PRINCETON UNIVERSITY

Differences-in-Differences (using Stata)

(*work in progress*)

Oscar Torres-Reyna
otorres@princeton.edu

August 2015

http://dss.princeton.edu/training/

Difference in differences (DID)
Estimation step-by-step

*** Getting sample data.**

```
use "http://dss.princeton.edu/training/Panel101.dta", clear
```

*** Create a dummy variable to indicate the time when the treatment started. Lets assume that treatment started in 1994. In this case, years before 1994 will have a value of 0 and 1994+ a 1. If you already have this skip this step.**

```
gen time = (year>=1994) & !missing(year)
```

*** Create a dummy variable to identify the group exposed to the treatment. In this example lets assumed that countries with code 5,6, and 7 were treated (=1). Countries 1-4 were not treated (=0). If you already have this skip this step.**

```
gen treated = (country>4) & !missing(country)
```

*** Create an interaction between time and treated. We will call this interaction 'did'**

```
gen did = time*treated
```

Difference in differences (DID)
Estimation step-by-step

*** Estimating the DID estimator**

```
reg y time treated did, r
```

```
. reg y time treated did, r

Linear regression                               Number of obs     =         70
                                                F(3, 66)          =       2.17
                                                Prob > F          =     0.0998
                                                R-squared         =     0.0827
                                                Root MSE          =     3.0e+09
```

y	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
time	2.29e+09	9.00e+08	2.54	0.013	4.92e+08	4.09e+09
treated	1.78e+09	1.05e+09	1.70	0.094	-3.11e+08	3.86e+09
did	-2.52e+09	1.45e+09	-1.73	0.088	-5.42e+09	3.81e+08
_cons	3.58e+08	7.61e+08	0.47	0.640	-1.16e+09	1.88e+09

*** The coefficient for ‘did’ is the differences-in-differences estimator. The effect is significant at 10% with the treatment having a negative effect.**

Difference in differences (DID)
Estimation step-by-step

*** Estimating the DID estimator (using the hashtag method, no need to generate the interaction)**

```
reg y time##treated, r
```

```
. reg y time##treated, r

Linear regression                               Number of obs     =         70
                                                F(3, 66)          =       2.17
                                                Prob > F          =     0.0998
                                                R-squared         =     0.0827
                                                Root MSE          =    3.0e+09
```

y	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
1.time	2.29e+09	9.00e+08	2.54	0.013	4.92e+08	4.09e+09
1.treated	1.78e+09	1.05e+09	1.70	0.094	-3.11e+08	3.86e+09
time#treated						
1 1	-2.52e+09	1.45e+09	-1.73	0.088	-5.42e+09	3.81e+08
_cons	3.58e+08	7.61e+08	0.47	0.640	-1.16e+09	1.88e+09

*** The coefficient for ‘time#treated’ is the differences-in-differences estimator (‘did’ in the previous example). The effect is significant at 10% with the treatment having a negative effect.**

Difference in differences (DID) Using the command `diff`

The command diff is user-defined for Stata. To install type

```
ssc install diff
```

Dummies for treatment and time, see previous slide

```
. diff y, t(treated) p(time)

Number of observations in the DIFF-IN-DIFF: 70
            Baseline       Follow-up
   Control: 16             24           40
   Treated: 12             18           30
            28             42

R-square:  0.08273
```

DIFFERENCE IN DIFFERENCES ESTIMATION

Outcome Variable	BASE LINE Control	Treated	Diff(BL)	FOLLOW UP Control	Treated	Diff(FU)	DIFF-IN-DIFF
y	3.6e+08	2.1e+09	1.8e+09	2.6e+09	1.9e+09	-7.4e+08	-2.5e+09
Std. Error	7.4e+08	8.5e+08	1.1e+09	6.0e+08	7.0e+08	9.2e+08	1.5e+09
t	0.49	3.6e+08	1.58	3.6e+08	4.4e+09	1.8e+09	-1.73
P>|t|	0.629	0.015	0.120	0.000	0.008	0.422	0.088*

* Means and Standard Errors are estimated by linear regression
**Inference: *** p<0.01; ** p<0.05; * p<0.1

p-value for the treatment effect, or DID estimator

Type `help diff` for more details/options

References

Introduction to econometrics, James H. Stock, Mark W. Watson. 2nd ed., Boston: Pearson Addison Wesley, 2007.

“Difference-in-Differences Estimation”, Imbens/Wooldridge, Lecture Notes 10, summer 2007.

http://www.nber.org/WNE/lect_10_diffindiffs.pdf

“Lecture 3: Differences-in-Differences”, Fabian Waldinger

http://www2.warwick.ac.uk/fac/soc/economics/staff/ffwaldinger/teaching/ec9a8/slides/lecture_3_-_did.pdf