

Development Data Boot Camp

Linear Regression: Export Regression Tables

Ge Sun

University of Notre Dame

May 22, 2023

Implement multi-variate OLS in Stata

- Now, it is time to try by your self! Use our South African Labour Force data, generate the `ln_wage_hr` and `age_sq` variables and label them, then:

```
reg ln_wage_hr edyears age age_sq
```

```
. reg ln_wage_hr age age_sq female black edyears
```

Source	SS	df	MS	Number of obs	=	6,550
Model	1887.00912	5	377.401824	F(5, 6544)	=	681.63
Residual	3623.25484	6,544	.553675862	Prob > F	=	0.0000
Total	5510.26396	6,549	.84139013	R-squared	=	0.3425
				Adj R-squared	=	0.3420
				Root MSE	=	.74409

ln_wage_hr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.079928	.0142388	5.61	0.000	.0520152	.1078408
age_sq	-.0006621	.0001906	-3.47	0.001	-.0010357	-.0002885
female	-.5108119	.0189834	-26.91	0.000	-.5480256	-.4735982
black	-1.205415	.0414284	-29.10	0.000	-1.286629	-1.124202
edyears	.0948038	.0026904	35.24	0.000	.0895297	.100078
_cons	.3875367	.2645774	1.46	0.143	-.1311213	.9061948

Figure 1: Regression results

Outline

How to export regression tables?

ssc install outreg2

- ▶ Sometimes we need to use some useful commands that are not built-in in the Stata.
- ▶ Once you know the name of the command, you can type

```
ssc install command_name
```

to install the command from SSC (Statistical Software Components)

- ▶ We will install a powerful command *outreg2* by typing

```
ssc install outreg2
```

- ▶ You only need to install it once!

Command: `outreg2`

- ▶ The command `outreg2` provides a fast and easy way to export types of regression tables that we see in the academic papers.
- ▶ First see a simple example of *outreg2*:

```
reg ln_wage_hr edyears age age_sq  
outreg2 using simple_reg_results, word
```

- ▶ *using simple_reg_results* means that we want Stata to export the regression results to a file named *simple_reg_results* in the working directory.
- ▶ *word* means that we want the results to be stored in the Word file.
- ▶ *word* can be replaced by *excel* and *tex*.

Command: outreg2

	(1)
VARIABLES	ln_wage_hr
age	0.0784*** (0.0158)
age_sq	-0.000634*** (0.000212)
edyears	0.109*** (0.00291)
Constant	-1.044*** (0.290)
Observations	6,550
R-squared	0.188
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Command: `outreg2`

- ▶ What if the output files already exist or we want to add more column to the original files?
- ▶ *replace*: this option replaces existing file that have the same names as the output files in the *outreg2* command.

```
reg ln_wage_hr age age_sq edyears female black  
outreg2 using simple_reg_results, word replace
```

- ▶ *append*: this option adds new columns to the existing ones in the file.

```
reg ln_wage_hr age age_sq edyears  
outreg2 using simple_reg_results, word replace  
reg ln_wage_hr age age_sq edyears female black  
outreg2 using simple_reg_results, word append
```

Command: outreg2

	(1)	(2)
VARIABLES	ln_wage_hr	ln_wage_hr
age	0.0784*** (0.0158)	0.0799*** (0.0142)
age_sq	-0.000634*** (0.000212)	-0.000662*** (0.000191)
edyears	0.109*** (0.00291)	0.0948*** (0.00269)
female		-0.511*** (0.0190)
black		-1.205*** (0.0414)
Constant	-1.044*** (0.290)	0.388 (0.265)
Observations	6,550	6,550
R-squared	0.188	0.342

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 3: Replacement and appending using *outreg2*

Command: outreg2

- ▶ To distinguish each column with outputs from different regressions, we can specify the name of each column by using option *ctitle()*.

```
reg ln_wage_hr age age_sq edyears
outreg2 using simple_reg_results, word ctitle(Model1) replace
reg ln_wage_hr age age_sq edyears female black
outreg2 using simple_reg_results, word ctitle(Model2) append
```

Command: outreg2

VARIABLES	(1) Model1	(2) Model2
age	0.0784*** (0.0158)	0.0799*** (0.0142)
age_sq	-0.000634*** (0.000212)	-0.000662*** (0.000191)
edyears	0.109*** (0.00291)	0.0948*** (0.00269)
female		-0.511*** (0.0190)
black		-1.205*** (0.0414)
Constant	-1.044*** (0.290)	0.388 (0.265)
Observations	6,550	6,550
R-squared	0.188	0.342

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 4: Specify column names using *outreg2*

Command: outreg2

- In addition, you want to give an additional heading to each regression output column that would help you refer to.

```
reg ln_wage_hr age age_sq edyears
outreg2 using simple_reg_results, word ///
ctitle(Model1, ln_wage_hr) replace
reg ln_wage_hr age age_sq edyears female black
outreg2 using simple_reg_results, word ///
ctitle(Model2,ln_wage_hr) append
```

Command: outreg2

	(1)	(2)
	Model1	Model2
VARIABLES	ln_wage hr	ln_wage hr
age	0.0784*** (0.0158)	0.0799*** (0.0142)
age_sq	-0.000634*** (0.000212)	-0.000662*** (0.000191)
edyears	0.109*** (0.00291)	0.0948*** (0.00269)
female		-0.511*** (0.0190)
black		-1.205*** (0.0414)
Constant	-1.044*** (0.290)	0.388 (0.265)
Observations	6,550	6,550
R-squared	0.188	0.342

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 5: Specify column names using *outreg2*: add additional headings

Command: outreg2

- ▶ You may have noticed that the names of variable in output file are short and ambiguous, so we want to replace them with the corresponding variable labels. This can be implemented by *label* option.

```
reg ln_wage_hr age age_sq edyears
outreg2 using simple_reg_results, word ///
ctitle(Model1, ln_wage_hr) label replace
reg ln_wage_hr age age_sq edyears female black
outreg2 using simple_reg_results, word ///
ctitle(Model2,ln_wage_hr) label append
```

Command: outreg2

	(1)	(2)
	Model1	Model2
VARIABLES	ln wage hr	ln wage hr
age	0.0784*** (0.0158)	0.0799*** (0.0142)
age squared	-0.000634*** (0.000212)	-0.000662*** (0.000191)
educational years	0.109*** (0.00291)	0.0948*** (0.00269)
female		-0.511*** (0.0190)
black		-1.205*** (0.0414)
Constant	-1.044*** (0.290)	0.388 (0.265)
Observations	6,550	6,550
R-squared	0.188	0.342

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 6: Using labels instead of variable names

Command: outreg2

- ▶ We always want to adjust the formats of statistics (like coefficients, standard errors, and R squared) in the table to make them concise and apparent. We can do this by we specify option *dec()*, with desired number of decimal in the brackets.

```
reg ln_wage_hr age age_sq edyears
outreg2 using simple_reg_results, word ///
ctitle(Model 1, OLS) label dec(3) replace
reg ln_wage_hr age age_sq edyears female black
outreg2 using simple_reg_results, word ///
ctitle(Model 2, Robust) label dec(3) append
```

Command: outreg2

	(1)	(2)
VARIABLES	Model 1	Model 2
	OLS	Robust
age	0.078*** (0.016)	0.080*** (0.014)
age squared	-0.001*** (0.000)	-0.001*** (0.000)
educational years	0.109*** (0.003)	0.095*** (0.003)
female		-0.511*** (0.019)
black		-1.205*** (0.041)
Constant	-1.044*** (0.290)	0.388 (0.265)
Observations	6,550	6,550
R-squared	0.188	0.342

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 7: Adjusting formats of statistics

Command: `outreg2`

- What if you want different statistics have different formats? You can specify them separately. Type *help outreg2* for details.

```
fixed decimals & formats

dec(#) or fmt(type) for everything, default dec(3) and fmt(fc)

where # is fixed decimals between 0 and 11,
and type is one of the following:

e    exponential
f    fixed
fc   fixed with commas for thousands
g    general
gc   general with commas for thousands

Use the following only if you want them different from dec(#) or fmt(type):

bdec(#) or bfmt(type) for coefficient only
sdec(#) or sfmt(type) for standard error
tdec(#) or tfmt(type) for t-statistics
pdec(#) or pfmt(type) for p-value
cdec(#) or cfmt(type) for conf. interval
rdec(#) or rfmt(type) for r-square
adec(#) or afmt(type) for addstat( ) contents
```

Figure 8: Different formats for different statistics

Command: outreg2

- ▶ When you want to add notes below the table to make your results more understandable, you can use *addnote()* option with notes in the brackets. When you want to add a title on the table, ye can use *title()* option with title in the brackets.

```
reg ln_wage_hr age age_sq edyears
outreg2 using simple_reg_results, word ///
ctitle(Model 1, OLS) label dec(3)///
addnote(Notes: You can add notes here!) ///
title(Regression Tables) replace
reg ln_wage_hr age age_sq edyears female black
outreg2 using simple_reg_results, word ///
ctitle(Model 2, Robust) label dec(3) append
```

Command: outreg2

Regression Tables		
	(1)	(2)
	Model 1	Model 2
VARIABLES	OLS	Robust
age	0.078*** (0.016)	0.080*** (0.014)
age squared	-0.001*** (0.000)	-0.001*** (0.000)
educational years	0.109*** (0.003)	0.095*** (0.003)
female		-0.511*** (0.019)
black		-1.205*** (0.041)
Constant	-1.044*** (0.290)	0.388 (0.265)
Observations	6,550	6,550
R-squared	0.188	0.342

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: You can add notes here!

Figure 9: Adding titles and notes

Command: outreg2

- Sometimes it is better to separate regression parts with exporting parts, that is, we can first do some regression, and then export the result in one time.

```
reg ln_wage_hr age age_sq edyears
```

```
est store Model1
```

```
reg ln_wage_hr age age_sq edyears female black
```

```
est store Model2
```

```
outreg2 [Model1 Model2] using simple_reg_results, word ///
```

```
ctitle(Model1, ln_wage_hr; Model2, ln_wage_hr) label dec(3) ///
```

```
addnote(Notes: You can add notes here!) ///
```

```
title(Regression Table)
```

Command: `esttab`

- Sometimes you just want to take a look at your results in your results window, then:

```
reg ln_wage_hr age age_sq edyears
```

```
est store Model1
```

```
reg ln_wage_hr age age_sq edyears female black
```

```
est store Model2
```

```
esttab Model1 Model2
```

```
esttab Model1 Model2, b(2) se(2) r2
```