Making regression tables simplified

Two new commands: esto and esta

Ben Jann, ETH Zürich, jann@soz.gess.ethz.ch

5th German Stata Users Group Meeting
April 2, 2007, Essen

Outline
• Introduction
• esto: Storing estimates simplified
• esta: Tabulating estimates simplified
• Use with Word, Excel, LaTeX, etc.
• Discussion

Introduction

• Output from estimation commands contains all sorts of details and we often only want to display selected and concisely arranged results in form of "regression tables".

Regression tables are useful
• to get an overview of results in daily research work,
• for results presentation in reports and publications.

• Stata features a command to produce regression tables called estimates table.

However, estimates table
• is only intended for displaying the models in the Results window or the log,
• is often not flexible enough to produce customized tables.
Introduction

- Since copy/paste by hand is not an option, various user programs for automatic results processing have been written.

One of these commands is estout (Jann 2005, SJ 5-3). It compiles regression tables containing results from one or more estimation commands for use in, say, LaTeX documents, spreadsheet programs, or word processors.

- Some alternatives:
  - outreg (John Luke Gallup)
  - outreg2 (Roy Wada)
  - xml_tab (Michael Lokshin and Zurab Sajaia)
  - outtex (Antoine Terracoli)
  - est2tex (Marc Muenelder)
  - mktab (Nicholas Winter)
  - parmest (Roger Newson)

Although estout turned out to be quite powerful, the motivational orientation towards functionality rather than ease-of-use brought with it some limitations that may make the use of estout somewhat clumsy in daily work.

- estout tables are usually not suitable for display in Stata’s results window.

- estout’s syntax is not as intuitive and user-friendly as it could be.

- The amount of typing required to compile even a simple table can be quite considerable. (The basic approach of estout is to provide a “clean desk” from which a fully-fledged end-product can be built up. Although, options may be pre-specified via so-called defaults files, this does not appear beneficial unless working on large reports with lots of similar tables.)
Introduction

An additional issue with estout is that the estimation sets have to be stored using official Stata's estimates store before they can be tabulated.

Drawbacks of estimates store are:

- The user is required to specify names under which to store the estimation sets. This can be distracting.

- The stored estimates consume a considerable amount of memory. (In order to preserve functionality of postestimation commands, an estimation sample indicator variable is stored for each estimation set. These indicators may blow up the dataset if it contains a large number of observations or if many estimation sets are stored. Additionally, storing the estimation samples has the side effect of slowing down cycling through the stored sets, which also slows down tabulation programs such as estout or official Stata's estimates table.)

Introduction

To summarize, there seems to be a need for

(1) an easy-to-use version of estout,

(2) a simplified procedure to hold on to estimates for tabulation.

In the remainder of this talk I will address these two points (in reverse order) and present

- a command called esto to overcome the limitations of estimates store,

- a user-friendly estout wrapper called esta.
**esto: Storing estimates simplified**

**Syntax:**

``` stata
esto [ name ] [ , options ] [: command ]
esto drop {#|name} [...] 
esto clear
```

<table>
<thead>
<tr>
<th>options</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>no</code>esample</td>
<td>do not/do store e(sample)</td>
</tr>
<tr>
<td><code>title(string)</code></td>
<td>specify a title for the stored set</td>
</tr>
<tr>
<td><code>addscalars(...)</code></td>
<td>add scalar statistics</td>
</tr>
<tr>
<td><code>refresh[()]</code></td>
<td>overwrite a previously stored set</td>
</tr>
<tr>
<td><code>nocopy</code></td>
<td>clear e(0) after storing the set</td>
</tr>
</tbody>
</table>

Usage is analogous to official Stata's `estimates store`. However, `name` is optional and the `e(sample)` may be dropped.

`_esto` is short for `esto`, `noesample`

**Basic example:**

``` stata
sysuse auto, clear
regress price weight mpg
esto
regress price weight mpg foreign
esto
estout, style(fixed)
<run>
macro dir
<run>
esto clear
<run>
```
```
price | Coef.  | Std. Err. | t     | P>|t| | [95% Conf.]
------|--------|-----------|-------|-----|----------------
weight| 3.464706| .630749   | 5.49  | 0.000| 2.206717
mpg   | 21.8536 | 74.22114  | 0.29  | 0.769| -126.1758
foreign| 3673.06 | 683.9783  | 5.37  | 0.000| 2308.909
_cons | -5853.696| 3376.987  | -1.73 | 0.087| -12588.88

estout, style(fixed)

est1 est2
b b
weight 1.746559 3.464706
mpg -49.51222 21.8536
foreign 3673.06
_cons 1946.069 -5853.696

end of do-file
```

```
end of do-file
```
esto: storing estimates simplified

Use `esto` as a prefix command:

```
esto: regress price weight mpg
est: regress price weight mpg foreign
estout, style(fixed)
<run>
```

Drop the `e(sample)`:

```
esto, noesample: reg price weight mpg
.est: regress price weight mpg
estimates dir
describe _est*
<run>
```

Add additional results while storing:

```
esto clear
regress price weight mpg
test weight = mpg
est, add(p_diff r(p))
estout, style(fixed) stats(p_diff)
<run>
```

(est4 stored)

```
estimates dir
```

<table>
<thead>
<tr>
<th>model</th>
<th>command</th>
<th>depvar</th>
<th>npar</th>
<th>title</th>
</tr>
</thead>
<tbody>
<tr>
<td>est1</td>
<td>regress</td>
<td>price</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>est2</td>
<td>regress</td>
<td>price</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>est3</td>
<td>regress</td>
<td>price</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>est4</td>
<td>regress</td>
<td>price</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

```
describe _est*
```

<table>
<thead>
<tr>
<th>_est_est1</th>
<th>byte</th>
<th>%8.0g</th>
<th>esample() from estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>_est_est2</td>
<td>byte</td>
<td>%8.0g</td>
<td>esample() from estimates</td>
</tr>
</tbody>
</table>

end of do-file

---

```
test weight = mpg

( 1) weight - mpg = 0

F( 1, 71) = 0.36
    Prob > F = 0.5514

esto, add(p_diff r(p))
(e(p_diff) = .55138216 added)
(estl stored)

    estout, style(fixed) stats(p_diff)

        estl
    b
weight      1.746559
mpg       -49.51222
    _cons    1946.069
    p_diff  .5513822

end of do-file

esta: Tabulating estimates simplified

The new esta command is a wrapper for estout. Some features:

• simplified syntax

• default: publication-style table that displays nicely in Stata’s results window

• provides full estout functionality since all estout options are allowed

• extends functionality (e.g. Word RTF and Excel CSV output modes, improved LaTeX support)

Basic syntax:

esta [ namelist ] [ using filename ] [ , opts estout_opts ]

If namelist is omitted, esta tabulates the estimation sets stored by esto.
esta: Tabulating estimates simplified

Overview of options:

b([fmt]) beta[([fmt])] main(name [fmt]) t(fmt) abs not
se([fmt]) p([fmt]) ci([fmt]) aux(name [fmt]) [no]constant
[no]start([list]) staraux
r2([fmt]) ar2([fmt]) pr2([fmt]) aic([fmt]) bic([fmt])
scalars([list]) sfmt(fmt ...) noobs obslast
wide [no]parentheses brackets [no]gaps [no]lines compress plain
label title(string) mtitles([list]) nomtitles [no]depvars
[no]numbers coeflabels([list]) [no]notes addnotes([list])
smcl | fixed | tab | csv | scsv | rtf | html | tex | booktabs
fragment page([packages]) alignment(string) width(string)
replace append type noisily
drop([list]) keep([list]) order([list]) equations([list])
eform margin unstack other_estout_options

esta: Tabulating estimates simplified

Basic usage:

esta clear
sysuse auto, clear
esta: regress price weight mpg
esta: regress price weight mpg foreign
esta
<run>

Display standard errors and add summary statistics:

esta, se ar2 nostar
<run>

Display beta coefficients:

esta, beta not
<run>
### esta

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>weight</td>
<td>1.747**</td>
<td>3.465***</td>
</tr>
<tr>
<td></td>
<td>(2.72)</td>
<td>(5.49)</td>
</tr>
<tr>
<td>mpg</td>
<td>-49.51</td>
<td>21.85</td>
</tr>
<tr>
<td></td>
<td>(-0.57)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>foreign</td>
<td>3673.1***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.37)</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>1946.1</td>
<td>-5853.7</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(-1.73)</td>
</tr>
</tbody>
</table>

N = 74

* t statistics in parentheses
  * p<0.05, ** p<0.01, *** p<0.001

### esta, se ar2 nostar

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>weight</td>
<td>1.747</td>
<td>3.465</td>
</tr>
<tr>
<td></td>
<td>(0.641)</td>
<td>(0.631)</td>
</tr>
<tr>
<td>mpg</td>
<td>-49.51</td>
<td>21.85</td>
</tr>
<tr>
<td></td>
<td>(86.16)</td>
<td>(74.22)</td>
</tr>
<tr>
<td>foreign</td>
<td>3673.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(684.0)</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>1946.1</td>
<td>-5853.7</td>
</tr>
<tr>
<td></td>
<td>(3597.0)</td>
<td>(3377.0)</td>
</tr>
</tbody>
</table>

N = 74

adj. R-sq = 0.273 0.478

Standard errors in parentheses
end of do-file

.do_example do/9.ihlp

.   esta, beta not

(1)                      (2)
<table>
<thead>
<tr>
<th>price</th>
<th>price</th>
</tr>
</thead>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>weight          0.460**        0.913***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mpg            -0.097          0.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td>foreign</td>
<td>0.573***</td>
<td></td>
</tr>
</tbody>
</table>

N   74    74

standardized beta coefficients
* p<0.05, ** p<0.01, *** p<0.001

end of do-file

wide format:

esta, wide compress

_labels and titles:

esta, se ar2 nostar brackets label
/* title(This is a regression table) */
/* nonumbers mtitles("Model A" "Model B") */
/* addnote('Source: auto.dta') */

plain table:

esta, plain

more

esta: Tabulating estimates simplified
```
end of do-file

. do_example do/10.ihelp
   . estab, wide compress

(1)          (2)
price
---          ---
weight       1.747** (2.72)  3.465*** (5.49)
mpg          -49.51 (-0.57)  21.85 (0.29)
foreign      3673.1*** (5.37)
_cons        1946.1   (0.54) -5853.7 (-1.73)

N            74          74

* t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

end of do-file
```

```
 */ nonumbers mtitles("Model A" "Model B") */
 */ addnote("Source: auto.dta") */

This is a regression table

<table>
<thead>
<tr>
<th>Model A</th>
<th>Model B</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight (lbs.)</td>
<td>1.747</td>
</tr>
<tr>
<td></td>
<td>[0.641]</td>
</tr>
<tr>
<td>Mileage (mpg)</td>
<td>-49.51</td>
</tr>
<tr>
<td></td>
<td>[86.16]</td>
</tr>
<tr>
<td>Car type</td>
<td>3673.1</td>
</tr>
<tr>
<td>Constant</td>
<td>1946.1</td>
</tr>
<tr>
<td></td>
<td>[3597.0]</td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.273</td>
</tr>
</tbody>
</table>

Standard errors in brackets
Source: auto.dta
```
Standard errors in brackets
Source: auto.dta

end of do-file

do_example do/12.ihlp

    esta, plain

    est1          est2
    weight        1.746559   3.464706
                  2.723238   5.493003
    mpg          -49.51222   21.8536
                  -.5746808   .2944391
    foreign      3673.06     5.370142
    _cons       1946.069    -5855.696
                  .541018   -1.733408
    N           74          74

end of do-file

esta: Tabulating estimates simplified

Numerical display formats may be specified as

  * official Stata's display formats such as %9.0g or %8.2f
  * integer values such as 0, 1, 2, etc. for fixed formats
  * a1, a2, ..., or a9 to cause esta choose a reasonable
display format depending on the scale of the displayed
number (the # in a# is the minimum number of significant
digits)

The default display format depends on type of displayed
statistic (e.g. a3 for point estimates and fixed format 3 for
p-values and the R-squared):

    esta, p r2 nostar wide
    <run>

Specifying alternative formats:

    esta, b(%9.0g) p(4) r2(4) nostar wide
    <run>
. end of do-file

. do_example do/13.ihlp

   esta, p r2 nostar wide

       (1)               (2)
    price    price

    weight  1.747   (0.008)   3.465   (0.000)
    mpg    -49.51   (0.567)   21.85   (0.769)
  foreign    3673.1   (0.000)
   _cons   1946.1   (0.590)  -5853.7   (0.087)

    N    74
    R-sq    0.293
   0.500

p-values in parentheses

. end of do-file

   --more--

. end of do-file

. do_example do/14.ihlp

   esta, b(%9.0g) p(4) r2(4) nostar wide

       (1)               (2)
    price    price

    weight  1.746559   (0.0081)   3.464706   (0.0000)
    mpg    -49.51222   (0.5673)   21.8536   (0.7693)
  foreign    3673.06   (0.0000)
   _cons   1946.069   (0.5902)  -5853.696   (0.0874)

    N    74
    R-sq    0.2934
   0.4996

p-values in parentheses

. end of do-file

   --more--
Use with Word, Excel, LaTeX, etc.

**esta** features various output formats:

- **smcl**: SMCL formatted (default unless using is specified)
- **fixed**: fixed-format ASCII (default if using is specified)
- **tab**: tab-delimited ASCII
- **csv**: CSV (Comma Separated Value format) for use with Excel
- **scsv**: “German” version of CSV (semicolon instead of comma)
- **rtf**: Rich Text Format for use with word processors
- **html**: HTML-formatted
- **tex**: LaTeX-formatted
- **booktabs**: LaTeX-formatted for use with *booktabs*

**Excel**: **csv** or **scsv**

esta using example.csv

<run>

esta using example.csv, scsv replace

<run>

**Use the plain option if you intend to do additional computations in excel:**

esta using example.csv, scsv replace wide plain

<run>

**Excel XML is not supported yet.**
. do_example do/15.ihlp
  . esta using example.csv
  (output written to example.csv)

. end of do-file

. do_example do/16.ihlp
  . esta using example.csv, scsv replace
  (output written to example.csv)

. end of do-file

—more—
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>est1</td>
<td>est2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>weight</td>
<td>1.746559</td>
<td>2.723238</td>
<td>3.464706</td>
<td>5.493003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>mpg</td>
<td>-49.51222</td>
<td>-0.5746808</td>
<td>21.8536</td>
<td>0.2944391</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>foreign</td>
<td></td>
<td></td>
<td>3673.06</td>
<td>5.370142</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>cons</td>
<td>1948.069</td>
<td>0.541018</td>
<td>-5853.896</td>
<td>-1.733408</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>N</td>
<td>74</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Use with Word, Excel, LaTeX, etc.

Word: rtf

esta using example.rtf
<run>

Appending is possible. Furthermore, use varwidth(#) and modelwidth(#) to change column widths:

esta using example.rtf, append wide label modelwidth(8)
<run>
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th></th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td></td>
<td>price</td>
<td></td>
</tr>
<tr>
<td>weight</td>
<td>1.747**</td>
<td>3.465***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.72)</td>
<td>(5.49)</td>
<td></td>
</tr>
<tr>
<td>mpg</td>
<td>-49.51</td>
<td>21.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.57)</td>
<td>(0.29)</td>
<td></td>
</tr>
<tr>
<td>foreign</td>
<td></td>
<td>3673.1***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.37)</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>1946.1</td>
<td>-5853.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(-1.73)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>74</td>
<td>74</td>
<td></td>
</tr>
</tbody>
</table>

*statistics in parentheses*

*"p<0.05, **p<0.01, ***p<0.001"*
Use with Word, Excel, LaTeX, etc.

**LaTeX:** tex

```latex
\begin{document}
\begin{table}[h]
\begin{tabular}{|c|c|c|}
\hline
Weight & 1.7474 & 3.4656 \\
\hline
\end{tabular}
\end{table}
```

**LaTeX:** booktabs

```latex
\begin{document}
\begin{table}[h]
\begin{tabular}{|c|c|c|}
\hline
Weight & 1.7474 & 3.4656 \\
\hline
\end{tabular}
\end{table}
```

!texify.exe --pdf example.tex

winexec SAcroRd example.pdf

<run>
Table 1: Regression table

<table>
<thead>
<tr>
<th></th>
<th>(1) Price</th>
<th>(2) Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (lbs.)</td>
<td>1.747</td>
<td>3.465</td>
</tr>
<tr>
<td></td>
<td>(2.72)</td>
<td>(5.49)</td>
</tr>
<tr>
<td>Mileage (mpg)</td>
<td>-49.51</td>
<td>21.85</td>
</tr>
<tr>
<td></td>
<td>(-0.57)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Car type</td>
<td></td>
<td>3673.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.37)</td>
</tr>
<tr>
<td>Constant</td>
<td>1946.1</td>
<td>-5853.7</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(-1.73)</td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
<td>74</td>
</tr>
</tbody>
</table>

*t statistics in parentheses*
Use with Word, Excel, LaTeX, etc.

Improved LaTeX table using the \texttt{dcolumn} package:

\begin{verbatim}
esta using example.tex, label replace booktabs /*
 */ alignment(D{.}{.}{-1}) /*
 */ page(dcolumn) /*
 */ title(Regression table\label{tab1})
\end{verbatim}

\texttt{!texify.exe --pdf example.tex}
\texttt{winexec $ACRORD example.pdf}
\texttt{<run>}

---

Table 1: Regression table

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th></th>
<th>(2)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (lbs.)</td>
<td></td>
<td></td>
<td>1.747**</td>
<td>3.465***</td>
</tr>
<tr>
<td></td>
<td>(2.72)</td>
<td>(5.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mileage (mpg)</td>
<td>-49.51</td>
<td></td>
<td>21.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.57)</td>
<td></td>
<td>(0.29)</td>
<td></td>
</tr>
<tr>
<td>Car type</td>
<td></td>
<td>3673.1***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1946.1</td>
<td></td>
<td>-5853.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(-1.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
<td></td>
<td>74</td>
<td></td>
</tr>
</tbody>
</table>

\textit{t} statistics in parentheses

* \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \)
Behind the scenes

Get to be an estout wizard using esta's noisily option:

```
esta, noisily
<run>

return list
<run>

`r(estout)`
<run>
```

```
estout

cells(b(fmt(a3) star) t(fmt(2) par("\{align 12:{txt}\{res\} "\{txt\}\{\}")
stats(N, fmt(\%18.0g) labels("N"))
starlevels(* 0.05 ** 0.01 *** 0.001)
varwidth(12)
modelwidth(12)
abbrev
delimiter(" ")
smc1tags
prehead("\{hline \width\}"")
posthead("\{hline \width\}"")
prefoot("\{hline \width\}"")
postfoot("\{hline \width\} ""t statistics in parentheses"" @starlee,
varlabels(, end("" "") nolast)
mlabels(, depvar)
numbers
collabels(, none)
level(95)
```

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td></td>
<td>price</td>
</tr>
</tbody>
</table>
Thank you for listening!