For fun, you're welcome to also try out using `"nevertreated"` units as the control.) **Hint:** replace missing values of `yexp2` to some large number (say, 3000) for the `did` package to incorporate the never-treated units as controls.

For R users, apply the summary command to the results from the `att_gt` command. For Stata users, this should already be reported as a result of `csdid` command. After applying the correct command, you should have a table with estimates of the ATT(g,t) – that is, average treatment effects for a given "cohort" first-treated in period g at each time t. For example, ATT(2014,2015) gives the treatment effect in 2015 for the cohort first treated in 2014.

3. Compare to DiD estimates calculated by hand

To understand how these ATT(g,t) estimates are constructed, we will manually compute one of them by hand. For simplicity, let's focus on ATT(2014, 2014), the treatment effect for the first treated cohort (2014) in the year that they're treated (2014). Create an indicator variable D for whether a unit is first-treated in 2014. Calculate the conditional mean of `dins` for the years 2013 and 2014 for units with D=1 and units with D=0 (i.e. calculate 4 means, for each combination of year and D). Manually compute the 2x2 DiD between D=1 and D=0 and 2013 and 2014. If you did it right, this should line up exactly with the ATT(g,t) estimate you got from the CS package! (Bonus: If you're feeling ambitious, you can verify by hand that the other ATT(g,t) estimates from the CS package also correspond with simple 2x2 DiDs that you can compute by hand)

4. Aggregate the ATT(g,t)

We are often interested in a summary of the ATT(g,t)'s.

In R, use the `aggte` command with option `type = "dynamic"` to compute "event-study" parameters. These are averages of the ATT(g,t) for cohorts at a given lag from treatment — for example, the estimate for event-time 3 gives an average of parameters of the form ATT(g,g+3), i.e. treatment effects 3 periods after units were first treated. You can use the `ggdid` command to plot the relevant event-study.

In Stata, use the commands `qui: estat event` followed by `csdid_plot`.

You can also calculate overall summary parameters. E.g., in R, using `aggte` with the option `type = "simple"` takes a simple weighted average of the ATT(g,t), weighting proportional to cohort sizes. In Stata, you can use `estat simple`.

5. Compare to TWFE estimates (part 1)

Estimate the OLS regression specification

$$Y_{it} = \alpha_i + \lambda_t + D_{it}\beta + \epsilon_{it},$$

where D_{it} is an indicator for whether unit i was treated in period t. How does the estimate for β compare to the simple weighted average you got from Callaway and Sant'Anna? (Don't forget to cluster your SEs at the state level!)

6. Explain this result using the Bacon decomposition

You probably noticed that the static TWFE estimate and the simple-weighted average from C&S were fairly similar. The reason for that is that in this example, there are a fairly large number of never-treated units, and so TWFE mainly puts weight on "clean comparisons". We can see this by using the Bacon decomposition, which shows how much weight static TWFE is putting on clean versus forbidden comparisons. In R, use the `bacon()` command to estimate the weights that TWFE puts on each of the types of comparisons. The first data-frame returned by the command shows how much weight OLS put on the three types of comparisons. How much weight is put on forbidden comparisons here (i.e. comparisons of "Later vs Earlier")? In Stata, use the command `ddtiming`.

7. Compare to TWFE estimates (part 2)

To see a situation where negative weights can matter (somewhat) more, drop from your dataset all the observations that are never-treated. Re-run the Callaway and Sant'Anna and TWFE estimates like you did before on this modified data-set. How does the TWFE estimate compare to the simple weighted average (or the average of the event-study coefficients) now?

8. Run the Bacon decomposition (part 2)

Re-run the Bacon decomposition on the modified dataset. How much weight is put on "forbidden comparisons" now?

9. Even bigger TWFE problems

In the last question, you saw an example where TWFE put a lot of weight on "forbidden comparisons". However, the estimates from the forbidden comparisons were not so bad because the treatment effects were relatively stable over time (the post-treatment event-study coefficients are fairly flat). To see how dynamic treatment effects can make the problem worse, create a variable `relativeTime` that gives the number of periods since a unit has been treated. Create a new outcome variable `dins2` that adds 0.01 times `relativeTime` to `dins` for observations that have already been treated (i.e., we add in some dynamic treatment effects that increase by 0.01 in each period after a unit is treated). Re-run the Callaway & Sant'Anna and TWFE estimates and the Bacon decomp using the dataset from the previous question and the `dins2` variable. How do the differences between C&S and TWFE compare to before?

Question 2

b) Pre-trend sensitivity analysis: This question is a continuation of the question before the question before use is given.

1. Run the baseline DiD

For simplicity, we will first focus on assessing sensitivity to violations of parallel trends in a non-staggered DiD. Load the same dataset on Medicaid as in the previous part. For simplicity, restrict the sample to the years 2015 and earlier, drop the three states which expand Medicaid in 2015 (this ensures states are either first-treated in 2014 or never-treated over our sample). We are now left with a panel dataset where some units are first treated in 2014 and the remaining units are not treated during the sample period.

Start by running the simple TWFE regression

$$Y_{it} = \alpha_i + \lambda_t + \sum_{s \neq 2013} \nVdash[s = t] \times D_i \times \beta_s + u_{it},$$

where $D_i = 1$ if a unit is first treated in 2014 and 0 otherwise. Note that since we do not have staggered treatment, the coefficients β_s are equivalent to DiD estimates between the treated and non-treated units between period s and 2013. I recommend using the `feols` command from the `fixest` package in R and `reghdfe` command in Stata; although feel free to use your favorite regression command. Don't forget to cluster your SEs at the state level.

2. Sensitivity analysis using relative magnitudes restrictions

We are now ready to apply the `HonestDiD` package to do sensitivity analysis. Suppose we're interested in assessing the sensitivity of the estimate for 2014 (the first year of treatment). We will use the "relative magnitudes" restriction that allows the violation of parallel trends between 2013 and 2014 to be no more than $\bar{M}$ times larger than the worst pre-treatment violation of parallel trends.

R instructions:

To create a sensitivity analysis, load the `HonestDiD` package, and call the `createSensitivityResults_r` function. You will need to input the parameters `betahat` and `sigma` calculated above, `numPrePeriods` (in this case, 5), and `numPostPeriods` (in this case, 2). I suggest that you also give the optional parameter `Mbarvec = seq(0.2, by=0.5)` to specify the values of M you wish to use. (Note: it may take a couple of minutes to calculate the sensitivity results.)

Look at the results of the sensitivity analysis you created. For each value of $\bar{M}$, it gives a robust confidence interval that allows for violations of parallel trends between 2013 and 2014 to be no more than $\bar{M}$ times the max pre-treatment violation of parallel trends. What is the "breakdown" value of $\bar{M}$ at which we can no longer reject a null effect? Interpret this parameter.

Stata instructions:

To create a sensitivity analysis, use the `honestdid` function. You will need to pass the options `pre` and `post` to specify the pre and post treatment estimates. I suggest that you also give the optional parameter `mvec = (0.5(0.5)2` to specify the values of $\bar{M}$ you wish to use. (Note: it may take a couple of minutes to calculate the sensitivity results.)

Look at the results of the sensitivity analysis you created. For each value of $\bar{M}$, it gives a robust confidence interval that allows for violations of parallel trends between 2013 and 2014 to be no more than $\bar{M}$ times the max pre-treatment violation of parallel trends. What is the "breakdown" value of $\bar{M}$ at which we can no longer reject a null effect? Interpret this parameter.

3. Create a sensitivity analysis plot

R Instructions:

We can also visualize the sensitivity analysis using the `createSensitivityPlot_relativeMagnitudes` command. To do this, we first have to calculate the CI for the original OLS estimates using the `constructOriginalCS` command. We then pass our sensitivity analysis and the original results to the `createSensitivityPlot_relativeMagnitudes` command.

Stata Instructions:

We can also visualize the sensitivity analysis using the `honestdid` command by adding the `coefplot` option. You can use the `cached` option to use the results from the previous `honestdid` call (for speed's sake).