ECON 104-920 : Econometrics

Artem Kuriksha

Summer 2020
(July 2 through August 7)

E-mail: `kuriksha@sas.upenn.edu`
Synchronous classes: TR 1:30PM-3:30PM
Asynchronous classes: videos uploaded twice a week
Office hours: Friday 3PM-5PM
Class Room and office hours: Zoom link is available on Canvas

Course Description

This course will introduce you to the field of econometrics, which is the part of statistics that economists use to understand causal relationships in the world. We will cover the most important methods and discuss the complications that arise while applying them. In the class, we will see what the mathematics behind econometrics is and some of the most interesting empirical applications.

After taking the course, you are expected to know the main econometric methods and why, how, and when they work. You will know how to handle economic data yourself and interpret the results obtained by others.

There are going to be 4 home assignments, which will include both practical and theoretical exercises. Theoretical exercises will help you to understand the math and the role of assumptions we impose. Practical exercises will help you with that too and, most importantly, will allow you to learn how to apply all the methods to real data.

Finally, there will be a midterm and a final.

Course Structure

We will have two types of classes: synchronous classes and asynchronous classes. Asynchronous classes will be for covering the material only. Synchronous classes will be for covering the material, answering your questions, and in-class exercises.

In-class exercises are exercises you will do online and at the time of the class, individually or in small groups (this depends on your preferences). They are going to be relatively simple but will require you to be up-to-date. Each of them will take around an hour. You are expected to share with me the progress you have made. Then I will show the way I solved the exercise. They are graded mainly based on your participation, but you should actually produce some results.

In addition to that, you are required to read the textbook and some additional materials I will provide.

Required Software

We are going to use Python as the programming language for this course. Python is a modern, simple, powerful, elegant, and widely accepted in industry programming language. You will be required to submit your codes with commenting texts in the form of iPython notebooks (you will appreciate how convenient and simple it is). For this, you will have to either install Jupyter on your working station or simply use Colab, an incredible online solution provided by Google.

Prerequisites

ECON 101, 103; MATH 104, MATH 114, MATH 115, or permission from instructor. Essentially, you should be comfortable with basic notions from statistics (e.g., random variables and hypothesis testing) and multivariate calculus.

Course Policies

Communication

These are strange times, and we have to experiment to conduct the course effectively in a new format. Some things might be subject to change if while going through the course we find some approaches to work better than others. We will actively communicate via emails and Canvas to manage all that. Please do not miss my emails and messages.

Grading Policy

The formula that will determine your grade is

$$\text{Final Grade} = 40\% \times \text{Problem Sets} + 5\% \times \text{ In-class exercises} +$$

$$25\% \times \text{ Midterm} + 30\% \times \text{Final exam.}$$

Attendance Policy

You are required to attend all exams. Attendance of synchronous classes is not required, and the synchronous classes will be recorded and uploaded on Canvas. However, you should attend synchronous classes with in-class exercises to get the corresponding part of the grade.

Contact me if you have problems with attending a class.

Academic Integrity and Honesty

Any student found in violation of academic integrity will receive no credit for the assessment in question. Multiple integrity violations will have more serious consequences.

Course Schedule

Home assignments

You will have a new home assignment on every Monday except for the last week, that is, on July 6, July 13, July 20, and July 27. The deadline for each one will be Sunday 11:59PM.

Classes and exams

The schedule is tentative and subject to change.

Class 1   Introduction (July 2)

- Course structure

- What is econometrics

- Probability and statistics review

- Python quick start

Class 2   OLS, single regressor (July 7)

- Properties of OLS

- Gauss-Markov Theorem

- Statistical tests

Class 3   OLS, multiple regressors (July 9)

Class 4   Discussion of OLS (July 14)

- Endogeneity

- Biases

- Non-linear models

Class 5   (July 16)
Midterm

Class 6   Addressing endogeneity (July 21)

- Instrumental variables approach

- Two stage least squares

- RCT, natural experiments, regression discontinuity

Class 7 Panel data (July 23)

- Fixed effects model, random effects model

- Diff-in-diff estimation

Class 8 Dummy/Limited Dependent Variable Regressions (July 28)

- Maximum Likelihood Estimation

- Logit, probit, censored, truncated

Class 9 Elements of Machine Learning (July 30)

- Difference between machine learning and econometrics

- Bias-variance trade-off, lasso, ridge regression, elastic net

- Perceptron, kernel regression, k-means, kNN

- Neural nets (briefly)

Class 10 (August 4)
Final exam

Readings

Required

Stock & Watson, Introduction to Econometrics, 4th Edition

Complimentary

This is if you have some questions about the probability and statistics part.

Casella & Berger, Statistical Inference

Additional

These are the books I recommend to look at if you are interested in economics and data analysis in general. By no means they are required for passing the course.

- *Angrist, Mostly Harmless Econometrics*

- *James, Witten, Hastie & Tibshirani, An Introduction to Statistical Learning with Applications in R*