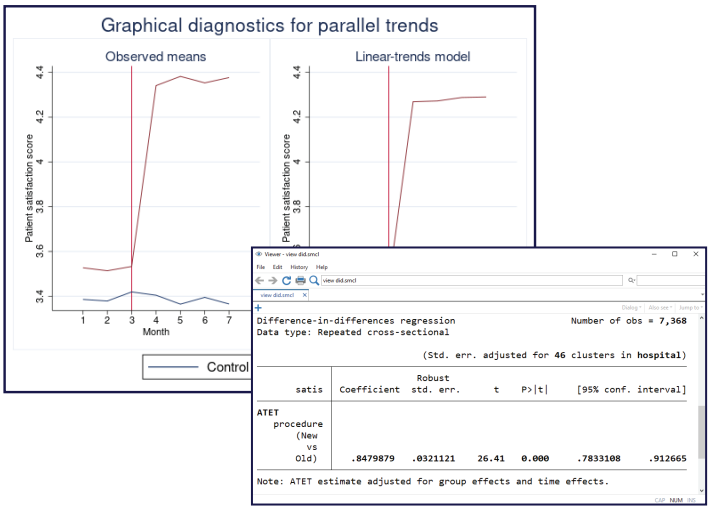
New in Stata 17

Difference in differences

Difference-in-differences (DID) estimation is one of the most popular methodologies for causal inference. Stata's new **didregress** and **xtdidregress** commands fit DID and difference-in-difference-in-differences (DDD) models that control for unobserved group and time effects.

didregress can be used with repeated cross-sectional data, where we sample different units of observations at different points in time. **xtdidregress** is for use with panel (longitudinal) data.



Highlights

- DID and DDD average treatment effect on the treated (ATET) estimators for repeated cross-sections and panel data
- Wild-bootstrap *p*-values and confidence intervals
- Bias-corrected standard errors using the Bell and McCaffrey degrees-of-freedom adjustment
- ATET estimates and standard errors using the Donald and Lang method
- Mean-outcome and parallel-trends graphical diagnostics
- Granger-type and parallel-trends tests

Fit a DID model

We study the effect of a new hospital admissions procedure (**procedure**) on patient satisfaction (**satis**) using monthly data on patients before and after the new procedure was implemented by some hospitals. To fit the DID model, we type

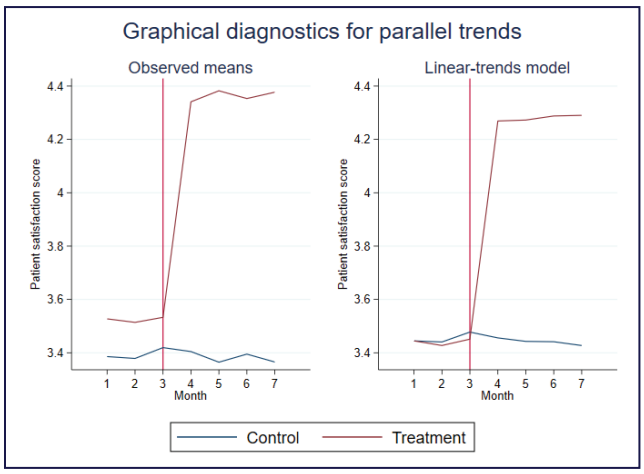
```
. didregress (satis) (procedure),
              group(hospital) time(month)
```

The ATET of **procedure** on **satis** is 0.85 (95%CI [0.78, 0.91]), accounting for **hospital** and **month** fixed effects. Treatment hospitals had a 0.85 increase in patient satisfaction relative to if they hadn't implemented the new procedure.

```
. 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)
------------------------------------------------------------------------------
             |               Robust
       satis | Coefficient  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.
```

Obtain graphical diagnostics

We assume the trajectories of **satis** are parallel for the control and treatment groups prior to implementation of the new procedure. We can obtain a graphical diagnostic of the assumption using **estat trendplot**.



Test for parallel trends

Prior to the policy implementation, control and treatment hospitals followed a parallel path. We can further evaluate this assumption using a parallel-trends test with **estat ptrends**.

```
. estat ptrends

Parallel-trends test (pretreatment time period)
H0: Linear trends are parallel

F(1, 45) =    0.55
Prob > F =  0.4615
```

We do not have sufficient evidence to reject the null hypothesis of parallel trends.

Perform Granger causality test

We also assume that the control or treatment groups do not change their behavior in anticipation of treatment. We compute a Granger causality test using **estat granger**.

```
. estat granger

Granger causality test
H0: No effect in anticipation of treatment

F(2, 45) =    0.33
Prob > F =  0.7239
```

We do not have sufficient evidence to reject the null hypothesis of no behavior change prior to treatment. Together with our previous diagnostics, these results suggest that we should trust the validity of our ATET estimate.

Compute appropriate standard errors

In this example, we had a sufficient number of hospitals (46) to make reliable inferences about our treatment effect. If we only had data from 15 hospitals, however, we may have considered alternative methods.

To use bias-corrected standard errors with the Bell and McCaffrey (2002) degrees-of-freedom adjustment, we can add the **vce(hc2)** option.

```
. didregress (satis) (procedure),
             group(hospital) time(month) vce(hc2)
```

To use the aggregation method proposed by Donald and Lang (2007), we can add the **aggregate(dlang)** option.

```
. didregress (satis) (procedure),
             group(hospital) time(month) aggregate(dlang)
```

We can also use the wild cluster bootstrap to obtain *p*-values and confidence intervals. As with all bootstrap-type methods, we need to set a seed to make our results replicable.

```
. didregress (satis) (procedure),
             group(hospital) time(month)
             wildbootstrap(rseed(111))
```

And more

What if we want to fit a DDD model? Just add another group to **group**, and define the new treated observations.

```
. didregress (satis) (new_treated),
             group(hospital another_group) time(month)
```

What if our data were longitudinal? Tell Stata using **xtset** and then just type

```
. xtdidregress (satis) (procedure),
               group(hospital) time(month)
```

References

Bell, R. M., and D. F. McCaffrey. 2002. Bias reduction in standard errors for linear regression with multi-stage samples. Survey Methodology 28: 169–181.

Donald, S. G., and K. Lang. 2007. Inference with difference-in-differences and other panel data. Review of Economics and Statistics 89: 221–233.

Learn more about difference in differences at **stata.com/did**.