

# ECON4274 Programing Econometrics with R

Hong Kong University of Science and Technology  
Department of Economics

## Course Information

Lecture: Friday 10:30 - 13:30 at LSK G021  
Tutorial section: Thursday 18:00 - 18:50 at LSK G021

Instructor: Yangguang (Sunny) HUANG  
Email: [huangyg@ust.hk](mailto:huangyg@ust.hk)  
Office hour: Thursday 15:00 – 17:00 or by appointment (LSK 6075)  
Course website: [canvas.ust.hk](http://canvas.ust.hk)

Teaching assistant: Hector CHENG  
Email: [echector@ust.hk](mailto:echector@ust.hk)  
Office hour: Friday 15:00-16:00 (LSK 6066)

Student helper: Hoi-Sing CHOI  
Email: [hsuchoi@connect.ust.hk](mailto:hsuchoi@connect.ust.hk)

## Learning Goal

This course puts statistical theory and econometric method into practice by working with data and cases. Students are required to write their own program with R for summary statistics, data visualization, estimation, and hypothesis testing. The students are also trained to produce table and write report for empirical study. The main objective is to help students develop skills in working with data to support a better business decision or analysis.

**Warning:** This is an advanced level course and require significant amount of time and commitment. The learning curve of programming is very steep. Students are assumed to have reasonable knowledge of Econometrics, Calculus, Probability, Statistics, and Linear Algebra. TA and instructor are not responsible to help line-by-line debugging of student's code.

Per Program Intended Learning Objective (PILO) for [BSc in Economics and Finance](#), after completing this course, students will

1. Understand the logic, scientific basis, and critical thinking of economic analyses. Solve business problems using appropriate quantitative and analytical techniques. (PILO 1)
2. Analyze qualitatively and quantitatively basic economic and financial problems. Apply economic knowledge to practical situations and make sound economic and finance decisions. Have a solid foundation for postgraduate studies. (PILO 4)
3. Work with others effectively and responsibly. (PILO 5)
4. Graduates will be effective users of information technology in business applications. Demonstrate proficiency in using IT applications in business and management. Use econometric or statistical software to deal with databases and conduct empirical analysis. (PILO 7)
5. Be lifelong users of economic analysis and econometric or statistical software. (PILO 9)

## Grading

1. Five problem sets (each counts 10%)  
Can be submitted by individual student or by a group with maximum three students. (Grouping can be changed across problem sets.) Submission will be through Canvas. Please submit both your answer (as word document or PDF) and your code.
2. Programming exam, 30%  
The programming exam tests for core materials of this course. You will have internet access but no communication is allowed. Absence in an exam without a prior notification and a reasonable and verifiable justification will result in zero point for that part.
3. Final project, 20%  
Students can choose whether or not to do the final project with attempting on fulfilling ECON 4670's paper/project requirement.  
If a student chooses no, then can submit it by individual or as a group with maximum three.  
If a student chooses yes for the 4670 option, then must submit it by individual.

## Other Information

Academic Integrity and Honesty (<http://www.ust.hk/vpao/ug-guide/integrity/>)

Learning Environment ([http://www.ust.hk/vpao/conduct/good\\_learning\\_experience.pps](http://www.ust.hk/vpao/conduct/good_learning_experience.pps))

Center for Language Education (<http://cle.ust.hk/>)

Math Support Center (<https://www.math.ust.hk/~support/>)

## Course Materials

### Books:

1. *Econometrics* (2019) by Bruce Hanse (Hansen)  
(<http://www.ssc.wisc.edu/~bhansen/econometrics/>)
2. *Mostly Harmless Econometrics* (2008) by Joshua Angrist and Jorn-Steffen Pischke. (Harmless)
3. *Discrete Choice Methods with Simulation* (2003) by Kenneth Train (Train)  
(<https://eml.berkeley.edu/books/choice2.html>)
4. *Data Analysis and Graphics Using R* (2010), by John Maindonald and John Braun.  
(<http://maths-people.anu.edu.au/~johnm/r-book/daagur3.html>)
5. *Applied Econometrics with R* (2008) by Achim Zeileis and Christian Kleiber.

### Papers:

1. David, H., Alan Manning, and Christopher L. Smith. "The Contribution of the Minimum Wage to US Wage Inequality Over Three Decades: a Reassessment." *American Economic Journal: Applied Economics* (2016).
2. Mankiw, N. G., David Romer, and David Weil. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* (1992).
3. Card, David, and Alan B. Krueger. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* (1994).

4. Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* (2001).
5. Islam, Nazrul. "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics* (1995).
6. Angrist, Joshua D., and Alan B. Krueger. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* (1991).
7. Angrist, Joshua D. "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants." *Econometrica* (1998).
8. Shum, Matthew. "Estimating Demand in Discrete-Choice Differentiated Product Markets". (<http://people.hss.caltech.edu/~mshum/gradio/china1.pdf>)
9. Petrin, Amil. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy* (2002).
10. Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica* (1995).

## Outline and Schedule

### Module 1. Programming boot camp (6 lectures)

- Introduction to R Programming
  - Basic commands
  - Matrix algebra
  - Function
  - Loop
  - Plot
  - Solving mathematical problem by computer programming

Readings: Hansen Appendix A

- Ordinary least square (OLS)
  - Linear regression and OLS
  - Unbiasedness and consistency
  - Monte-Carlo simulation

Readings: Hansen 3, 4

### Problem set 1

### Module 2. Linear regression (5 lectures)

- Data
  - Importing/exporting data
  - Organizing and visualizing data
  - Summary statistic table

Readings: paper 1, 2, 3

- Linear Regression
  - Coefficient estimation and standard error
  - Hypothesis tests
  - Model specification and diagnostic
  - Reporting estimation results in tables

Readings: Hansen 5, 7; paper 6

### Problem set 2

### **Module 3. Causal inference** (5 lectures)

- Causal inference
  - Endogeneity and causality
  - Treatment effect
  - Difference-in-difference (DID)
  - Instrumental variable (IV)
  - Two-stage least square (TSLS)
  - Using Monte-Carlo simulation: bootstrap for standard error

Readings: Harmless 1, 2, 3, 4, 5; Hansen 12; paper 3, 4, 6

- Panel data
  - Structure of panel data
  - Fixed effect estimation
  - Random effect and Hausman test
  - Time series\*

Readings: Hansen 16; paper 5, 7

### Problem set 3

### **Start final project!!**

### **Module 4. Discrete choice model** (4 lecture)

- Extremum estimators
  - Use of “optim” function
  - Non-linear least square (NLS)
  - Maximum likelihood estimation (MLE)
  - Generalized method of moments (GMM)

Readings: Hansen 5, 20

- Limited dependent variable and discrete choice
  - Limited dependent variable
  - Binary outcome: Logit and Probit
  - Multinomial choice
  - Ordered choice
  - Count data
  - Demand estimation based on discrete choice

Readings: Train 1, 2, 3, 4, 13; paper 8, 9, 10

#### Problem set 4

**Programming exam** (Lecture time on April 26<sup>th</sup>, subject to adjustment)

#### **Module 5. Selected topics** (2 lecture)

- Bootstrap method
  - Find critical value of test statistic
  - Find standard error of estimator

Readings: Hansen 10.

- Nonparametric estimation
  - Estimate CDF and PDF
  - Nonparametric regression

Reading: Hansen 18.

- Big data
  - High-dimensional data and Lasso
  - Tree-based classification
  - Clustering

#### Problem set 5

**Final project** is due in the final exam period.