Econometrics II
1st Tutorial:
Introduction to Stata

Fabian T. Dehos

Ruhr Graduate School in Economics

Summer Semester 2020

Outline

- contact: fabian.dehos@rwi-essen.de

- weekly tutorial: Tuesday 16:15-17:45.

$\rightarrow$ length of the tutorial might be subject to adjustments.

  - hands-on applications using STATA.
  - theoretical problems with applied focus.
  - discussion of selected research papers.

  - all tutorials are supposed to be interactive.
  - $\rightarrow$ i.e. prepare the exercises in advance.
  - $\rightarrow$ collect questions, doubts & concerns for discussion.

- three (or two) assignments to be handed in.

  - your can earn bonus points.
  - will be announced in advance.

Introduction to Stata

- Stata is a statistical software package for data preparation, data analysis and graphic representation.

- contains many tools that make life of an applied econometrician comfortable.

- possible to write your own commands / programms (ado-files).

- new features are implemented by other users (see **statalist**) or through StataCorp (see **Stata Journal**).

- a few basic commands will be introduced in the tutorial.

- there are several introductory books:

  - Kohler, Kreuter (2016), *Datenanalyse mit Stata: Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung*.
  - Cameron, Trivedi (2010), *Microeconometrics using Stata*.

- further tools: **Stata Manual** and Stata **"help"**-command.

Do-files, Log-files and Data-files

- results-window:
  main window you will see when you open Stata. Displays the
  results. Note: you can move the different windows and change
  their size.

- command-line:
  Here you can directly type in short commands and execute them
  by typing return. *(For example type in: edit)*.

- **Log file**: *filename.scml* (only readable in stata) or *filename.log*
  (text file). Documentation file. All output appearing in the
  results window can be can be captured in a log file.

- **Data-files**: *filename.dta*
  If your data is stored in different format, like .xls , it can be read
  into data format.

- **Do-file**: *filename.do*
  file that will contain your program code.

Documentation of your research

- Use a master Do-file and document everything in detail.

- Make your research transparent and replicable.

Why?

- After a few weeks you won't remember what your code is about.

- Journals often require you to handle in data and code.

- $\rightarrow$ One key aspect to fight the current *replication crisis* in the field of economics.

A few basic Stata commands I

- It is useful to work with commands, but Stata provides you click and drag options as well.

- Obtaining HELP in STATA
    - → type **help** plus some command or **search** plus some keyword to search into the command line.

- Load the data-file into Stata
    - → **use mydata.dta**
    - → or, alternatively, using Menu → **File** → **Open**.

- **browse** → displays data in the data editor.

- **describe** → describes the data in memory if a data set description has been attached to the data set by its provider.

- All Stata commands and variable names may be abbreviated as long as there is no ambiguity.

A few basic Stata commands II

In case you want to know more about the following commands
type in **help command**.

- **Datasets and data handling**

    $\rightarrow$ use, import, export, merge, sort, reshape.

- **Finding out what is in your dataset**

    $\rightarrow$ browse, describe, list, lookfor.

- **Describing data and tabulating**

    $\rightarrow$ sum, "sum, detail", tab.

- **Making new variables / Changing variables**

    $\rightarrow$ gen, replace, egen, "tab , gen(group)", drop, recode, label.

- **Regression etc.**

    $\rightarrow$ regress, ivreg, probit, logit, heckman, qreg, xtreg.

A few basic Stata commands III

Logical operators

- $a == b$ true if a equals b
- $a > b$ true if a greater than b
- &   and
- | or
- !   not
- $a! = b$ true if a not equal to b