Advanced Econometrics (ECON 904)

Spring 2024 Syllabus

Prof. Antonio Rodriguez Andres
Economics, German University in Cairo

Course Info

When:	Monday: 11:45 am–14:15 pm
Websites:	`https://github.com/antonio1970`
Instructor:	Antonio Rodriguez Andres
Email:	`antonio.andres@guc.edu.eg`
Office Hours:	Thursday, 10:00–11:00 am.

Course Description

This is the final course in the graduate econometrics sequence in the Department of Economics. Following previous courses on quantitative and qualitative methods. The goal of this course is to provide students with an in-depth understanding of the most common estimation methods in modern empirical economics and with the technical ability to apply these methods to their own research. The course will focus on the application of all these methods, which underlie many economic decisions studied in applied microeconomics and related fields. In all statistical techniques, we will be using R programming.

Prerequisites

Students should have a strong familiarity with statistics, linear algebra, and the classical linear regression model at the level of undergraduate courses in econometrics. If you have not taken that course but think your background in these topics is sufficient, please see me and we can discuss whether this course is appropriate for you.

Course Structure

This course will be based on lecturing. The general structure of the course is that we will cover one block each month approximately (whenever possible); most topics will include both "theory" and "practical applications.". Lecture materials will encourage you to think about applying the material to your own research and give you the opportunity to ask questions about the material. Most importantly, we will also work through applications of the material in the R statistical programming language. I will provide you with the materials for the classes on the CMS system. No need to be familiar with R as we will start from the scratch.

Textbook (s)

The textbooks for this graduate course are:

> Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press.
>
> Greene, William. 2019. *Econometric Analysis Global Edition.* 1168. Pearson Deutschland.
>
> Wooldridge, J.M . 2002. *Econometric analysis of cross section and panel data.* Cambridge, MA.
>
> Kleiber, Christian, and Achim Zeileis. 2008. *Applied Econometrics with R.* Springer.

This list is not exhaustive, and many similar textbooks exist. Most students already have an applicable textbook from a previous course. I will not assign reading from these textbooks, but you will likely find one of these "more rigorous" textbooks to be a helpful reference at times.

Software: R

We will use the R statistical programming language in this course. R is a free and powerful software environment for statistical analysis. It can be used for almost all analysis in applied economics and related fields: basic statistics, data cleaning, linear regression, structural estimation, data visualization, etc. I will work through examples in class using R, and I will provide example R code for your own use. You may use another programming language in this course if you would like, but I strongly recommend that you learn and use R. I will not provide any support for this course in other programming languages, nor will I provide partial credit on the research report or exams written in a programming language other than R.

I will give an introduction to R early in this course to ensure all students have a basic understanding of key features of the R language. If you have never used R, this introduction may not be sufficient to implement the methods we will cover in this course. Several free online R resources that you may find useful are:

> Introduction to R (`www.datacamp.com/courses/free-introduction-to-r`)
>
> R for Data Science (`r4ds.had.co.nz`)
>
> Advanced R (`adv-r.hadley.nz`)

Grades

Grades will be based on the categories listed below with the corresponding weights.

- Research Project (40%): The empirical project is a very important part of your learning experience in the class. It provides you with the opportunity to use the tools you learn in the class to answer a question that you come up with and that you care about. It is designed to guide you through the steps of answering a research question the way applied economists do. Each student will identify and submit their research question on March 10, 2024. Find an interesting data to analyze (you can also search either on the web or Kaggle's webpage (https://www.kaggle.com/), where you can find lots of open datasets with statistical analysis.

The week after, each student will pass in a document that quantitatively describes their sample and the relevant variables in it. During the last week of the classes, we will hold a poster session, and by the end of that week, students will submit a written project report that includes a discussion of the empirical analysis. You can employ one of the statistical techniques we have learned in the course (regression analysis, instrumental variables, panel data models or other machine learning techniques).

The research question is worth 5% of your final grade, the data description another 5%, and the final analysis (presentation and report) is worth 5%. You will get a fair bit of written and verbal feedback guidance throughout the process.

The empirical analysis can be carried out by using R. You can also take a look at the template provided on the CMS system. What you need to turn in:

  - One spreadsheet and/or R script with the data and empirical analysis
  - One word document with the written paper
  - Write up approx. 10 page (12pt font, double spaced, excluding References, Figures, and Tables) summary of your findings, including discussion about what prior studies of the same topic have found, as well as citations to prior studies

- Final exam (45%). To be defined later.
- Referee report (10%) : You will write a referee report on an unpublished paper of your choosing. The paper must use some method we discussed in the syllabus.
- Class participation (5%): I am defining as "active participation" as keeping up with assigned readings, lecture slides, class discussion and homework each week.;

The tentative schedule of topics is shown below:

Tentative lecture schedule

Block (3 hours)	Dates	Topic
1	March. 04	Syllabus & Introduction & Introduction to R
2	March. 18	Data wrangling & data visualization using Tidyverse
3	April. 01	Basic concepts in Machine Learning & Regression & IV approach
4	April. 15	Time Series: Forecasting & Panel data models
5	May. 06	Logistic regression (classification models) & Other classification models
6	May. 13	Practical Project using a large data set (Hands on programming with R)