

DIDの計量経済手法の近年の展開 ～問題と対策～

【演習編】

小西祥文

慶應義塾大学

日本経済学会 春季大会
共催：日本学術会議 数量的経済・政策分析分科会
Tutorial Session

2023.5.28

Stata/Rによる演習

(1) 基礎編

- Baker *et al.* の仮想例
(Heterogeneous, Dynamic Effects)
 - PackageのDL
 - Diagnosisの方法
 - Goodman-Bacon decomposition
 - Jakiela's diagnosis
 - 三つのアプローチによる推定
 - Fully saturated TWFE
 - Rolling methods
 - Imputation methods

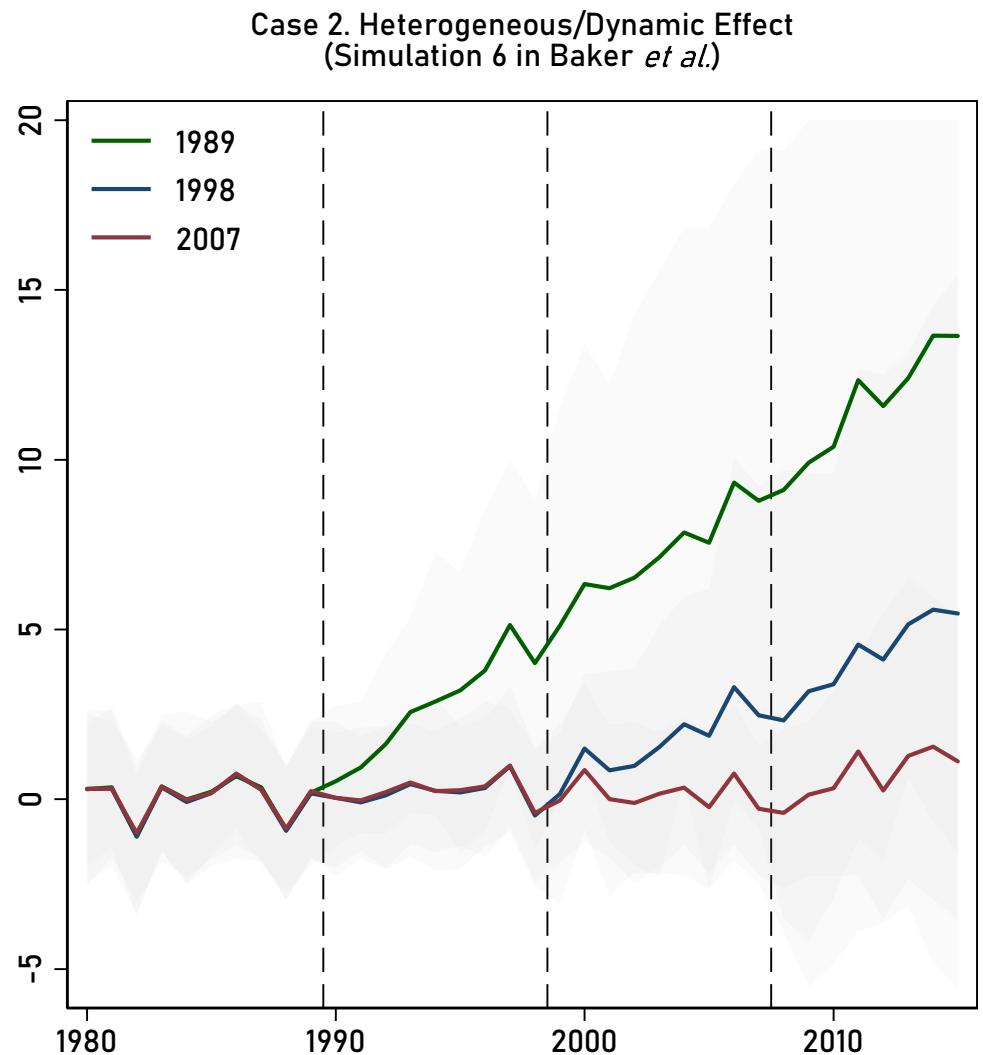
仮想例 : Baker *et al.* (JF, 2022)

- Baker *et al.* (2022) の仮想例
 - $T = 36 : 1980 \sim 2015$
 - 処置タイミング : 1989, 1998, 2007
 - $N = 1000$ firms \Rightarrow
 i randomly assigned $\{G_{89}, G_{98}, G_{07}\}$
 - $\alpha_i, \lambda_t, \epsilon_{it} \sim N(0, 0.5)$
- 処置効果は異質的, Dynamic Treatment Effect

DGP: $y_{it} = \alpha_i + \lambda_t + \tau_k D_{it} + \epsilon_{it}$

$$\tau_k \sim N(\delta_k \times [T - k], 0.2^2 \times [T - k])$$

$$\delta_{89} = 0.5, \delta_{98} = 0.3, \delta_{07} = 0.1$$



仮想例 : Baker *et al.* (JF, 2022)

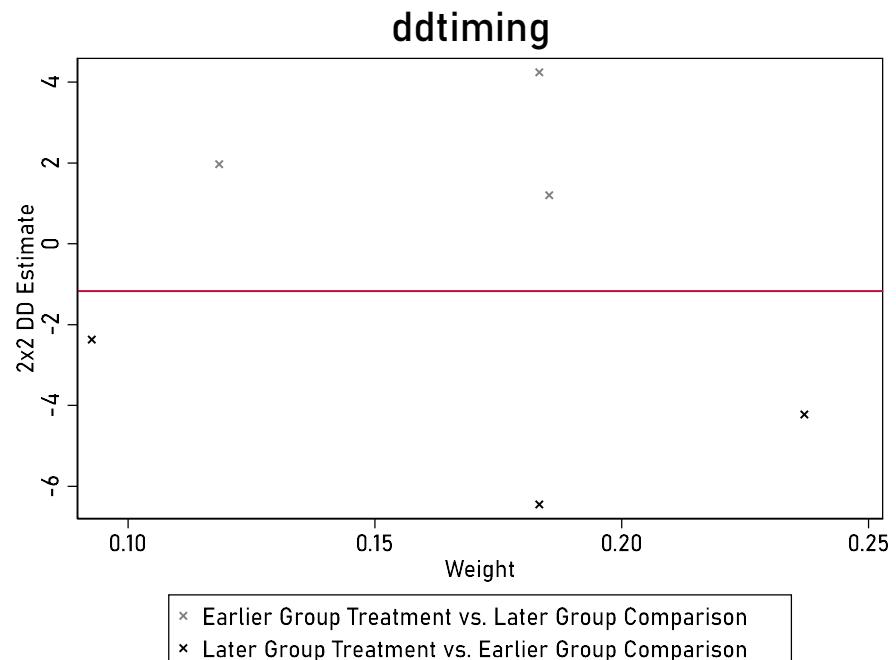
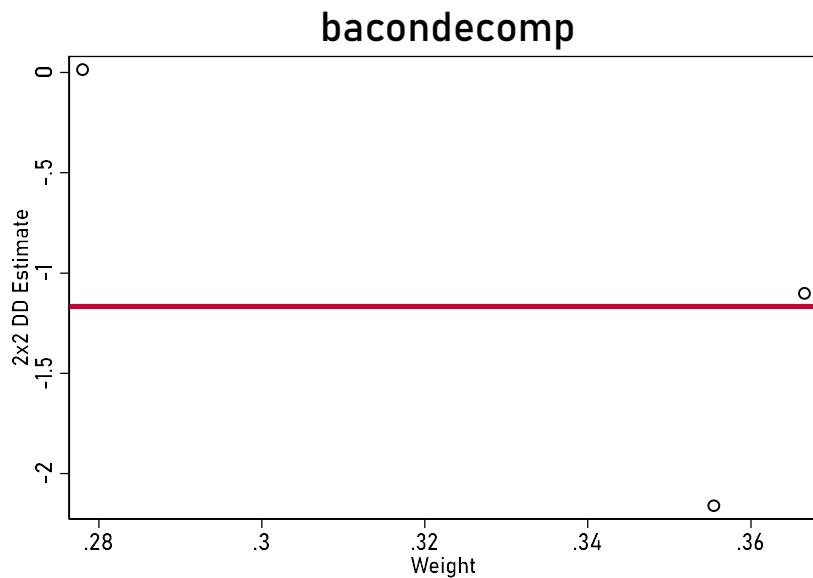
```
1 ****
2 ***   Staggered DID: Problems & Solutions      ***
3 ***   Part I. Basics using Simulated Data      ***
4 ***   Written by Yoshifumi Konishi               ***
5 ***   Date: 2023.5.22                            ***
6 ****
7
8 /***** Install Packages ****/
9
10
11 *** (1) Utility packages ***
12 ssc install reghdfe, replace
13 ssc install avar, replace /* eventsudydyinteract dependency */
14 ssc install ftools, replace
15 ssc install event_plot, replace
16 ssc install addplot, replace
17 ssc install estout, replace
18
19 *** (2) Decomposition packages ***
20 ssc install bacondecomp, replace /* Goodman-Bacon decomp */
21 net install ddtiming, from(https://tgoldring.com/code/) /* Old Goodman-Bacon
22 ssc install eventstudyweights, replace /* Sun-Abraham decomp */
23 ssc install twoawayfeweights, replace /* deCDH decomp */
24
25 *** (3) DID packages ***
26 ssc install stackedev, replace /* Stacked regression */
27 ssc install eventstudyinteract, replace /* Sun-Abraham*/
28 ssc install csdid, replace /* Callaway-Sant'Anna */
29 ssc install did_multiplegt, replace /* deCDH */
30 ssc install did_imputation, replace /* Borusyak et al. */
31 ssc install did2s, replace /* Gardner */
32
```

仮想例 : Baker *et al.* (JF, 2022)

```
38 *** Case 2. Heterogeneous, Dynamic Effect ***
39 clear
40 set obs 1000
41 gen id = _n
42 set seed 100000
43 gen ai = rnormal(0,0.5)
44 forvalues t = 0(1)35 {
45 scalar t = rnormal(0,0.5)
46 gen lt`t' = t
47 }
48 gen rn = runiform()
49 gen group = 1 if rn <= 17/50
50 replace group = 2 if rn > 17/50
51 replace group = 3 if rn > 35/50
52
53 gen tr_time = 1989 if group == 1
54 replace tr_time = 1998 if group == 2
55 replace tr_time = 2007 if group == 3
56
57 reshape long lt, i(id) j(time)
58 replace time = time + 1980
59 gen d89 = (time >= 1989)
60 gen d98 = (time >= 1998)
61 gen d07 = (time >= 2007)
62
63 gen treat = d89 if group == 1
64 replace treat = d98 if group == 2
65 replace treat = d07 if group ==3
66
67 gen y = ai + lt + rnormal(0,0.5)
68 replace y = y + rnormal(0.5,0.2)*(time-1989)*treat if group == 1
69 replace y = y + rnormal(0.3,0.2)*(time-1998)*treat if group == 2
70 replace y = y + rnormal(0.1,0.2)*(time-2007)*treat if group == 3
```

仮想例 : Baker *et al.* (JF, 2022)

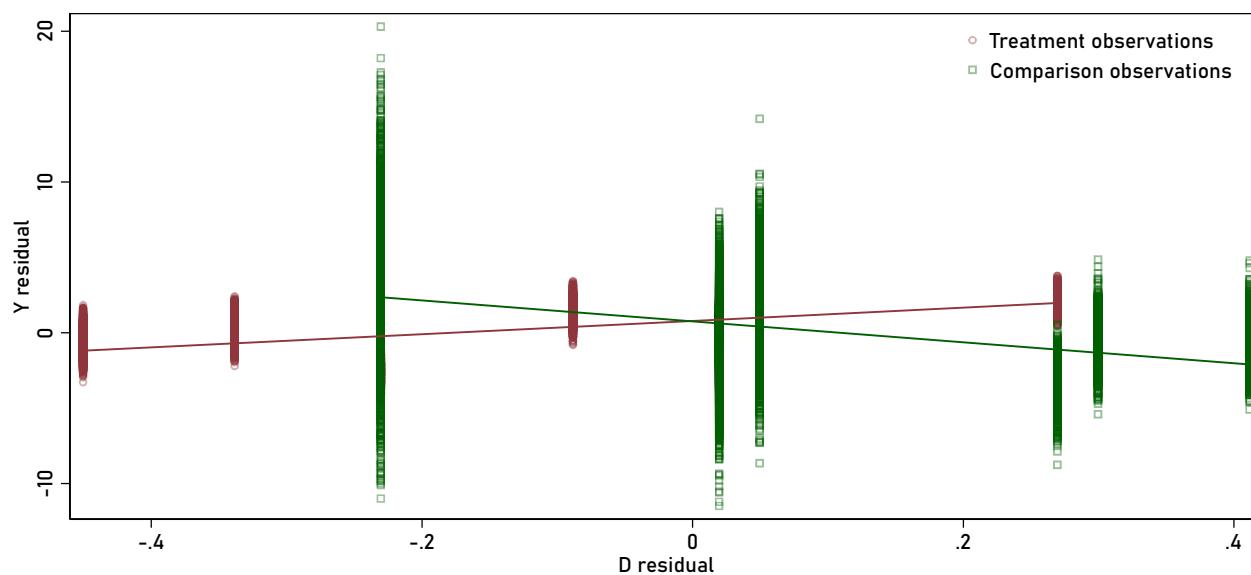
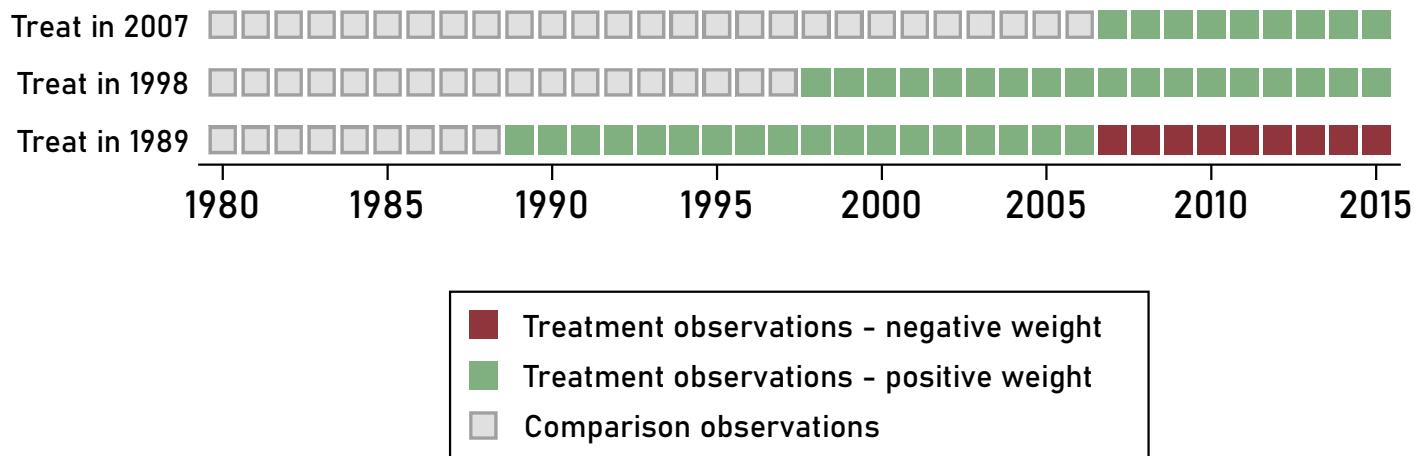
```
105 //***** II. Diagnosis *****/
106
107 *** II-A. Goodman-Bacon Decomposition ***
108
109 xtset id time
110 bacondecomp y treat, msymbols(oh t) /* Current version, aggregates early vs. late */
111 * ddtiming y treat, i(id) t(time) savedata(nfd2) replace /* Old version, each timing comparison */
112 * do "sim_bacondecomp.do"
113
114 * Memo: Similar packages by Sun-Abraham and deCDH, but Jakiela diagnosis
115 *       is more intuitive and easier to customize for quick diagnosis.
116
```



仮想例 : Baker *et al.* (JF, 2022)

```
116 *** II-B. Jakiela Diagnosis ***
117
118 *** prep weight & residualized outcomes ***
119 qui: reg treat i.time i.id
120 predict tr_resid, resid
121 gen tr_resid2 = tr_resid^2
122 egen denom = total(tr_resid2)
123 gen w = tr_resid/denom
124 bysort group time: egen wm = mean(w)
125 * Note: Confirm distribution of weights over time is not uniform across groups
126 *       by running : scatter w time if group == `i' & time >= tr_time
127
128 *** Diagnosis (1): Weight ***
129 tw (scatter group time if time < tr_time, ms(s) mfc(gs10) mfc(gs14) msiz(vlarge)) ///
130 (scatter group time if time >= tr_time & wm > 0, ms(s) mc(dkgreen*0.5) msiz(vlarge)) ///
131 (scatter group time if time >= tr_time & wm < 0, ms(s) mc(maroon) msiz(vlarge)), ///
132     aspect(0.1) plotregion(style(none)) ///
133     ylabel(1 "Treat in 1989" 2 "Treat in 1998" 3 "Treat in 2007" , ///
134         angle(0) noticks labsiz(small)) ytitle(" ") yscale(lstyle(none)) ///
135     legend(order(3 2 1) label(3 "Treatment observations - negative weight") ///
136             label(2 "Treatment observations - positive weight") ///
137             label(1 "Comparison observations") size(small) col(1)) xlabel(1980(5)2015)
138
139 *** Diagnosis (2): Heterogeneity ***
140 qui: reg y i.time i.id
141 predict y_resid, resid
142
143 tw (scatter y_resid tr_resid if treat == 0, m(oh) mc(maroon%40)) ///
144     (scatter y_resid tr_resid if treat == 1, m(sh) mc(dkgreen%40)) ///
145     (lfit y_resid tr_resid if treat == 0, lc(maroon)) ///
146     (lfit y_resid tr_resid if treat == 1, lc(dkgreen)) ///
147     , xtitle("D residual") ytitle("Y residual") ///
148     title("Case 2. Heterogeneous/Dynamic Effect" ///
149             "(Simulation 6 in Baker {it:et al.})", size(medsmall) margin(b=3)) ///
150     legend(order(1 "Treatment observations" 2 "Comparison observations")) ///
151     legend(ring(0) pos(1) col(1) region(style(off))) xsize(8.5)
152 graph save sim2_jak, replace
```

仮想例 : Baker *et al.* (JF, 2022)



仮想例 : Baker *et al.* (JF, 2022)

```
148 **** III. Alternative Estimators ****/
149
150 gen rel_time = time - tr_time
151 sum tr_time
152 gen lastcohort = (tr_time == r(max)) // dummy for the latest- or never-treated cohort
153
154 forvalues l = 0/27 {
155     gen lag`l' = (rel_time == `l')
156 }
157 forvalues l = 1/27 {
158     gen lead`l' = (rel_time == -`l')
159 }
160 replace lead1 = 0 // normalize first_treat = -1 to be zero
161 save "sim2_full.dta", replace
162
163 *** 0. Benchmark TWFE ***
164 use "sim2_full.dta", clear
165
166 reghdfe y lead* lag*, a(id time) cluster(id)
167 * Note 1: lead1 coef is set to zero. This will drop lags 18-27.
168 * Note 2: Runnin reghdfe y lag* lead* instead will drop leads 19-27 and lag 27,
169 *           and it will get us unbiased estimates since it "happens" to drop wrong comparisons.
170
171 event_plot, default_look stub_lag(lag#) stub_lead(lead#) together ///
172     graph_opt(xtitle("Periods since the event") ytitle("Estimates") ///
173     title("Naive TWFE Event Study", size(medsmall) margin(b=3)) xlabel(-5(1)5) ///
174     name(g0, replace)) trimlead(5) trimlag(5)
175 addplot: (scatteri 0 -5 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
176
```

仮想例 : Baker *et al.* (JF, 2022)

```
177 *** 1. Fully-saturated TWFE ***
178
179 * (1-A) Stacked regression *
180
181 * subsetting & stacking *
182 use "sim2_full.dta", clear
183
184 keep if (time >= 1989 - 5) & (time <= 1989 + 5)
185 save "sim2_sub1.dta", replace
186
187 use "sim2_full.dta", clear
188 keep if (time >= 1998 - 5) & (time <= 1998 + 5)
189 keep if group >= 2
190 save "sim2_sub2.dta", replace
191
192 use "sim2_full.dta", clear
193 keep if (time >= 2007 - 5) & (time <= 2007 + 5)
194 keep if group == 3
195
196 append using "sim2_sub1.dta"
197 append using "sim2_sub2.dta"
198
199 reghdfe y lag* lead*, a(id time) cluster(id)
200
201 event_plot, default_look stub_lag(lag#) stub_lead(lead#) together trimlead(5) trimlag(5) ///
202 graph_opt(xtitle("Periods since the event") ytitle("Estimates") xlabel(-5(1)5) ///
203 title("Stacked Regression", size(medsmall) margin(b=3)) xlabel(-5(1)5) ///
204 name(g1, replace))
205 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
206
207 * Note 1: "stackedev" package requires never-treated units.
208 *           Hence cannot be used in this example b/c all units eventually receive treatment.
209 * Note 2: Runnding reghdfe y lead* lag* removes "correct" comparions due to multicolinarity.
```

仮想例 : Baker *et al.* (JF, 2022)

```
212 * (1-B) Sun and Abraham *
213
214 use "sim2_full.dta", clear
215
216 timer clear
217 timer on 1
218 eventstudy interact y lead* lag*, vce(cluster i) absorb(id time) cohort(tr_time) ///
219     control_cohort(lastcohort)
220 timer off 1
221
222 event_plot e(b_iw)#e(V_iw), default_look stub_lag(lag#) stub_lead(lead#) ///
223     together trimlead(5) trimlag(5) ///
224     graph_opt(xtitle("Periods since the event") ytitle("Estimates") xlabel(-5(1)5) ///
225         title("Sun and Abraham", size(medsmall) margin(b=3)) name(g2, replace))
226 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
227
228 /*
229 eventstudyweights y lead* lag*, vce(cluster i) absorb(id time) ///
230     cohort(tr_time) rel_time(rel_time) control_cohort(lastcohort) saveweights("weights")
231 import excel "weights.xlsx", clear firstrow
232 keep L1event first_treat rel_time
233 reshape wide L1event, i(rel_time) j(first_treat)
234 twoway line L1event* rel_time, xline(1, lp(dash) lw(thin) lc(black)) ///
235     yline(0, lp(dash) lw(thin) lc(black)) xlabel(-30(5)30) ///
236     ytitle("weights") xtitle("time relative to treatment") ///
237     title("Sun-Abraham Decomp""(Weights for t = 1)", size(medsmall) margin(b=2))
238 */
239
```

仮想例 : Baker *et al.* (JF, 2022)

```
239 * (1-C) Wooldridge *
240 use "sim2_full.dta", clear
241
242 forvalues g = 1/2 {
243     forvalues l = 1/27 {
244         gen lead`l'_g`g' = lead`l'*(group == `g')
245     }
246 }
247 forvalues g = 1/2 {
248     forvalues l = 0/27 {
249         gen lag`l'_g`g' = lag`l'*(group == `g')
250     }
251 }
252
253 reghdfe y lag*_`* lead*_`*, a(time group) cluster(id)
254
255 * Note: The following estimate the same.
256 *          ssc install jwdid, replace
257 *          jwdid y, ivar(id) tvar(time) gvar(tr_time)
258
259 /*
260 lincom (lag0_g1*(1/2) + lag0_g2*(1/2))
261 lincom (lag1_g1*(1/2) + lag1_g2*(1/2))
262 lincom (lag2_g1*(1/2) + lag2_g2*(1/2))
263 lincom (lag3_g1*(1/2) + lag3_g2*(1/2))
264 lincom (lag4_g1*(1/2) + lag4_g2*(1/2))
265 lincom (lag5_g1*(1/2) + lag5_g2*(1/2))
266 */
```

仮想例 : Baker *et al.* (JF, 2022)

```
268 gen grid = _n in 1/11
269 gen beta = .
270 gen var = .
271
272 forvalues i = 0/5 {
273     reghdfe y lag*_`i' lead*_`i', a(time group) cluster(id)
274     margins, expression({_b[lag`i'_g1]+_b[lag`i'_g2])/2) post
275     mat blag`i' = e(b)
276     mat vlag`i' = e(V)
277     replace beta = blag`i'[1,1] if grid == `i'+6
278     replace var = vlag`i'[1,1] if grid == `i'+6
279 }
280 forvalues i = 1/5 {
281     reghdfe y lag*_`i' lead*_`i', a(time group) cluster(id)
282     margins, expression(_b[lead`i'_g1]+_b[lead`i'_g2]/2) post
283     mat blead`i' = e(b)
284     mat vlead`i' = e(V)
285     replace beta = blead`i'[1,1] if grid == 6-`i'
286     replace var = vlead`i'[1,1] if grid == 6-`i'
287 }
288
289 gen lower = beta - 1.96*sqrt(var) in 1/11
290 gen upper = beta + 1.96*sqrt(var) in 1/11
291 replace grid = grid - 6
292 tw (rarea upper lower grid, lc(navy) fc(navy) fi(30)) ///
293 (connected beta grid, lc(navy) mc(navy)) ///
294 , xtitle("Periods since the event") ytitle("Estimates") ///
295 title("Wooldridge", size(medsmall) margin(b=3)) xlabel(-5(1)5) ///
296 name(g3, replace) legend(off) ///
297 yline(0, lc(black*.5)) xline(0, lc(black*.5) lp(dash))
298 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
```

仮想例 : Baker *et al.* (JF, 2022)

```
300 *** 2. Rolling methods ***
301
302 * (2-A) Callaway-Sant'Anna *
303 use "sim2_full.dta", clear
304
305 timer on 2
306 csdid y, ivar(id) time(time) gvar(tr_time) notyet ///
307     agg(event) wboot driwp
308     * Memo: Default uses "uniform" SEs to account for multiple testing.
309 timer off 2
310 estat event, estore(cs)
311
312 event_plot cs, default_look stub_lag(Tp#) stub_lead(Tm#) together ///
313     trimlead(5) trimlag(5) ///
314     graph_opt(xtitle("Periods since the event") ytitle("Estimates") ///
315     title("Callaway and Sant'Anna", size(medsmall) margin(b=3)) xlabel(-5(1)5)
316     name(g4, replace))
317 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
318
```

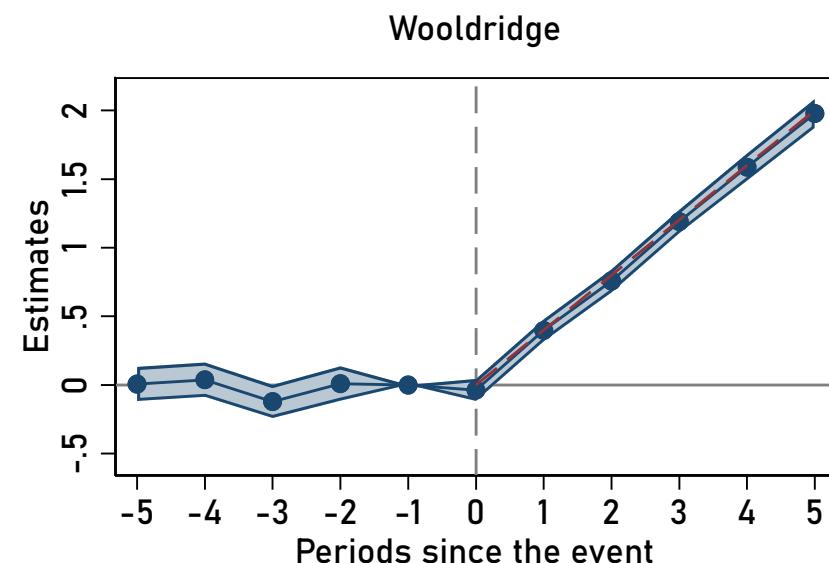
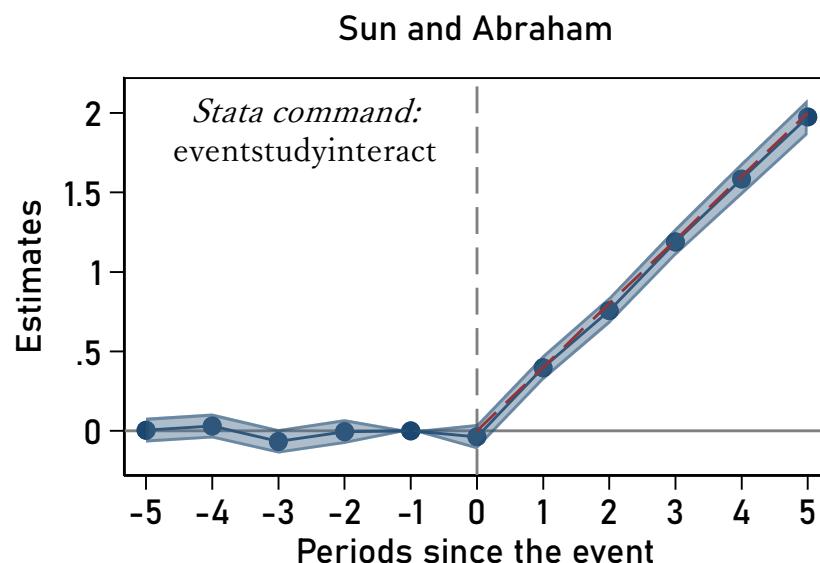
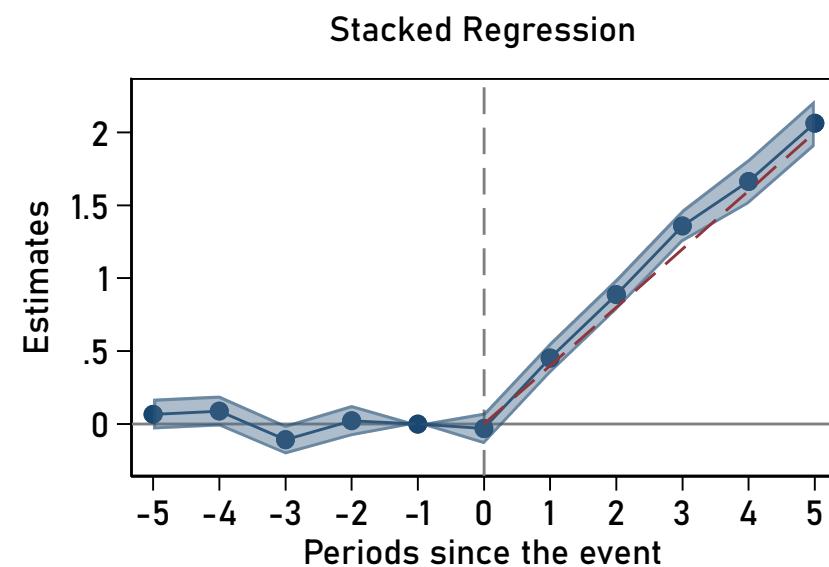
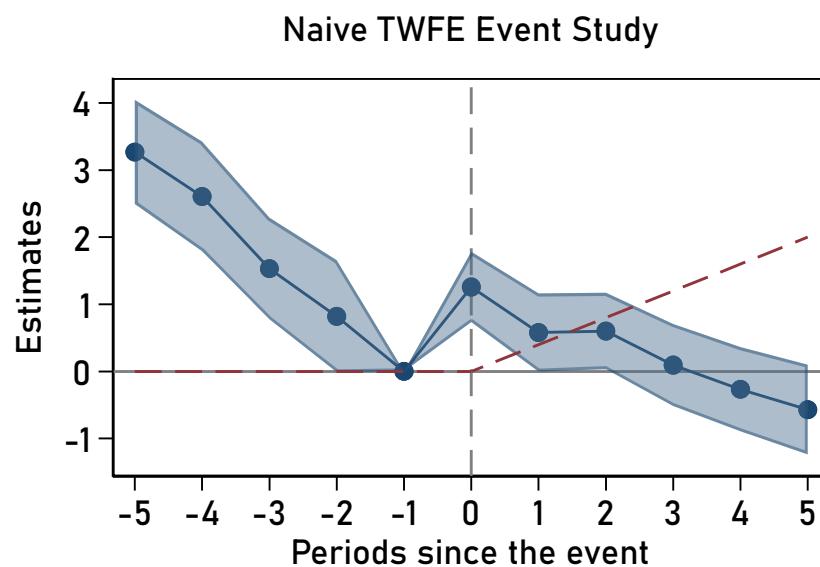
仮想例 : Baker *et al.* (JF, 2022)

```
319 * (2-B) de Chaisemartin and D'Haultfoeuille *
320 use "sim2_full.dta", clear
321
322 timer on 3
323 did_multiplegt y id time treat, robust_dynamic ///
324     dynamic(5) placebo(5) breps(100) cluster(id)
325 timer off 3
326
327 event_plot e(estimates)#e(variances), default_look ///
328     stub_lag(Effect_#) stub_lead(Placebo_#) together ///
329     trimlead(5) trimlag(5) ///
330     graph_opt(xtitle("Periods since the event") ytitle("Estimates") ///
331     title("de Chaisemartin and D'Haultfoeuille", size(medsmall) margin(b=3)) ///
332     xlabel(-5(1)5) name(g5, replace))
333 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
334
335 /*
336 twowayfeweights y id time treat, type(feTR) path("dCDH_decomp.dta")
337 * twowayfeweights y id time treat, type(feS) path("dCDH_decomp.dta")
338 use "dCDH_decomp.dta", clear
339 egen group = group(weight)
340 tw (scatter weight Time_TWFE if group == 1, mc(maroon*0.5)) ///
341 (scatter weight Time_TWFE if group == 2, mc(green*0.1)) ///
342 (scatter weight Time_TWFE if group == 3, mc(green*0.3)) ///
343 (scatter weight Time_TWFE if group == 4, mc(green*0.5)) ///
344 (scatter weight Time_TWFE if group == 5, mc(green*0.7)) ///
345 (scatter weight Time_TWFE if group == 6, mc(green)), ///
346 yline(0, lc(grey) lw(thin)) xline(1998 2007, lc(red*0.5) lp(thin)) ///
347 xtitle(Time, size(medsmall)) ytitle(Weights, size(medsmall)) ///
348 title(TWFE Regression Weight to Each Obs, size(medsmall) margin(b=2)) legend(off)
```

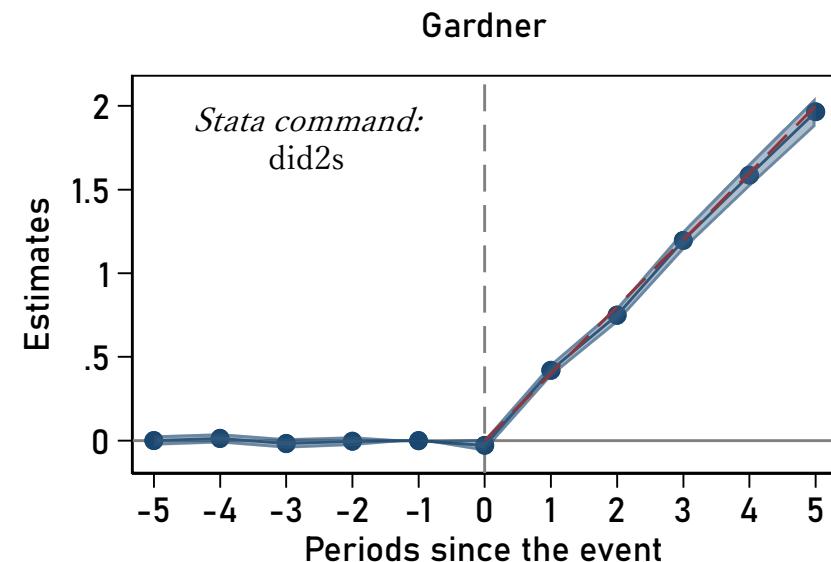
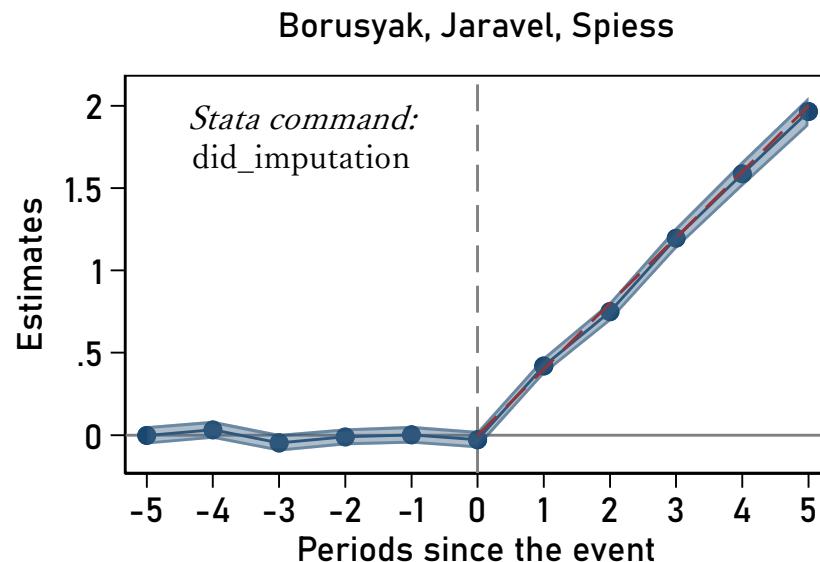
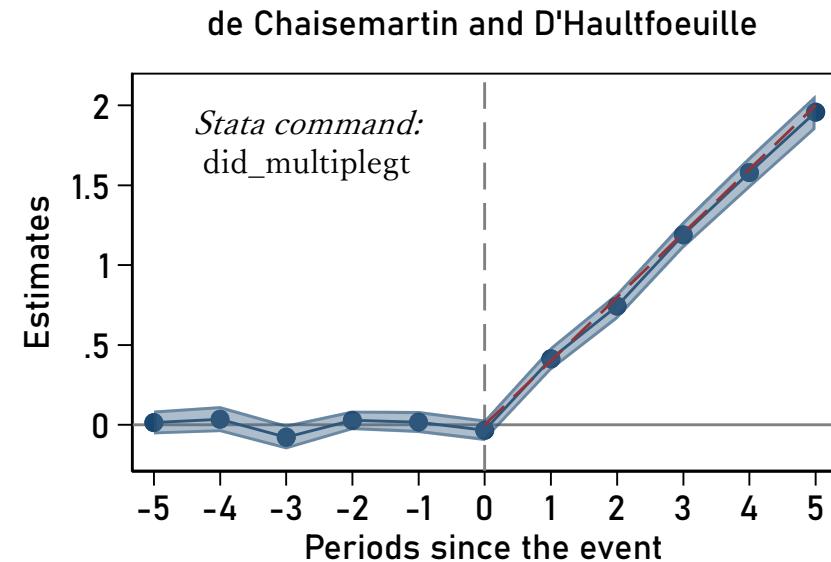
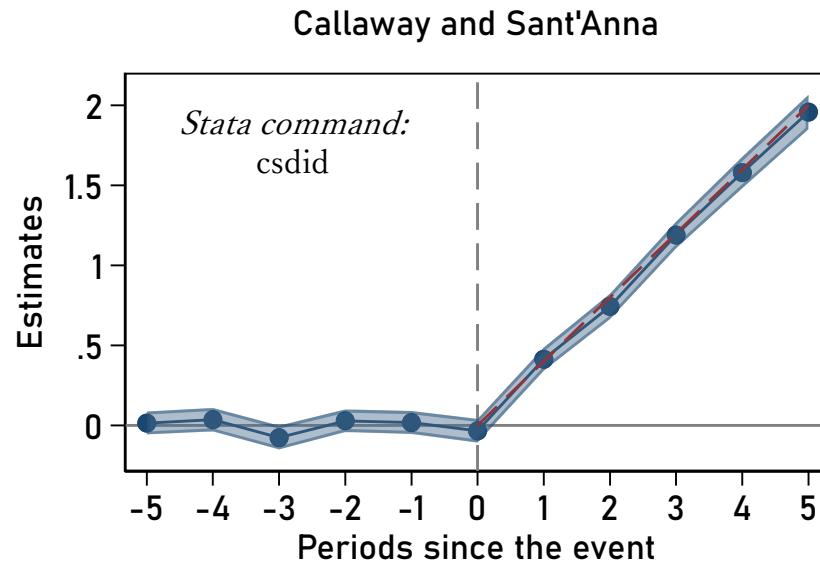
仮想例 : Baker *et al.* (JF, 2022)

```
351 *** 3. Imputation methods ***
352
353 * (3-A) Borusyak et al. *
354 use "sim2_full.dta", clear
355
356 timer on 4
357 did_imputation y id time tr_time, autosample horizon(0/5) pretrend(5)
358 timer off 4
359
360 event_plot, default_look stub_lag(tau#) stub_lead(pre#) together trimlead(5) trimlag(5) ///
361     graph_opt(xtitle("Periods since the event") ytitle("Estimates") ///
362     title("Borusyak, Jaravel, Spiess", size(medsmall) margin(b=3)) ///
363     xlabel(-5(1)5) name(g6, replace))
364 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
365
366 * (3-B) Gardner *
367 use "sim2_full.dta", clear
368
369 timer on 5
370 reg y i.id i.time if treat == 0, nocons
371 predict yhat, residual
372 reg yhat lead* lag* if d89 != 1 | time < 2007, nocons cluster(id)
373 timer off 5
374
375 event_plot, default_look stub_lag(lag#) stub_lead(lead#) together trimlead(5) trimlag(5) ///
376     graph_opt(xtitle("Periods since the event") ytitle("Estimates") xlabel(-5(1)5) ///
377     title("Gardner", size(medsmall) margin(b=3)) xlabel(-5(1)5) ///
378     name(g7, replace))
379 addplot: (scatteri 0 0 2 5, xlabel(-5(1)5) recast(line) lp(dash) lc(maroon))
380
```

Alternative Estimators in Stata



Alternative Estimators in Stata



仮想例 : Baker *et al.* (JF, 2022)

Method	Computing Time (hh:mm:ss)
Sun and Abraham	00:00:04
Callaway and Sant'Anna	00:00:11
de Chaisemartin and D'Haultfoeuille	00:08:29
Borusyak, Jaravel, Spiess	00:00:10
Gardner	00:00:04

Stata/Rによる演習

(2) 実践編

- Cicala (AER, 2022) のデータ
(電力自由化が経済厚生に与えた影響)
 - 実証研究の文脈と研究デザイン
 - 実証研究で生じうる問題は？
 - Diagnosis
 - Jakiela's diagnosis
 - 三つのアプローチによる推定
 - SA vs. CS vs. deCDH vs. BJS

Cicala (AER, 2022)

- 1999～2013年の電力自由化の経済効果：

$$\ln(Y_{it}) = \alpha_{im} + \lambda_{my} + \delta D_{it} + \gamma_i \ln(L_{it}) + \epsilon_{it}$$

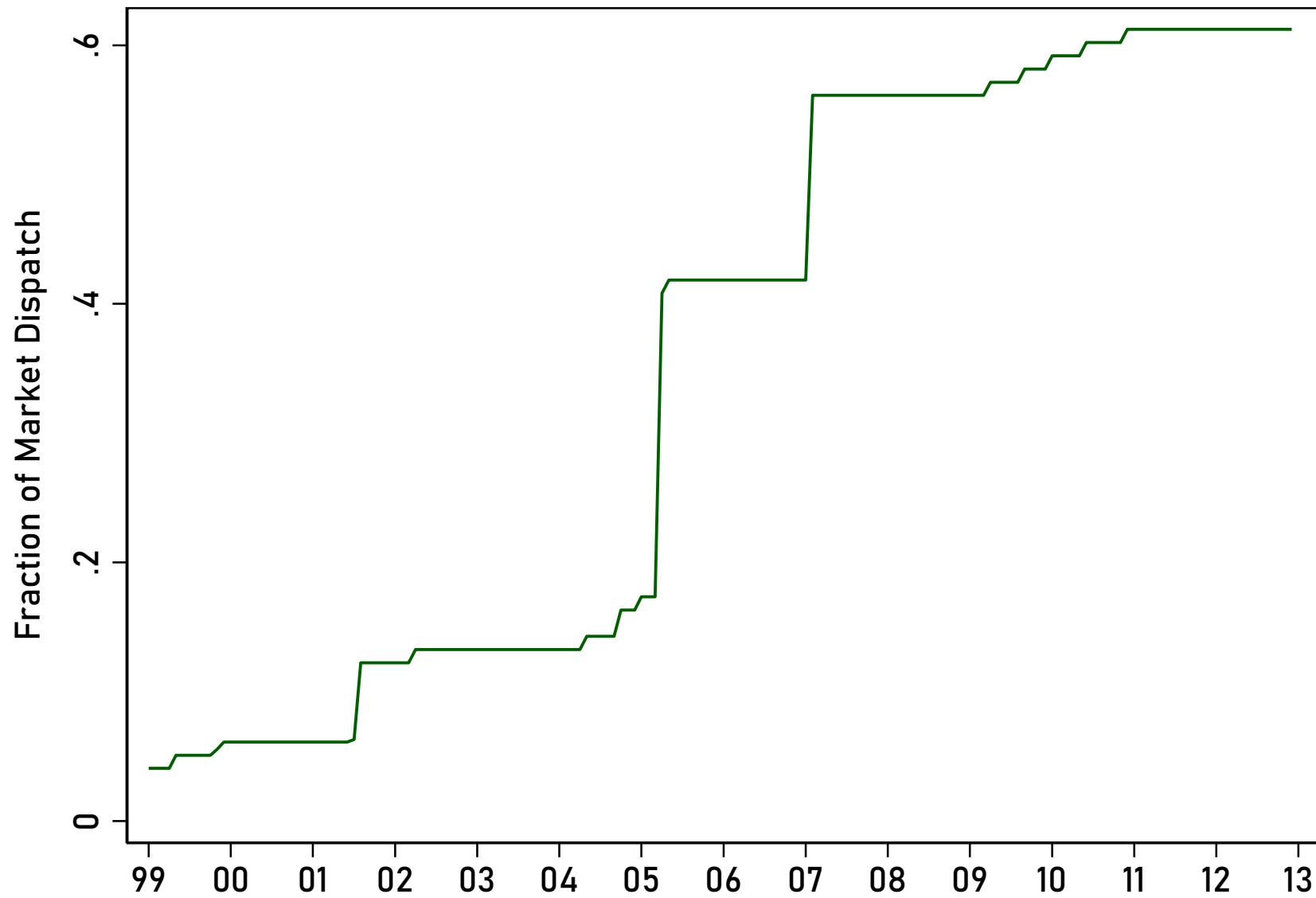
- i : Power Control Area (電力地域)
- t : 日(d)・月(m)・年(y)
- Y_{it} : 電力取引による経済余剰
- D_{it} : 電力自由化(送電分離)の有無
- L_{it} : 電力需要量

- ※1. 簡略版の推定式. 元論文では, Hourly dataを用い, さらに多くの共変量を入れている.
- ※2. 元論文では, PCAの規模の違いを考慮しWLS-TWFEによって推定.

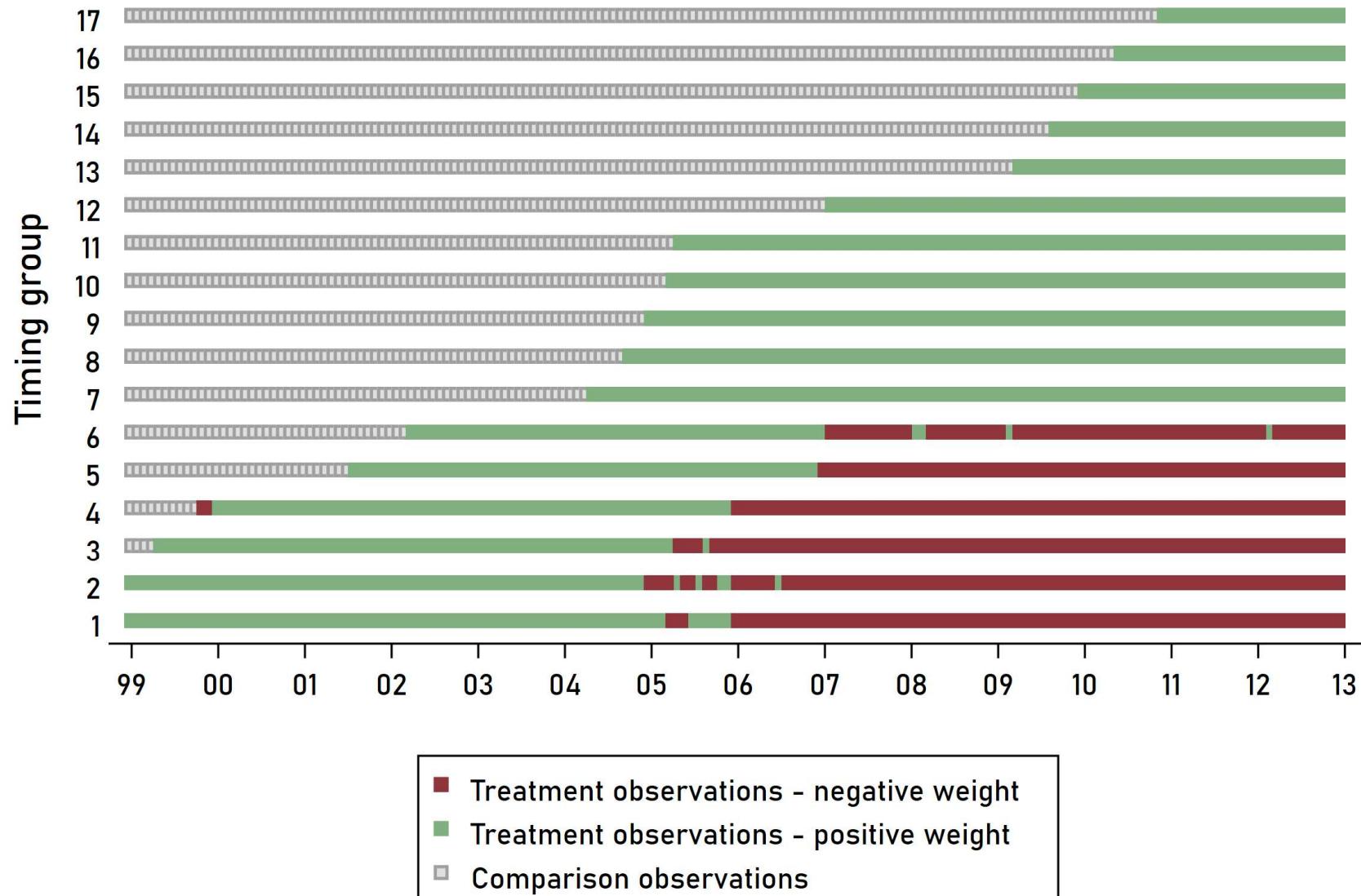
Cicala (AER, 2022)

# of PCAs	98
# of periods (months)	168
# of timing groups	17
# of obs. (daily)	501172
Fraction treated	0.349
Outcome var.	Treat = 0 Treat = 1
Out-of-merit costs (\$US)	14,527 25,431 (29,706.42) (44,032.51)
Gains from trade (\$US)	7,181 6,755 (21,640.67) (25,775.62)

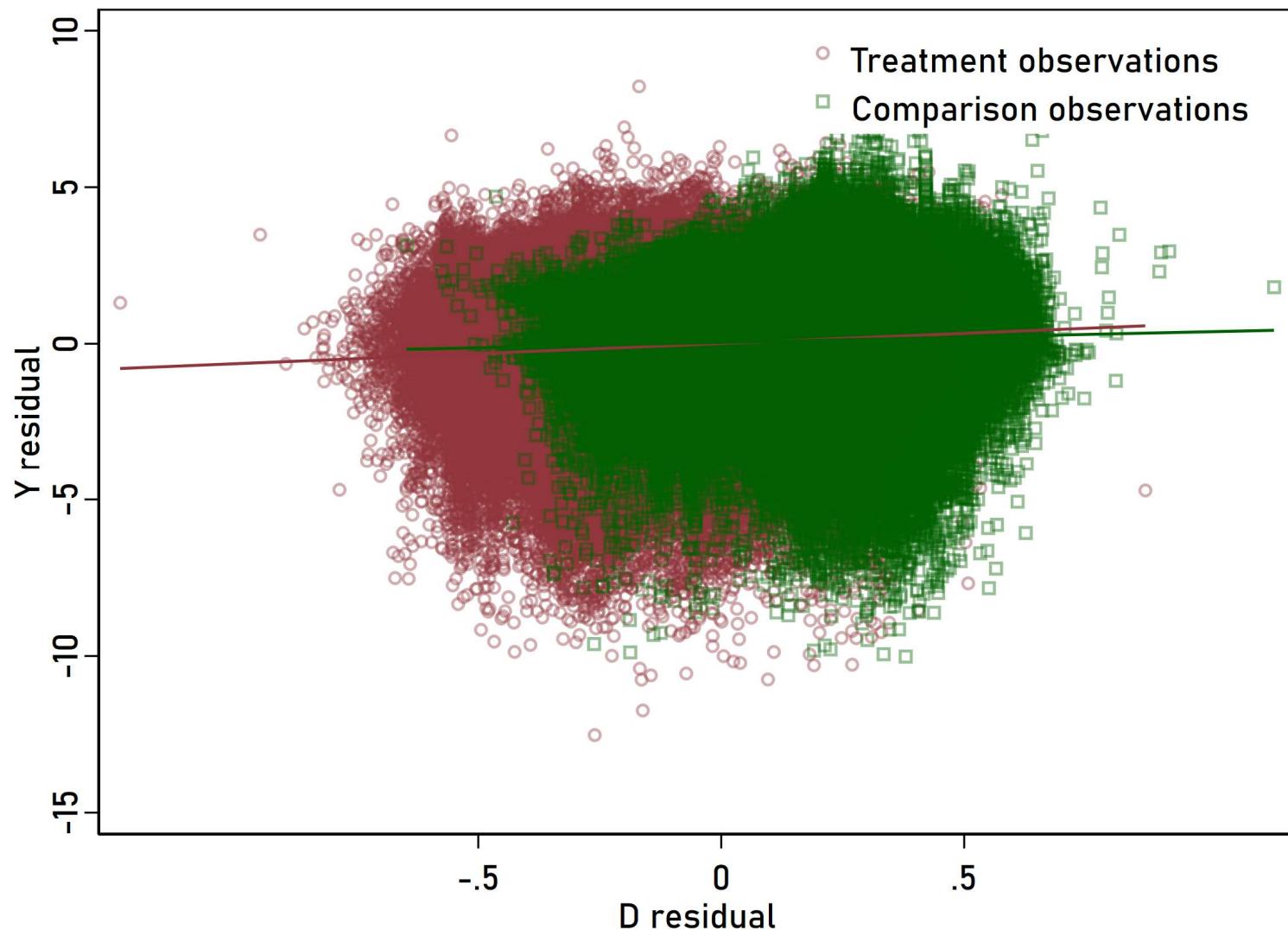
Cicala (AER, 2022)



Cicala (AER, 2022)



Cicala (AER, 2022)



Cicala (AER, 2022)

```
31 **** I. Overview of the Data ****/
32
33 cd "$hdir"
34 use "cicala_aer_2022_ready", clear
35
36 gen y = log_ideal_trade_surplus /* Gains from trade in log */
37 gen treat = market_operation /* Market dispatch */
38 gen tr_time = mkt_ym /* Treatment timing */
39 egen group = group(mkt_ym) /* Timing group */
40 gen time = ym(year,month) /* Time var */
41 egen pca_id = group(pca_abbrev99) /* PCA ID var */
```

Cicala (AER, 2022)

```
120
127 **** III. Alternative Estimators ****/
128
129 * Define relative time *
130
131 * Memo: Cicala (2022) estimate only within 24-month window.
132 gen rel_time = time - tr_time
133 sum tr_time
134 gen lastcohort = (tr_time == r(max)) // dummy for the latest-treated cohort
135 gen never = mi(mkt_ym) // dummy for never-treated cohort
136 forvalues l = 0/24 {
137     gen lag`l' = (rel_time == `l')
138 }
139 forvalues l = 1/24 {
140     gen lead`l' = (rel_time == -`l')
141 }
142 replace lead1 = 0 // normalize first_treat = -1 to be zero
143 save "cicala_alt.dta", replace
144
145 *** 0. Benchmark TWFE ***
146 use "cicala_alt.dta", clear
147
148 reghdfe y lead* lag* [aweight=wgt], ///
149     a(year i.pca_id##i.month i.pca_id#c.log_load) cluster(pca_modate)
150 * Note: lead1 coef is set to zero.
151 *      Unlike in simulated data, all included leads/lags are estimated.
152
153 event_plot, default_look stub_lag(lag#) stub_lead(lead#) together ///
154     trimlead(12) trimlag(18) ///
155     graph_opt(xtitle("Months since the event") ytitle("Estimates") ///
156     title("Naive TWFE Event Study", size(medsmall) margin(b=3)) ///
157     xlabel(-12(3)18) name(cg0, replace))
158 graph save cicala_g0, replace
```

Cicala (AER, 2022)

```
--  
164 * (1-B) Sun and Abraham *  
165 use "cicala_alt.dta", clear  
166  
167 timer clear  
168 timer on 1  
169 eventstudy interact y lead* lag* [aweight=wgt], vce(cluster pca_modade) ///  
170     absorb(year i.pca_id##i.month i.pca_id#c.log_load) cohort(tr_time) ///  
171     control_cohort(never)  
172 timer off 1  
173  
174 event_plot e(b_iw)#e(V_iw), default_look stub_lag(lag#) stub_lead(lead#) ///  
175     together trimlead(12) trimlag(18) ///  
176     graph_opt(xtitle("Months since the event") ytitle("Estimates") ///  
177     title("Sun and Abraham", size(medsmall) margin(b=3)) xlabel(-12(3)18) ///  
178     name(cg1, replace))  
179 graph save cicala_g1, replace
```

Cicala (AER, 2022)

```
179 *** 2. Rolling methods ***
180
181 * (2-A) Callaway-Sant'Anna *
182 use "cicala_alt.dta", clear
183
184 * Make sure to code "never treated group" = 0 *
185 replace tr_time = 0 if mi(tr_time)
186
187 * Absorb FEs and Time-varying Covariates a la Caetano et al. (2022) *
188 qui reg y i.year i.pca_id##i.month i.pca_id#c.log_load ///
189     if treat == 0, nocons /* weighting does not matter here */
190 predict ycs, residual
191
192 timer on 2
193 csdid ycs, time(time) gvar(tr_time) agg(event) ///
194     wboot(reps(50)) rseed(1) cluster(pca_modate)
195 * Memo 1: Drop ivar() to allow for multiple obs within a panel.
196 * Memo 2: Use wgt as covariate to weight obs.
197 timer off 2
198 * estat event, estore(cs)
199
200 event_plot, default_look stub_lag(Tp#) stub_lead(Tm#) together ///
201     trimlead(12) trimlag(18) ///
202     graph_opt(xtitle("Months since the event") ytitle("Estimates") ///
203     title("Callaway and Sant'Anna", size(medsmall) margin(b=3)) ///
204     xlabel(-12(6)18) name(cg2, replace))
205 graph save cicala_g2_2, replace
206
```

Cicala (AER, 2022)

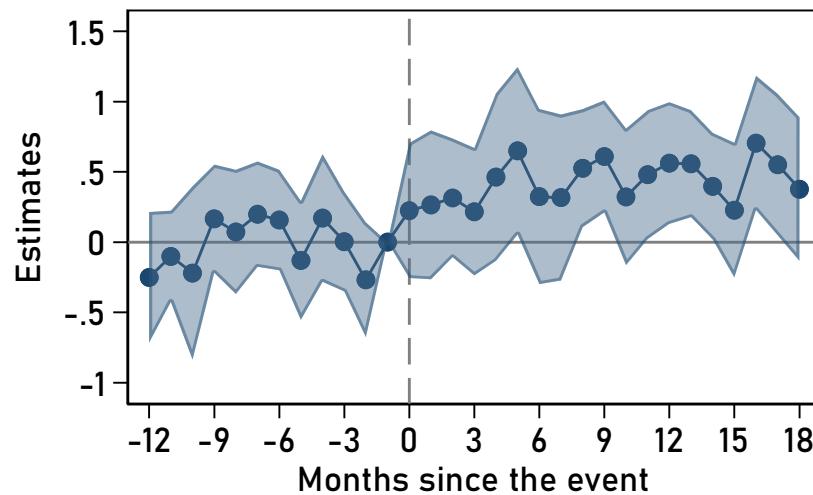
```
223
224 * (2-B) de Chaisemartin and D'Haultfoeuille *
225 use "cicala_alt.dta", clear
226
227 * Absorb FEs and Time-varying Covariates a la Caetano et al. (2022) *
228 qui reg y i.year i.pca_id##i.month i.pca_id#c.log_load ///
229     if treat == 0, nocons /* weighting does not matter here */
230 predict ydc, residual
231
232 timer on 3
233 did_multiplegt ydc pca_id time treat, robust_dynamic dynamic(18) placebo(12) ///
234     breps(50) seed(1) cluster(pca_modate) firstdiff_placebo weight(wgt)
235 * Memo: Package allows for time-varying controls, but not factor variables.
236 * Hence, residualize manually.
237 timer off 3
238
239 event_plot e(estimates)#e(variances), default_look ///
240     stub_lag(Effect_#) stub_lead(Placebo_#) together ///
241     trimlead(12) trimlag(18) ///
242     graph_opt(xtitle("Months since the event") ytitle("Estimates") ///
243     title("de Chaisemartin and D'Haultfoeuille", size(medsmall) ///
244     margin(b=3)) xlabel(-12(6)18) name(cg3, replace))
245 graph save cicala_g3, replace
```

Cicala (AER, 2022)

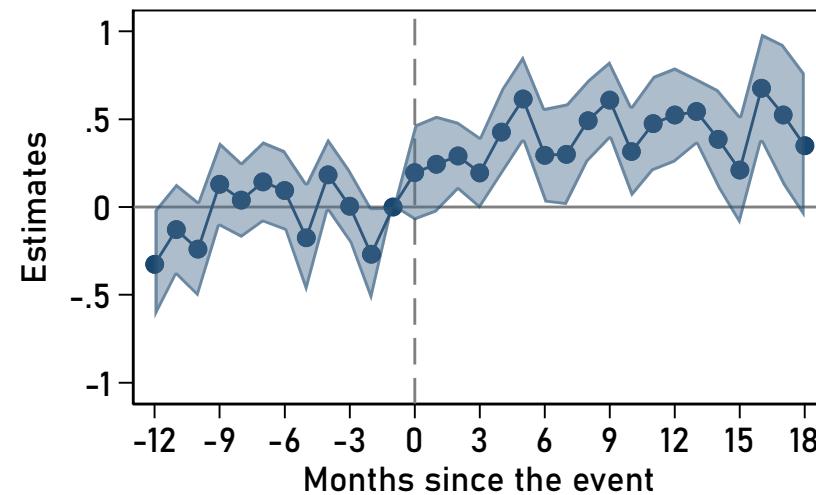
```
246
247 *** 3. Imputation methods ***
248
249 * (3-A) Borusyak et al. *
250 use "cicala_alt.dta", clear
251
252 timer on 4
253 did_imputation y pca_id time tr_time [aweight=wgt], ///
254     horizon(0/18) fe(year month pca_id#month) unitc(log_load) ///
255     autosample pretrend(12)
256     * Full model get the error message: Could not run imputation
257     * for some observations b/c some absorbed variables/FEs
258     * are collinear in the D==0 subsample but in the full sample.
259 timer off 4
260
261 event_plot, default_look stub_lag(tau#) stub_lead(pre#) together ///
262     trimlead(12) trimlag(18) ///
263     graph_opt(xtitle("Months since the event") ytitle("Estimates") ///
264     title("Borusyak, Jaravel, Spiess", size(medsmall) margin(b=3)) ///
265     xlabel(-12(6)18) ///
266     name(cg4, replace))
267 graph save cicala_g4, replace
```

Cicala (AER, 2022)

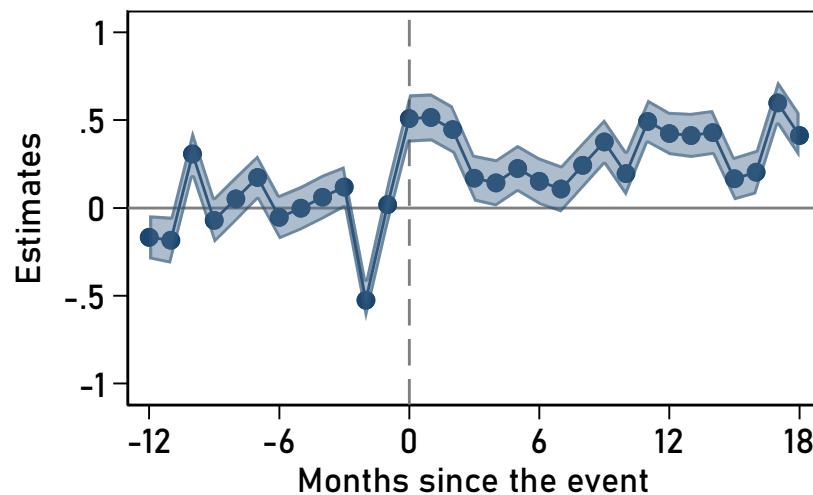
Naive TWFE Event Study



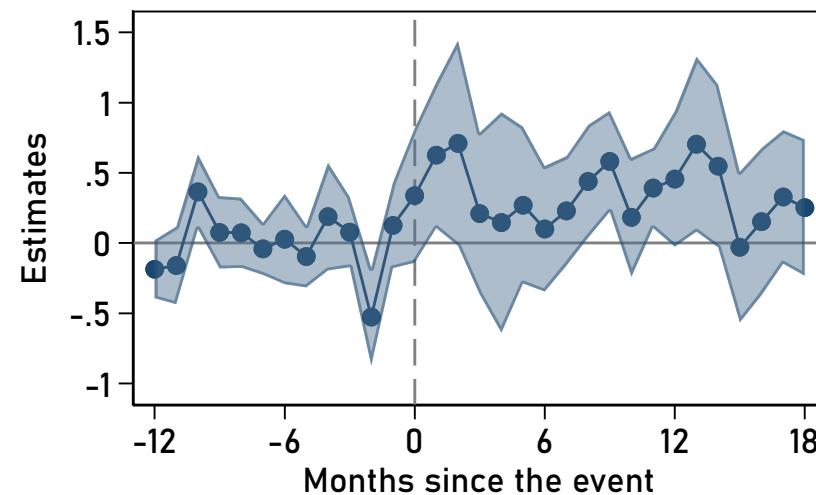
Sun and Abraham



Callaway and Sant'Anna

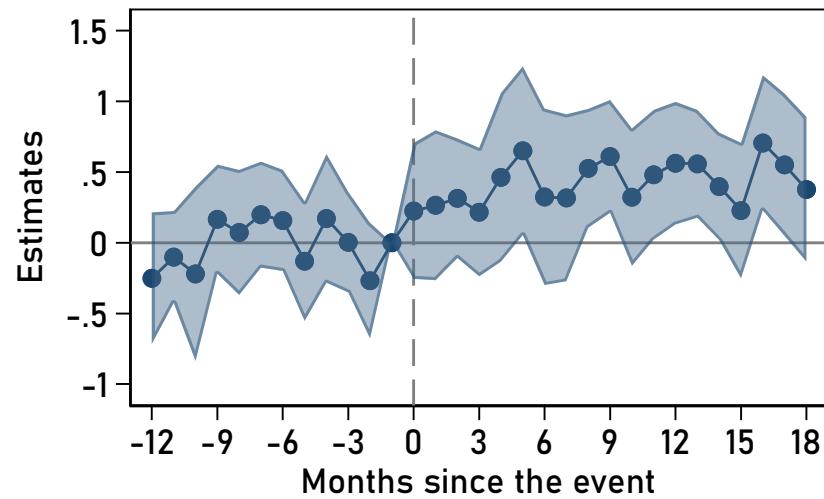


de Chaisemartin and D'Haultfoeuille

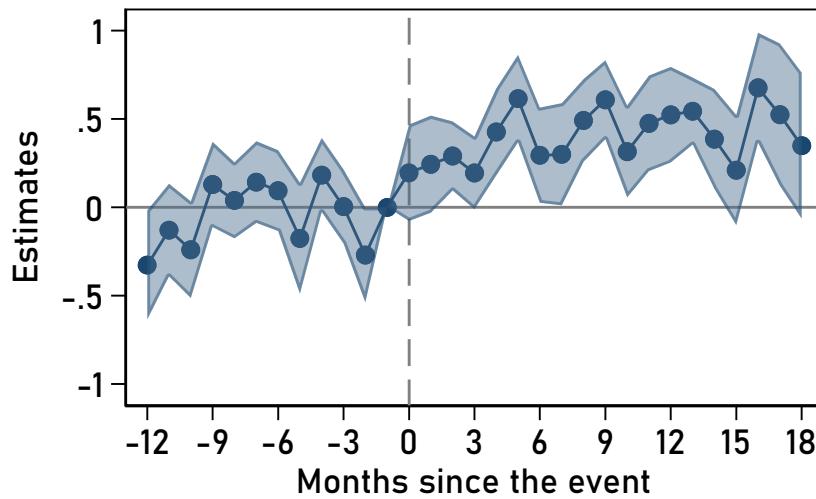


Cicala (AER, 2022)

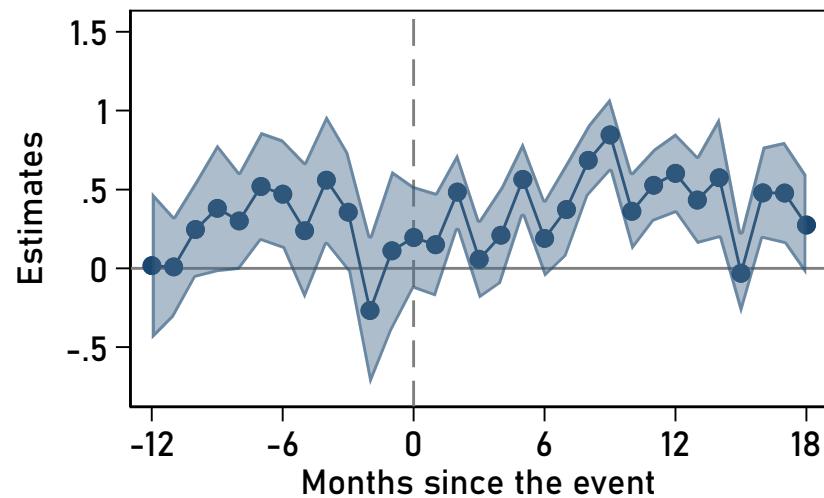
Naive TWFE Event Study



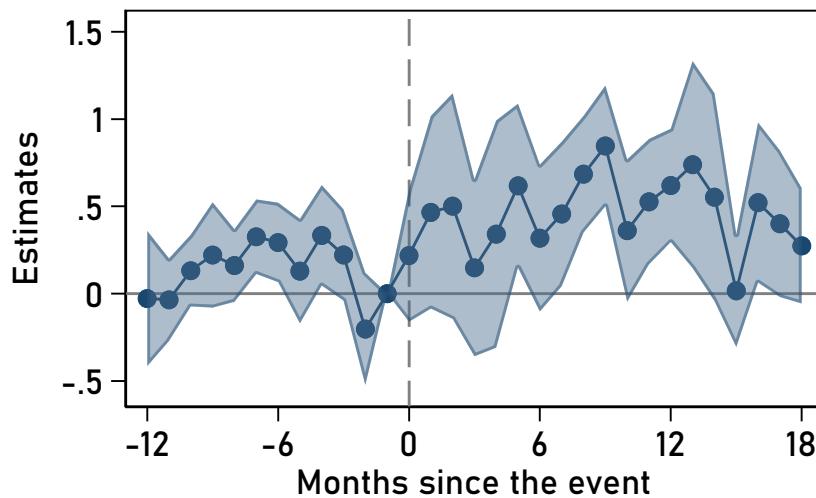
Sun and Abraham



Borusyak, Jaravel, Spiess



Gardner



Cicala (AER, 2022)

Method	Computing Time (hh:mm:ss)
Sun and Abraham	06:00:40
Callaway and Sant'Anna	07:13:47
de Chaisemartin and D'Haultfoeuille	00:30:41
Borusyak, Jaravel, Spiess	00:13:00
Gardner	00:00:58