



Munich Personal RePEc Archive

Simplifying the estimation of difference in differences treatment effects with Stata

Juan M. Villa

Brooks World Poverty Institute, University of Manchester

November 2012

Online at <https://mpra.ub.uni-muenchen.de/43943/>

MPRA Paper No. 43943, posted 22. January 2013 22:59 UTC

Simplifying the Estimation of Difference in Differences Treatment Effects with Stata*

Juan M. Villa
Brooks World Poverty Institute
University of Manchester
Manchester, UK.
juan.villalora@postgrad.manchester.ac.uk

*** DRAFT VERSION ***

Abstract. This paper explains the insights of the Stata's user written command `diff` for the estimation of Difference in Differences treatment effects (DID). The options and the formulas are detailed for the single DID, Kernel Propensity Score DID, Quantile DID and the balancing properties. An example of the features of `diff` is presented by using the dataset from Card and Krueger (1994).

Keywords: Difference in differences, causal inference, kernel propensity score, quantile treatment effects, quasi-experiments.

1. Introduction

Difference in Differences treatment effects (DID) have been widely used when the evaluation of a given intervention entails the collection of panel data or repeated cross sections. DID integrates the advances of the fixed effects estimators with the causal inference analysis when unobserved events or characteristics confound the interpretations (Angrist and Pischke, 2008).

Despite the existence of other plausible methods based on the availability of observational data for quasi-experimental causal inference -i.e. matching methods, instrumental variable, regression discontinuity-, DID estimations offer an alternative reaching the unconfoundedness by controlling for unobserved characteristics and combining it with observed or complementary information. Additionally, the DID is a flexible form of causal inference because it can be combined with some other procedures, such as the Kernel

* A previous version of this paper was presented at the 2012 UK Stata Users Group Meeting in London, UK. This version: November, 2012.

Propensity Score (Heckman et al., 1997, 1998) and the quintile regression (Meyer et al., 1995).

In this paper, the Stata's command `diff` is explained and some details on its implementation are given by using the datasets from the Card and Krueger (1994) article on the effects of the increase in the minimum wage. Similarly, it is explain how the balancing properties can be tested when observational data is provided.

In the next section the equations behind the estimation of the DID are explained along with the features of the `diff` command. In the third section an example is provided and, in the fourth section, the balancing properties are tested with the options that can be specified with the command.

2. `diff` syntax and equations

`diff` can be installed or updated from the SSC archive by running the command:

```
ssc install diff, replace
```

The `diff` syntax is detailed as follows:

```
diff outcome_var [if] [in] [weight] , [ options]
```

The command requests the specification of the outcome variable (`outcome_var`) and allows the use of weights, except for some options. The initial required option is the `period(varname)`, which contains a dummy variable indicating the baseline (`period==0`) and a follow-up (`period==1`) periods. Additionally, the option `treated(varname)`, is need, containing a dummy variable with the indicator of the control (`treated==0`) and `treated (treated==1)` individuals.

For the individual i , this initial setting performs the following linear regression:

$$outcome_var_i = \beta_0 + \beta_1 \cdot period_i + \beta_2 \cdot treated_i + \beta_3 \cdot period_i \cdot treated_i + e_i$$

The estimated coefficients have the following interpretation:

- $\widehat{\beta}_0$: Is the mean outcome for the control group on the baseline.
- $\widehat{\beta}_0 + \widehat{\beta}_1$: Is the mean outcome for the control group in the follow-up.
- $\widehat{\beta}_2$: Is the single difference between treated and control groups on the baseline.
- $\widehat{\beta}_0 + \widehat{\beta}_2$: Is the mean outcome for the treated group on the baseline.
- $\widehat{\beta}_0 + \widehat{\beta}_1 + \widehat{\beta}_2 + \widehat{\beta}_3$: Is the mean outcome for the treated group in the follow-up.

- $\widehat{\beta}_3$: Is the DID or impact.

The `diff` command arranges these coefficients in the output table. The number of observations, r-squared, standard errors, t-statistic -or the z-stat when standard errors are bootstrapped- and the p-value are also reported:

```
Number of observations in the DIFF-IN-DIFF: #
      Baseline      Follow-up
Control: #          #
Treated: #          #

R-square: 0.0

                DIFFERENCE IN DIFFERENCES ESTIMATION
-----
Outcome Variable | Control | Treated | Diff(BL) | Control | Treated | Diff(FU) | DIFF-IN-DIFF
-----
outcome_variable |  $\widehat{\beta}_0$  |  $\widehat{\beta}_0 + \widehat{\beta}_2$  |  $\widehat{\beta}_2$  |  $\widehat{\beta}_0 + \widehat{\beta}_1$  |  $\widehat{\beta}_0 + \widehat{\beta}_1 + \widehat{\beta}_2 + \widehat{\beta}_3$  |  $\widehat{\beta}_2 + \widehat{\beta}_3$  |  $\widehat{\beta}_3$ 
Std. Error       |         |         |         |         |         |         |
t/z              |         |         |         |         |         |         |
P>|t/z|          |         |         |         |         |         |         |
-----
* Means and Standard Errors are estimated by linear regression
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

2.1 Options

`cov(varlist)` - Specifies the pre-treatment covariates of the model. These variables are also known as controls or observable characteristics. If we denote $X_{k,i}$ as the k th covariate, `diff` runs the following regression with this option:

$$outcome_var_i = \beta_0 + \beta_1 \cdot period_i + \beta_2 \cdot treated_i + \beta_3 \cdot period_i \cdot treated_i + \beta_k \cdot X_{k,i} + e_i$$

The coefficients β_k are not reported in the output table. However, it is possible to request them if option `report` is specified.

`kernel` - Performs the Kernel-based Propensity Score DID. At a first stage, this option runs a `probit` model -or `logit` if this option is selected- of the `treated(varname)` on the `cov(varlist)`. It generates the variables `_weights` that contains the weights derived from the kernel density function and `_ps` when the Propensity Score is not specified in `pscore(varname)`. This option requires the `id(varname)` of each individual, hence it is not compatible with repeated cross section. It also allows the estimation of the DID on the common support by specifying the option `support`.

In a second stage, `diff` runs a regression applying the Stata's average weights option `[av=_weights]`, obtained from the propensity score:

$$outcome_var_i \cdot weights_i = \beta_0 + \beta_1 \cdot period_i + \beta_2 \cdot treated_i + \beta_3 \cdot period_i \cdot treated_i + e_i$$

Option `kernel` can be customized by selection the bandwidth, `bw(#)` and the kernel type, `ktype(kernel)`, according to the Stata's `kdensity` choices. Finally, the first stage is explicitly showed if `report` is specified.

`qdid(quantile)` - Performs the Quantile Difference in Differences estimation at the specified quantile from 0.1 to 0.9 (quantile 0.5 performs the QDID at the medeian). It may be combined with `kernel` and `cov(varlist)` options. `qdid(quantile)` does not support weights nor robust standard errors. This option uses Stata's `qreg` and `bsqreg` for bootstrapped standard errors. See Angrist and Pischke (2008) for detailed information on Quantile Treatment Effects and Meyer et al. (1995) for a illustrative example.

`cluster(varname)` - Calculates clustered standard errors by `varname`.

`robust` - Calculates robust Std. Errors.

`bs` - Performs a Bootstrap estimation of coefficients and standard errors. `reps(int)` specifies the number of repetitions when the `bs` is selected. The default are 50 repetitions.

`nostar` - Removes the inference stars from the p-values.

2.2 Option: balancing test

`test` - Performs a balancing t-test of difference in means of the specified covariates between the control and treated groups in `period == 0`. The option `test` combined with `kernel` performs the balancing t-test with the weighted covariates. Stata's `ttest` command is used to estimate the t-statistics and standard errors.

For each variable in `cov(varlist)`, `test` option runs the command:

```
ttest cov(varname) if period == 0, by(treated)
```

When combined with `kernel`, the differences, t-statistics and standard errors are generated with linear regression.

3. Example

`diff` offers an example with the dataset from Card and Krueger (1994). It can be downloaded into the working directory by running `net get diff` and then, use `cardkrueger1994, clear`. In this case, the authors study the impact of the increase in the minimum wage in the state of New Jersey -the treated group- on the employment level at the fast food industry. They compare the changes in the number of employees at the restaurants in this treated group to the ones of the neighbor state, Pennsylvania -the control group-. They collect a baseline in February, 1992, and a follow-up in November.

The description of the variables in the dataset are is the following:

```

Contains data from cardkrueger1994.dta
  obs:          820                Dataset from Card&Krueger (1994)
  vars:         8
  size:        18,860 (99.9% of memory free)
-----
storage display      value
variable name  type   format      label      variable label
-----
id             int    %8.0g                Store ID
t              byte   %8.0g                Feb. 1992 = 0; Nov. 1992 = 1
treated        long   %8.0g      treated   New Jersey = 1; Pennsylvania = 0
fte            float  %9.0g                Output: Full Time Employment
bk             byte   %8.0g                Burger King == 1
kfc            byte   %8.0g                Kentucky Fried Chicken == 1
roys           byte   %8.0g                Roy Rogers == 1
wendys         byte   %8.0g                Wendy's == 1
-----
Sorted by:  id t

```

With 820 observations, the number of individuals or stores are 331 and 79 in the treated and control groups, respectively. The outcome variable is fte, while some covariates are defined as dummy variable indicating whether the observation belongs to a given fast food restaurant. The basic statistic are show as follows:

Variable	Obs	Mean	Std. Dev.	Min	Max
id	820	246.5073	148.1413	1	522
t	820	.5	.5003052	0	1
treated	820	.8073171	.3946469	0	1
fte	801	17.59457	9.022517	0	80
bk	820	.4170732	.4933761	0	1
kfc	820	.195122	.3965364	0	1
roys	820	.2414634	.4282318	0	1
wendys	820	.1463415	.3536639	0	1

3.1 DID with no covariates

diff fte, t(treated) p(t)

The output table of this initial setting is:

```

Number of observations in the DIFF-IN-DIFF: 801
      Baseline      Follow-up
Control: 78         77         155
Treated: 326       320         646
      404           397

R-square:  0.00805

                DIFFERENCE IN DIFFERENCES ESTIMATION
-----
Outcome Variable | Control | Treated | Diff(BL) | Control | Treated | Diff(FU) | DIFF-IN-DIFF
-----
fte              | 19.949 | 17.065  | -2.884   | 17.542  | 17.573  | 0.030    | 2.914
Std. Error       | 1.019  | 0.499   | 1.135    | 1.026   | 0.503   | 1.143    | 1.611
t                | 19.57  | 14.17   | -2.54    | 17.60   | 20.45   | -0.33    | 1.81
P>|t|            | 0.000  | 0.000   | 0.011**  | 0.000   | 0.000   | 0.979    | 0.071*

```

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

The baseline information contains the columns with the mean outcome for each group and its difference (-2.88 in this case). These estimators are presented along with standard errors, t-statistics and p-values. The same information is showed for the baseline (with a difference of 0.03). The last column is the difference in differences, that is, $0.03 - (-2.88) = 2.94$. The p-value is accompanied by a star interpreted as the statistical inference at different significant levels.

Alternatively, bootstrapped standard errors can be requested by adding the option `bs`:

```
diff fte, t(treated) p(t) bs rep(50)
```

```
Bootstrap replications (50)
----- 1 ----- 2 ----- 3 ----- 4 ----- 5
..... 50
```

Number of observations in the DIFF-IN-DIFF: 801

	Baseline	Follow-up	
Control:	78	77	155
Treated:	326	320	646
	404	397	

R-square: 0.00805

Bootstrapped Standard Errors

Outcome Variable	DIFFERENCE IN DIFFERENCES ESTIMATION			DIFFERENCE IN DIFFERENCES ESTIMATION			DIFF-IN-DIFF
	BASE LINE			FOLLOW UP			
	Control	Treated	Diff(BL)	Control	Treated	Diff(FU)	
fte	19.949	17.065	-2.884	17.542	17.573	0.030	2.914
Std. Error	1.330	0.494	1.381	0.830	0.477	0.920	1.792
z	15.00	14.12	-2.09	17.05	20.76	0.28	1.63
P> z	0.000	0.000	0.037**	0.000	0.000	0.974	0.104

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

3.2 DID with covariates

```
diff fte, t(treated) p(t) cov(bk kfc roys)
```

DIFFERENCE-IN-DIFFERENCES WITH COVARIATES

Number of observations in the DIFF-IN-DIFF: 801

	Baseline	Follow-up	
Control:	78	77	155
Treated:	326	320	646
	404	397	

R-square: 0.18784

Outcome Variable	DIFFERENCE IN DIFFERENCES ESTIMATION			DIFFERENCE IN DIFFERENCES ESTIMATION			DIFF-IN-DIFF
	BASE LINE			FOLLOW UP			
	Control	Treated	Diff(BL)	Control	Treated	Diff(FU)	

fte	21.161	18.837	-2.324	18.758	19.369	0.611	2.935
Std. Error	1.142	0.851	1.031	1.158	0.853	1.037	1.460
t	18.53	18.43	-2.25	19.09	19.87	0.51	2.01
P> t	0.000	0.000	0.024**	0.000	0.000	0.556	0.045**

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

Option report allows the output table of the coefficients from the `cov(varlist)`:

Covariates and Coefficients:

Variable(s)	Coeff.	Std. Err.	t	P> t
bk	0.917	0.889	1.032	0.303
kfc	-9.205	1.006	-9.154	0.000
roys	-0.897	0.967	-0.927	0.354

3.3 Kernel Propensity Score DID

The Kernel Propensity Score DID can be estimated on the common support of the propensity score. If you have previously estimated the propensity score you can provide it with the option `pscore(varname)`. The basic syntax is:

```
diff fte, t(treated) p(t) cov(bk kfc roys) kernel id(id)
```

The full options are:

```
diff fte, t(treated) p(t) cov(bk kfc roys) kernel id(id) report
```

With the following output table:

KERNEL PROPENSITY SCORE DIFFERENCE-IN-DIFFERENCES

Report - Propensity score estimation:

```
Iteration 0: log likelihood = -198.21978
Iteration 1: log likelihood = -196.7657
Iteration 2: log likelihood = -196.7636
```

Probit regression	Number of obs	=	404
	LR chi2(3)	=	2.91
	Prob > chi2	=	0.4053
Log likelihood = -196.7636	Pseudo R2	=	0.0073

treated	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
bk	.1812529	.2090916	0.87	0.386	-.2285591 .5910649
kfc	.3888298	.246799	1.58	0.115	-.0948873 .8725469
roys	.2997977	.2318227	1.29	0.196	-.1545664 .7541618
_cons	.6476036	.1777446	3.64	0.000	.2992305 .9959767

Number of observations in the DIFF-IN-DIFF: 800
 Baseline Follow-up


```
Control: 78      76      154
Treated: 326    320     646
          404     396
```

R-square: 0.02819

DIFFERENCE IN DIFFERENCES ESTIMATION							
Outcome Variable	BASE LINE			FOLLOW UP			
	Control	Treated	Diff(BL)	Control	Treated	Diff(FU)	DIFF-IN-DIFF
fte	21.656	17.065	-4.591	18.914	17.573	-1.341	3.250
Std. Error	0.572	1.093	1.234	0.576	1.103	1.245	1.752
t	37.88	17.46	-3.72	16.89	17.27	-1.98	1.85
P> t	0.000	0.000	0.000***	0.000	0.000	0.282	0.064*

* Means and Standard Errors are estimated by linear regression

Inference: * p<0.01; ** p<0.05; * p<0.1

3.4 Quantile DID

The Quantile DID is obtained when specifying the option `qdid(quantile)`. For example, estimating the treatment effects on the median requires the following syntax:

```
diff fte, t(treated) p(t) qdid(0.50)
```

It may be combined with covariates:

```
diff fte, t(treated) p(t) qdid(0.50) cov(bk kfc roys)
```

With the following output:

QUANTILE DIFFERENCE-IN-DIFFERENCES WITH COVARIATES

Number of observations in the DIFF-IN-DIFF: 801

```
Baseline      Follow-up
Control: 78    77      155
Treated: 326  320     646
          404     397
```

R-square: 0.14861

DIFFERENCE IN DIFFERENCES ESTIMATION							
Outcome Variable	BASE LINE			FOLLOW UP			
	Control	treated	Diff(BL)	Control	treated	Diff(FU)	DIFF-IN-DIFF
fte	17.750	17.250	-0.500	17.750	17.750	-0.000	0.500
Std. Error	1.124	0.835	1.013	1.132	0.840	1.007	1.426
t	15.79	17.15	-0.49	17.75	17.85	-0.00	0.35
P> t	0.000	0.000	0.622	0.000	0.000	1.000	0.726

* Values are estimated at the .5 quantile

Inference: * p<0.01; ** p<0.05; * p<0.1

Quantile DID is combinable with the option `kernel`:

```
diff fte, t(treated) p(t) qdid(0.50) cov(bk kfc roys) kernel id(id)
report
```

KERNEL PROPENSITY SCORE QUANTILE DIFFERENCE-IN-DIFFERENCES

Report - Propensity score estimation:

```
Iteration 0: log likelihood = -198.21978
Iteration 1: log likelihood = -196.7657
Iteration 2: log likelihood = -196.7636
```

```
Probit regression                        Number of obs =      404
                                         LR chi2(3)      =       2.91
                                         Prob > chi2     =     0.4053
Log likelihood = -196.7636                Pseudo R2       =     0.0073
```

treated	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
bk	.1812529	.2090916	0.87	0.386	-.2285591	.5910649
kfc	.3888298	.246799	1.58	0.115	-.0948873	.8725469
roys	.2997977	.2318227	1.29	0.196	-.1545664	.7541618
_cons	.6476036	.1777446	3.64	0.000	.2992305	.9959767

Number of observations in the DIFF-IN-DIFF: 800

```
          Baseline      Follow-up
Control:  78            76      154
Treated: 326           320      646
         404            396
```

R-square: 0.00477

DIFFERENCE IN DIFFERENCES ESTIMATION

Outcome Variable	BASE LINE			FOLLOW UP			DIFF-IN-DIFF
	Control	Treated	Diff(BL)	Control	Treated	Diff(FU)	
fte	18.500	16.000	-2.500	18.500	17.500	-1.000	1.500
Std. Error	1.578	0.732	1.739	1.614	0.727	1.770	2.482
t	11.72	15.08	-1.44	18.50	18.06	-1.65	0.60
P> t	0.000	0.000	0.151	0.000	0.000	0.572	0.546

* Values are estimated at the .5 quantile
Inference: * p<0.01; ** p<0.05; * p<0.1

3.5 Balancing test

The balancing test is obtained only on the baseline. The syntax is similar to the one presented before, except for the supply of option `test`.

```
diff fte, t(treated) p(t) cov(bk kfc roys wendys) test
```

TWO-SAMPLE T TEST

```
Number of observations (baseline): 404
          Baseline      Follow-up
Control:  78            -          78
Treated: 326           -         326
         404            -
```

t-test at period = 0:

Variable(s)	Mean Control	Mean Treated	Diff.	t	Pr(T > t)
fte	19.949	17.065	-2.884	2.44	0.0150**
bk	0.443	0.411	-0.032	0.52	0.6035
kfc	0.152	0.205	0.054	1.08	0.2818
roys	0.215	0.248	0.033	0.61	0.5448

```
wendys          | 0.190          | 0.136          | -0.054         | 1.22          | 0.2241
```

```
-----
*** p<0.01; ** p<0.05; * p<0.1
```

When combined with option `kernel`, the covariates are weighted and the differences obtained by linear regression:

```
diff fte, t(treated) p(t) cov(bk kfc roys wendys) test id(id) kernel
```

```
TWO-SAMPLE T TEST
```

```
Number of observations (baseline): 404
```

	Baseline	Follow-up
Control:	78	78
Treated:	326	326
	404	-

```
t-test at period = 0:
```

Weighted Variable(s)	Mean Control	Mean Treated	Diff.	t	Pr(T > t)
fte	21.656	17.065	-4.591	3.22	0.0014***
bk	0.618	0.408	-0.210	3.55	0.0004***
kfc	0.104	0.209	0.104	2.60	0.0097***
roys	0.183	0.252	0.068	1.42	0.1570
wendys	0.095	0.132	0.037	1.01	0.3123

```
-----
*** p<0.01; ** p<0.05; * p<0.1
```

```
Attention: option kernel weighs variables in cov(varlist)
```

```
Means and t-test are estimated by linear regression
```

4. Saved results

`diff` saves in the memory each number of the output table as return-type scalars:

- `r(mean_c0)`: mean of *output_var* of the control group in period == 0.
- `r(mean_t0)`: mean of *output_var* of the treated group in period == 0.
- `r(diff0)`: difference of the mean of *output_var* between treated and control groups in period t=0.
- `r(mean_c1)`: mean of *output_var* of the control group in period == 1.
- `r(mean_t1)`: mean of *output_var* of the treated group in period == 1.
- `r(diff1)`: difference of the mean of *output_var* between treated and control groups in period == 1.
- `r(diffdiff)`: DID - Treatment Effect.
- `r(se_c0)`: Standard Error of the mean of *output_var* of the control group in period == 0.
- `r(se_t0)`: Standard Error of the mean of *output_var* of the treated group in period ==0.
- `r(se_d0)`: Standard Error of the difference of *output_var* between the treated and control groups in period ==0.

- `r(se_c1)`: Standard Error of the mean of `output_var` of the control group in `period == 1`.
- `r(se_t1)`: Standard Error of the mean of `output_var` of the treated group in `period == 1`.
- `r(se_d1)`: Standard Error of the difference of `output_var` between the treated and control groups in `period == 0`.
- `r(se_dd)`: Standard Error of the difference in difference.

5. Acknowledgements

I thank David Card for his useful suggestions on DID and Kit Baum for his valuable comments on the Stata's ado file. I also thank the comments from the attendees at the 2012 UK Stata Users Meeting Group in London, UK.

6. References

- Angrist, J.D., Pischke, J.-S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.
- Card, D., Krueger, A.B., 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review* 84, 772–93.
- Heckman, J.J., Ichimura, H., Todd, P., 1998. Matching as an Econometric Evaluation Estimator. *The Review of Economic Studies* 65, 261–294.
- Heckman, J.J., Ichimura, H., Todd, P.E., 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies* 64, 605–654.
- Meyer, B.D., Viscusi, W.K., Durbin, D.L., 1995. Workers' Compensation and Injury Duration: Evidence from a Natural Experiment. *The American Economic Review* 85, 322–340.