Applied Econometrics

Difference-in-differences and synthetic control

Prof. Bruno Lanz
University of Neuchâtel

MSc in Applied Economics

Where we stand

- To estimate a treatment effect with observational data, we need to deal with selection bias
- Main insight from the omitted variable bias: selection bias materializes as confounding factors which are
  1. correlated with treatment status
  2. affect the outcome of interest
- If units who benefit more from the treatment self-select into the treatment group, the regression coefficient is subject to selection bias
- Control variables: adjust the comparison between treated and non-treated units within clusters
  - Unobserved heterogeneity (e.g. ability) can be captured with fixed effects when it is constant within-cluster

Objectives for today

- While control variables and fixed effects are useful, they do not guarantee causal identification
  - Instead approximate experimental random assignment with "quasi-experiments" or "natural experiments"
- Difference in differences: Compare before and after treatment among treated and control units
  - Exploit control units to estimate a counterfactual outcome for treated units
  - MM chapter 5, MHE chapter 5.2
- Extension for quantitative case studies: Synthetic control method
  - Construct a hypothetical unit that is observationally similar to the treated unit

Toward difference-in-differences: simple differences

- "Natural" experiment: exploit a change that was not made explicitly to quantify the effect of the intervention
  - Key issue: what is a valid counterfactual?
- Consider a country that introduces a policy
  - We want to quantify the impact of the policy on Y_{it} for all units i observed over time t in that country
- An obvious approach is to compare the year before ($T_{it} = 0$) and after ($T_{it} = 1$) the change:

$$Y_{it} = \alpha^D + \beta^D T_{it} + u_{it}$$

- β^D measures the difference in average outcome use before and after the policy is introduced

Data for today

- Consider a country that introduces a carbon tax in time t^*
- Assume we design a survey to measure energy use (or CO_2 emissions) and administer it to 28 firms 6 months before the tax is introduced
  - We administer the survey again 6 months after the tax is introduced
  - Our outcome variable is energy use and it is measure both before and after the intervention
- For now we assume that these are not the same firms, and we just stack the data as if this were two cross-sections
  - In our data firms in the second wave of the survey are identified with a dummy variable *after*
- We consider the log of energy use as our outcome variable

```
. gen lnenergy = ln(energy)
```

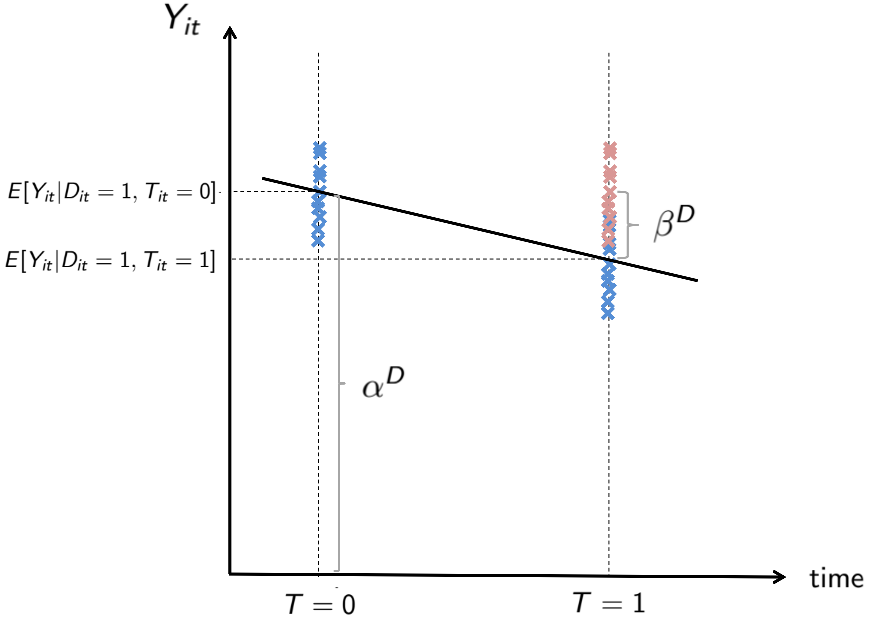
	country	after	energy	lnenergy		
22	1	0	6.97233	1.941949		
23	1	0	8.26153	2.11161		
24	1	0	7.00745	1.946974		
25	1	0	7.18769	1.97237		
26	1	0	9.16809	2.215729		
27	1	0	6.62426	1.890738		
28	1	0	7.88208	2.064592		
29	1	1	7.46592	2.010348		
30	1	1	6.69519	1.901389		
31	1	1	7.34096	1.993469		
32	1	1	4.99213	1.607863		
33	1	1	6.16065	1.818183		
34	1	1	6.90568	1.932344		
35	1	1	7.43338	2.005981		
36	1	1	6.29807	1.840243		
37	1	1	6.00559	1.792691		

```
.
. mean lnenergy, over(after)

Mean estimation                    Number of obs     =         56

            0: after = 0
            1: after = 1
```

Over	Mean	Std. Err.	[95% Conf.	Interval]
lnenergy				
0	1.991856	.0274746	1.936796	2.046916
1	1.911473	.0257771	1.859815	1.963131

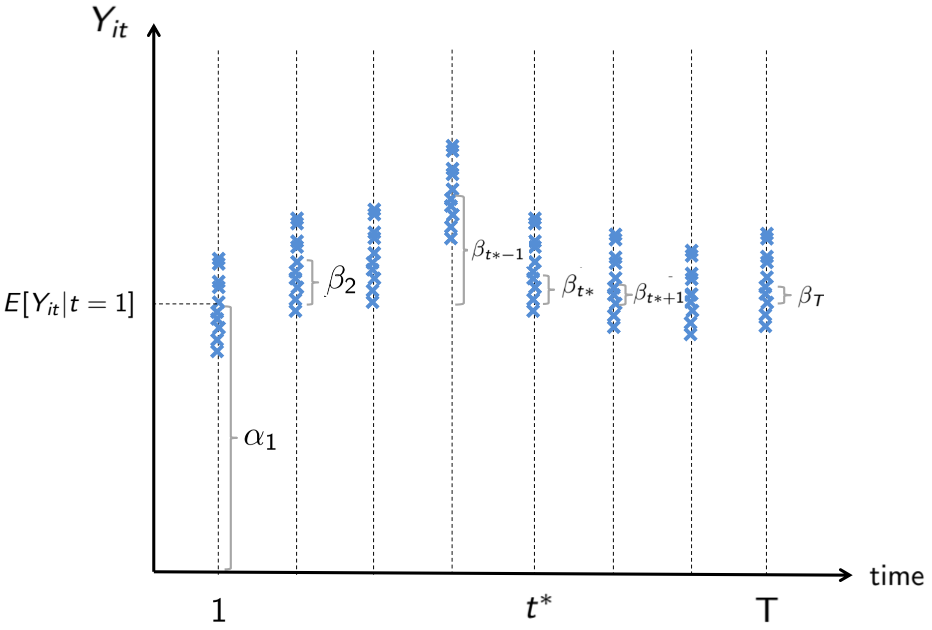
```
. reg lnenergy after, robust

Linear regression                               Number of obs     =         56
                                                F(1, 54)          =       4.55
                                                Prob > F          =     0.0374
                                                R-squared         =     0.0778
                                                Root MSE          =     .14096
```

lnenergy	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
after	-.0803831	.0376738	-2.13	0.037	-.1559144	-.0048518
_cons	1.991856	.0274746	72.50	0.000	1.936773	2.046939


Y_{it}
$E[Y_{it}|D_{it}=1, T_{it}=0]$
$E[Y_{it}|D_{it}=1, T_{it}=1]$
β^D
α^D
time
$T=0$
$T=1$

Event study with multiple periods

- Problem: with just one period before and one after we cannot distinguish the impact of policy from the passage of time
  - A dummy equal to one if the policy is introduced, zero otherwise, is redundant
- Implicitly: we assume that pre-treatment observations are a good control group
  - $\hat{\beta}$ measures the impact of the policy only if we assume that there would have been no change in $E[Y_{it}]$ absent the policy change
- If we have more than two periods (before/after): Event study analysis
  - Years 1,...,T are available and the policy starts in year t^*
    $Y_{it} = \alpha_1 + \sum_{\tau=2}^{T} \beta_\tau I_\tau + u_{it}$
    where I_τ is a dummy variable equal to 1 if $t = \tau$, 0 otherwise (we only include $T - 1$ dummies)
  - $\beta_\tau = E[Y_{it}|t = \tau] - E[Y_{it}|t = 1]$: look for a break in the β's after t^*



Y_{it}
$E[Y_{it}|t=1]$
α_1
β_2
β_{t*-1}
β_{t*}
β_{t*+1}
β_T
time
1
t^*
T

```
. tab year, gen(yr)
```

year	Freq.	Percent	Cum.
1	28	12.50	12.50
2	28	12.50	25.00
3	28	12.50	37.50
4	28	12.50	50.00
5	28	12.50	62.50
6	28	12.50	75.00
7	28	12.50	87.50
8	28	12.50	100.00
Total	224	100.00	

	country	year	after	energy	lnenergy	yr1	yr2	yr3	yr4	yr5	yr6	yr7	yr8
103	1	4	0	8.26153	2.11161	0	0	0	1	0	0	0	0
104	1	4	0	9.3191	2.232066	0	0	0	1	0	0	0	0
105	1	4	0	7.82549	2.057386	0	0	0	1	0	0	0	0
106	1	4	0	9.8664	2.289135	0	0	0	1	0	0	0	0
107	1	4	0	8.02937	2.083107	0	0	0	1	0	0	0	0
108	1	4	0	6.19587	1.823882	0	0	0	1	0	0	0	0
109	1	4	0	6.31098	1.84229	0	0	0	1	0	0	0	0
110	1	4	0	9.16809	2.215729	0	0	0	1	0	0	0	0
111	1	4	0	7.18769	1.97237	0	0	0	1	0	0	0	0
112	1	4	0	6.90775	1.932644	0	0	0	1	0	0	0	0
113	1	5	1	6.29807	1.840243	0	0	0	0	1	0	0	0
114	1	5	1	6.16065	1.818183	0	0	0	0	1	0	0	0
115	1	5	1	6.55919	1.880867	0	0	0	0	1	0	0	0
116	1	5	1	7.34096	1.993469	0	0	0	0	1	0	0	0
117	1	5	1	5.33834	1.674914	0	0	0	0	1	0	0	0
118	1	5	1	7.15222	1.967423	0	0	0	0	1	0	0	0
119	1	5	1	6.34695	1.847974	0	0	0	0	1	0	0	0
120	1	5	1	6.57788	1.883713	0	0	0	0	1	0	0	0
121	1	5	1	7.13153	1.964526	0	0	0	0	1	0	0	0

```
. mean lnenergy, over(year)

Mean estimation                   Number of obs   =        224

          1: year = 1
          2: year = 2
          3: year = 3
          4: year = 4
          5: year = 5
          6: year = 6
          7: year = 7
          8: year = 8
```

Over	Mean	Std. Err.	[95% Conf.	Interval]
lnenergy				
1	$\boxed{1.641535}$	.0372066	1.568214	1.714857
2	1.864199	.023859	1.817181	1.911217
3	1.783069	.031601	1.720795	1.845344
4	1.991856	.0274746	1.937713	2.045999
5	1.911473	.0257771	1.860675	1.962271
6	1.732166	.0423748	1.648659	1.815672
7	1.829234	.0324509	1.765285	1.893184
8	1.965037	.0220784	1.921528	2.008546

$$\ln(energy)_{it} = \alpha_1 + \sum_{\tau=2}^{8} \beta_\tau I_\tau + u_{it}$$

```
. reg lnenergy yr2-yr8, robust

Linear regression                         Number of obs     =        224
                                          F(7, 216)         =      14.02
                                          Prob > F          =     0.0000
                                          R-squared         =     0.3218
                                          Root MSE          =     .16425
```

	lnenergy	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
β_2 →	yr2	.222664	.0441994	5.04	0.000	.1355467	.3097813
	yr3	.1415339	.0488155	2.90	0.004	.0453182	.2377497
	yr4	.3503208	.0462514	7.57	0.000	.259159	.4414826
β_{t*} →	yr5	.2699377	.0452636	5.96	0.000	.1807228	.3591525
	yr6	.0906304	.0563911	1.61	0.109	-.0205169	.2017777
	yr7	.1876989	.04937	3.80	0.000	.0903903	.2850075
β_8 →	yr8	.3235017	.0432642	7.48	0.000	.2382276	.4087757
α_1 →	_cons	$\boxed{1.641535}$	.0372066	44.12	0.000	1.568201	1.71487

Adding a control group

- Even though adding more time periods alleviates some of the concerns, the interpretation of the β's remains problematic
  - Has there been another shock in t^*?
- One way to improve the simple difference method is to compare the before ($T = 0$) and after ($T = 1$) situation for
  - A treatment group ($D = 1$) and a control group ($D = 0$)
- In $T = 0$ no treatment and in period $T = 1$ only one group is treated

  $\rightarrow$ Use the control group to correct for the influence of time (trend)
- Difference in differences (DD) estimator:

$$\beta^{DD} = E[Y_{it}|D_{it} = 1, T_{it} = 1] - E[Y_{it}|D_{it} = 1, T_{it} = 0] - \\ (E[Y_{it}|D_{it} = 0, T_{it} = 1] - E[Y_{it}|D_{it} = 0, T_{it} = 0])$$

Data example

- Let's assume that we also run our survey among 22 firms located in an adjacent country
- Country 2 does not introduce the carbon tax
  - Firms in country 1 are treated (*treated*=1)
  - Firms in country 2 are control (*treated*=0)
- For now we assume that we have just one survey observation before (*after*=0) and one after (*after*=1)
- Intuitively: the DD estimator compares averages across four clusters (2x2 box)
  - This approach can be applied with repeated cross-sections or panel data

	country	year	treated	after	energy	lnenergy
49	1	4	1	0	9.16809	2.215729
50	1	5	1	1	8.44407	2.133464
51	1	4	1	0	8.97073	2.193968
52	1	5	1	1	6.29807	1.840243
53	1	4	1	0	9.8664	2.289135
54	1	5	1	1	8.08022	2.089419
55	1	4	1	0	9.3191	2.232066
56	1	5	1	1	7.75673	2.048561
57	2	4	0	0	10.7603	2.375861
58	2	5	0	1	10.929	2.391421
59	2	4	0	0	10.506	2.351944
60	2	5	0	1	10.8149	2.380924
61	2	4	0	0	9.41996	2.242831
62	2	5	0	1	8.99351	2.196503

```
tab treated after
```

treated	after 0	1	Total
0	22	22	44
1	28	28	56
Total	50	50	100

```
. mean lnenergy, over(treated after)

Mean estimation                   Number of obs   =        100

        Over: treated after
    _subpop_1: 0 0
    _subpop_2: 0 1
    _subpop_3: 1 0
    _subpop_4: 1 1
```

Over	Mean	Std. Err.	[95% Conf.	Interval]
lnenergy				
_subpop_1	2.335661	.0242895	2.287466	2.383857
_subpop_2	2.355709	.0161819	2.323601	2.387818
_subpop_3	1.991856	.0274746	1.93734	2.046372
_subpop_4	1.911473	.0257771	1.860326	1.96262

$$\beta^{DD} = E[\ln(energy)_{it} | treated_{it} = 1, after_{it} = 1] - E\ln(energy)_{it} | treated_{it} = 1, after_{it} = 0] - (E[\ln(energy)_{it} | treated_{it} = 0, after_{it} = 1] - E[\ln(energy)_{it} | treated_{it} = 0, after_{it} = 0])$$

```
. di 1.911473 - 1.991856 - (2.355709 - 2.335661)
-.100431
```

DD regression: notation

- Potential outcomes: $Y_{1,igt}$ and $Y_{0,igt}$

  where i are units of observations $(1, ..., N)$, g are groups (0 if always untreated, and 1 if treated in $t = 1$), t is time

- Assume homogeneous treatment effect: $\beta = Y_{1,igt} - Y_{0,igt}$

- Denote conditional averages as follows:

$$E[Y_{0,igt}|D_{it} = 0, T_{it} = 0] = \alpha$$
$$E[Y_{0,igt}|D_{it} = 0, T_{it} = 1] = \alpha + \lambda$$
$$E[Y_{0,igt}|D_{it} = 1, T_{it} = 0] = \alpha + \gamma$$
$$E[Y_{1,igt}|D_{it} = 1, T_{it} = 1] = \alpha + \gamma + \lambda + \beta^{DD}$$

  - Interpretation: λ is a period effect, γ is a group effect
    You can check that the DD estimator gives β^{DD}

DD: Causal effect identification

- Crucial identifying assumption: parallel trend
  - DD estimation identifies the impact of the intervention if treated and control units *would* follow the same trend *without* treatment
  - $E[Y_{0,igt}|D_{it}=1, T_{it}=1] - E[Y_{0,igt}|D_{it}=1, T_{it}=0] =$
    $E[Y_{0,igt}|D_{it}=0, T_{it}=1] - E[Y_{0,igt}|D_{it}=0, T_{it}=0]$

- With the notation above:

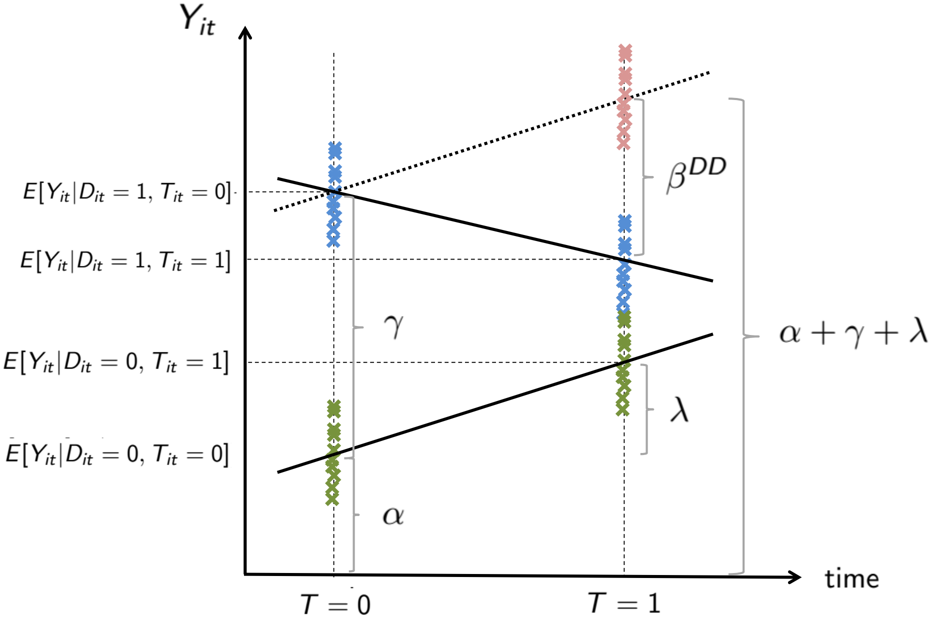
$$
\begin{aligned}
\beta^{DD} &= E[Y_{1,igt}|D_{it}=1, T_{it}=1] - E[Y_{0,igt}|D_{it}=1, T_{it}=0] - \\
&\quad (E[Y_{0,igt}|D_{it}=0, T_{it}=1] - E[Y_{0,igt}|D_{it}=0, T_{it}=0]) = \\
&E[Y_{1,igt}|D_{it}=1, T_{it}=1] - E[Y_{0,igt}|D_{it}=1, T_{it}=0] \\
&\quad -(E[Y_{0,igt}|D_{it}=1, T_{it}=1] - E[Y_{0,igt}|D_{it}=1, T_{it}=0]) = \\
&E[Y_{1,igt}|D_{it}=1, T_{it}=1] - E[Y_{0,igt}|D_{it}=1, T_{it}=1] = \beta
\end{aligned}
$$

DD regression

- We are effectively comparing conditional means: We can just run a regression to estimate β^{DD}
- With repeated cross-sections (i.e. not the same firms are surveyed in $T = 0$ and $T = 1$), we have:

$$Y_{igt} = \alpha + \lambda after_{igt} + \gamma treated_{igt} + \beta^{DD}(after_{igt} \cdot treated_{igt}) + e_{it}$$

- Under the parallel trend assumption, β^{DD} estimates the causal effect of the intervention β
  - We use the evolution observed in the control group to construct a counterfactual outcome for the treated group



Y_{it}
$E[Y_{it}|D_{it}=1, T_{it}=0]$
$E[Y_{it}|D_{it}=1, T_{it}=1]$
$E[Y_{it}|D_{it}=0, T_{it}=1]$
$E[Y_{it}|D_{it}=0, T_{it}=0]$
β^{DD}
γ
$\alpha + \gamma + \lambda$
λ
α
time
$T=0$
$T=1$

```
. reg lnenergy i.treated##i.after, robust

Linear regression                               Number of obs     =        100
                                                F(3, 96)          =     101.52
                                                Prob > F          =     0.0000
                                                R-squared         =     0.7278
                                                Root MSE          =      .1236

------------------------------------------------------------------------------
             |               Robust
    lnenergy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   1.treated |  -.3438051   .0366722    -9.38   0.000    -.4165989   -.2710112
     1.after |   .0200481   .0291031     0.69   0.493    -.0377212    .0778174
             |
treated#after |
        1 1  |  -.1004312   .0476722    -2.11   0.038    -.1950598   -.0058026
             |
       _cons |   2.335661   .0242204    96.43   0.000     2.287584    2.383738
------------------------------------------------------------------------------
```

DD with panel data

- Now consider a case where the same firms participate in the first and second waves of the survey
  - Repeated observations: panel dataset
- Instead of group fixed effect (*treated=1*), we can introduce firm-level fixed effects
  - Recall: FEs control for all characteristics that are time-invariant
  - Same average as group-dummies, but more precise β^{DD}
- We have:

  $$Y_{igt} = \alpha + \sum_{j=1}^{N-1} \alpha_j I_j + \lambda after_{igt} + \beta^{DD}(after_{igt} \cdot treated_{igt}) + e_{it}$$

  or simplifying notation:

  $$Y_{igt} = \alpha_i + \lambda after_{igt} + \beta^{DD}(after_{igt} \cdot treated_{igt}) + e_{it}$$

- In Stata, *xtset* your data and use *xtreg* with the option *fe*
  - Adjust standard errors for clusters (*cluster(firms)*)

	country	firm	year	treated	after	energy	lnenergy	
49	1	25	4	1	0	9.16809	2.215729	
50	1	25	5	1	1	8.44407	2.133464	
51	1	26	4	1	0	8.97073	2.193968	
52	1	26	5	1	1	6.29807	1.840243	
53	1	27	4	1	0	9.8664	2.289135	
54	1	27	5	1	1	8.08022	2.089419	
55	1	28	4	1	0	9.3191	2.232066	
56	1	28	5	1	1	7.75673	2.048561	
57	2	29	4	0	0	10.7603	2.375861	
58	2	29	5	0	1	10.929	2.391421	
59	2	30	4	0	0	10.506	2.351944	
60	2	30	5	0	1	10.8149	2.380924	
61	2	31	4	0	0	9.41996	2.242831	
62	2	31	5	0	1	8.99351	2.196503	
63	2	32	4	0	0	10.1891	2.321315	
64	2	32	5	0	1	9.75351	2.277627	

```
xtset firm after
     panel variable:  firm (strongly balanced)
      time variable:  after, 0 to 1
              delta:  1 unit
```

```
. xtreg lnenergy i.treated##i.after, fe cluster(firm)
note: 1.treated omitted because of collinearity

Fixed-effects (within) regression               Number of obs     =        100
Group variable: firm                            Number of groups  =         50

R-sq:                                           Obs per group:
     within  = 0.2124                                         min =          2
     between = 0.7744                                         avg =        2.0
     overall = 0.4010                                         max =          2

                                                F(2,49)           =       5.62
corr(u_i, Xb)  = 0.5350                         Prob > F          =     0.0064

                                 (Std. Err. adjusted for 50 clusters in firm)
------------------------------------------------------------------------------
             |               Robust
    lnenergy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   1.treated |          0  (omitted)
     1.after |   .0200481   .0226421     0.89   0.380    -.0254528     .065549
             |
treated#after|
         1 1 |  -.1004312   .0336318    -2.99   0.004    -.1680168   -.0328456
             |
       _cons |    2.14313   .0085614   250.32   0.000     2.125925    2.160335
------------------------------------------------------------------------------

. mean lnenergy if after==0

Mean estimation                   Number of obs    =         50

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
    lnenergy |    2.14313    .030636      2.081565    2.204696
--------------------------------------------------------------
```

DD with more than two periods

- Consider again a case where we observe firms four years before the interventions and four years after
- Use period fixed effects to flexibly measure how the outcome evolves among control units
  - Recall: time FEs capture all phenomena that affect all units at a given point in time (macro shocks)
- Regression equation:

  $Y_{igt} = \alpha_i + \sum_{\tau=2}^{T} \lambda_\tau I_\tau + \beta^{DD}(after_{igt} \cdot treated_{igt}) + e_{it}$
  or simplifying notation:
  $Y_{igt} = \alpha_i + \lambda_t + \beta^{DD}(after_{igt} \cdot treated_{igt}) + e_{it}$

- Identifying dynamic effects: instead of an average impact we can estimate per-period impact β_τ^{DD}

  $Y_{igt} = \alpha_i + \lambda_t + \sum_{\tau=t^*}^{T} \beta_\tau^{DD}(I_\tau \cdot treated_{igt}) + e_{it}$

```
. tab year, gen(yr)
```

year	Freq.	Percent	Cum.
1	50	12.50	12.50
2	50	12.50	25.00
3	50	12.50	37.50
4	50	12.50	50.00
5	50	12.50	62.50
6	50	12.50	75.00
7	50	12.50	87.50
8	50	12.50	100.00
Total	400	100.00	

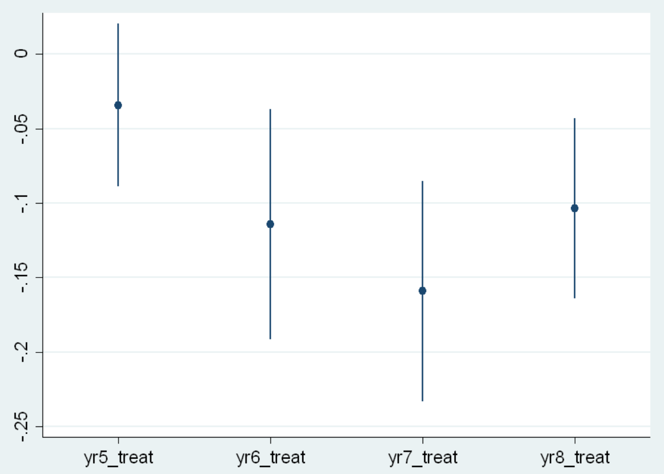
	country	firm	year	treated	after	energy	lnenergy	yr1	yr2	yr3	yr4	yr5	yr6	yr
216	1	27	8	1	1	7.85563	2.06123	0	0	0	0	0	0	
217	1	28	1	1	0	6.2386	1.830756	1	0	0	0	0	0	
218	1	28	2	1	0	5.71418	1.742951	0	1	0	0	0	0	
219	1	28	3	1	0	6.61309	1.889051	0	0	1	0	0	0	
220	1	28	4	1	0	9.3191	2.232066	0	0	0	1	0	0	
221	1	28	5	1	1	7.75673	2.048561	0	0	0	0	1	0	
222	1	28	6	1	1	5.39693	1.685831	0	0	0	0	0	1	
223	1	28	7	1	1	7.59459	2.027437	0	0	0	0	0	0	
224	1	28	8	1	1	9.04881	2.202633	0	0	0	0	0	0	
225	2	29	1	0	0	8.47281	2.136862	1	0	0	0	0	0	
226	2	29	2	0	0	8.53564	2.144251	0	1	0	0	0	0	
227	2	29	3	0	0	9.73931	2.27617	0	0	1	0	0	0	
228	2	29	4	0	0	10.7603	2.375861	0	0	0	1	0	0	
229	2	29	5	0	1	10.929	2.391421	0	0	0	0	1	0	

```
. xtset firm year
       panel variable:  firm (strongly balanced)
        time variable:  year, 1 to 8
                delta:  1 unit
```

```
. xtreg lnenergy after_treat yr2-yr8, fe cluster(firm)

Fixed-effects (within) regression               Number of obs      =       400
Group variable: firm                            Number of groups   =        50

R-sq:                                           Obs per group:
     within  = 0.5042                                          min =         8
     between = 0.8677                                          avg =       8.0
     overall = 0.2960                                          max =         8

                                                F(8,49)            =     56.24
corr(u_i, Xb)  = 0.2157                         Prob > F           =    0.0000

                                    (Std. Err. adjusted for 50 clusters in firm)
------------------------------------------------------------------------------
             |               Robust
    lnenergy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 after_treat |  -.1027022   .0211873    -4.85   0.000    -.1452797   -.0601248
         yr2 |   .1723477   .0246111     7.00   0.000     .1228899    .2218055
         yr3 |   .1106416   .0227017     4.87   0.000     .0650207    .1562624
         yr4 |   .2843666     .02411    11.79   0.000     .2359158    .3328175
         yr5 |   .3056865   .0242081    12.63   0.000     .2570385    .3543346
         yr6 |   .1615744   .0242957     6.65   0.000     .1127503    .2103985
         yr7 |   .2784237   .0193299    14.40   0.000     .2395788    .3172686
         yr8 |   .3897303   .0244912    15.91   0.000     .3405135    .4389471
       _cons |   1.858764   .0166451   111.67   0.000     1.825314    1.892213
-------------+----------------------------------------------------------------
     sigma_u |  .22459994
     sigma_e |  .11279323
         rho |  .79859384   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
. gen yr5_treat = yr5*treated

. gen yr6_treat = yr6*treated

. gen yr7_treat = yr7*treated

. gen yr8_treat = yr8*treated

. xtreg lnenergy yr5_treat-yr8_treat yr2-yr8, fe cluster(firm)

Fixed-effects (within) regression               Number of obs      =       400
Group variable: firm                            Number of groups   =        50

R-sq:                                           Obs per group:
     within  = 0.5154                                         min =          8
     between = 0.8677                                         avg =        8.0
     overall = 0.2984                                         max =          8

                                                F(11,49)           =     43.59
corr(u_i, Xb)  = 0.2134                         Prob > F           =    0.0000

                                    (Std. Err. adjusted for 50 clusters in firm)
------------------------------------------------------------------------------
             |               Robust
    lnenergy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   yr5_treat |  -.0341507   .0272059    -1.26   0.215     -.088823    .0205217
   yr6_treat |  -.1141396   .0383434    -2.98   0.005    -.1911935   -.0370857
   yr7_treat |  -.1590959   .0367717    -4.33   0.000    -.2329914   -.0852004
   yr8_treat |  -.1034228   .0300996    -3.44   0.001    -.1639103   -.0429354
         yr2 |   .1723477    .024706     6.98   0.000     .1226991    .2219964
         yr3 |   .1106416   .0227893     4.85   0.000     .0648447    .1564384
         yr4 |   .2843666    .024203    11.75   0.000     .2357288    .3330045
         yr5 |   .2672976   .0211897    12.61   0.000     .2247154    .3098799
         yr6 |   .1679793   .0220262     7.63   0.000     .1237161    .2122425
         yr7 |   .3100041   .0222452    13.94   0.000     .2653008    .3547074
         yr8 |   .3901338   .0229851    16.97   0.000     .3439435    .4363241
       _cons |   1.858764   .0167093   111.24   0.000     1.825185    1.892342
```

```
coefplot, keep( yr5_treat yr6_treat yr7_treat yr8_treat) vertical
```



Parallel trends?

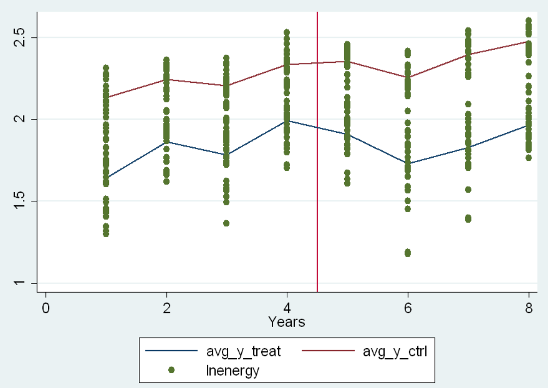
- The assumption of parallel trends cannot be directly tested
  - We would need to observe both potential outcomes

- However, we should document whether the parallel trend assumption is plausible before treatment

- Two suggestions to provide evidence about *pre-treatment* trends
  1. Plot average outcomes for treated and control units over time
  2. Placebo test: introduce DD estimates in all years before $t^* - 2$ and show that δ_τ^{DD} estimates are not statistically significant
     $Y_{igt} = \alpha_i + \lambda_t + \sum_{\tau=1}^{t^*-2} \delta_\tau^{DD} (I_\tau \cdot treated_{igt}) + \beta^{DD} (after_{igt} \cdot treated_{igt}) + e_{it}$

- Note functional form dependence: the parallel trend may hold for logs but not for levels (and conversely)
  - Can select preferred functional form in light of pre-treatment trends

```
. bysort year: egen avg_y_treat = mean(lnenergy) if country==1
(176 missing values generated)

. bysort year: egen avg_y_ctrl = mean(lnenergy) if country==2
(224 missing values generated)

. twoway (line avg_y_treat year, sort) (line avg_y_ctrl year, sort) (scatter lnenergy year, sort), xtitl
> e(Years) xline(4.5)
```

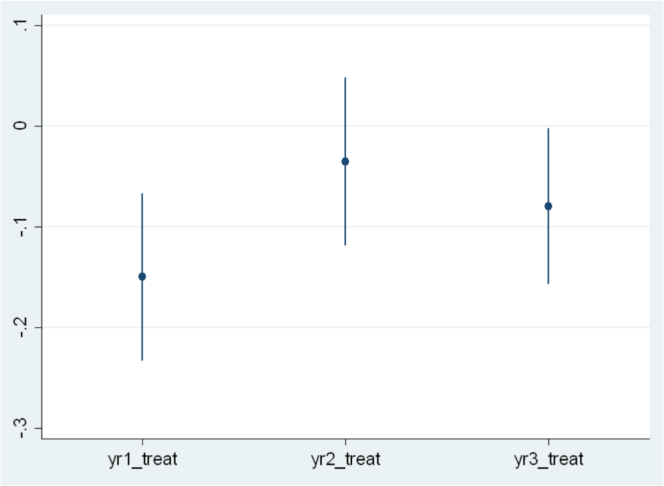


```
. gen yr1_treat = yr1*treated

. gen yr2_treat = yr2*treated

. gen yr3_treat = yr3*treated
```

```
. xtreg lnenergy after_treat yr1_treat-yr3_treat yr2-yr8, fe cluster(firm)

Fixed-effects (within) regression               Number of obs      =       400
Group variable: firm                            Number of groups   =        50

R-sq:                                           Obs per group:
     within  = 0.5217                                         min =         8
     between = 0.8677                                         avg =       8.0
     overall = 0.5033                                         max =         8

                                                F(11,49)           =     53.06
corr(u_i, Xb)  = 0.4243                         Prob > F           =    0.0000

                                 (Std. Err. adjusted for 50 clusters in firm)
```

lnenergy	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
after_treat	-.1689828	.032901	-5.14	0.000	-.2350998	-.1028657
yr1_treat	-.1498957	.0413091	-3.63	0.001	-.2329094	-.0668821
yr2_treat	-.0355406	.0414754	-0.86	0.396	-.1188886	.0478074
yr3_treat	-.0796858	.0384724	-2.07	0.044	-.156999	-.0023725
yr2	.1083089	.0181604	5.96	0.000	.0718141	.1448036
yr3	.071324	.0183359	3.89	0.000	.0344767	.1081712
yr4	.200425	.0233704	8.58	0.000	.1534605	.2473895
yr5	.258862	.0194715	13.29	0.000	.2197325	.2979915
yr6	.1147499	.0218893	5.24	0.000	.0707617	.1587381
yr7	.2315992	.020547	11.27	0.000	.1903084	.27289
yr8	.3429058	.021195	16.18	0.000	.3003128	.3854988
_cons	1.942705	.0203383	95.52	0.000	1.901834	1.983577

```
. coefplot, keep(yr1_treat yr2_treat yr3_treat) vertical
```



DD: Remarks and pitfalls

- In this case the results suggests that pre-treatment trends are not parallel
  - Try linear specification instead
- Alternatively, control for firm-level time-varying characteristics
  - Avoid bad controls: only variable that are not affected by the treatment can be included
- Heterogeneous treatment effect: Exploit within-treated variation to quantify differences in β^{DD} across clusters
  - Interact $after_{igt} \cdot treated_{igt}$ with pre-treatment variables for treated units
- Other pitfall: there should be no spillover across treatment and control groups (e.g. trade among countries?)

Comparative case study: Synthetic control

- In some settings only one or a very small number of units are treated
  - Finding a control group with parallel trend is likely to fail
- Synthetic control (Abadie and Gardeazabal, 2003): DD-style estimator for "case studies"
  - Growing in popularity: 2000+ citations on Google scholar
- Intuition: construct a weighted average of control units to match pre-treatment trajectory for the treated unit
  - Then compare the post-treatment trajectory for the treated unit and the synthetic counterpart

Synthetic control: Notation

- Outcome variable for treated unit: Y_t
  - Outcome for the synthetic unit:
    $Y_t^{SCM} = \sum_i \omega_i Y_{it}$
    where i is set of untreated units (the "donor pool") and ω_i is the weight attributed to each unit i
- Weights ω_i minimize the distance between Y_t and Y_t^{SCM} before t^*:

  $\min_{\omega_i} \sum_{t=0}^{t^*-1} (Y_t - \sum_i \omega_i Y_{it})^2$
  s.t. $\sum_i \omega_i = 1$, $\omega_i \geq 0$
- Treatment effect: $\beta_t = Y_t - Y_t^{SCM}$

Synthetic control: Example

- Sweden implemented a carbon tax in 1991
  - Starting at US$30/t$CO_2$ up to US$132 in 2019
  - Mainly affect transportation sector (exceptions for industries)
  - What is the control group?
- Andersson (AEJ: EconPolicy, 2019) constructs a “synthetic” Sweden
  - Donor pool: all OECD countries
  - Compute weights given to each OECD country
  - Matching period: 1960 to 1990
- The impact of the 1990 carbon tax: compare CO_2 emissions in Sweden against its synthetic counterfactual

TABLE 2—COUNTRY WEIGHTS IN SYNTHETIC SWEDEN

Country	Weight	Country	Weight
Australia	0.001	Japan	0
Belgium	0.195	New Zealand	0.177
Canada	0	Poland	0.001
Denmark	0.384	Portugal	0
France	0	Spain	0
Greece	0.090	Switzerland	0.061
Iceland	0.001	United States	0.088

Note: With the synthetic control method, extrapolation is not allowed so all weights are between $0 \leq w_j \leq 1$ and $\sum w_j = 1$.

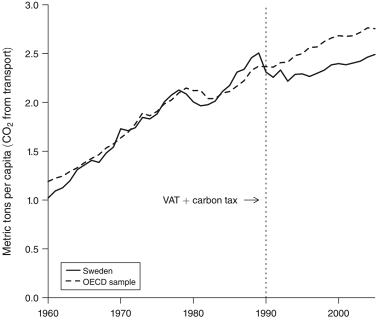


FIGURE 3. PATH PLOT OF PER CAPITA CO_2 EMISSIONS FROM TRANSPORT DURING 1960–2005: SWEDEN VERSUS THE OECD AVERAGE OF MY 14 DONOR COUNTRIES

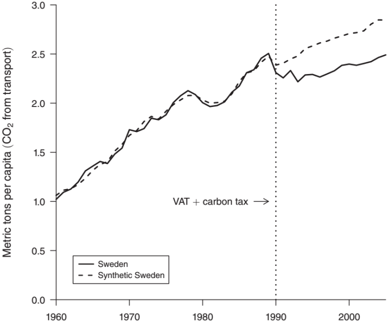


FIGURE 4. PATH PLOT OF PER CAPITA CO_2 EMISSIONS FROM TRANSPORT DURING 1960–2005: SWEDEN VERSUS SYNTHETIC SWEDEN

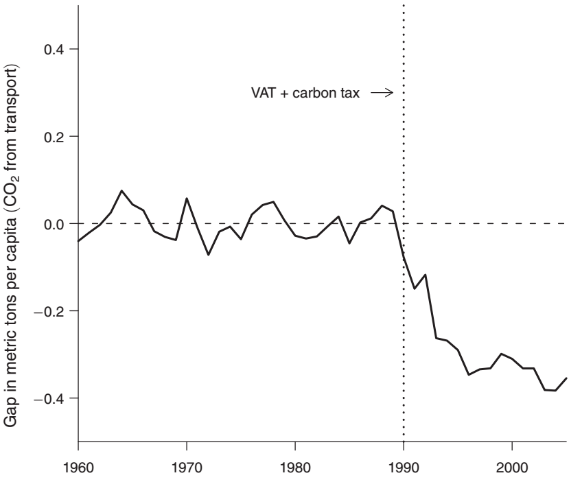


Figure 5. Gap in Per Capita CO_2 Emissions from Transport between Sweden and Synthetic Sweden

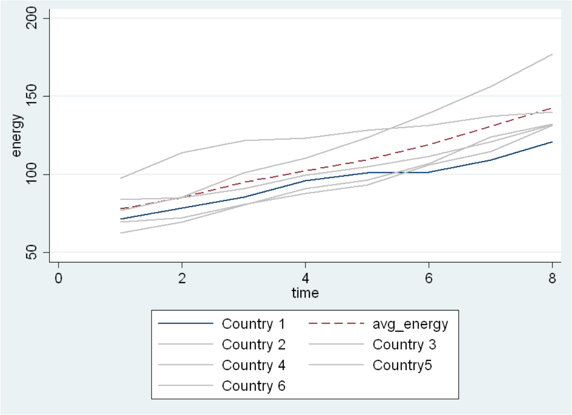
Implementation in Stata

- Synthetic control in Stata: user written command "synth"
  - Written for the paper Abadie et al. (JASA, 2010), also available for R (and Matlab)
- Implementation: Consider simulated data on energy use observed over 8 years:
  - 1 country imposes a carbon tax in period 5 and later
  - 5 countries have no tax and are included in the donor pool
- Our objective is to combine the 5 countries into a synthetic control
  - First we need to install the package and declare group/time dimensions

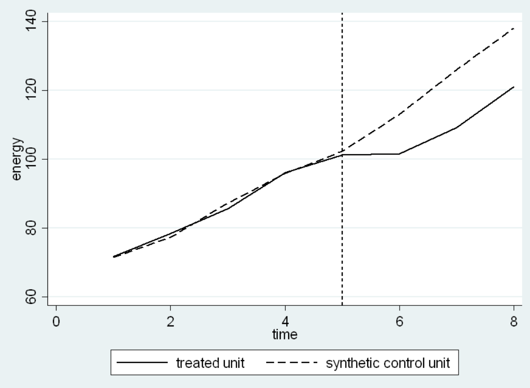
```
. net from "https://web.stanford.edu/~jhain/Synth"
--------------------------------------------------------------------------------
https://web.stanford.edu/~jhain/Synth/
Synthetic Control Methods for Comparative Case Studies
--------------------------------------------------------------------------------

Alberto Abadie, Kennedy School of Government, Harvard University and NBER
Jens Hainmueller, Department of Political Science, MIT
Alexis Diamond, IFC

Also see the homepage for the paper that describes the method.

PACKAGES you could -net describe-:
    synth            Synthetic Control Methods
--------------------------------------------------------------------------------

.
. net install synth, all replace force
checking synth consistency and verifying not already installed...
installing into c:\ado\plus\...
installation complete.


. tsset country time
       panel variable:  country (strongly balanced)
        time variable:  time, 1 to 8
                delta:  1 unit
```

	country	time	energy
1	1	1	71.6902
2	1	2	78.5058
3	1	3	85.7271
4	1	4	96.1347
5	1	5	101.161
6	1	6	101.536
7	1	7	109.227
8	1	8	121.079
9	2	1	97.479
10	2	2	114.125
11	2	3	121.943
12	2	4	123.15
13	2	5	128.524
14	2	6	131.489
15	2	7	137.319
16	2	8	140.237
17	3	1	62.4368
18	3	2	69.548
19	3	3	80.4012
20	3	4	91.067
21	3	5	96.4389
22	3	6	107.073

```
. bysort time: egen avg_energy = mean(energy) if country!=1
(8 missing values generated)

. twoway (line energy time if country==1, sort) (line avg_energy time, sort lpattern(dash)) (line energy
>  time if country==2, lcolor(gs12) sort) (line energy time if country==3, lcolor(gs12) sort) (line ener
> gy time if country==4, lcolor(gs12) sort) (line energy time if country==5, lcolor(gs12) sort) (line en
> ergy time if country==6, lcolor(gs12) sort)
```



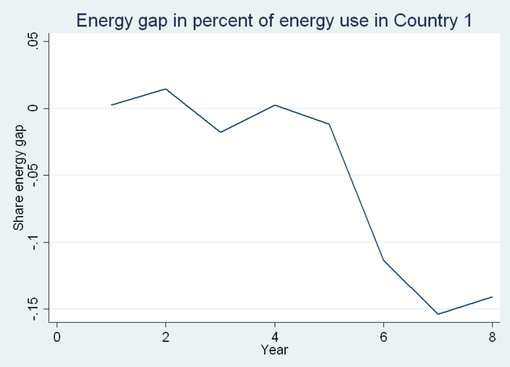
```
synth energy energy, trunit(1) trperiod(5) keep(synth_out, replace) fig
```



Unit Weights:

Co_No	Unit_Weight
2	.063
3	.393
4	.126
5	.274
6	.144

```
. use "synth_out.dta", clear

. gen tr_effect_1=_Y_treated-_Y_synthetic

. gen share_energy_gap=tr_effect_1/_Y_treated

. twoway (line share_energy_gap _time, sort), ytitle(Share energy gap) xtitle(Year) title(Energy gap in
> percent of energy use in Country 1)
```



Synthetic control: Placebo experiments

Make a mistake to check reliability

- One may wonder if these results are just derived "by chance"
  - Placebo experiments: try to convince the reader that the treatment effect is not spurious
  - Abadie: Akin to making a mistake in the estimation code!

1. In-time placebo: Shift the treatment to a prior time period

   We would want to see the treatment effect only start with a delay

   Check and look for a good fit
2. In-space placebo: Re-assign the treatment to each unit in the donor pool, and re-run the algorithm
   - Use only non-treated units
   - Allows a form of inference: p-values as the fraction of donor pool units with a treatment effect that is larger than that of the treated unit

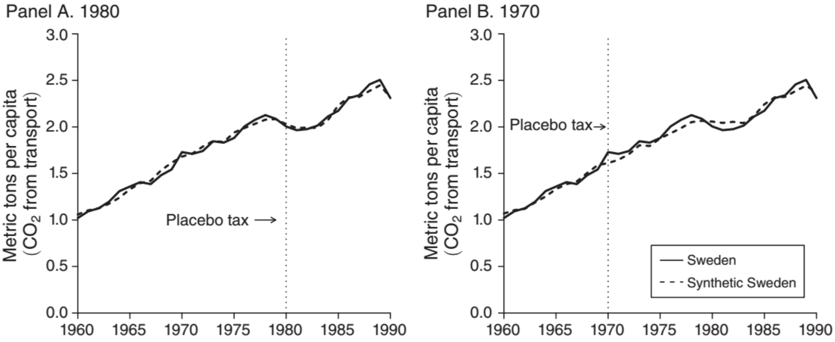


FIGURE 6. PLACEBO IN-TIME TESTS

Notes: In panel A, the placebo tax is introduced in 1980, ten years prior to the actual policy changes. In panel B, the placebo tax is introduced in 1970.

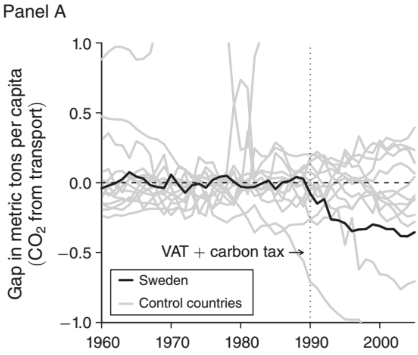
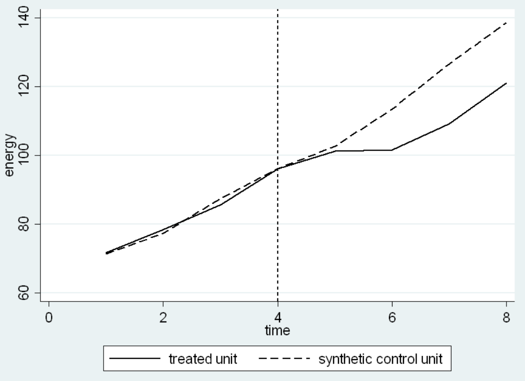


Figure 7. Permutation Test: Per Capita CO_2 Emissions Gap in Sweden and Placebo Gaps for the Control Countries

Notes: Panel A shows per capita CO_2 emissions gap in Sweden and placebo gaps in all 14 OECD control countries. Panel B shows per capita gap in Sweden and placebo gaps in nine OECD control countries (countries with a pretreatment MSPE 20 times higher than Sweden's are excluded).

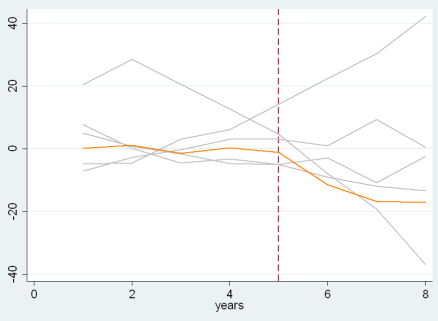
```
synth energy energy, trunit(1) trperiod(4) fig
```



Unit Weights:

Co_No	Unit_Weight
2	.062
3	.416
4	.134
5	.249
6	.14

```
drop if country==1
forval i=2/6{
qui synth energy energy, trunit(`i') trperiod(5) keep(synth_`i', replace)
}
forval i=2/6{
use synth_`i', clear
rename _time years
gen tr_effect_`i' = _Y_treated - _Y_synthetic
keep years tr_effect_`i'
save synth_`i', replace
}
use synth_1, clear

forval i=1/6{
qui merge 1:1 years using synth_`i', nogenerate
}
local lp

forval i=2/6 {
   local lp `lp' line tr_effect_`i' years, lcolor(gs12) ||
}

twoway `lp' || line tr_effect_1 years, lcolor(orange) legend(off) xline(5, lpattern(dash))
```



See also the user written command “synth_runner”

Wrapping up

- DD estimation: Illustrates how variability in treatment assignment identifies the treatment effect in a panel FE model
  - Use units in the control group to estimate a counterfactual trajectory
- Need to document whether the parallel trend assumption is plausible *before* the treatment
- With synthetic control, we impose both level and trend equality prior to treatment
  - Downside: this is only a case study: can we generalize?