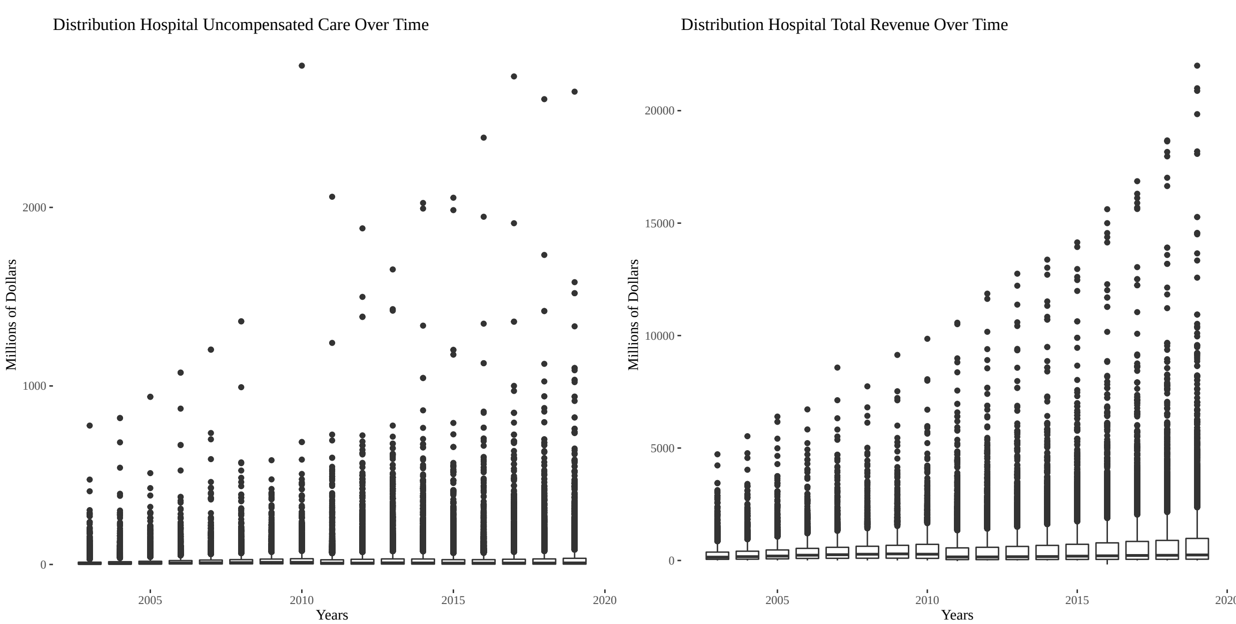
Econ771 - Empirical Exercise 1

Nixon Torres Candiales

September 23, 2022

Summary Statistics

Provide and discuss a table of simple summary statistics showing the mean, standard deviation, min, and max of hospital total revenues and uncompensated care over time. We start first by downloading and merging all the Data from the Github repository. We present the distribution of the variables of interest, Uncompensated care and Hospital Revenue over time



(a) Uncompensated Care (b) Revenue

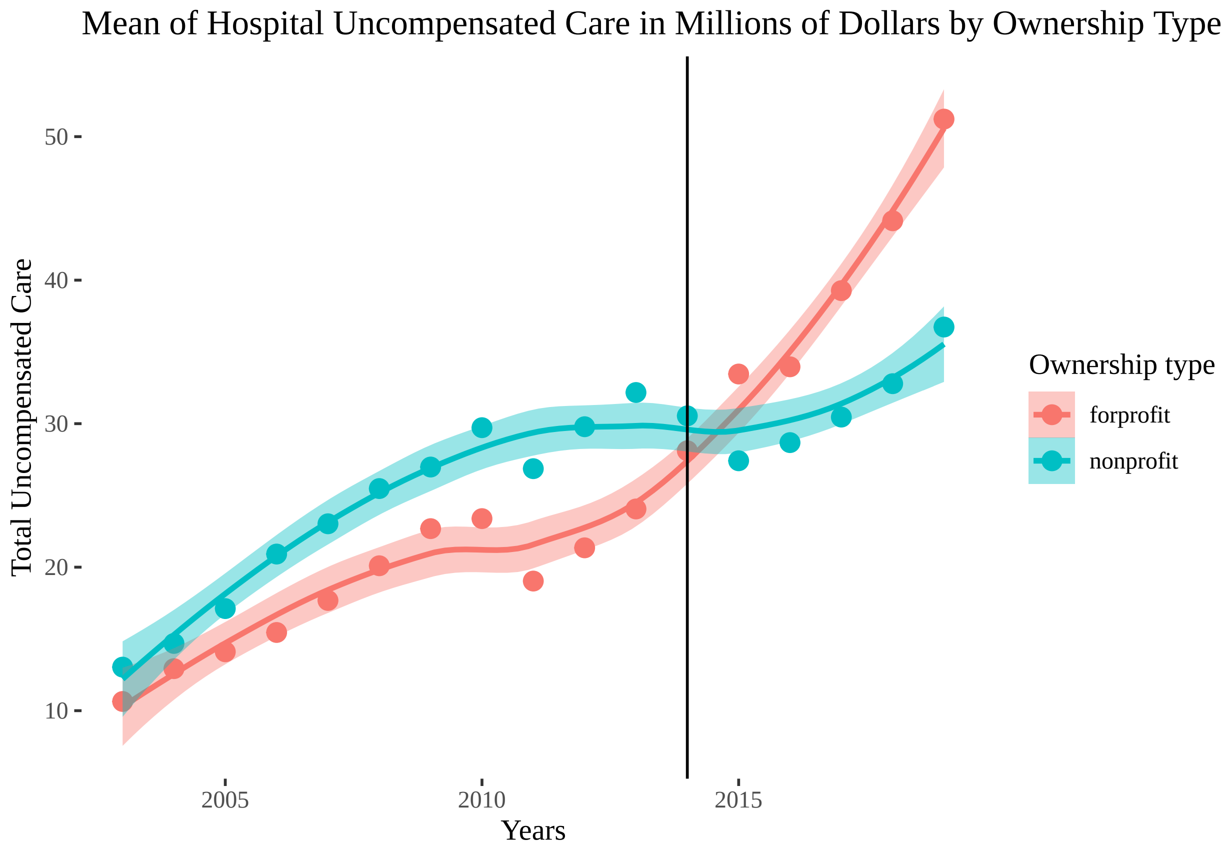
Figure 1: Box-plot

We see evidence of negative entries in uncompensated care as well as extreme atypical values that might be caused by mistyping. As such, we subset the data by not including the top and bottom 0.5% of the observations. The mean, standard deviation, and minimum and maximum values across the years are presented in table 1 of summary statistics.

Table 1: Hospital Summary Statistics

	Uncompensated Care				Total Revenue			
year	Mean	Sd	Min	Max	Mean	Sd	Min	Max
2003	13.57	32.05	0	777.99	292.35	397.70	1.66	4722.76
2004	15.33	36.66	0	820.25	327.69	445.42	0.27	5525.73
2005	17.41	37.81	0	939.13	380.14	514.68	1.14	6398.55
2006	20.97	47.16	0	1074.62	433.73	558.58	1.33	6718.17
2007	23.56	51.28	0	1203.37	484.24	646.76	0.99	8577.05
2008	26.43	57.06	0	1361.81	513.43	655.58	0.97	7743.08
2009	27.44	46.42	0	583.98	552.80	718.35	0.89	9139.32
2010	29.89	72.41	0	2793.92	576.62	779.98	0.84	9857.53
2011	26.82	63.17	0	2059.70	480.04	776.84	-27.58	10572.29
2012	29.87	72.54	0	1882.62	505.10	830.36	0.85	11865.32
2013	31.93	72.63	0	1652.58	539.20	903.93	0.95	12751.71
2014	31.79	77.39	0	2024.85	577.45	980.59	1.09	13376.35
2015	29.83	74.67	0	2054.15	623.38	1048.30	1.05	14143.53
2016	31.14	80.95	0	2390.67	677.54	1157.37	-177.03	15618.75
2017	33.38	87.36	0	2733.60	727.36	1263.39	1.00	16863.43
2018	35.90	90.49	0	2606.35	782.42	1386.47	1.07	18677.25
2019	39.82	99.48	0	2648.26	855.33	1538.23	0.72	22000.93
2003-2019	28.77	71.96	0	2793.92	574.79	993.11	-177.03	22000.93

Next, we create a figure showing the mean hospital uncompensated care from 2003 to 2019. We show this trend separately by hospital ownership type in figure 3. We present an smooth trend to easily identify shifts after the adoption of medicare expansion in 2014. We can see an abrupt bump from years 2010 to 2011, this might be due to the adoption of the new form and a change in the way to measure uncompensated care ^[after 2010 uncompensated care = total uncompensated care - total uncompensated partial payments + bad debt]. Also, in recent years, it seems for-profit hospitals get to provide more uncompensated care than not for profit hospitals.



Mean of Hospital Uncompensated Care in Millions of Dollars by Ownership Type
Total Uncompensated Care
50
40
30
20
10
2005
2010
2015
Years
Ownership type
forprofit
nonprofit

Table 2: Two Way Fixed Effects

	1	2	3	4
Treatment	-28.363***	-31.518***	-12.173***	-12.153***
	(1.893)	(2.185)	(1.848)	(1.550)
Num.Obs.	79557	69824	74768	77624
R2	0.699	0.708	0.690	0.691
RMSE	38.21	38.97	39.48	38.91
Std.Errors	by: pn	by: pn	by: pn	by: pn

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TWFE Specification

Using a simple DD identification strategy, we estimate the effect of Medicaid expansion on hospital uncompensated care using a traditional two-way fixed effects (TWFE) estimation:

$$y_{it} = \alpha_i + \gamma_t + \delta D_{it} + \varepsilon_{it}, \tag{1}$$

where $D_{it} = 1(E_i \leq t)$ in Equation 1 is an indicator set to 1 when a hospital is in a state that expanded as of year t or earlier, γ_t denotes time fixed effects, α_i denotes hospital fixed effects, and y_{it} denotes the hospital's amount of uncompensated care in year t. We present four estimates from this estimation in table **??**: one based on the full sample (1); one when limiting to the 2014 treatment group (2); one when limiting to the 2015 treatment group (3); and one when limiting to the 2016 treatment group (3).

A first appreciation of our results indicate that the ATE of medicaid expansion on uncompensated care is negative. The point estimate varies when limiting the treatment sample for 2014 to 2016 as well as the confidence intervals but we get a consistent trend across the samples. We see the effect is larger when limiting the sample to the states that expanded in 2014.

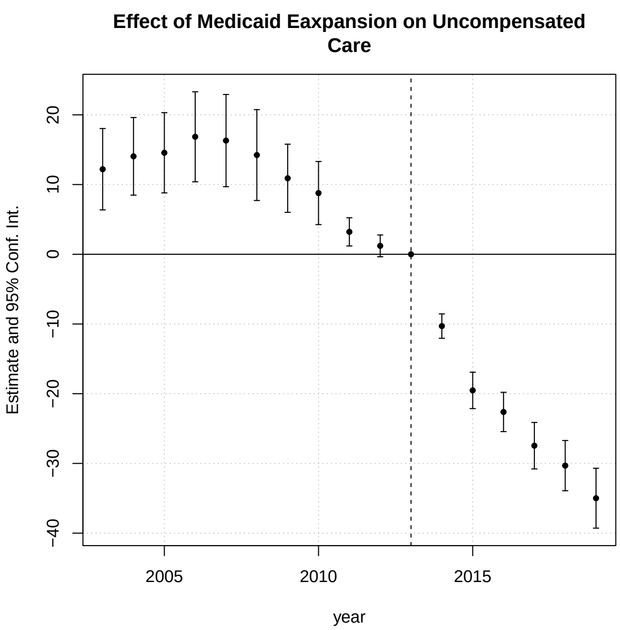
Event Study Specification

We estimate an event study version of the specification in part 3:

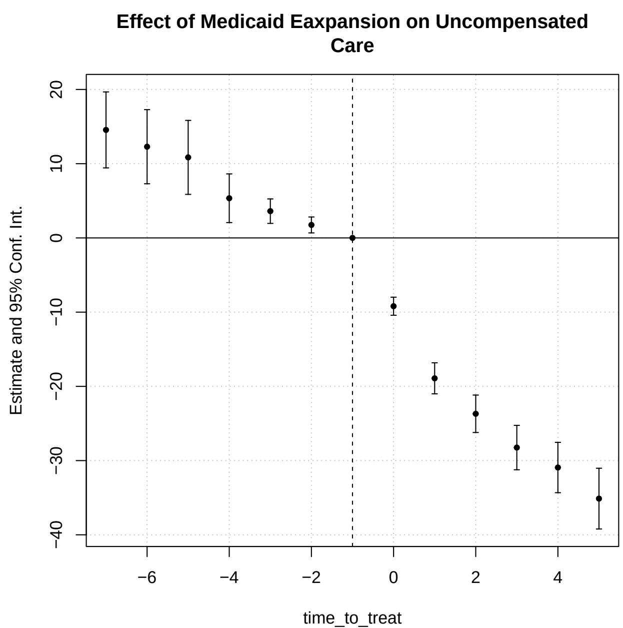
$$y_{it} = \alpha_i + \gamma_t + \sum_{\tau<-1} D_{it}^{\tau}\delta_{\tau} + \sum_{\tau>=0} D_{it}^{\tau}\delta_{\tau} + \varepsilon_{it}, \tag{2}$$

where $D_{it}^{\tau} = 1(t - E_i = \tau)$ in Equation 2 is an interaction between the treatment indicator and a relative time indicator. τ denotes years relative to Medicaid expansion. In this case $\tau = -1$ denotes the year before a state expanded Medicaid and the control group for those never treated. $\tau = 0$ denotes the year of expansion, and so on.

Table **??** presents the estimates for the common with two different samples: one based on the full sample and one based only on those that expanded in 2014 (with never treated as the control group).



(a) Sample 2014



(b) Full Sample

Figure 2: Event Study

Table 3: Event Study Sample 2014

	Model 1
year = 2003 × treated	12.199***
	(2.978)
year = 2004 × treated	14.048***
	(2.840)
year = 2005 × treated	14.556***
	(2.936)
year = 2006 × treated	16.852***
	(3.294)
year = 2007 × treated	16.307***
	(3.373)
year = 2008 × treated	14.229***
	(3.324)
year = 2009 × treated	10.901***
	(2.492)
year = 2010 × treated	8.781***
	(2.309)
year = 2011 × treated	3.213**
	(1.033)
year = 2012 × treated	1.204
	(0.797)
year = 2014 × treated	-10.301***
	(0.892)
year = 2015 × treated	-19.522***
	(1.331)
year = 2016 × treated	-22.619***
	(1.438)
year = 2017 × treated	-27.455***
	(1.700)
year = 2018 × treated	-30.317***
	(1.838)
year = 2019 × treated	-34.989***
	(2.191)
Num.Obs.	69824
RMSE	38.82
Std.Errors	by: pn
FE: pn	X
FE: year	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Event Estudy Full Sample

	Model 1
time_to_treat = -7 × treated	14.553***
	(2.611)
time_to_treat = -6 × treated	12.287***
	(2.547)
time_to_treat = -5 × treated	10.850***
	(2.542)
time_to_treat = -4 × treated	5.346**
	(1.674)
time_to_treat = -3 × treated	3.606***
	(0.842)
time_to_treat = -2 × treated	1.750**
	(0.548)
time_to_treat = 0 × treated	-9.201***
	(0.622)
time_to_treat = 1 × treated	-18.916***
	(1.067)
time_to_treat = 2 × treated	-23.692***
	(1.287)
time_to_treat = 3 × treated	-28.244***
	(1.526)
time_to_treat = 4 × treated	-30.934***
	(1.730)
time_to_treat = 5 × treated	-35.122***
	(2.088)
Num.Obs.	79557
RMSE	38.02
Std.Errors	by: pn
FE: pn	X
FE: year	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

SA Specification

Sun and Abraham(SA) show that the δ_τ coefficients in Equation @ref(eq:event) can be written as a non-conxvex average of all other group-time specific average treatment effects. They propose an interaction weighted specification:

$$y_{it} = \alpha_i + \gamma_t + \sum_e \sum_{\tau \neq -1} \left(D_{it}^{\tau} \times 1(E_i = e)\right) \delta_{e,\tau} + \varepsilon_{it}.(\#eq:iwevent) \tag{3}$$

Re-estimate your event study using the SA specification in Equation @ref(eq:iwevent). Show your results for $\hat{\delta}_{e,\tau}$ in a Table, focusing on states with $E_i = 2014$, $E_i = 2015$, and $E_i = 2016$.

Present an event study graph based on the results in part 5. Hint: you can do this automatically in `R` with the `fixest` package (using the `sunab` syntax for interactions), or with `eventstudyinteract` in `Stata`. These packages help to avoid mistakes compared to doing the tables/figures manually and also help to get the standard errors correct.

CS Specification

Callaway and Sant'Anna (CS) offer a non-parametric solution that effectively calculates a set of group-time specific differences, $ATT(g,t) = E[y_{it}(g) - y_{it}(\infty)|G_i = g]$, where g reflects treatment timing and t denotes time. They show that under the standard DD assumptions of parallel trends and no anticipation, $ATT(g,t) = E[y_{it} - y_{i,g-1}|G_i = g] - E[y_{it} - y_{i,g-1}|G_i = \infty]$, so that $\hat{ATT}(g,t)$ is directly estimable from sample analogs. CS also propose aggregations of $\hat{ATT}(g,t)$ to form an overall ATT or a time-specific ATT (e.g., ATTs for τ periods before/after treatment). With this framework in mind, provide an alternative event study using the CS estimator. Hint: check out the `did` package in `R` or the `csdid` package in `Stata`.

RR Specification

Rambachan and Roth (RR) show that traditional tests of parallel pre-trends may be underpowered, and they provide an alternative estimator that essentially bounds the treatment effects by the size of an assumed violation in parallel trends. One such bound RR propose is to limit the post-treatment violation of parallel trends to be no worse than some multiple of the pre-treatment violation of parallel trends. Assuming linear trends, such a violation is reflected by

$$\Delta(\bar{M}) = \left\{ \delta : \forall t \geq 0, |(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1})| \leq \bar{M} \times \max_{s<0} |(\delta_{s+1} - \delta_s) - (\delta_s - \delta_{s-1})| \right\}.$$

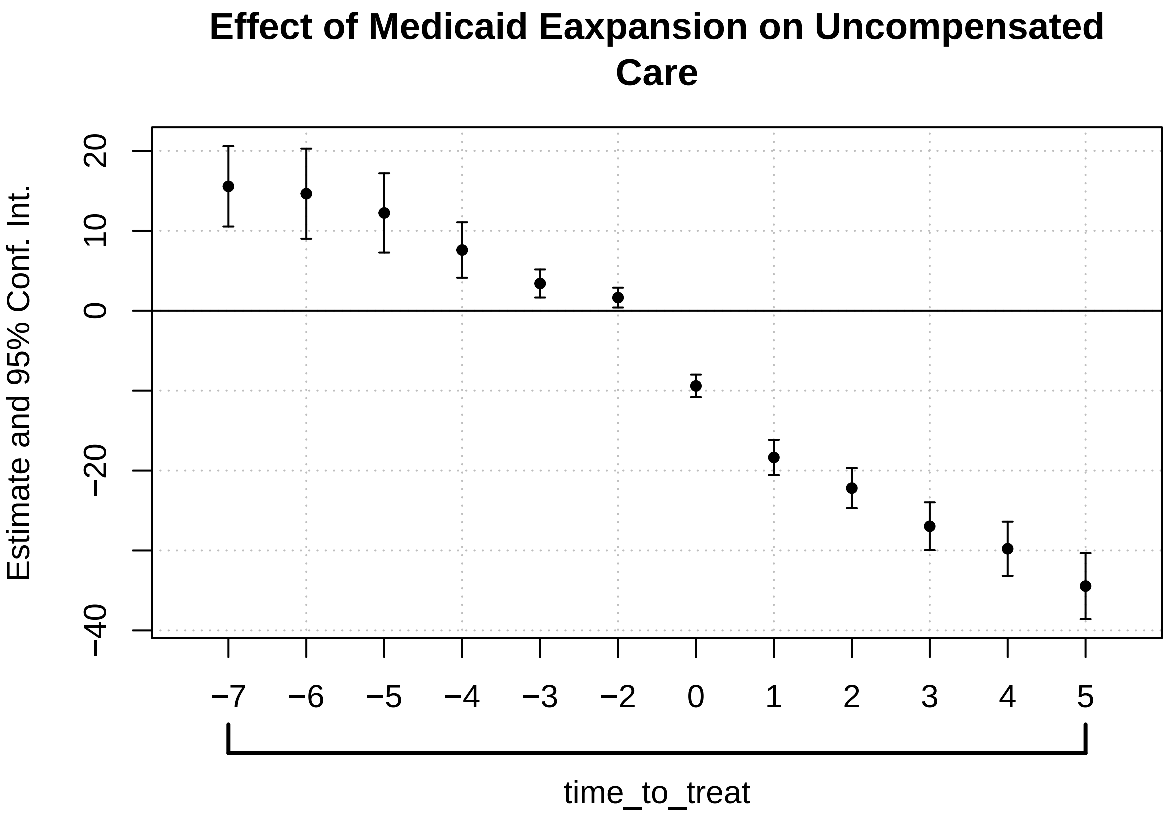
Using the `HonestDiD` package in `R` or `Stata`, present a sensitivity plot of your CS ATT estimates using $\bar{M} = \{0, 0.5, 1, 1.5, 2\}$. Check out the GitHub repo here for some help in combining the `HonestDiD` package with CS estimates.

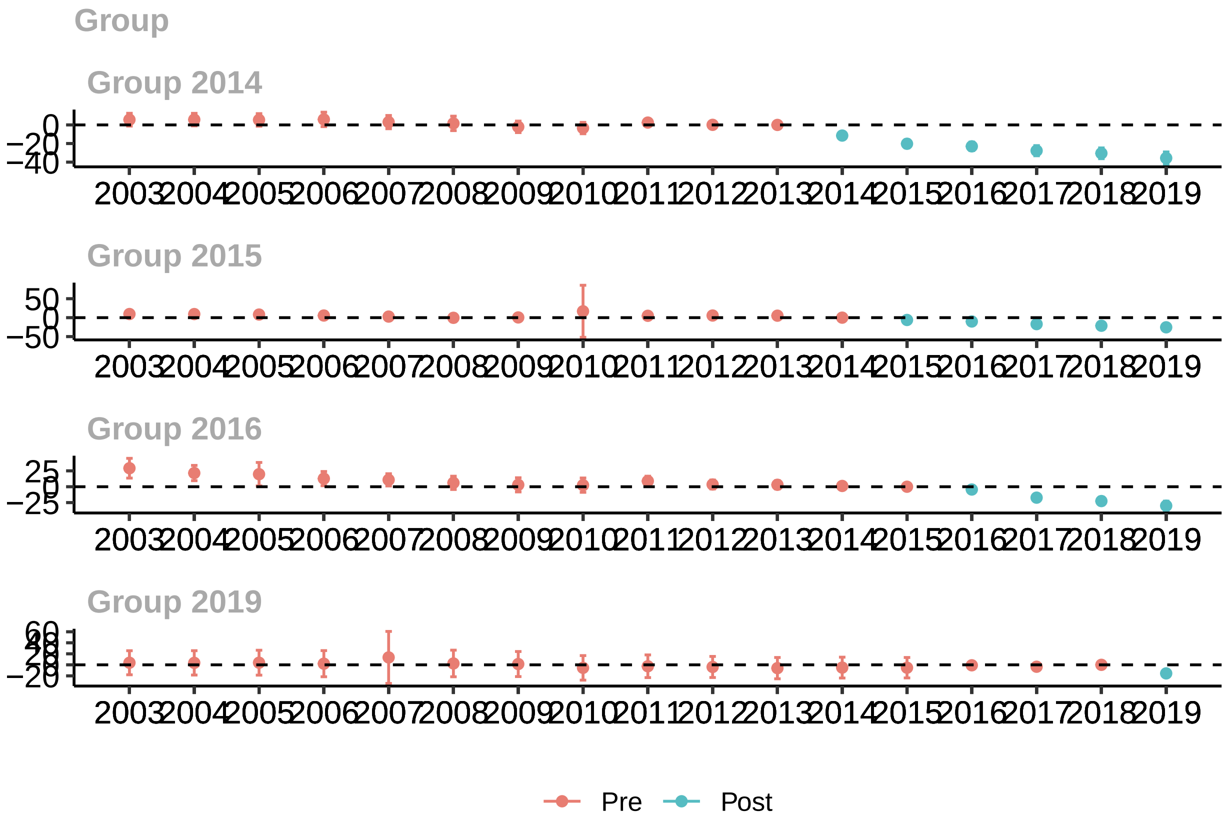
Discussion

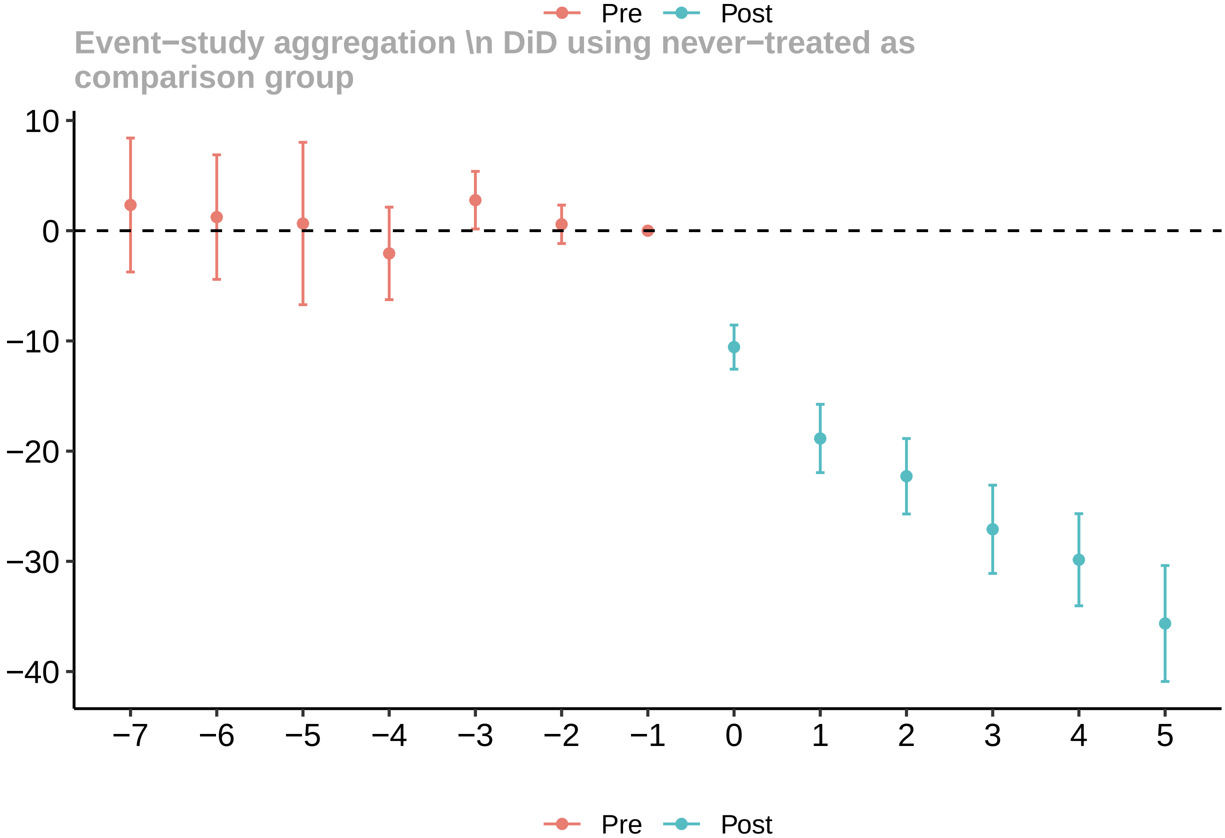
Discuss your findings and compare estimates from different estimators (e.g., are your results sensitive to different specifications or estimators? Are your results sensitive to violation of parallel trends assumptions?).

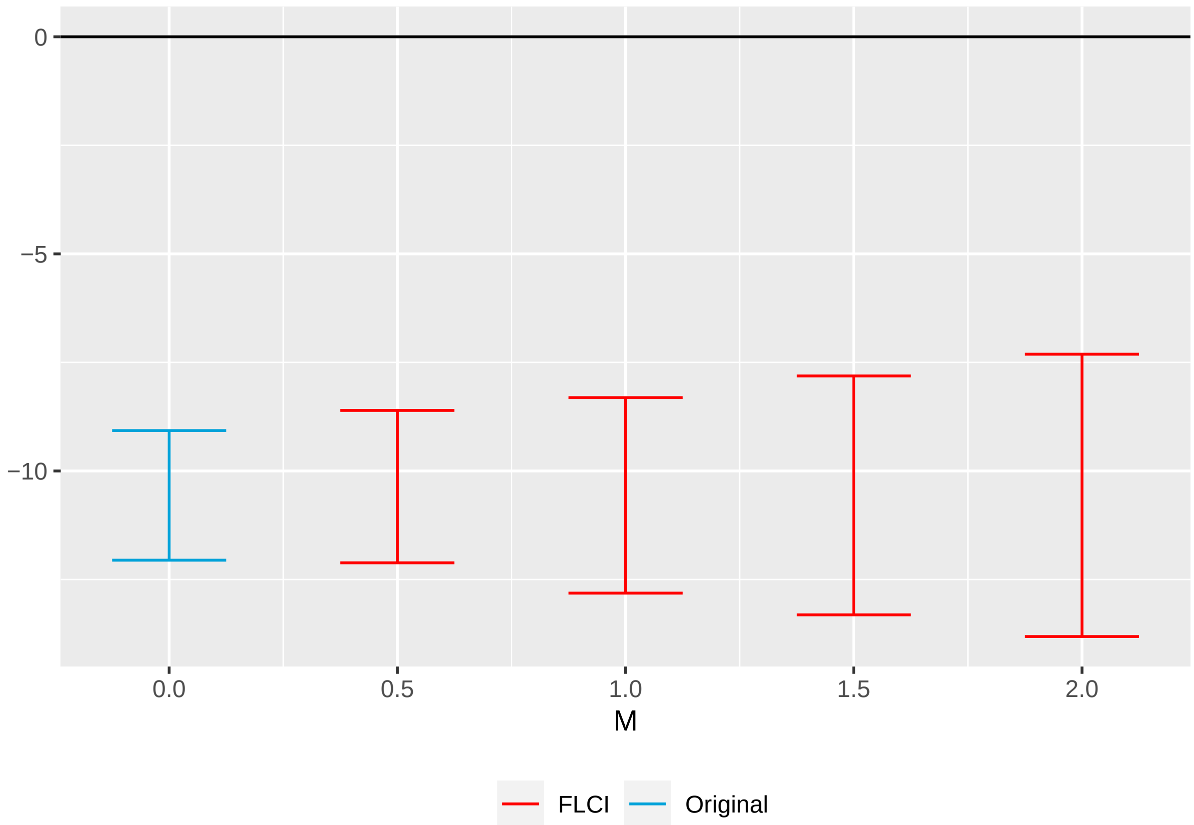
Reflection

Reflect on this assignment. What did you find most challenging? What did you find most surprising?




Group
Group 2014
Group 2015
Group 2016
Group 2019
Pre
Post


Event−study aggregation \n DiD using never−treated as comparison group
Pre
Post


0
−5
−10
0.0
0.5
1.0
1.5
2.0
M
FLCI
Original

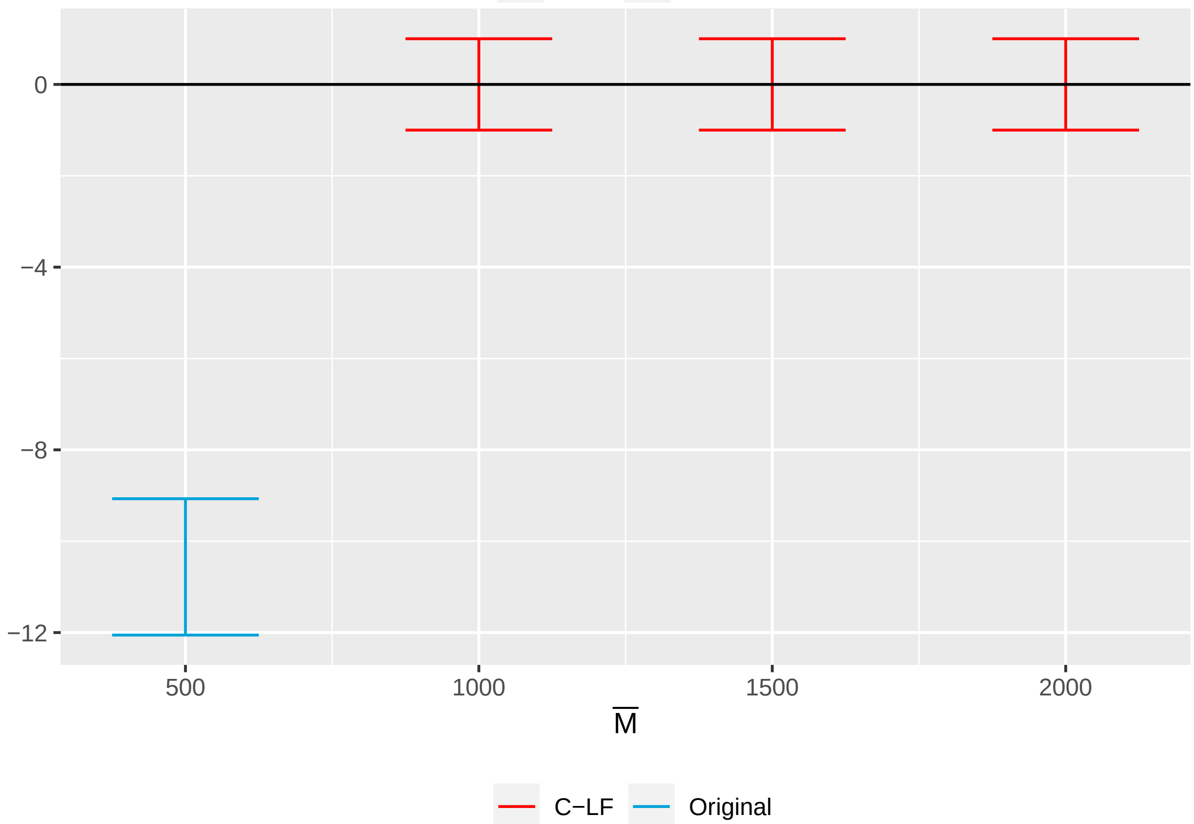

0
−4
−8
−12
500
1000
1500
2000
$\overline{M}$
C−LF
Original

Table 5: Two Way Fixed Effects long format

	1	2	3	4
Treatment	-28.363***	-31.518***	-12.173***	-12.153***
	(1.893)	(2.185)	(1.848)	(1.550)
Num.Obs.	79557	69824	74768	77624
R2	0.699	0.708	0.690	0.691
R2 Adj.	0.675	0.684	0.666	0.667
AIC	817114.7	720035.7	772873.5	800104.3
BIC	871269.6	767580.1	823742.8	852930.5
RMSE	38.21	38.97	39.48	38.91
Std.Errors	by: pn	by: pn	by: pn	by: pn

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Event Study Common Treatment

	Model 1
year = 2003 × treated	12.199***
	(2.978)
year = 2004 × treated	14.048***
	(2.840)
year = 2005 × treated	14.556***
	(2.936)
year = 2006 × treated	16.852***
	(3.294)
year = 2007 × treated	16.307***
	(3.373)
year = 2008 × treated	14.229***
	(3.324)
year = 2009 × treated	10.901***
	(2.492)
year = 2010 × treated	8.781***
	(2.309)
year = 2011 × treated	3.213**
	(1.033)
year = 2012 × treated	1.204
	(0.797)
year = 2014 × treated	-10.301***
	(0.892)
year = 2015 × treated	-19.522***
	(1.331)
year = 2016 × treated	-22.619***
	(1.438)
year = 2017 × treated	-27.455***
	(1.700)
year = 2018 × treated	-30.317***
	(1.838)
year = 2019 × treated	-34.989***
	(2.191)
Num.Obs.	69824
AIC	709145.8
BIC	709301.4
RMSE	38.82
Std.Errors	by: pn
FE: pn	X
FE: year	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Even Study Staggered

	Model 1
time_to_treat = -7 × treated	14.553***
	(2.611)
time_to_treat = -6 × treated	12.287***
	(2.547)
time_to_treat = -5 × treated	10.850***
	(2.542)
time_to_treat = -4 × treated	5.346**
	(1.674)
time_to_treat = -3 × treated	3.606***
	(0.842)
time_to_treat = -2 × treated	1.750**
	(0.548)
time_to_treat = 0 × treated	-9.201***
	(0.622)
time_to_treat = 1 × treated	-18.916***
	(1.067)
time_to_treat = 2 × treated	-23.692***
	(1.287)
time_to_treat = 3 × treated	-28.244***
	(1.526)
time_to_treat = 4 × treated	-30.934***
	(1.730)
time_to_treat = 5 × treated	-35.122***
	(2.088)
Num.Obs.	79557
AIC	804661.4
BIC	804782.1
RMSE	38.02
Std.Errors	by: pn
FE: pn	X
FE: year	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: SA especification

	Model 1
time_to_treat = -7	15.552***
	(2.563)
time_to_treat = -6	14.639***
	(2.875)
time_to_treat = -5	12.228***
	(2.528)
time_to_treat = -4	7.590***
	(1.767)
time_to_treat = -3	3.401***
	(0.895)
time_to_treat = -2	1.638**
	(0.634)
time_to_treat = 0	-9.411***
	(0.723)
time_to_treat = 1	-18.360***
	(1.129)
time_to_treat = 2	-22.200***
	(1.280)
time_to_treat = 3	-26.974***
	(1.528)
time_to_treat = 4	-29.782***
	(1.729)
time_to_treat = 5	-34.456***
	(2.107)
Num.Obs.	79557
AIC	804610.3
BIC	804731.0
RMSE	38.00
Std.Errors	by: pn
FE: pn	X
FE: year	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: CS specification

	Model 1
ATT(2014,2003)	5.619
	(2.250)
ATT(2014,2004)	5.711
	(2.203)
ATT(2014,2005)	5.374
	(2.212)
ATT(2014,2006)	5.922
	(2.542)
ATT(2014,2007)	3.030
	(2.332)
ATT(2014,2008)	1.638
	(2.538)
ATT(2014,2009)	-2.067
	(2.018)
ATT(2014,2010)	-3.365
	(1.997)
ATT(2014,2011)	2.530
	(1.080)
ATT(2014,2012)	0.090
	(0.963)
ATT(2014,2013)	0.000
ATT(2014,2014)	-11.446
	(0.930)
ATT(2014,2015)	-20.250
	(1.402)
ATT(2014,2016)	-23.016
	(1.488)
ATT(2014,2017)	-27.688
	(1.785)
ATT(2014,2018)	-30.492
	(1.897)
ATT(2014,2019)	-35.639
	(2.161)
ATT(2015,2003)	9.581
	(3.110)
ATT(2015,2004)	9.444
	(3.227)
ATT(2015,2005)	8.175
	(3.261)
ATT(2015,2006)	5.645
	(3.355)
ATT(2015,2007)	2.815
	(3.227)
ATT(2015,2008)	-0.399
	(3.403)
ATT(2015,2009)	0.605
	(3.169)

Table 10: CS specification Event Study

	Model 1
ATT(-7)	2.333
	(2.315)
ATT(-6)	1.235
	(2.153)
ATT(-5)	0.653
	(2.807)
ATT(-4)	-2.060
	(1.599)
ATT(-3)	2.769
	(0.995)
ATT(-2)	0.581
	(0.665)
ATT(-1)	0.000
ATT(0)	-10.562
	(0.763)
ATT(1)	-18.852
	(1.181)
ATT(2)	-22.281
	(1.305)
ATT(3)	-27.093
	(1.525)
ATT(4)	-29.855
	(1.593)
ATT(5)	-35.639
	(2.004)
Num.Obs.	5815
Std.Errors	by: pn_id
type	dynamic
ngroup	4.000
ntime	17.000
control.group	nevertreated
est.method	dr