**2024 Causal Inference Workshop: Stata, R, and Python Sessions**

**Day 7. Advanced Difference-in-differences (with staggered adoption dates)**

Dataset: Causal Inference Workshops (All Years)\Stata and R Materials\Cheng-Hoekstra-castle-doctrine-simplified.dta [same as simple DiD from day 3]

Underlying article: Cheng Cheng and Mark Hoekstra, *Does Strengthening Self-Defense Law Deter Crime or Escalate Violence? Evidence from Expansions to Castle Doctrine*, **Journal of Human Resources** 48(3), 821-853 (2013). [same as day 3]

**Answer Sheet**

**Question 1**. Run the csdid.ado package in Stata (or the did package in R).

1. What is the overall ATT, across all treatment years?
2. How does this compare to the ATT estimated last week, using only the states that adopted rules in 2006?

c. Generate the group-level ATTs (specific to each treatment year). Which are significant/insignificant? Is there a noteworthy pattern? It will be harder, of course, to find significance for a treatment year with fewer treated states.

**STATA Code for Step 1a-c**:

**\*Load the Castle Doctrine dataset**

use "https://github.com/scunning1975/mixtape/raw/master/castle.dta", clear

**\*Install packages**

ssc install csdid, replace

ssc install drdid, replace

**\*Generate log-variables**

gen homicide\_log=log(homicide)

**\*The variable "treatment\_date" measures the year in which the Stand Your Ground law was implemented in a state.**

**\*Set treatment\_date equal zero if treatment is missing (control group). This variable should be coded as zero for never-treated states. Once a state is treated, it remains treated.**

replace treatment\_date=0 if treatment\_date==.

\* Define control group to include only never-treated states

gen control\_group=1 if treatment\_date==0

**\*Construct Event-time variable (year-treatment\_year) so that we can estimate outcomes in event time, for states with different treatment years**

gen event\_time = year-treatment\_date //this is going to be 200x-0 if control, and 200x-200y for treated.

replace event\_time=-1 if control\_group==1 //set to -1 for control units (event\_time=200x-0).

**\* Callaway-Sant’Anna requires non-negative values for the event-time variable. So we need to construct a revised event\_time variable that satisfies this requirement**

sum event\_time

**\*we learn that min -9, max 5. So we have to add 9 to get non-negative values for event time**

replace event\_time=event\_time+9

**\*Define a variable to denote post-treatment period for each treated state**

gen treatment\_active= (year >= treatment\_date)

**\*Now we’re ready to run the csdid package and generate the Callaway Sant'Anna Group-level DID estimator, Control group: never treated. In the dataset “sid” is an indicator for the 51 states (50 states plus DC)**

csdid homicide\_log, ivar(sid) time(year) treatment(treatment\_active) gvar(treatment\_date) agg(event) asinr

\* The output is annual point estimates in event time. We’ll use these below to generate leads-and-lags graphs

estat all

R will give you a warning message for this step but will still run.

**Solution:** The overall ATT is 0.0974, with standard error 0.0409 (and thus *t*-statistic = 2.38). The group ATTs are):

|  |  |  |
| --- | --- | --- |
| **Group** | **Coefficient** | **S.E.** |
| Overall ATT | 0.0974 | 0.0409 |
| Group Average | 0.0999 | 0.0358 |
| Group 2005 | -0.0149 | 0.0774 |
| Group 2006 | 0.1261 | 0.0574 |
| Group 2007 | 0.1284 | 0.0513 |
| Group 2008 | 0.1221 | 0.0567 |
| Group 2009 | -0.0028 | 0.0385 |

Recall that for the day 3 DiD problems, we used only the states that adopted laws in 2006, and found an ATT estimate of 0.0894 (s.e. = 0.095), without population weights. The overall ATT estimate of 0.0994 is close to that estimate, but is substantially more precise.

***Comment 1***: For years with a small number of treated states (all years except 2006), the year-specific s.e.’s are based primarily on the control states, so may be unreliable.

***Comment 2***. The overall t-statistic of 2.38. This meets the standard 5% level for considering a result to be statistically significant, but is at a level where the observed result could plausibly arise by chance.

d. Create an event study plot. How does this differ from the simple DiD event study plot from last week’s question?

**STATA Code for Step 1d**:

**\*Plot Dynamic ATTs, Control Group: Never treated**

csdid\_plot, xlab(-8(1)5) xtitle("Time to Treatment", height(5)) ylabel(-0.6 "-0.6" -.2 "-0.2" 0.2 "0.2" 0.6 "0.6") ymtick(-0.8(0.1)0.8) ytitle("Marginal effect", height(5)) title ("Event-Study Plot: Dynamic ATT") subtitle("Control Group: Never Treated")

**\*Save the graph**

graph export "\att\_csdid\_never\_treated.png", as(png) replace width(3050) height(1350)

**Solution:A graph showing a number of patients

Description automatically generated**

**Comment 1:** The dataset starts in 2000, so the estimates for years -7 and -8 are based on a small number of treated states. The estimate for year -8 is based only on Montana.

**Comment 2:** The leads-and-lags graph above suggests that there may have been a treatment effect for a few years, which then faded away. In contrast the original paper, which used classic two-way fixed effects in event time, did not. Assuming that the fade out is real, and there is no basis from this study to assume otherwise, this substantially changes the policy takeaway from the study.

**Warning**. Deeply hidden in the Callaway-Sant’Anna code is the following quirk: In the pre-treatment period, the individual event study coefficients are all relative to the one before. That’s weird, and not the way that any other package works. We can make all of the pre-treatment coefficients relative to time -1 using a “long2” option. The coefficients form periods 0 and after are already relative to time -1. Let’s fix that:

csdid homicide\_log, ivar(sid) time(year) treatment(treatment\_active) gvar(treatment\_date) agg(event) asinr long2

csdid\_plot, xlab(-8(1)5) xtitle("Time to Treatment", height(5)) ylabel(-0.6 "-0.6" -.2 "-0.2" 0.2 "0.2" 0.6 "0.6") ymtick(-0.8(0.1)0.8) ytitle("Marginal effect", height(5)) title ("Event-Study Plot: Dynamic ATT") subtitle("Control Group: Never Treated")



Not the same for the pre period! Note too that now there is no estimate for period -1, which is assumed to be zero.

**Let’s compare this to the original paper.**



**Not the same!** In the original paper, the effect appears to survive through year 4. Although I wish that they hadn’t averaged across years +3 nd +4. With CS, there is a short term effect, which peters out.

e. The CS estimator allows you to use two different control groups: (i) never treated; and (ii) both never treated and not-yet treated. How does this choice affect the overall ATT estimate?

**STATA Code for Step 1e**:

**\*Callaway Sant'Anna DID estimator, Control group: Not Yet + Never Treated**

\* Include both never treated and not yet treated in control group

\* Relative to the code above, you need to redefine the control group.

replace control\_group = 1 if treatment\_date == 0 | treatment\_date > 2006

csdid homicide\_log, ivar(sid) time(year) treatment(treatment\_active) gvar(treatment\_date) agg(event) asinr notyet long2

estat all

**Solution:** The ATT, including the not-yet treated states in the control group, is 0.0968 (s.e. = 0.0411). This is very similar to the result above using only never-treated states as controls: ATT: 0.0974 (s.e. = 0.0409).

f. The CS estimator allows you to use multiple different estimation methods. By default it uses the Sant’Anna and Zhao doubly robust DiD approach. Assess by how much results would differ if you instead used inverse propensity weights (IPW). The ATT weights are 1 for treated states, and p/(1-p) for control states.

**STATA Code for Step 1f**:

**\*Code using never treated states as the control group.**

csdid homicide\_log, ivar(sid) time(year) treatment(treatment\_active) gvar(treatment\_date) agg(event) asinr method(ipw)

estat all

**Solution:** In this case, we use IPW. The ATT is 0.0974 (s.e. = 0.0409). Both the coefficient and s.e. are very close to those using the default weights. There would be larger, but still modest differences between the default approach and IPW weights.

g. Compare the overall ATT, obtained using Callaway-Sant’Anna, to the estimate from a standard TWFE regression.

**STATA Code for Step 1g:**

**\*Install packages**

ssc install reghdfe, replace

ssc install ftools, replace

**\*Overall ATT**

reghdfe homicide\_log i.treatment\_active, absorb(sid year) cluster(sid)

**Solution:** The overall ATT using the TWFE model is 0.0788 (s.e. =.0581) (*t* = 1.36). So Callaway-Sant’Anna (using the never-treated control group) provides both a larger coefficient estimate, and a smaller s.e, leading to statistical significance ( *t* = 2.43) , which is not found with the TWFE model.

h. Compare the TWFE results you obtained from those reported by Cheng and Hoekstra.

**Solution:** The TWFE estimate of 0.0788 (s.e. 0.0581) is somewhat smaller than the same as the Cheng and Hoekstra coefficient of 0.0877 (s.e. 0.0331). Our estimate is closer to Cheng and Hoekstra model specification 2, where they include Region-by-Year fixed effects (coeff. 0.0811, s.e. 0.0769).

i. Bonus 1: re-generate the event study plot using Sun and Abraham instead. Are there any major differences versus Callaway-SantAnna?

**\*Install package**

ssc install eventstudyinteract, replace

ssc install avar, replace

**\*Restore Original Values of event\_time**

replace event\_time=event\_time-9

**\*Generate Relative Time Indicators**

sum event\_time

**\*Lags, Drop -1**

forvalues k = `r(min)'(1)-2 {

local k = -`k'

gen event\_time\_lag\_`k' = event\_time == -`k'

lab var event\_time\_lag\_`k' "-`k'"

}

**\*Leads**

forvalues k = 0(1)`r(max)' {

gen event\_time\_lead\_`k' = event\_time == `k'

lab var event\_time\_lead\_`k' "`k'"

}

**\*Generate never treated**

gen never\_treated = treatment\_date == 0

**\*Replace values in event\_time to be positive**

replace event\_time=event\_time+9

**\*Staggered DiD**

eventstudyinteract homicide\_log event\_time\_lag\_\* event\_time\_lead\_\*, cohort(event\_time) control\_cohort(never\_treated) absorb(sid year) vce(cluster sid)

**\*event study plot**

ssc install coefplot

matrix C = e(b\_iw)

mata st\_matrix("A",sqrt(diagonal(st\_matrix("e(V\_iw)"))))

matrix C = C \ A'

matrix list C

matrix new\_col = (0, 0)'

matrix C = C, new\_col

local colnames : colnames C

local colnames = subinstr("`colnames'", "c1", "-1", .)

matrix colnames C = `colnames'

matrix C\_part1 = C[1..., 1..8]

matrix C\_minus1 = C[1..., colsof(C)]

matrix C\_part2 = C[1..., 9..(colsof(C)-1)]

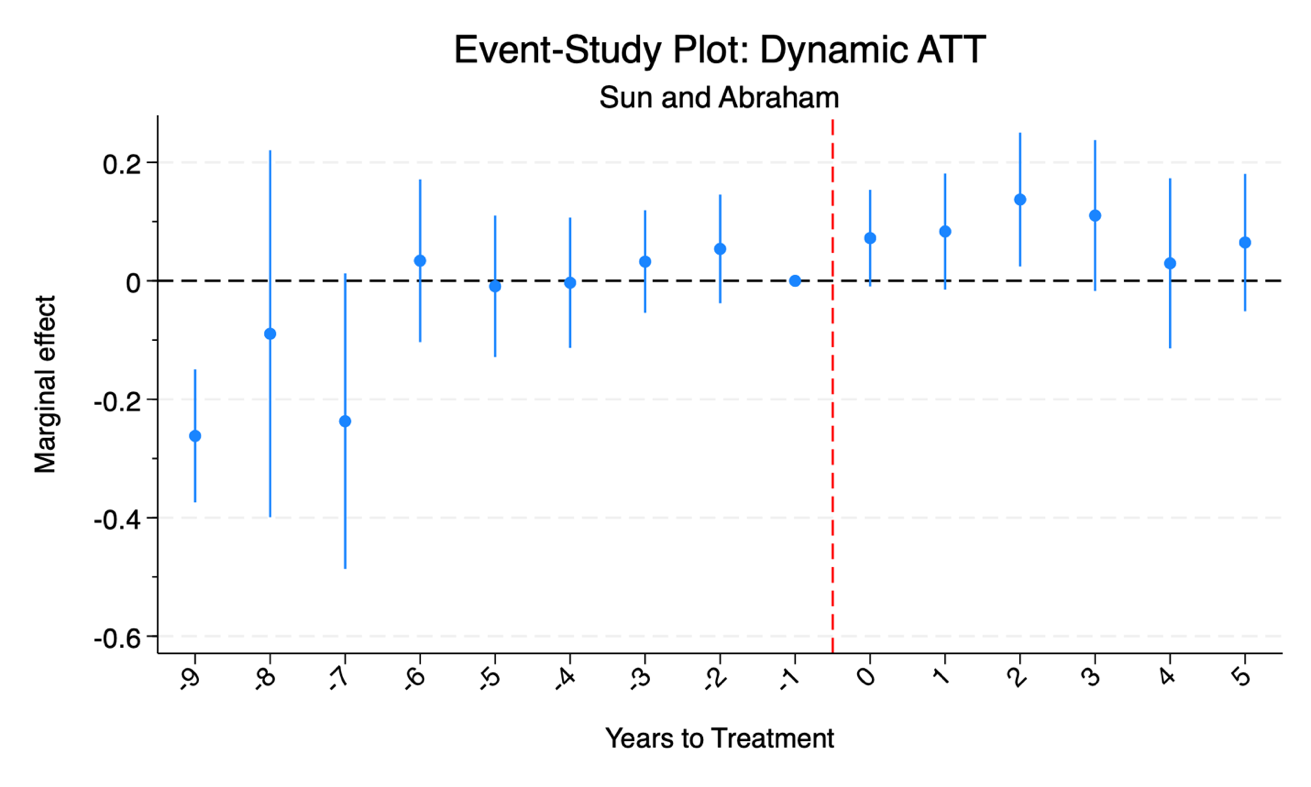
matrix C\_new = (C\_part1, C\_minus1, C\_part2)

matrix C = C\_new

matrix list C

coefplot matrix(C[1]), se(C[2]) vertical lab xtitle("Years to Treatment", height(5)) ytitle("Marginal effect", height(5)) ylab(-0.6 "-0.6" -0.4 "-0.4" -0.2 "-0.2" 0 "0" 0.2 "0.2") ymtick(-0.6(0.1)0.2) xlab(,angle(45)) title("Event-Study Plot: Dynamic ATT") subtitle("Sun and Abraham") yline(0, lcol(black) lpattern(dash)) xline(9.5, lcol(red) lpattern(dash))

**Solution:** Note that the Sun and Abraham packageproduces only dynamic treatment effects, and not an overall ATT.



Note that for Sun-Abraham, similar to Callaway-Sant’Anna, the treatment effect shrinks to insignificance in years +4 and +5

i. Bonus 2: Examine treatment group heterogeneity using the Goodman-Bacon decomposition package. Run the decomposition and plot the different group weights. Are there any groups driving the TWFE overall estimate?

S**TATA code for Step Bonus 2: (requires Stata18)**

**\*Goodman-Bacon Decomposition**

use "https://github.com/scunning1975/mixtape/raw/master/castle.dta", clear

gen homicide\_log=log(homicide)

gen treatment\_active= (year >= treatment\_date)

xtdidregress (homicide\_log) (treatment\_active), group(sid) time(year)

estat bdecomp, graphA graph with blue dots and red text

Description automatically generated

The y-axis provides 2x2 estimates. There are 5 timing groups (2005, 2006, 2007, 2008, and 2009) and one never treated group. There are thus 25 2x2 estimates.