Econ 617 Software for Econometrics and Big Data 2022-2023

COURSE AIMS & OBJECTIVES, KEY SKILLS AND LEARNING OUTCOMES

Course Aims & Objectives: The aim of this course is to provide students with programming skills that can be used to perform data management operations and several types of statistical and econometric analysis in an efficient and reproducible way. Students will be taught the fundamentals of programming in Stata. The module will prepare students for PhD research or work as an econometrician consultant in public and private organisations

Key Skills: By the end of this course, students should have some knowledge and understanding of:

- Understand the principles of Stata programming;

- Write Stata do-files for automating the workflow;

- Write ADO-files to extend Stata by creating new commands;

- Understand the basics of Mata Programming Language.

Desired Outcomes: By the end of this course, students should be able to:

- Use automatic model selection to build econometric models
- Use saturation estimation
- Present, interpret and analyse the results of automatic model selection
- Utilise effectively PcGive and the R gets package packages

COURSE STRUCTURE

Econ 617 is a 10 credits course and therefore students are expected to input approximately 100 hours of study into the course. The total number of contact hours on Econ 617 is 15 hours. This leaves 85 hours for private study. Course Delivery comes in the form of Lectures with 15 hours delivered over the first 3 weeks of the term (10 hours of lectures and 5 hours of tutorials). There will be optional clinics on the last day of the course.

During your private study you should strike a balance between reading the course material (which is the primary source of information) and the recommended textbooks, thinking critically about how these fit in to the body of knowledge on the subject and

about how our level of knowledge can be improved, performing exercises, completing coursework and revising for examinations. You can expect to perform well on this course only if you work consistently through the year.

COURSE CONVENOR

Dr Malvina Marchese and Professor Kit Baum (special guest lecturer)

LECTURERS CONTACT INFORMATION (Including Office Hours)

Email: malvinamarchese@timberlake.co.uk

Available by appointment (please email to arrange a convenient time)

COURSEWORK ASSESSMENT

The final mark for the course will depend on a written exam. Timetable for details of time and venues will be communicated via Moodle and by Timberlake well in advance.

The CWA mark will be calculated as 100% coursework. The coursework will be assigned at the beginning of the module.

Coursework must be submitted electronically through the Moodle site for this course: https://mle.lancs.ac.uk/course. Login using your regular Lancaster University access details. This opens a page headed MLE: My home.

The format of the submission is as follows.

- The submitted file must be in pdf format with the following name

  stud#_studname_cw**_cw#.pdf**

  where: **stud#** is your student number, **studname** is your name in the format *surname_name*, **cw#** is either 1 or 2 according to the piece of coursework submitted. Eg a student with student number 111 would submit a file named *111_surname_firstname_cw_1.pdf*.

- Maximum file size is 2MB: figures resolution must be adjusted accordingly.

Note that your work will be screened using software designed to detect plagiarism.

Do not rely upon someone else to submit your coursework.
*Word counts are inclusive of all material submitted apart from the Bibliography.

FEEDBACK ON COURSEWORK:

The coursework will be marked and returned to students within 4 weeks of the submission deadline. Feedback will consist of marker's notes appended to the pdf of your coursework.

MARKING CRITERIA AND PENALTIES

Marking criteria can be found in the Economics Undergraduate Handbook and the general course information paper. An electronic copy of this can be found via the Current Student page of the university website then follow the Academic Regulations link
https://gap.lancs.ac.uk/ASQ/QAE/MARP/Documents/UG-Assess-Regs.pdf

FINAL MARK INFORMATION

This course is assessed 100% by means of coursework. The final mark is the average of the marks obtained in the two pieces of coursework.

COURSE TEXT AND RECOMMENDED READING

The main recommended textbook are:

- Mitchell, M. N. (2010). Data Management using Stata A Practical Handbook. Stata Press

- Baum, C. F. (2016). An Introduction to Stata Programming, second edition 2016

COURSE OUTLINE/LECTURE SCHEDULE

Day 1:

Lecture 1: Programming: Some Fundamental Concepts (

- Do and ado files in Stata

- Data types: Working with dates and time, time series operators, factor variables

- Debugging – the capture command

- Protecting your data with preserve and restore

Lecture 2: Guidelines for writing programs in Stata

- Basic guidelines for writing do-files

- Enhancing speed and efficiency

- How to seek help for Stata programming

DAY 2

Lecture 3: Programming Stata Do-Files I

- Transformation of string and numeric variables

- Recoding discrete and continuous variables

- Recoding missing variables

- Generate and egen functions
- Local and global macros
- Extended macro functions and list functions

Lecture 4: Programming Stata Do-Files II

- Scalars and matrices
- Temporary variables and tempnames
- Looping commands: forvalues and foreach
- By, statsby, rolling and other prefix commands
- Transforming data with reshape
- Combining datasets with append and merge

DAY 3:

Lecture 5: Programming Concepts and Tools in Action I

- Tabulating and computing statistics across a set of variables and groups
- Computing durations and aggregating transactions data
- Summarising group characteristics over observations

Lecture 6: Programming Concepts and Tools in Action II

- Adding aggregate characteristics to micro observations
- Reusing results: return and ereturn commands
- Accessing estimation results and margins

Day 4

Lecture 7: Automation - Do-File Programming I

- Storing results in matrices
- Stata's post and postfile commands

Lecture 8: Automation - Do-File Programming II

- Automating estimation output – the user-written commands estout and tabout, saveresults
- Automating graphical output

Day 5:

Lecture 9: ADO-File Programming I

- Structure of an ado-file

- Syntax and return statements
- Using program options
- Including a subset of observations
- Generalizing commands to handle multiple variables
- Rendering commands byable

Lecture 10: ADO-File Programming II

- Writing an e-class program
- Programs for egen, ml, nl
- Programs for gmm
- Guidelines for ado-file programming style

Lecture 11: Fundamentals of Mata programming

- Mata Fundamentals
- A simple Mata function
- Components of a Mata function
- Calling Mata functions
- Mata's st_interface functions