

## Statistics / Data analysis (R)

```

name: <unnamed>
log: d:\Temp\MCP\Did\Did-toolbox\Examples\medicaid.smcl
log type: smcl
opened on: 18 Jan 2026, 14:15:49

1 . * just 2013 vs 2014
2 . import excel "D:\Projekte\basic_econometrics\Diff_in_Diff\DiDToolbox\Medicaid_own.xlsx", sheet("Sheet1") firstrow
(32 vars, 28,644 obs)

3 .
4 . tsset county_code year_code

Panel variable: county_code (strongly balanced)
Time variable: year_code, 2009 to 2019
    Delta: 1 unit

5 . keep if yaca==.|(yaca==2014|yaca>2019)
(4,444 observations deleted)

6 .
7 . xtreg crude_rate_20_64 treatPost i.year if year_code>=2013 & year_code<=2014 , fe vce(cluster county_code)

Fixed-effects (within) regression                               Number of obs     =      4,400
Group variable: county_code                                Number of groups  =      2,200

R-squared:                                                 Obs per group:
    Within = 0.0108                                         min =           2
    Between = 0.0350                                         avg =          2.0
    Overall = 0.0008                                         max =           2

corr(u_i, Xb) = -0.0012                                     F(2, 2199)        =     12.09
                                                               Prob > F        =  0.0000

                                                (Std. err. adjusted for 2,200 clusters in county_code)



|           | Robust      |           |        |       |                                   |
|-----------|-------------|-----------|--------|-------|-----------------------------------|
|           | Coefficient | std. err. | t      | P> t  | [95% conf. interval]              |
| treatPost | .1216303    | 3.748009  | 0.03   | 0.974 | -7.228378 7.471639                |
| year 2014 | 9.148031    | 2.598863  | 3.52   | 0.000 | 4.051548 14.24452                 |
| _cons     | 449.6517    | .9387732  | 478.98 | 0.000 | 447.8108 451.4927                 |
| sigma_u   | 145.45523   |           |        |       |                                   |
| sigma_e   | 62.271173   |           |        |       |                                   |
| rho       | .84510841   |           |        |       | (fraction of variance due to u_i) |



8 .
9 . * all years
10 . import excel "D:\Projekte\basic_econometrics\Diff_in_Diff\DiDToolbox\Medicaid_own.xlsx", sheet("Sheet1") firstrow
(32 vars, 28,644 obs)

11 . tsset county_code year_code

Panel variable: county_code (strongly balanced)
Time variable: year_code, 2009 to 2019
    Delta: 1 unit

12 .
13 .
14 . xtdidregress (crude_rate_20_64) (treatPost), group(county_code) time(year_code)

Treatment and time information

Time variable: year_code
Control: treatPost = 0
Treatment: treatPost = 1



|                   | Control | Treatment |
|-------------------|---------|-----------|
| Group county_code | 1222    | 1382      |
| Time              |         |           |
| Minimum           | 2009    | 2014      |
| Maximum           | 2009    | 2019      |



Difference-in-differences regression                               Number of obs = 28,644
Data type: Longitudinal

                                                (Std. err. adjusted for 2,604 clusters in county_code)



|                    | Robust      |           |       |       |                      |
|--------------------|-------------|-----------|-------|-------|----------------------|
|                    | Coefficient | std. err. | t     | P> t  | [95% conf. interval] |
| crude_rat~64       |             |           |       |       |                      |
| ATET               |             |           |       |       |                      |
| treatPost (1 vs 0) | -1.011909   | 1.723404  | -0.59 | 0.557 | -4.39129 2.367473    |



Note: ATET estimate adjusted for panel effects and time effects.
Note: Treatment occurs at different times.

```

15 . estat bdecomp

DID treatment-effect decomposition

ATET = <b>-1.011909</b>	Number of obs = 28,644
	Number of groups = 2,604
	Number of cohorts = 5

ATET decomposition summary	ATET component	Weight
Treated vs never treated	.6843215	<b>0.845990</b>
Treated earlier vs later	-12.038776	<b>0.109981</b>
Treated later vs earlier	-6.0596275	<b>0.044629</b>

Full ATET decomposition	2x2 coefficient	Weight
Treated vs never treated		
2014 vs never treated	.9977025	<b>0.645146</b>
2015 vs never treated	1.3537912	<b>0.112802</b>
2016 vs never treated	-10.960388	<b>0.057258</b>
2019 vs never treated	13.322783	<b>0.030784</b>
Treated earlier vs later		
2014 vs 2015	-1.239248	<b>0.015046</b>
2014 vs 2016	-5.2519698	<b>0.016366</b>
2014 vs 2019	-14.656263	<b>0.061593</b>
2015 vs 2016	-6.3446361	<b>0.001717</b>
2015 vs 2019	-15.556798	<b>0.010339</b>
2016 vs 2019	-29.468659	<b>0.004920</b>
Treated later vs earlier		
2015 vs 2014	-2.6411471	<b>0.015046</b>
2016 vs 2014	-19.112871	<b>0.013093</b>
2019 vs 2014	2.9920635	<b>0.012319</b>
2016 vs 2015	-19.578119	<b>0.001145</b>
2019 vs 2015	5.515887	<b>0.001723</b>
2019 vs 2016	-1.0893875	<b>0.000703</b>

Note: Number of cohorts includes never treated.  
Note: The ATET reported by `xtdidregress` is a weighted average of the ATET components. If any component is substantially different from the ATET reported by `xtdidregress` and the weight is large, consider accounting for treatment-effect heterogeneity by using `xtdidregress`.

16 .  
17 . \*\* Wooldridge  
18 . xtdidregress twfe (crude\_rate\_20\_64) (treatPost), group(county\_code)  
note: variable `_did_cohort`, containing cohort indicators formed by treatment variable `treatPost` and group variable `county_code`, was added to the dataset.

Treatment and time information

Time variable: <code>year_code</code>	
Time interval: 2009 to 2019	
Control: <code>_did_cohort = 0</code>	
Treatment: <code>_did_cohort &gt; 0</code>	
	<code>_did_cohort</code>
Number of cohorts	5
Number of obs	
Never treated	13442
2014	10758
2015	1881
2016	1023
2019	1540

Heterogeneous-treatment-effects regression                          Number of obs = 28,644  
    Number of panels = 2,604

Estimator: Two-way fixed effects  
Panel variable: `county_code`  
Treatment level: `county_code`  
Control group: Never treated  
Heterogeneity: Cohort and time

(Std. err. adjusted for 2,604 clusters in `county_code`)

Cohort	Robust					
	ATET	std. err.	t	P> t	[95% conf. interval]	
<b>2014</b>						
<code>year_code</code>						
2014	-4.786314	2.733836	-1.75	<b>0.080</b>	<b>-10.14703</b>	<b>.5743998</b>
2015	-3.787322	3.04012	-1.25	<b>0.213</b>	<b>-9.74862</b>	<b>2.173976</b>
2016	4.659988	3.142988	1.48	<b>0.138</b>	<b>-1.503021</b>	<b>10.823</b>
2017	.3071711	3.38638	0.09	<b>0.928</b>	<b>-6.3331</b>	<b>6.947442</b>
2018	-.9385941	3.290743	-0.29	<b>0.775</b>	<b>-7.391332</b>	<b>5.514144</b>
2019	2.548081	3.389093	0.75	<b>0.452</b>	<b>-4.097508</b>	<b>9.193671</b>
<b>2015</b>						
<code>year_code</code>						
2015	-4.99783	4.951455	-1.01	<b>0.313</b>	<b>-14.70702</b>	<b>4.711357</b>
2016	6.868282	4.718382	1.46	<b>0.146</b>	<b>-2.383878</b>	<b>16.12044</b>
2017	3.839901	5.340868	0.72	<b>0.472</b>	<b>-6.632878</b>	<b>14.31268</b>
2018	-5.471354	5.520333	-0.99	<b>0.322</b>	<b>-16.29604</b>	<b>5.353333</b>
2019	-1.503282	5.454547	-0.28	<b>0.783</b>	<b>-12.19897</b>	<b>9.192407</b>
<b>2016</b>						
<code>year_code</code>						
2016	-.1452888	8.213822	-0.02	<b>0.986</b>	<b>-16.25157</b>	<b>15.961</b>
2017	-12.30152	8.515601	-1.44	<b>0.149</b>	<b>-28.99955</b>	<b>4.396519</b>
2018	-27.32965	8.230132	-3.32	<b>0.001</b>	<b>-43.46791</b>	<b>-11.19138</b>
2019	-10.19354	10.44771	-0.98	<b>0.329</b>	<b>-30.68021</b>	<b>10.29313</b>
<b>2019</b>						
<code>year_code</code>						

2019	13.32278	7.265116	1.83	0.067	-.923206	27.56877
------	----------	----------	------	-------	----------	----------

19 . estat aggregation, cohort

ATET over cohort Number of obs = 28,644

(Std. err. adjusted for 2,604 clusters in county\_code)

Cohort	Robust				
	ATET	std. err.	t	P> t	[95% conf. interval]
2014	-.3328317	1.931198	-0.17	0.863	-4.119672 3.454008
2015	-.2528568	3.241521	-0.08	0.938	-6.609077 6.103363
2016	-12.4925	5.83558	-2.14	0.032	-23.93534 -1.049649
2019	13.32278	7.265116	1.83	0.067	-.923206 27.56877

20 . estat aggregation, time

ATET over time Number of obs = 28,644

(Std. err. adjusted for 2,604 clusters in county\_code)

Time	Robust				
	ATET	std. err.	t	P> t	[95% conf. interval]
2014	-4.786314	2.733836	-1.75	0.080	-10.14703 .5743998
2015	-3.967476	2.877608	-1.38	0.168	-9.610109 1.675156
2016	4.604213	2.905074	1.58	0.113	-1.092276 10.3007
2017	-1.1505674	3.1131485	-0.05	0.962	-6.291021 5.989887
2018	-3.538811	3.087589	-1.15	0.252	-9.59319 2.515567
2019	2.280863	3.113742	0.73	0.464	-3.824798 8.386525

21 .

22 .

23 . \*\* Callaway/Sant'Anna DR

24 . xthdidregress twfe (crude\_rate) (treatPost), group(county\_code)  
note: variable \_did\_cohort, containing cohort indicators formed by treatment variable treatPost and group variable county\_code, was added to the dataset.

Treatment and time information

Time variable: year\_code  
 Time interval: 2009 to 2019  
 Control: \_did\_cohort = 0  
 Treatment: \_did\_cohort > 0

	_did_cohort
Number of cohorts	5
Number of obs	
Never treated	13442
2014	10758
2015	1881
2016	1023
2019	1540

Heterogeneous-treatment-effects regression Number of obs = 28,644  
 Number of panels = 2,604

Estimator: Two-way fixed effects

Panel variable: county\_code

Treatment level: county\_code

Control group: Never treated

Heterogeneity: Cohort and time

(Std. err. adjusted for 2,604 clusters in county\_code)

Cohort	Robust				
	ATET	std. err.	t	P> t	[95% conf. interval]
2014					
year_code					
2014	-4.786314	2.733836	-1.75	0.080	-10.14703 .5743998
2015	-3.787322	3.04012	-1.25	0.213	-9.74862 2.173976
2016	4.659988	3.142988	1.48	0.138	-1.503021 10.823
2017	-3.071711	3.38638	0.09	0.928	-6.3331 6.947442
2018	-9.9385941	3.290743	-0.29	0.775	-7.391332 5.514144
2019	2.548081	3.389093	0.75	0.452	-4.097508 9.193671
2015					
year_code					
2015	-4.99783	4.951455	-1.01	0.313	-14.70702 4.711357
2016	6.868282	4.718382	1.46	0.146	-2.383878 16.12044
2017	3.839901	5.340868	0.72	0.472	-6.632878 14.31268
2018	-5.471354	5.520333	-0.99	0.322	-16.29604 5.353333
2019	-1.503282	5.454547	-0.28	0.783	-12.19897 9.192407
2016					
year_code					
2016	-1452888	8.213822	-0.02	0.986	-16.25157 15.961
2017	-12.30152	8.515601	-1.44	0.149	-28.99955 4.396519
2018	-27.32965	8.230132	-3.32	0.001	-43.46791 -11.19138
2019	-10.19354	10.44771	-0.98	0.329	-30.68021 10.29313
2019					
year_code					
2019	13.32278	7.265116	1.83	0.067	-.923206 27.56877

25 . estat aggregation, cohort

ATET over cohort						Number of obs = 28,644
(Std. err. adjusted for 2,604 clusters in county_code)						
Cohort	ATET	Robust std. err.	t	P> t	[95% conf. interval]	
2014	-.3328317	1.931198	-0.17	0.863	-4.119672	3.454008
2015	-.2528568	3.241521	-0.08	0.938	-6.609077	6.103363
2016	-12.4925	5.83558	-2.14	0.032	-23.93534	-1.049649
2019	13.32278	7.265116	1.83	0.067	-.923206	27.56877

26 . estat aggregation, time

ATET over time						Number of obs = 28,644
(Std. err. adjusted for 2,604 clusters in county_code)						
Time	ATET	Robust std. err.	t	P> t	[95% conf. interval]	
2014	-4.786314	2.733836	-1.75	0.080	-10.14703	.5743998
2015	-3.967476	2.877608	-1.38	0.168	-9.610109	1.675156
2016	4.684213	2.905074	1.58	0.113	-1.092276	10.3007
2017	-.1505674	3.131485	-0.05	0.962	-6.291021	5.989887
2018	-3.538811	3.087589	-1.15	0.252	-9.59319	2.515567
2019	2.280863	3.113742	0.73	0.464	-3.824798	8.386525

27 .

```
28 . xthdidregress aipw (crude_rate) (treatPost perc_female perc_white perc_hispanic unemp_rate_pc poverty_rate median_income_k), group(county_code)
> )
note: variable _did_cohort, containing cohort indicators formed by treatment variable treatPost and group variable county_code, was added to
the dataset.
```

Computing ATET for each cohort and time:  
Cohort 2014 (10): .....10 done  
Cohort 2015 (10): .....10 done  
Cohort 2016 (10): .....10 done  
Cohort 2019 (10): .....10 done

Treatment and time information

Time variable: year_code	
Time interval: 2009 to 2019	
Control: _did_cohort = 0	
Treatment: _did_cohort > 0	
	_did_cohort
Number of cohorts	5
Number of obs	
Never treated	13442
2014	10758
2015	1881
2016	1023
2019	1540

Heterogeneous-treatment-effects regression Number of obs = 28,644  
Number of panels = 2,604  
Estimator: Augmented IPW  
Panel variable: county\_code  
Treatment level: county\_code  
Control group: Never treated

(Std. err. adjusted for 2,604 clusters in county_code)						
Cohort	ATET	Robust std. err.	z	P> z	[95% conf. interval]	
2014						
year_code						
2010	-5.604832	3.9404	-1.42	0.155	-13.32788	2.11821
2011	1.565364	4.02757	0.39	0.698	-6.328529	9.459258
2012	4.333278	4.041783	1.07	0.284	-3.58847	12.25503
2013	-7.134687	4.832863	-1.48	0.140	-16.60692	2.337545
2014	-.8585637	4.578522	-0.19	0.851	-9.832302	8.115175
2015	6.322813	4.612823	1.37	0.170	-2.718155	15.36378
2016	13.24828	4.413365	3.00	0.003	4.598245	21.89832
2017	13.71379	4.809485	2.85	0.004	4.287371	23.14021
2018	11.20397	5.128611	2.18	0.029	1.15208	21.25586
2019	18.11959	4.906644	3.69	0.000	8.50274	27.73643
2015						
year_code						
2010	8.266803	6.439382	1.28	0.199	-4.354154	20.88776
2011	1.740605	6.766555	0.26	0.797	-11.5216	15.00281
2012	-6.266669	6.442608	-0.97	0.331	-18.89395	6.360611
2013	-1.1612685	8.60929	-0.02	0.985	-17.03517	16.71263
2014	1.748804	7.842482	0.22	0.824	-13.62218	17.11979
2015	.5232453	7.120012	0.07	0.941	-13.43172	14.47821
2016	7.24529	7.158168	1.01	0.311	-6.784461	21.27504
2017	11.72394	7.677091	1.53	0.127	-3.322878	26.77077
2018	.5067276	7.981191	0.06	0.949	-15.13612	16.14957
2019	7.980641	7.337396	1.09	0.277	-6.40039	22.36167
2016						
year_code						
2010	-12.64943	13.80444	-0.92	0.359	-39.70563	14.40677
2011	-7.384062	16.60646	-0.44	0.657	-39.93212	25.164
2012	-7.495301	20.6746	-0.36	0.717	-48.01678	33.02617
2013	1.179399	11.02206	0.11	0.915	-20.42344	22.78223

2014	-3.01771	12.4699	-0.24	0.809	-27.45826	21.42284
2015	-2.327461	12.49634	-0.19	0.852	-26.81983	22.16491
2016	.6638016	12.32086	0.05	0.957	-23.48464	24.81224
2017	-16.99212	12.07182	-1.41	0.159	-40.65244	6.668212
2018	-34.36462	12.05456	-2.85	0.004	-57.99113	-10.73811
2019	-19.19046	14.48101	-1.33	0.185	-47.57272	9.191802
<b>2019</b>						
year_code						
2010	-.7287782	7.932366	-0.09	0.927	-16.27593	14.81837
2011	9.486301	7.534617	1.26	0.208	-5.281277	24.25388
2012	-1.979087	7.365776	-0.27	0.788	-16.41574	12.45757
2013	-6.536168	7.619386	-0.93	0.352	-20.29391	7.221575
2014	8.193503	7.315425	1.12	0.263	-6.144466	22.53147
2015	7.68665	8.340697	0.92	0.357	-8.660816	24.03412
2016	12.11781	9.013375	1.34	0.179	-5.548083	29.7837
2017	-1.915545	9.061966	-0.21	0.833	-19.67667	15.84558
2018	-5.344114	8.427639	-0.63	0.526	-21.86198	11.17375
2019	3.413603	8.620258	0.40	0.692	-13.48179	20.309

Note: ATET computed using covariates.

Note: Base time for pretreatment ATETs is the previous period.

29 . estat aggregation, cohort

ATET over cohort Number of obs = 28,644

(Std. err. adjusted for 2,604 clusters in county\_code)

Cohort	Robust				
	ATET	std. err.	z	P> z	[95% conf. interval]
2014	10.29165	3.631284	2.83	0.005	3.17446 17.40883
2015	5.59597	5.746254	0.97	0.330	-5.666481 16.85842
2016	-17.47085	10.46041	-1.67	0.095	-37.97288 3.831183
2019	3.413603	8.620258	0.40	0.692	-13.48179 20.309

30 . estat aggregation, time

ATET over time Number of obs = 28,644

(Std. err. adjusted for 2,604 clusters in county\_code)

Time	Robust				
	ATET	std. err.	z	P> z	[95% conf. interval]
2014	-.8585637	4.578522	-0.19	0.851	-9.832302 8.115175
2015	5.459692	4.213269	1.30	0.195	-2.798164 13.71755
2016	11.47947	3.921092	2.93	0.003	3.794267 19.16467
2017	11.14059	4.323841	2.58	0.010	2.666017 19.61516
2018	6.319022	4.657509	1.36	0.175	-2.809527 15.44757
2019	12.86457	4.086501	3.15	0.002	4.855176 20.87396

31 .

32 .

33 . \*\* CH?

34 . gen Dhat = treatPost

35 . replace Dhat=. if Dhat==0  
(21,409 real changes made, 21,409 to missing)

36 . \*did\_imputation\_crude\_rate\_20\_64 county\_code year\_code Dhat, controls(perc\_female perc\_white perc\_hispanic unemp\_rate\_pc poverty\_rate median\_i &gt; ncome\_k) \*autosample

37 .

38 . log close  
name: <unnamed>  
log: d:\Temp\MCP\Did\Did-toolbox\Examples\medicaid.smcl  
log type: smcl  
closed on: 18 Jan 2026, 14:15:59