

esttab/estout **csv tables demo**

Luiza Andrade
World Bank, Development Impact Evaluations
lcardoso@worldbank.org

Contents

1 Set up	2
2 Creating simple tables.....	3
3 Table with two panels.....	8

List of tables

1 - Simple esttab table	3
2 - Table with labels and standards errors	4
3- Removing omitted variables.....	5
4- Selecting variables to display.....	6
5- Customizing model titles.....	7
6- Table with two panels.....	9

Visit <https://github.com/bbdaniels/stata-tables> to see the latest version of this material.

1 Set up

We'll start by loading a built-in data set and running some regressions to add to a table. These are mock regressions, so we suggest you don't spend a lot of time trying to interpret them.

```
// Install user-written commands *****

ssc install estout

// Load the data *****

local output "..\outputs\Raw"      // Replace ".." with your folder path

sysuse census.dta, clear
xtset region

// Run regressions *****

// Regression 1: nothing interesting
reg death marriage pop

est sto reg1
estadd local region "No"

// Regression 2: a different regression
reg death popurban

est sto reg2
estadd local region "No"

// Regression 3: indicator expansion
reg divorce marriage pop

est sto reg3
estadd local region "No"

// Regression 4: categorical control
reg divorce marriage pop i.region

est sto reg4
estadd local region      "Yes"

// South region only
reg death marriage if region == 3
est sto s1

reg death marriage pop if region == 3
est sto s2

// West region only
reg death marriage if region == 4
est sto w1

reg death marriage pop if region == 4
est sto w2
```

2 Creating simple tables

These regressions will be added to the same table. The table below is the simplest one that esttab can create.

```
// Export tables *****

local regressions reg1 reg2 reg3 reg4

*-----
* The simplest esttab tables will not compile for these regressions. This is
* due to the occurrence of the special character # on categorical variables
*-----

esttab `regressions' using "`output'/t1_esttab_basic.csv", ///
    replace
```

	A	B	C	D	E
1					
2		(1)	(2)	(3)	(4)
3		death	death	divorce	divorce
4					
5	marriage	-0.0849		0.260***	0.190**
6		(-1.63)		(4.36)	(3.20)
7					
8	pop	0.00949***		0.00271***	0.00342***
9		(19.08)		(4.74)	(5.98)
10					
11	popurban		0.00992***		
12			(28.57)		
13					
14	1.region				0
15					(.)
16					
17	2.region				6362.7*
18					(2.06)
19					
20	3.region				9215.6**
21					(3.01)
22					
23	4.region				9627.3**
24					(3.03)
25					
26	_cons	650.9	6464.6***	-963.3	-7786.2**
27		(0.48)	(3.55)	(-0.62)	(-3.00)
28					
29	N	50	50	50	50
30					
31	t statistics in parentheses				
32	=** p<0.05	** p<0.01	*** p<0.001"		

1 - Simple esttab table

The previous table can certainly look better. Let's start by adding variable labels (using the `label` option), and replacing t-statistics with standard errors (using the `se` option).

```
*-----
* The issue above can be solved by displaying variable labels instead of
* variables names with the 'label' option
*-----

esttab `regressions' using "`output'/t2_esttab_label.csv", ///
    label ///    Add variable labels
    se    /// Display standard errors instead of t-statistics
    replace
```

	A	B	C	D	E
1					
2		(1)	(2)	(3)	(4)
3		Number of deaths	Number of deaths	Number of divorces	Number of divorces
4					
5	Number of marriages	-0.0849		0.260***	0.190**
6		(0.0520)		(0.0597)	(0.0593)
7					
8	Population	0.00949***		0.00271***	0.00342***
9		(0.000497)		(0.000571)	(0.000572)
10					
11	Urban population		0.00992***		
12			(0.000347)		
13					
14	NE				0
15					(.)
16					
17	N Cntrl				6362.7*
18					(3085.9)
19					
20	South				9215.6**
21					(3061.1)
22					
23	West				9627.3**
24					(3180.5)
25					
26	Constant	650.9	6464.6***	-963.3	-7786.2**
27		(1347.4)	(1819.7)	(1547.9)	(2591.7)
28					
29	Observations	50	50	50	50
30					
31	Standard errors in parentheses				
32	=*	**	***	p<0.001"	

2 - Table with labels and standards errors

Since we have a categorical variable and a constant, one of the categories is dropped. The code below removes the base level coefficient through the `nobaselevel`. Option `noomit` would do the same for other omitted variables. The base category is then indicated through the `refcat` option, that adds the text specified inside the parentheses above the indicated coefficient (`2.region`, in this case). The code also adds two custom notes.

```
*-----
* Remove omitted variables and levels from regression
*-----
esttab `regressions' using "`output'/t3_esttab_omitted.csv", ///
    se /// Display standard errors instead of t-statistics
    noomit /// Remove omitted variables
    nobaselevels /// Remove categorical variables base levels
    /// Add note indicating the omitted region -- refcat() can be used more
    /// broadly to add lines before a variable
    refcat(_cons "Omitted category: NE region", nolabel) ///
    addnotes("Add a note here." "Other custom note here.") /// Custom footnotes
    label /// Show variable and value labels instead of names
    replace
```

	A	B	C	D	E
1					
2		(1)	(2)	(3)	(4)
3		Number of deaths	Number of deaths	Number of divorces	Number of divorces
4					
5	Number of marriages	-0.0849		0.260***	0.190**
6		(0.0520)		(0.0597)	(0.0593)
7					
8	Population	0.00949***		0.00271***	0.00342***
9		(0.000497)		(0.000571)	(0.000572)
10					
11	Urban population		0.00992***		
12			(0.000347)		
13					
14	N Cntrl				6362.7*
15					(3085.9)
16					
17	South				9215.6**
18					(3061.1)
19					
20	West				9627.3**
21					(3180.5)
22					
23	Omitted category: NE region				
24					
25	Constant	650.9	6464.6***	-963.3	-7786.2**
26		(1347.4)	(1819.7)	(1547.9)	(2591.7)
27					
28	Observations	50	50	50	50
29					
30	Standard errors in parentheses				
31	Add a note here.				
32	Other custom note here.				
33	=** p<0.05	** p<0.01	*** p<0.001"		

3- Removing omitted variables

You may also prefer to not show the control coefficients in the table at all, and just add a line in the footer indicating that they were used. The example below removes all region variable coefficients using the `drop()` option. The options `scalar` adds the footnote line. Remember we added a region scalar to each model when the regressions results were stored. This option uses the scalar name followed by its label in the table.

```
*-----
* The option 'drop' allows you to remove variables from the final table.
* Use variable names to list all variables to be removed. If you want to
* remove specific categories within categorical variables, type
* 'mat list r(table)' to see how Stata refers to each category (on the
* column names)
* Option 'keep' would do the opposite, keeping only listed variables.
* Option 'order' allows you to specify the order of the rows.
*-----

esttab `regressions' using "`output'/t4_esttab_scalar.csv", ///
    se /// Display standard errors instead of t-statistics
    drop(*.region*) /// Remove fixed effects estimates from table
    scalars("region Region controls") /// Add row with FEs used instead
    label /// Show variable and value labels instead of names
    replace
```

	A	B	C	D	E
1					
2		(1)	(2)	(3)	(4)
3		Number of deaths	Number of deaths	Number of divorces	Number of divorces
4					
5	Number of marriages	-0.0849		0.260***	0.190**
6		(0.0520)		(0.0597)	(0.0593)
7					
8	Population	0.00949***		0.00271***	0.00342***
9		(0.000497)		(0.000571)	(0.000572)
10					
11	Urban population		0.00992***		
12			(0.000347)		
13					
14	Constant	650.9	6464.6***	-963.3	-7786.2**
15		(1347.4)	(1819.7)	(1547.9)	(2591.7)
16					
17	Observations	50	50	50	50
18	Region controls	No	No	No	Yes
19					
20	Standard errors in parentheses				
21	Add a note here.				
22	Other custom note here.				
23	** p<0.05	** p<0.01	*** p<0.001"		

4- Selecting variables to display

Now let's add model titles. Custom column titles are added `mtitles`.

```
*-----
* Add custom model titles
*-----
esttab `regressions' using "`output'/t5_esttab_titles.csv", ///
    mtitles("Title 1" "Title 2" "Title 3" "Title 4") /// List titles here
    se /// Display standard errors instead of t-statistics
    drop(*,region*) /// Remove fixed effects estimates from table
    scalars("region Region controls") /// Add row indicating FE used instead
    label /// Show variable and value labels instead of names
    replace
```

	A	B	C	D	E
1					
2		(1)	(2)	(3)	(4)
3		Title 1	Title 2	Title 3	Title 4
4					
5	Number of marriages	-0.0849		0.260***	0.190**
6		(0.0520)		(0.0597)	(0.0593)
7					
8	Population	0.00949***		0.00271***	0.00342***
9		(0.000497)		(0.000571)	(0.000572)
10					
11	Urban population		0.00992***		
12			(0.000347)		
13					
14	Constant	650.9	6464.6***	-963.3	-7786.2**
15		(1347.4)	(1819.7)	(1547.9)	(2591.7)
16					
17	Observations	50	50	50	50
18	Region controls	No	No	No	Yes
19					
20	Standard errors in parentheses				
21	Add a note here.				
22	Other custom note here.				
23	=** p<0.05	** p<0.01	*** p<0.001"		

5- Customizing model titles

3 Table with two panels

The `estout` and `esttab` packages can also be used to make more complex tables, appending results from different regressions into separate panels. To do this, all panels but the top one must use `append` instead of `replace`. Finally, in addition, the `posthead()` option lets you add your panel titles.

```
*-----
* Create a two-panel table. This is done using the option 'prehead' to add
* a line before the first regression variable, and hardcoding the LaTeX
* text and formatting. Both panels are created as 'fragment', which means
* Stata won't add the lines to begin and end the tabular object
*-----

* Top panel
esttab s1 s2 using "`output'/t7_esttab_panel.csv", ///
    posthead("Panel A: South")  /// This is the panel header
    nomtitles nonotes          /// Group title replaces model title
    label                      ///
    replace

* Bottom panel
esttab w1 w2 using "`output'/t7_esttab_panel.csv", ///
    posthead("Panel B: West")  /// Panel header
    append /// Appending bottom panel to top panel instead of replacing
    nomtitles nonumbers ///
    label
```


	A	B	C
1			
2		(1)	(2)
3	Panel A: South		
4	Number of marriages	0.622***	-0.203
5		(9.27)	(-1.12)
6			
7	Population		0.0107***
8			(4.68)
9			
10	Constant	6626.2	1783.3
11		(1.46)	(0.58)
12			
13	Observations	16	16
14			
15			
16	Panel B: West		
17	Number of marriages	0.737***	0.00548
18		(5.98)	(0.35)
19			
20	Population		0.00788***
21			(52.78)
22			
23	Constant	-5555.9	-1297.2*
24		(-0.64)	(-2.37)
25			
26	Observations	13	13
27			
28	t statistics in parentheses		
29	=** p<0.05	** p<0.01	*** p<0.001"

6- Table with two panels