`esttab`/`estout` LaTeX tables demo

Luiza Andrade

World Bank, Development Impact Evaluations
lcardoso@worldbank.org

May 28, 2020

Contents

1 Set up **2**

2 Creating simple tables **3**

3 Table with two panels **8**

List of Tables

1 Basic `esttab` table 3
2 Adding variable labels 4
3 Removing omitted levels 5
4 Removing variables from table 6
5 Customizing model titles and notes 7
6 Customizing header 8
7 Table with two panels 10

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. The regressions themselves are not very important.

```
.
.
. // Load some data ******************************
.
.         sysuse census.dta, clear
(1980 Census data by state)

.         xtset region
       panel variable:  region (unbalanced)

.
. // Run some regressions ************************
.
.         // Regression 1: nothing interesting
.         qui reg death marriage pop

.
.         est sto reg1

.         qui estadd local region "No"

.
.         // Regression 2: a different regression
.         qui reg death popurban

.
.         est sto reg2

.         qui estadd local region "No"

.
.         // Regression 3: indicator expansion
.         qui reg divorce marriage pop

.
.         est sto reg3

.         qui estadd local region "No"

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

.
.         est sto reg4

.         qui estadd local region         "Yes"

.
.         local regressions reg1 reg2 reg3 reg4
```

.

2 Creating simple tables

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

```
.
. esttab `regressions´ using "${tex}/Raw/t1_esttab_basic.tex", ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t1_esttab_basic.tex)
.
```

Table 1: Basic `esttab` table

	(1) death	(2) death	(3) divorce	(4) divorce
marriage	-0.0849		0.260***	0.190**
	(-1.63)		(4.36)	(3.20)
pop	0.00949***		0.00271***	0.00342***
	(19.08)		(4.74)	(5.98)
popurban		0.00992***		
		(28.57)		
1.region				0
				(.)
2.region				6362.7*
				(2.06)
3.region				9215.6**
				(3.01)
4.region				9627.3**
				(3.03)
_cons	650.9	6464.6***	-963.3	-7786.2**
	(0.48)	(3.55)	(-0.62)	(-3.00)
N	50	50	50	50

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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).

```
.
. esttab `regressions´ using "${tex}/Raw/t2_esttab_label.tex", ///
>         label  ///
>         se      ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t2_esttab_label.tex)
.
```

Table 2: Adding variable labels

	(1) Number of deaths	(2) Number of deaths	(3) Number of divorces	(4) Number of divorces
Number of marriages	-0.0849		0.260***	0.190**
	(0.0520)		(0.0597)	(0.0593)
Population	0.00949***		0.00271***	0.00342***
	(0.000497)		(0.000571)	(0.000572)
Urban population		0.00992***		
		(0.000347)		
NE				0
				(.)
N Cntrl				6362.7*
				(3085.9)
South				9215.6**
				(3061.1)
West				9627.3**
				(3180.5)
Constant	650.9	6464.6***	-963.3	-7786.2**
	(1347.4)	(1819.7)	(1547.9)	(2591.7)
Observations	50	50	50	50

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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 add the text specified inside the parentheses above the indicated coefficient `2.region`, in this case.

The next table also shows confidence intervals instead of standard errors (option `ci` does that), and adds two custom notes.

```
.
. esttab `regressions´ using "${tex}/Raw/t3_esttab_omitted.tex", ///
>         ci ///
>         nobaselevels ///
>         refcat(2.region "Omitted category: NE region", nolabel) ///
>         addnotes("Add a note here." "Other custom note here.") ///
>         label ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t3_esttab_omitted.tex)

.
```

Table 3: Removing omitted levels

	(1) Number of deaths	(2) Number of deaths	(3) Number of divorces	(4) Number of divorces
Number of marriages	-0.0849		0.260***	0.190**
	[-0.189,0.0196]		[0.140,0.380]	[0.0702,0.309]
Population	0.00949***		0.00271***	0.00342***
	[0.00849,0.0105]		[0.00156,0.00386]	[0.00226,0.00457]
Urban population		0.00992***		
		[0.00922,0.0106]		
Omitted category: NE region				
N Cntrl				6362.7*
				[143.6,12581.9]
South				9215.6**
				[3046.4,15384.8]
West				9627.3**
				[3217.5,16037.1]
Constant	650.9	6464.6***	-963.3	-7786.2**
	[-2059.8,3361.6]	[2805.9,10123.4]	[-4077.2,2150.5]	[-13009.4,-2562.9]
Observations	50	50	50	50

95% confidence intervals in brackets
Add a note here.
Other custom note here.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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 by throught the `drop()`[1] 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.

```
.
```

[1] `drop()` takes coefficient names as input. Because region is a categorical variable, the coefficient for each level is stored as its code plus the variable name, separated by a dot: `1.region`, `2.region`, etc.

```
. esttab `regressions´ using "${tex}/Raw/t4_esttab_scalar.tex", ///
>         ci ///
>         drop(*.region) ///
>         scalars("region Region controls") ///
>         addnotes("Add a note here." "Other custom note here.")  ///
>         label ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t4_esttab_scalar.tex)

.
```

Table 4: Removing variables from table

	(1) Number of deaths	(2) Number of deaths	(3) Number of divorces	(4) Number of divorces
Number of marriages	-0.0849		0.260***	0.190**
	[-0.189,0.0196]		[0.140,0.380]	[0.0702,0.309]
Population	0.00949***		0.00271***	0.00342***
	[0.00849,0.0105]		[0.00156,0.00386]	[0.00226,0.00457]
Urban population		0.00992***		
		[0.00922,0.0106]		
Constant	650.9	6464.6***	-963.3	-7786.2**
	[-2059.8,3361.6]	[2805.9,10123.4]	[-4077.2,2150.5]	[-13009.4,-2562.9]
Observations	50	50	50	50
Region controls	No	No	No	Yes

95% confidence intervals in brackets
Add a note here.
Other custom note here.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Now let's add model titles. The next two tables do that. The first one adds individual titles to each model using `mtitles`, while the second groups columns through `mgroups`.

`mgroups` first two inputs are the group names. `pattern` defines which columns are part of which group (1 marks the beginning of a new group). The next line centralize group titles to include both groups, and the last line inside `mgroups` adds a line under group names.

```
.
. esttab `regressions´ using "${tex}/Raw/t5_esttab_titles.tex", ///
>         mtitles("Title 1" "Title 2" "Title 3" "Title 4") ///
>         se ///
>         drop(*.region*) ///
>         scalars("region Region controls") ///
>         addnotes("Add a note here." "Other custom note here.") ///
>         label ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t5_esttab_titles.tex)
```

.

Table 5: Customizing model titles and notes

	(1) Title 1	(2) Title 2	(3) Title 3	(4) Title 4
Number of marriages	-0.0849		0.260***	0.190**
	(0.0520)		(0.0597)	(0.0593)
Population	0.00949***		0.00271***	0.00342***
	(0.000497)		(0.000571)	(0.000572)
Urban population		0.00992***		
		(0.000347)		
Constant	650.9	6464.6***	-963.3	-7786.2**
	(1347.4)	(1819.7)	(1547.9)	(2591.7)
Observations	50	50	50	50
Region controls	No	No	No	Yes

Standard errors in parentheses

Add a note here.

Other custom note here.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

```
.
. esttab `regressions´ using "${tex}/Raw/t6_esttab_header.tex", ///
>         nomtitles ///
>         drop(*.region*) ///
>         scalars("region Region fixed effects") ///
>         se ///
>         label ///
>         mgroups("Number of deaths" "Number of divorces", ///
>                         pattern(1 0 1 0) ///
>                         span prefix(\multicolumn{@span}{c}{) suffix(}) ///
>                         erepeat(\cmidrule(lr){@span})) ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t6_esttab_header.tex)
.
```

Table 6: Customizing header

	Number of deaths		Number of divorces	
	(1)	(2)	(3)	(4)
Number of marriages	-0.0849		0.260***	0.190**
	(0.0520)		(0.0597)	(0.0593)
Population	0.00949***		0.00271***	0.00342***
	(0.000497)		(0.000571)	(0.000572)
Urban population		0.00992***		
		(0.000347)		
Constant	650.9	6464.6***	-963.3	-7786.2**
	(1347.4)	(1819.7)	(1547.9)	(2591.7)
Observations	50	50	50	50
Region fixed effects	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3 Table with two panels

The `estout` and `esttab` packages can also be used to make more complex tables. However, this requires hardcoding some LaTeX into your Stata do-file.

The example below creates a table with two panels. `fragment` is used to avoid adding the whole tabular environment into LaTeX. This means that it must be added manually. That is what `prehead()` on the top panel and `postfoot()` in the bottom panel do.

The advantage of using the `fragment` option is that it allows you to append multiple tables, therefore creating the different panels. To do this, all panels but the top one must use `append` instead of `replace`. Finally, `posthead()` lets you add your panel titles.

```
.
. // South region only
. qui reg death marriage          if region == 3

. est sto s1

.
. qui reg death marriage pop      if region == 3

. est sto s2
```

```
.
. // West region only
. qui reg death marriage          if region == 4

. est sto w1

.
. qui reg death marriage pop      if region == 4

. est sto w2

.
. * Top panel
. esttab s1 s2 using "${tex}/Raw/t7_esttab_panel.tex", ///
>         prehead("\begin{tabular}{l*{2}{c}} \hline\hline") ///
>         posthead("\hline \\ \multicolumn{3}{c}{\textbf{Panel A: South}} \\\\[-1ex]") ///
>         fragment ///
>         mgroups("Number of deaths", ///
>                         pattern(1 0) ///
>                         span prefix(\multicolumn{@span}{c}{) suffix(})) ///
>         nomtitles ///
>         label ///
>         replace
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t7_esttab_panel.tex)

.
. * Bottom panel
. esttab w1 w2 using "${tex}/Raw/t7_esttab_panel.tex", ///
>         posthead("\hline \\ \multicolumn{3}{c}{\textbf{Panel B: West}} \\\\[-1ex]") ///
>         fragment ///
>         append ///
>         nomtitles nonumbers nolines ///
>         prefoot("\hline") ///
>         postfoot("\hline\hline \end{tabular}") ///
>         label
(output written to C:/Users/wb501238/Documents/GitHub/stata-tables/do/tex/Raw/t7_esttab_panel.tex)

.
```

Table 7: Table with two panels

	Number of deaths	
	(1)	(2)
Panel A: South		
Number of marriages	0.622***	-0.203
	(9.27)	(-1.12)
Population		0.0107***
		(4.68)
Constant	6626.2	1783.3
	(1.46)	(0.58)
Observations	16	16
Panel B: West		
Number of marriages	0.737***	0.00548
	(5.98)	(0.35)
Population		0.00788***
		(52.78)
Constant	-5555.9	-1297.2*
	(-0.64)	(-2.37)
Observations	13	13