

# An introduction to Stata, part II

## Advanced use and programming

Matthijs de Zwaan

ACED

December 3, 2010

- 1 Loops
- 2 Using your results
- 3 “Pretty” output and graphs
- 4 Make your own data
- 5 Packages
- 6 Programs and .ado files
- 7 Capita Selecta

## foreach, forval and while

- `foreach{}`, `forval{}` and `while{}`  allow you to repeat manipulations easily.
- useful for data manipulation or doing the same analyses over different subsets of your data.
- For example,

```
foreach i in 0 1 {  
  summarize wage if gender == 'i'  
}
```

summarizes the “wage” variable for men and women (presumably)

- take care of the syntax; see the help files!

# Macros, scalars and matrices

- Macros can store a list of things (both text and number)
  - ▶ can be used, for example, in a `foreach{}` loop
- Scalars hold a single number
  - ▶ can be used in further calculations; scalars are very precise!
- Matrices hold arrays of scalars
  - ▶ can be used in further calculations

# Access to results

- Stata stores results of commands in its memory
- you can access those results for further calculation/manipulation
- useful if you want to verify something, if or if the statistic you want is not directly shown and you need to calculate it from the available information.
- at the bottom of any help file, there is a list of things that are accessible
- see `help return`, and links therein
- Results are returned as matrices, scalars or macros

# Outputting your output

- export your results to Excel/Word/L<sup>A</sup>T<sub>E</sub>X
- Useful functions and SSC-packages:
  - ▶ estimate functions — see `help estimates`
  - ▶ `outreg2` — SSC
  - ▶ `estout`, `esttab`, `estpost`, `eststo` — SSC

# Publication quality graphs

- You can customize Stata's `graph twoway` commands in **many** ways
- Multiple axes, legend, labeling of data points, colours, add text and notes, etc
- see `help graph twoway` and `help twoway options`
- see Mitchell (2004), *A visual guide to Stata graphics*.

# Random number generation

You can use Stata to generate (pseudo-)random number for a number of distributions

- You know the **true** coefficients!
- You need to use `set obs #` to define the number of observations your data will hold
- use `set seed` to get the same “random” number each time: for reproducibility
- useful for testing, teaching
- see `help random number functions`



# Don't invent the wheel again!

- There are many user-written packages available
- Statistical Software Components (SSC) archive or <http://www.stata.com/stb/>: triple-checked by high quality Stata users, with good quality help files
- **DO NOT install manually!**
- use the `findit` command to search for a package and install it
- or use `ssc install <package name>` to install
- see also `help ssc`

# Programs

- If you use sequence of commands regularly, you can make a program, which you can then easily use again.
- If you save your program as an .ado file, you can then call the program any time, as you would any Stata command
- type `sysdir` to see where your .ado-files are saved.
- or install `adoedit` from SSC.
- see `help program` for syntax. Have a look at any of the .ado-files in your directory for inspiration!

# Capita Selecta

- Linking with data base programs: see `help odbc` and `http://www.stata.com/support/faqs/data/plugin_database.html`