Difference-in-differences estimation using Stata

Joerg Luedicke

Senior Social Scientist and Software Developer

StataCorp LLC

June 10, 2022
Frankfurt

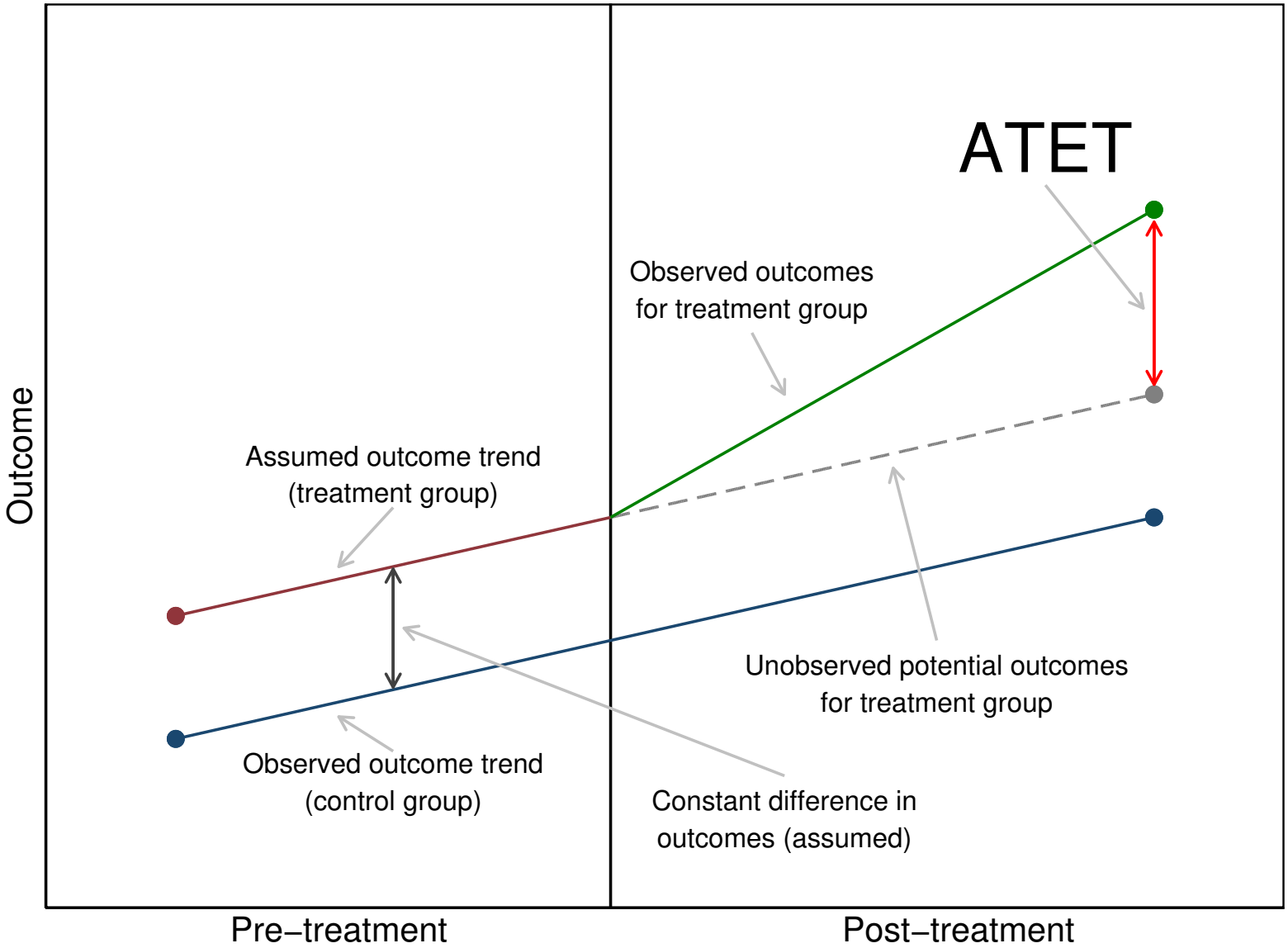
What is DID?

- Difference-in-differences models (DID) are used in before/after scenarios
- Examples: public policy evaluation, intervention studies
- Do increased cigarette taxes lead to a reduction in smoking?
- Simply looking at smokers who are exposed to the tax increase before and after the increase would not be enough to answer this question because there could be a general trend of reduced smoking and we could not say how much of the reduction, is attributible to the tax increase, if any
- What we also need is a group of smokers that were not exposed to the tax increase

The basic idea of DID

- The basic idea of DID here is, that we can estimate the general trend of smoking from the group of non-exposed smokers, and then the difference between that and the trend among exposed smokers is the effect that can be attributed to the tax increase
- In other words, we have a before/after difference for both a treament and a control group, and then the difference between these two differences is our treatment effect, hence the name DID
- An important assumption here is that the trends are the same in both groups prior to the intervention, and that the trends would be the same in both groups had the intervention not taken place
- This is known as the parallel trends assumption. This is an effect identifying assumption, and treatment effects cannot be estimated if this assumption is violated

DID estimates ATET



A simple 2x2 design

- We start with the simplest of cases where we have one treatment group, one control group, one period prior to intervention, and one period post intervention:

```
. list in 1/5
```

	id	post	treat	state	age	gender	income	ncigs
1.	1	1	0	3	58	1	60.47076	9.226687
2.	2	0	1	29	61	0	35.18789	9.694955
3.	3	1	1	49	29	0	48.33061	5.798977
4.	4	0	1	42	57	0	76.48076	4.361449
5.	5	0	1	40	40	1	65.70543	9.136448

A simple DID model

- We estimate our first DID model using a linear model with both the treatment and pre/post variable as well as their interaction:

```
. regress ncigs i.treat##i.post
      Source |       SS           df       MS      Number of obs   =     1,000
-------------+----------------------------------   F(3, 996)       =     13.64
       Model |  591.8889           3  197.2963      Prob > F        =    0.0000
    Residual |  14408.5533       996  14.466419     R-squared       =    0.0395
-------------+----------------------------------   Adj R-squared   =    0.0366
       Total |  15000.4422       999  15.0154577    Root MSE        =    3.8035

------------------------------------------------------------------------------
       ncigs | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    1.treat |    .1928871   .3442581     0.56   0.575    -.4826673    .8684415
     1.post |   -.7306656   .3352332    -2.18   0.030     -1.38851   -.0728211
            |
  treat#post |
        1 1 |   -1.242638   .4814626    -2.58   0.010    -2.187436   -.2978409
            |
      _cons |    10.15872   .2386512    42.57   0.000     9.690402    10.62704
------------------------------------------------------------------------------
```

- The coefficient on the interaction term is the estimate of the difference-in-differences

Cell means and DID

- More intuitively, with the simple 2x2 design, we can look at the expected outcome means for each cell:

```
. margins treat#post, post coeflegend
Adjusted predictions                                  Number of obs = 1,000
Model VCE: OLS
Expression: Linear prediction, predict()
```

	Margin	Legend
treat#post		
0 0	10.15872	_b[0bn.treat#0bn.post]
0 1	9.428053	_b[0bn.treat#1.post]
1 0	10.35161	_b[1.treat#0bn.post]
1 1	8.378302	_b[1.treat#1.post]

- Manually calculating the differences:

```
. display (_b[1.treat#1.post]   - _b[1.treat#0bn.post])  - ///
>         (_b[0bn.treat#1.post] - _b[0bn.treat#0bn.post])
-1.2426383
```

Including state fixed effects to account for group level unobservables

```
. regress ncigs i.state i.post 1.treat#1.post
      Source |       SS           df       MS      Number of obs   =     1,000
-------------+----------------------------------   F(51, 948)      =      1.86
       Model |  1363.45884        51  26.7344872   Prob > F        =    0.0003
    Residual |  13636.9834       948  14.3850035   R-squared       =    0.0909
-------------+----------------------------------   Adj R-squared   =    0.0420
       Total |  15000.4422       999  15.0154577   Root MSE        =    3.7928

------------------------------------------------------------------------------
       ncigs | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       state |
          2  |   1.122329   1.176153     0.95   0.340    -1.185834    3.430493
          3  |   2.860078    1.14564     2.50   0.013     .6117944    5.108363
        <snip>
         50  |    .946973   1.175844     0.81   0.421    -1.360585    3.254531
             |
      1.post |  -.7429488   .3422385    -2.17   0.030    -1.414582   -.0713161
             |
 treat#post  |
        1 1  |  -1.283547   .4919954    -2.61   0.009    -2.249073   -.3180208
             |
       _cons |    8.57291   .8021031    10.69   0.000     6.998807    10.14701
------------------------------------------------------------------------------
```

Stata's DID commands

- Usually we have data with multiple pre/post observations, and treatment may be administered at different times
- Data could be repeated cross-sections or panel data
- The commands in Stata to fit DID models are `didregress` for repeated cross-sections, and `xtdidregress` for panel data
- Variety of methods for standard errors including aggregation methods, bias-corrected standard errors, and wild bootstrap
- Postestimation commands for tests and diagnostics to evaluate parallel-trends assumption

Treatment as time-varying variable

- Both commands require the treatment variable to be specified as a time varying variable:

```
. gen policy = treat*post
. list id treat post policy in 1/10, sep(0)
```

	id	treat	post	policy
1.	1	0	1	0
2.	2	1	0	0
3.	3	1	1	1
4.	4	1	0	0
5.	5	1	0	0
6.	6	1	0	0
7.	7	1	1	1
8.	8	0	0	0
9.	9	0	1	0
10.	10	1	1	1

Using didregress

```
. didregress (ncigs) (policy), group(state) time(post)
Number of groups and treatment time
Time variable: post
Control:       policy = 0
Treatment:     policy = 1
```

	Control	Treatment
Group		
state	25	25
Time		
Minimum	0	1
Maximum	0	1

```
Difference-in-differences regression                 Number of obs = 1,000
Data type: Repeated cross-sectional
                                  (Std. err. adjusted for 50 clusters in state)
```

ncigs	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
policy (1 vs 0)	-1.283547	.5672038	-2.26	0.028	-2.423386	-.1437081

```
Note: ATET estimate adjusted for group effects and time effects.
```

The model

- We use the following (two-way fixed-effects) model:

$$y_{ist} = \gamma_i + \gamma_t + \mathbf{z}_{ist}\beta + D_{st}\delta + \varepsilon_{ist}$$

where:

y_{ist}	is the outcome of person i in group s at time t
γ_s	are group fixed effects
γ_t	are time fixed effects
$\mathbf{z}_{ist}$	are covariates
β	are the coefficients on the covariates
D_{st}	is the (time-varying) treatment indicator
δ	is the coefficient on the treatment indicator, i.e. the ATET
ε_{ist}	are the residual errors

Example with repeated cross-sections

```
. use https://www.stata-press.com/data/r17/hospdd
(Artificial hospital admission procedure data)
. describe
Contains data from https://www.stata-press.com/data/r17/hospdd.dta
 Observations:         7,368                  Artificial hospital admission
                                                procedure data
    Variables:             5                  7 Mar 2021 19:52
```

Variable name	Storage type	Display format	Value label	Variable label
hospital	byte	%9.0g		Hospital ID
frequency	byte	%9.0g	size	Hospital visit frequency
month	byte	%8.0g	mnth	Month
procedure	byte	%9.0g	pol	Admission procedure
satis	float	%9.0g		Patient satisfaction score

```
Sorted by: hospital
. list in 1/5
```

	hospital	frequency	month	proced~e	satis
1.	1	High	July	New	4.106527
2.	1	Medium	March	Old	3.319475
3.	1	Very high	February	Old	3.41172
4.	1	Medium	April	New	3.004025
5.	1	Low	March	Old	3.11072

DID model using didregress

```
. didregress (satis) (procedure), group(hospital) time(month)
Number of groups and treatment time
Time variable: month
Control:       procedure = 0
Treatment:     procedure = 1
```

	Control	Treatment
Group		
hospital	28	18
Time		
Minimum	1	4
Maximum	1	4

```
Difference-in-differences regression                 Number of obs = 7,368
Data type: Repeated cross-sectional
                                (Std. err. adjusted for 46 clusters in hospital)
```

satis	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
procedure (New vs Old)	.8479879	.0321121	26.41	0.000	.7833108	.912665

```
Note: ATET estimate adjusted for group effects and time effects.
```

DID model using areg

```
. areg satis i.month i.procedure, absorb(hospital) vce(cluster hospital)
Linear regression, absorbing indicators                 Number of obs     =  7,368
Absorbed variable: hospital                             No. of categories =     46
                                                        F(7, 45)          = 138.73
                                                        Prob > F          = 0.0000
                                                        R-squared         = 0.5333
                                                        Adj R-squared     = 0.5299
                                                        Root MSE          = 0.7238
                                      (Std. err. adjusted for 46 clusters in hospital)
```

satis	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
month						
February	-.0096077	.0184317	-0.52	0.605	-.0467311	.0275158
March	.0219686	.018251	1.20	0.235	-.0147907	.0587279
April	-.0032839	.0221028	-0.15	0.883	-.0478013	.0412335
May	-.0094027	.0232399	-0.40	0.688	-.0562103	.0374048
June	-.0038375	.0190634	-0.20	0.841	-.0422332	.0345581
July	-.0111941	.0230029	-0.49	0.629	-.0575244	.0351361
procedure						
New	.8479879	.0321121	26.41	0.000	.7833108	.912665
_cons	3.444675	.011354	303.39	0.000	3.421807	3.467543

Panel data

```
. use https://www.stata-press.com/data/r17/patents
(Excerpt from Moser and Voena (2012))
. describe
Contains data from https://www.stata-press.com/data/r17/patents.dta
 Observations:       471,120                  Excerpt from Moser and Voena
                                                (2012)
    Variables:             5                  7 Mar 2021 23:17
                                              (_dta has notes)
```

Variable name	Storage type	Display format	Value label	Variable label
year	int	%9.0g		Year
uspatents	byte	%9.0g		Number of US patents
fpatents	byte	%9.0g		Number of foreign patents
classid	float	%9.0g		Class ID
gotpatent	byte	%9.0g	gp	Subclass got patent post 1918

```
Sorted by:
```

Panel data DID model using xtdidregress

```
. xtset classid
Panel variable: classid (balanced)
. xtdidregress (uspatents fpatents) (gotpatent), group(classid) time(year)
Number of groups and treatment time
Time variable: year
Control:       gotpatent = 0
Treatment:     gotpatent = 1
```

	Control	Treatment
Group		
classid	6912	336
Time		
Minimum	1875	1919
Maximum	1875	1919

```
Difference-in-differences regression                 Number of obs = 471,120
Data type: Longitudinal
                                (Std. err. adjusted for 7,248 clusters in classid)
```

uspatents	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
gotpatent (Patent vs None)	.150516	.0356081	4.23	0.000	.0807137	.2203183

```
Note: ATET estimate adjusted for covariates, panel effects, and time effects.
```

Small number of groups

- Sometimes the number of groups is small. For example, consider the case where cigarette taxes are raised in one state, and the control population consists of people from just two or three other states
- The default cluster-robust standard errors do not perform well when the number of clusters is small.
- `didregress` and `xtdidregress` offer three ways to deal with this issue:
  - Wild cluster bootstrap
  - bias-corrected clustered standard errors
  - Aggregated estimation with or without bias-corrected clustered standard errors

Example with a small number of groups

- Data:

```
. use https://www.stata-press.com/data/r17/smallg
(Simulated data with a small number of groups)
. tab county
```

County	Freq.	Percent	Cum.
1	715	7.15	7.15
2	2,570	25.70	32.85
3	3,410	34.10	66.95
4	2,285	22.85	89.80
5	920	9.20	99.00
6	100	1.00	100.00
Total	10,000	100.00	

DID model with default standard errors

```
. didregress (outcome x i.b) (treated), group(county) time(year)
Number of groups and treatment time
Time variable: year
Control:       treated = 0
Treatment:     treated = 1
```

	Control	Treatment
Group		
county	4	2
Time		
Minimum	2011	2013
Maximum	2011	2013

```
Difference-in-differences regression                 Number of obs = 10,000
Data type: Repeated cross-sectional
                                  (Std. err. adjusted for 6 clusters in county)
```

outcome	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
treated						
(Treated						
vs						
Untreated)	-.9394987	.0884134	-10.63	0.000	-1.166773	-.7122247

```
Note: ATET estimate adjusted for covariates, group effects, and time effects.
```

DID model with wild cluster bootstrap

```
. didregress (outcome x i.b) (treated), group(county) time(year) wildbootstrap(
> rseed(123))
computing 1000 replications
Confidence interval lower bound
.................................
Confidence interval upper bound
.............................
Number of groups and treatment time
Time variable: year
Control:       treated = 0
Treatment:     treated = 1
```

	Control	Treatment
Group		
county	4	2
Time		
Minimum	2011	2013
Maximum	2011	2013

```
DID with wild-cluster bootstrap inference          Number of obs    = 10,000
                                                   No. of clusters =      6
                                                   Replications     =  1,000

Data type:    Repeated cross-sectional
Error weight: rademacher
```

outcome	Coefficient	t	P>|t|	[96.40% conf.	interval]
ATET					
treated					
(Treated vs Untreated)	-.9394987	-10.63	0.000	-1.214059	-.6689265

```
Note: 96.40% confidence interval is wider than requested.
Note: ATET estimate adjusted for covariates, group effects, and time effects.
```

Parallel trends assumption

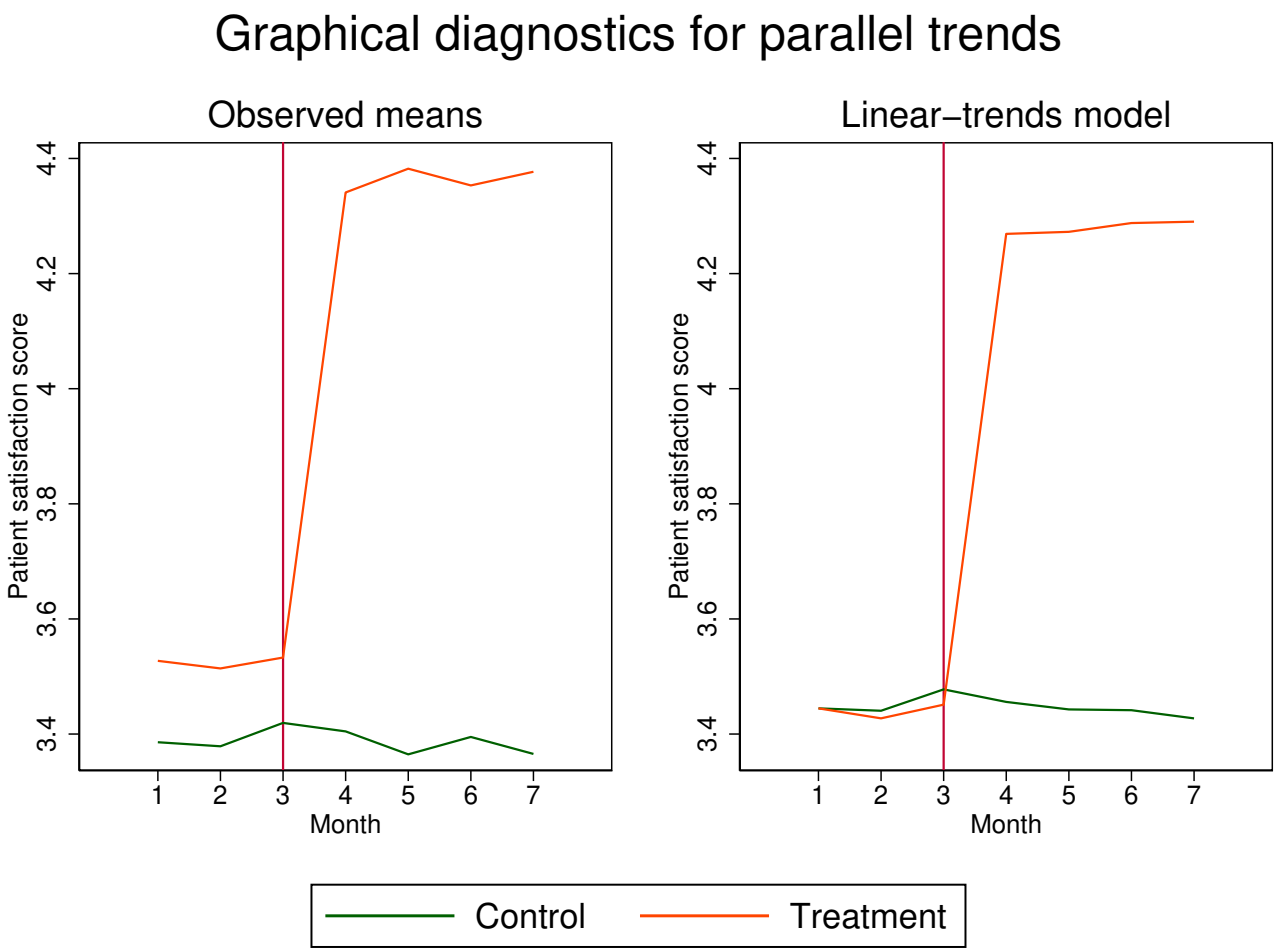
- As noted earlier, the parallel trends assumption has to hold for the ATET to be identified
- The assumption is that the trends are the same between the treatment and control groups prior to the intervention date, and that they would be the same past this date if there had not been an intervention
- It is not possible to directly test this assumption because we do not know what would have happened if there was no intervention
- A common strategy to indirectly evaluate this assumption is to check whether the trends prior to the intervention are the same, or at least similar
- The rationale here is that, if the pre-treatment trends are the same, then they would keep following these same trends in the absence of an intervention

Stata commands

- Postestimation commands to assess the parallel trends assumption:
  - `estat trendplots`
  - `estat ptrends`
  - `estat granger`
  - `estat grangerplot`

Graphical diagnostics

```
. use https://www.stata-press.com/data/r17/hospdd
(Artificial hospital admission procedure data)
. qui didregress (satis) (procedure), group(hospital) time(month)
. estat trendplots
```



The linear-trends model

- Let's rewrite $y_{ist} = \gamma_i + \gamma_t + \mathbf{x}_{ist}\beta + D_{st}\delta + \varepsilon_{ist}$ as $y_{ist} = \text{DID}_{ist} + \epsilon_{ist}$
- The linear-trends model augments the above model with two more terms:

  $y_{ist} = \text{DID}_{ist} + w_i d_{t,0} t\zeta_1 + w_i d_{t,1} t\zeta_2 + \varepsilon_{ist}$
- The augmentation terms consist of two 3-way interactions between $d_{t,0}$, w_i, and t, and $d_{t,1}$, w_i, and t
- $d_{t,0} = 1(d_t = 0)$ is a variable indicating pretreatment time periods
- $d_{t,1} = 1(d_t = 1)$ indicating posttreatment time periods
- w_i variable that is 1 if ever treated, and 0 if never treated
- The coefficient ζ_1 captures the differences in slopes between treatment group and control group in pretreatment periods, while ζ_2 captures the differences in slopes in posttreatment periods.
- If ζ_1 is 0, the linear trends in the outcome are parallel during pretreatment periods.
- `estat ptrends` uses a Wald test of ζ_1 against 0 to assess whether the linear trends are parallel prior to treatment

The Granger model

- `estat granger` performs a Granger-type causality test to assess whether treatment effects are observed prior to the treatment
- The Granger-type test augments the DID model with counterfactual treatment-time indicators
- The augmentation terms are referred to as leads in the DID literature
- The model used by `estat granger` augments the DID model with all leads (leaving out one for identification purposes):

$$y_{ist} = \text{DID}_{ist} + \sum_{j=2}^{J-1} 1(t_{it} \geq j) w_i \lambda_j + \nu_{ist}$$

- The test result is obtained by performing a joint Wald test on the coefficients λ_j.

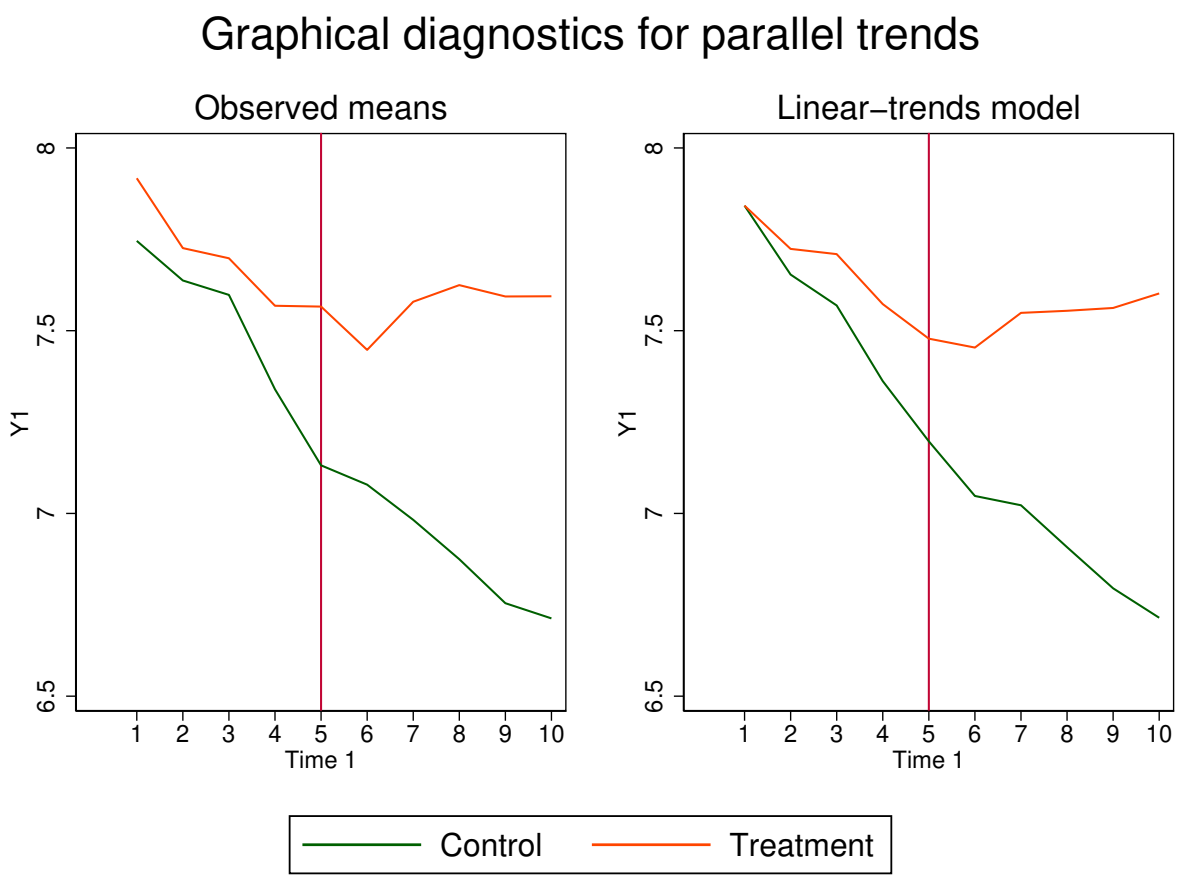
Testing for parallel trends

```
. estat ptrends
Parallel-trends test (pretreatment time period)
H0: Linear trends are parallel
F(1, 45) =   0.55
Prob > F = 0.4615

. estat granger
Granger causality test
H0: No effect in anticipation of treatment
F(2, 45) =   0.33
Prob > F = 0.7239
```

Non-parallel trends

```
. use https://www.stata-press.com/data/r17/parallelt
(Simulated data to test parallel-trends assumption)
. qui xtset id1
. qui xtdidregress (y1 c.x1##c.x2) (treated1), group(id1) time(t1)
. estat trendplots
```



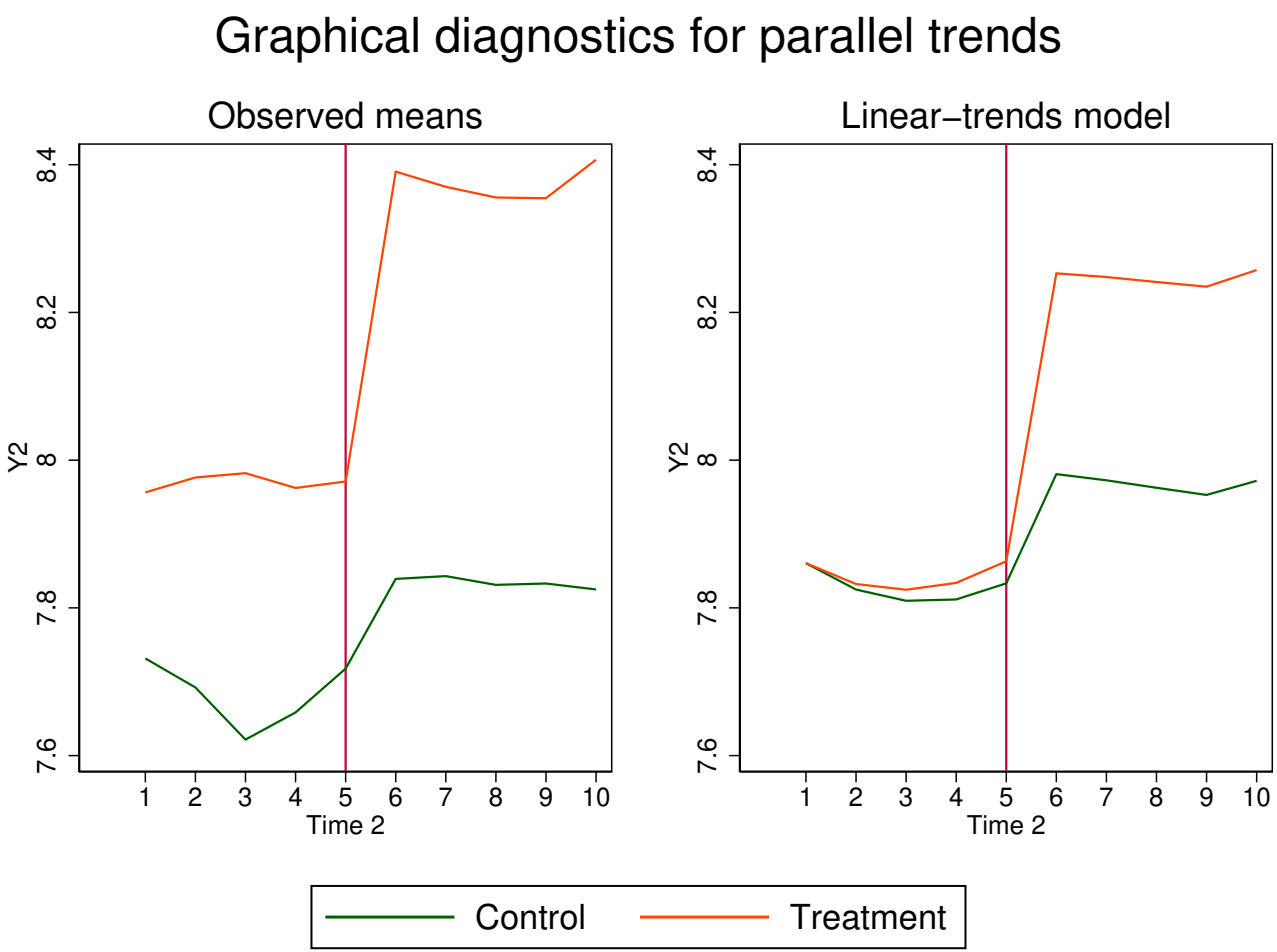
Testing

```
. estat ptrends
Parallel-trends test (pretreatment time period)
H0: Linear trends are parallel
F(1, 199) =  39.97
 Prob > F = 0.0000

. estat granger
Granger causality test
H0: No effect in anticipation of treatment
F(4, 199) =  18.17
 Prob > F = 0.0000
```

Nonlinear differences

```
. qui xtset id2
. qui xtdidregress (y2 c.z1##c.z2) (treated2), group(id2) time(t2)
. estat trendplots
```



The Granger test is more flexible

```
. estat ptrends
Parallel-trends test (pretreatment time period)
H0: Linear trends are parallel
F(1, 999) =    2.13
 Prob > F = 0.1446

. estat granger
Granger causality test
H0: No effect in anticipation of treatment
F(4, 999) =    9.86
 Prob > F = 0.0000
```

Time-varying effects

- `estat grangerplot` fits a generalization of the DID model and plots the estimated coefficients (including their 95% confidence intervals)
- The model is similar to the Granger model above, but uses a different parameterization, and includes lags in addition to leads
- Let I_s be the time of treatment, $m < 0$ be the number of time periods prior to I_s, $q \geq 0$ be the number of periods after I_s, and b be the baseline period, the model is

$$y_{ist} = \gamma_i + \gamma_t + \mathbf{x}_{ist}\beta + \sum_{k=m,k\neq b}^{q} B_{st}^k w_i \lambda_k + \epsilon_{ist}$$

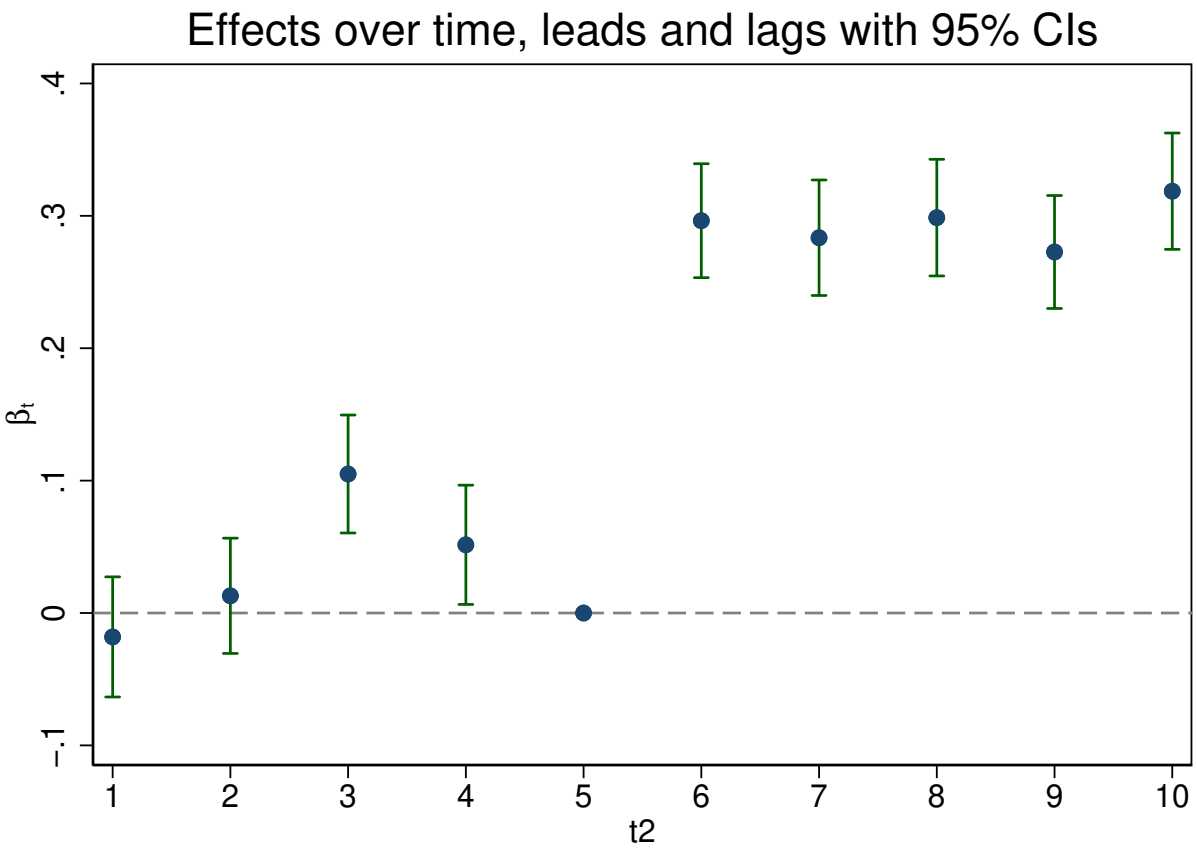
$$B_{st}^k = \begin{cases} 1(t_{it} \leq I_s + k), \text{ if } & k = m \\ 1(t_{it} = I_s + k), \text{ if } & m < k < q \\ 1(t_{it} \geq I_s + k), \text{ if } & k = q \end{cases}$$

- This model is also known as event-study model

estat grangerplot

```
. use https://www.stata-press.com/data/r17/parallelt
(Simulated data to test parallel-trends assumption)
. xtset id2
Panel variable: id2 (balanced)
. xtdidregress (y2 c.z1##c.z2) (treated2), group(id2) time(t2)
<output omitted>

. estat grangerplot
```



Numeric model results

```
. estat grangerplot, verbose nodraw
Fixed-effects (within) regression               Number of obs      =     10,000
Group variable: id2                             Number of groups   =      1,000
R-squared:                                      Obs per group:
     Within  = 0.5412                                         min =         10
     Between = 0.6513                                         avg =       10.0
     Overall = 0.5342                                         max =         10
                                                F(21,999)          =     466.16
corr(u_i, Xb) = 0.2086                          Prob > F           =     0.0000
                                    (Std. err. adjusted for 1,000 clusters in id2)
```

y2	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
z1	.8076361	.0217797	37.08	0.000	.7648969	.8503752
z2	-.0899957	.0230564	-3.90	0.000	-.1352401	-.0447513
c.z1#c.z2	-.174932	.0359373	-4.87	0.000	-.2454532	-.1044109
t2						
2	-.0480015	.0161614	-2.97	0.003	-.0797157	-.0162874
<snip>						
10	.0847351	.0159384	5.32	0.000	.0534585	.1160117
_lead5	-.018056	.0231218	-0.78	0.435	-.063429	.0273169
_lead4	.0129962	.0221836	0.59	0.558	-.0305356	.056528
_lead3	.1049961	.0226993	4.63	0.000	.0604523	.14954
_lead2	.0515103	.0229605	2.24	0.025	.0064539	.0965667
_lag0	.2963356	.0219243	13.52	0.000	.2533126	.3393587
_lag1	.2834751	.0222323	12.75	0.000	.2398477	.3271026
_lag2	.2986596	.0224389	13.31	0.000	.2546268	.3426925
_lag3	.2727031	.0217387	12.54	0.000	.2300443	.3153618
_lag4	.3186162	.0223891	14.23	0.000	.2746811	.3625514
_cons	7.555064	.0181931	415.27	0.000	7.519363	7.590765
sigma_u	.17808737					
sigma_e	.24913922					
rho	.33816659	(fraction of variance due to u_i)				

Granger test

```
. xtdidregress (y2 c.z1##c.z2) (treated2), group(id2) time(t2)
Number of groups and treatment time
Time variable: t2
Control:       treated2 = 0
Treatment:     treated2 = 1
```

	Control	Treatment
Group		
id2	480	520
Time		
Minimum	1	6
Maximum	1	6

```
Difference-in-differences regression                 Number of obs = 10,000
Data type: Longitudinal
                                  (Std. err. adjusted for 1,000 clusters in id2)
```

y2	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
ATET						
treated2 (Treated vs Untreated)	.2636651	.0097188	27.13	0.000	.2445936	.2827367

```
Note: ATET estimate adjusted for covariates, panel effects, and time effects.

. estat granger
Granger causality test
H0: No effect in anticipation of treatment
F(4, 999) =    9.86
 Prob > F = 0.0000
```

Reproducing Granger test

```
. estat grangerplot, post nodraw nlags(0)
Fixed-effects (within) regression               Number of obs      =     10,000
Group variable: id2                             Number of groups   =      1,000
R-squared:                                      Obs per group:
     Within  = 0.5410                                         min =         10
     Between = 0.6513                                         avg =       10.0
     Overall = 0.5340                                         max =         10
                                                F(17,999)          =     573.04
corr(u_i, Xb) = 0.2086                          Prob > F           =     0.0000
                                    (Std. err. adjusted for 1,000 clusters in id2)
```

y2	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
z1	.8078464	.0217724	37.10	0.000	.7651215	.8505714
z2	-.0899448	.0230523	-3.90	0.000	-.1351811	-.0447084
c.z1#c.z2	-.1750781	.035931	-4.87	0.000	-.2455871	-.1045692
t2						
2	-.0480031	.0161587	-2.97	0.003	-.0797119	-.0162942
<snip>						
10	.0975545	.0142977	6.82	0.000	.0694975	.1256114
_lead5	-.0180553	.0231174	-0.78	0.435	-.0634195	.0273089
_lead4	.0129911	.0221788	0.59	0.558	-.0305313	.0565136
_lead3	.1049946	.0226945	4.63	0.000	.0604603	.1495289
_lead2	.0515093	.0229561	2.24	0.025	.0064615	.096557
_lag0	.2939576	.0170635	17.23	0.000	.2604733	.3274419
_cons	7.554972	.0181973	415.17	0.000	7.519263	7.590681
sigma_u	.17808747					
sigma_e	.24915053					
rho	.33814654	(fraction of variance due to u_i)				

```
. test _lead2 _lead3 _lead4 _lead5
 ( 1)  _lead2 = 0
 ( 2)  _lead3 = 0
 ( 3)  _lead4 = 0
 ( 4)  _lead5 = 0
       F(  4,   999) =    9.86
            Prob > F =    0.0000
```

Reproducing ATET

```
. xtdidregress (y2 c.z1##c.z2) (treated2), group(id2) time(t2)
Difference-in-differences regression                 Number of obs = 10,000
Data type: Longitudinal
                                  (Std. err. adjusted for 1,000 clusters in id2)
------------------------------------------------------------------------------
             |               Robust
          y2 | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
ATET         |
    treated2 |
  (Treated vs |
 Untreated) |   .2636651   .0097188    27.13   0.000     .2445936    .2827367
------------------------------------------------------------------------------
Note: ATET estimate adjusted for covariates, panel effects, and time effects.
. estat grangerplot, verbose nodraw nleads(1) nlags(0)
Fixed-effects (within) regression               Number of obs     =     10,000
Group variable: id2                             Number of groups  =      1,000
R-squared:                                      Obs per group:
     Within  = 0.5390                                         min =         10
     Between = 0.5931                                         avg =       10.0
     Overall = 0.5097                                         max =         10
                                                F(13,999)         =     744.05
corr(u_i, Xb) = 0.1818                          Prob > F          =     0.0000
                                  (Std. err. adjusted for 1,000 clusters in id2)
------------------------------------------------------------------------------
             |               Robust
          y2 | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
          z1 |     .80816   .021798    37.07   0.000     .7653848    .8509351
          z2 |  -.0894164   .0231268   -3.87   0.000    -.1347991   -.0440338
             |
   c.z1#c.z2 |  -.1751386   .0360169   -4.86   0.000    -.2458161    -.104461
             |
          t2 |
           2 |   -.031873   .0110175   -2.89   0.004    -.0534932   -.0102528
       <snip>
          10 |   .1227044   .0122365   10.03   0.000     .0986921    .1467167
             |
       _lag0 |   .2636651   .0097188   27.13   0.000     .2445936    .2827367
       _cons |   7.545177   .0130655  577.49   0.000     7.519538    7.570816
-------------+----------------------------------------------------------------
     sigma_u |  .18910858
     sigma_e |   .2496288
         rho |  .36463394   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Final remarks

- `didregress` and `xtdidregress` implement the TWFE DID estimator
- Also fits triple-difference models (DDD)
- https://www.stata.com/manuals/tedidregress.pdf
- `didregress` and `xtdidregress` can handle multiple treatment times, but TWFE DID can be problematic due to timing effects (see Goodman-Bacon 2021; `ssc describe bacondecomp`)
- DID is a fast growing literature, e.g.:
  - Nonlinear DID (Wooldridge 2021)
  - Doubly-robust DID (Sant'Anna & Zhao 2020)
  - Heterogenous DID (Callaway & Sant'Anna 2021; De Chaisemartin & D'Haultfoeuille 2022)

Thank you!