ECON-C4100 - Econometrics I
Session 1: Introduction

Antto Jokelainen, Ellen Sahlström

Contact

Doctoral candidate Antto Jokelainen

Economics dept. Aalto

antto.jokelainen@aalto.fi

Doctoral candidate Ellen Sahlström

Economics dept. Aalto

ellen.sahlstrom@aalto.fi

Problem sets

- 5 graded problem sets.

- Deadlines on Thursdays at 18:00.

  - First one is out, DL next Thursday!

- Detailed instructions and deadlines are found on MyCourses.

Exercise sessions

- Two types of exercise sessions:

  1. Monday sessions
     - A "laskutupa"-type tutorial session.
     - Working on problem sets independently or in groups. TA is there to assist you when stuck.
     - Try to do as much as you can before the session - more time for the difficult questions.
     - U261 at 10:15-12:00 (No PCs, remember to bring your laptops!).
  2. Friday sessions
     - Going through model solutions.
     - No Stata lecturing like in previous years. You are expected to use online material (and tutorial sessions for advice).
     - U271 at 10:15-12:00.

- Attendance is not mandatory but highly recommended!

Communication

- Any questions related to the problem sets can be asked during the exercise sessions or on the MyCourses forum (or Zulip?).

- Questions by email should be avoided but can be used if necessary, such as regarding personal matters.

Our principles for the problem sets

- The exercises will be time consuming and probably a bit difficult.

- However, we are lenient in the grading process.

- Always return something, otherwise there is not much we can do for you.

- On the other hand, do not write novels for us. Most of the time a few bullet points are enough.

- Help will be provided when asked! **Attend the exercises**, use MyCourses etc.

- Plagiarism is strictly forbidden!

Objectives

- In addition to the theory of econometrics, we want you to learn how to apply what you've learned $\rightarrow$ empirical exercises.

- There are two ways to interpret empirical results:

  1. Statistical significance
  2. Economic significance

- If we ask you to interpret the results, we expect you to provide answers from both angles if not specified otherwise.

- Results without interpretation will give you **less than half the points.**

Stata

- The main tool for the empirical exercises and applications in the course is Stata.

- Stata is a statistical software and a programming language used widely especially in economics.

- You can also use any other statistical software (like R or Python). However, support is given only for Stata.

- Aalto offers a free Stata licence to all students. You can download the software directly from this website (Aalto credentials required!).

Why Stata and not R?

- A common question without a definite answer.

- Stata is still the *lingua franca* in empirical economics. Most of faculty in Aalto and around the world use Stata.

  - *"Because we use it, so should you!"*

  - Lots of built in and user written packages useful in applied econometrics. For instance: rdrobust and reghdfe

- Stata code is relatively easier to get started with.

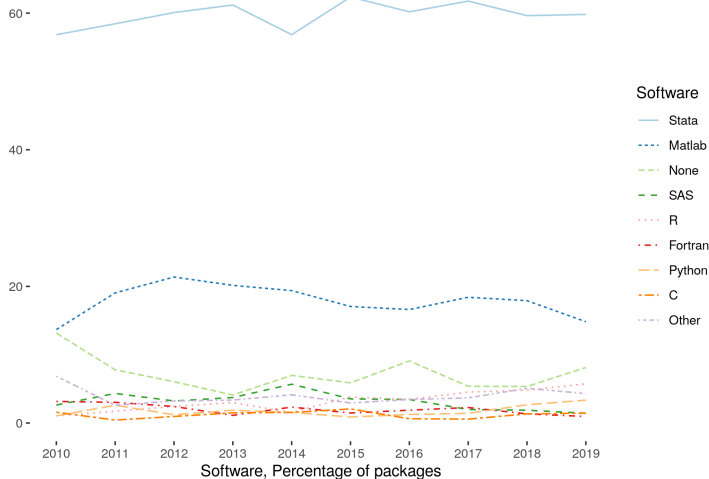**Figure:** AEA Data Editor: Software by Supplement across Years, Files

COMPASS LEXECON

Analyst   APPLY FOR THIS JOB

HELSINKI, FINLAND / COMPETITION PRACTICE – HELSINKI /
FULL TIME, PERMANENT

Figure: Economics-specific job listings often require some knowledge of Stata.

Why Stata and not R?

- However, R definitely deserves some love: ggplot2 is probably the best graphics package available anywhere.

- There are lots of MOOCs available both for R and Python/Pandas. But for now, we're stuck with Stata.

- Which ever software you choose, the difference is marginal!

Using Stata

- Stata has an easy-to-use GUI for most operations. It also prints the used command which makes the GUI useful in learning Stata.

- However, we recommend you start using Stata directly by writing the code to the command line or most ideally into the do-file editor.

- Do-files are the files where the user's code is written and which Stata executes.

Using Stata

- The most important command of Stata is the `help` command. No joke, it opens the Stata documentation.

  - Besides the function options and etc, the Stata documentation offers excellent documentation on statistical concepts. For instance, type in `help regress` .

- As usual, Google search is a great tool. Unorthodox to other programming languages, most online help about Stata are not on Stack Overflow but on Statalist.

Course Stata materials

- A template file is provided on MyCourses.

- Last year's TA, Jaakko Markkanen, has created online material for this course here.

- You can request Stata topics & tricks to be covered in the exercise sessions.

- Be proactive in finding learning material from the web.

Auxiliary tools

- In addition to Stata, we also encourage you to learn some of the auxiliary tools that are frequently used in applied empirical work.

- Topics to be covered: LaTeX, Markdown and outputting results with HTML-files.

- Using these is not mandatory, but they will make your life a lot easier.

LaTeX

- LaTeX is a typesetting system that makes it easy to to write elegant publication-like documents → widely used in academia.

- More controlled and sophisticated input for equations, graphs, and tables than in MS Office.

- Automated managing and formatting of citations and references with with BibLaTeX and Zotero or Mendeley.

- LaTeX takes time to master but it's totally worth it. Start now and you will never regret it!

Overleaf

- There are multiple editors for LATEX. The "best" editor available is Overleaf which is cloud-based. Aalto offers a premium subscription plan for students who register with their Aalto email.

- Other LATEX desktop-based clients are also available, one example is Texmaker.

- There are various templates available for homework problem sets or thesis documents.

- LATEX is all about flexibility. You can spend a lot of time on formatting margins, fonts, spacing and basically everything. Sticking with the default options can often be the right choice.

Markdown and dynamic documents

- Oftentimes you want to present both your code and your results. How to do this most efficiently?

- The answer: Markdown and dynamic documents.

- Modern IDEs for data scientists allow the user to write both the input and the output directly to html or pdf.

- This is the current best practice to present your empirical work. Example platforms include Jupyter and RMarkdown.

Markdown and dynamic documents with Stata

- Stata has a built in functionality for creating dynamic Markdown files.
- There is also the excellent Markstat package.
  - Requires some extra work to install but is easier to use notation-wise.
- You can also use Stata within Jupyter.