Course Syllabus

ECN 301E
Econometrics I
FALL 2024

Instructor: Osman DOGAN.
Email: osmandogan@itu.edu.tr.
Office Location: TBA.
Office Hours: Wednesday, 15:00-17:00.
Meeting Time/Classroom: Tuesday 14:30-18:29 / MSB-101.
TA: Gülen Derya Zayim Gider, zayim17@itu.edu.tr.

Course Description: This course is a one-semester undergraduate level study of the theory and practice of econometrics. In this course, we will cover the following topics: (i) probability and statistics needed for econometrics, (ii) linear regression with single and multiple regressors, (iii) hypothesis tests and confidence intervals for linear regression models, (iv) nonlinear regression functions, (v) assessing studies based on multiple regressions, and (vi) regression models with binary dependent variables. We will strike a balance between theory and application. Upon completion of the course, students will be ready to perform basic regression analysis and read simple applied work in the literature.

Course Objectives: The objectives of this course consist of the followings.

1. Students will understand the fundamentals of econometric modeling.
2. Students will master the foundations of the linear regression analysis.
3. Students will understand what is meant by *causal* analysis and limitations to it using observational data.
4. Extensions of the classical linear regression model will be explored.
5. Potential problems in linear regression analysis such as threats to internal validity and external validity will be explored.
6. Regression models with binary dependent variables will be explored.

Textbook information: The main reference for this course is the following textbook.

> James H. Stock and Mark W. Watson. *Introduction to Econometrics*. Pearson, fourth edition, 2020.

Assignments and Grading: There will be 10 to 13 problem sets throughout the semester. You may collaborate on the problem sets, but each student must submit their own separate solution. Submissions that are highly similar will not be accepted and will be considered plagiarism. Assignments must be submitted through the Ninova system, and late submissions will not be accepted. There are two exams: (i) a midterm exam and (ii) a final exam. Your course grade will be determined by your performance on the problem sets (15 percent), class participation (5 percent), the midterm exam (40 percent), and the final exam

(40 percent). The midterm exam is scheduled for the **7th week** during regular lecture hours.

Attendance and Absences: Attendance is expected and will be taken each class. Students are responsible for all missed work, regardless of the reason for absence. It is also the absentee's responsibility to get all missing notes or materials.

Computations: In this class, we will use R and Python for statistical computations, with a *primary focus* on R implementations. These languages are open-source, high-level programming languages that are widely used for statistics and econometrics. R can be downloaded from http://www.R-project.org, and Python can be installed by downloading Anaconda from https://www.anaconda.com/download. Lecture notes and related scripts will be available to students during lab sessions. There are no required textbooks for R and Python; however, the following sources are recommended as references:

1. Christoph Hanck, Martin Arnold, Alexander Gerber, and Martin Schmelzer. *Introduction to Econometrics with R*. 2021. URL https://www.econometrics-with-r.org/index.html.

2. H. Wickham and G. Grolemund. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, second edition, 2022. URL https://r4ds.hadley.nz/.

3. Florian Heiss. *Using R for Introductory Econometrics*. 2020. URL http://www.urfie.net/downloads/PDF/URfIE_web.pdf.

4. Florian Heiss and Daniel Brunner. *Using Python for Introductory Econometrics*. 2020. URL https://www.urfie.net/downloads/PDF/UPfIE_web.pdf.

5. Kevin Sheppard. *Introduction to Python for Econometrics, Statistics and Data Analysis*. Fifth edition, 2021. URL https://www.kevinsheppard.com/files/teaching/python/notes/python_introduction_2021.pdf.

6. Wes McKinney. *Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter*. OReilly Media, third edition edition, 2022. URL https://wesmckinney.com/book/.

Lab Session: The "Lab session" is a one-hour study session that starts immediately after the lecture. In the lab sessions, the fundamentals of R and Python programming will be covered, along with solutions to some of the questions in the assignments.

Course Policies:

- Exams are cumulative and may cover any of the material presented in the course before the exam date.
- No makeup exams will be given except in the case of an emergency, documented with either a doctor's note or a letter from the chair of the department.
- You can read on the academic regulations at https://www.sis.itu.edu.tr/TR/mevzuat/mevzuat.php.

Guidelines for AI Tools: The use of artificial intelligence (AI) language models, such as ChatGPT, can be a valuable tool for students in this course, but it is important to use it ethically and appropriately. The following guidance is provided to assist you in using the AI tools.

- The AI-generated texts cannot be used in your theses and assignments, as doing so would be considered plagiarism. Tools such as AI, Turnitin, and Copyleaks can identify the AI-generated text.
- Please note that AI models are machine learning models and may not always offer the most accurate and relevant information. You should critically evaluate the text provided by AI and use other scientific sources for verification. All sources should always be cited properly.

Important Dates: Sept 30 semester starts, Sept 30–Oct 4 adding and dropping a class, Nov 18-Nov 22 fall semester break, Jan 10 fall semester ends, Jan 13–26 final exams.

Tentative Course Outline

1. **Introduction to Econometric Analysis**
   (a) What is econometrics?
   (b) What type of questions econometric analysis can help answer?
   (c) Causal analysis or causal effect
   (d) Experimental data and observational data
   (e) Different kinds of data sets: Cross-sectional data, time series data and panel data
   (f) Summation and product operators
   (g) Some useful results from linear algebra

   Readings: Stock and Watson [2020, Chapter 1].

2. **Review of Probability**
   (a) Random variable and probability distributions
   (b) Expected values, mean and variance
   (c) The normal, chi-squared, student t and F distributions
   (d) Random sampling and sampling distributions

   Readings: Stock and Watson [2020, Chapter 2].

3. **Review of Statistics**
   (a) Estimation of population mean
   (b) Hypothesis tests and confidence intervals for the population mean
   (c) Comparing means from different populations
   (d) Scatter plots, sample covariance and sample correlation

Readings: Stock and Watson [2020, Chapter 3].

4. **Linear regression with one regressor**

   (a) Estimating the coefficients of the linear regression model
   (b) Measure of fit and prediction accuracy
   (c) The least squares assumptions for causal inference
   (d) The sampling distribution of the OLS estimator

   Readings: Stock and Watson [2020, Chapter 4].

5. **Hypothesis tests and confidence intervals: Part 1**

   (a) Testing hypotheses about one of the regression coefficients
   (b) Confidence intervals for a regression coefficient
   (c) Regression with dummy variables
   (d) Heteroskedasticity and homoskedasticity
   (e) Theoretical foundation of ordinary least squares
   (f) Applications with the test score data

   Readings: Stock and Watson [2020, Chapter 5].

6. **Linear regression with multiple regressors**

   (a) Omitted variable bias
   (b) The multiple linear regression model
   (c) The least squares assumptions for causal inference in multiple regression
   (d) Multicollinearity
   (e) Control variables and conditional mean independence
   (f) Applications with the test score data

   Readings: Stock and Watson [2020, Chapter 6].

7. **Hypothesis tests and confidence intervals: Part 2**

   (a) Hypothesis tests and confidence intervals for a single coefficient
   (b) Tests of joint hypotheses
   (c) Testing single restrictions involving multiple coefficients
   (d) Confidence sets for multiple coefficients
   (e) Applications with the test score data

   Readings: Stock and Watson [2020, Chapter 7].

8. **Nonlinear regression functions**

   (a) Modeling nonlinear regression functions: polynomials and logarithms
   (b) Interaction between independent variables

(c) Applications with the test score data

Readings: Stock and Watson [2020, Chapter 8].

9. **Assessing studies based on multiple regression**

   (a) Internal and external validity
   (b) Threats to internal validity of multiple regression analysis
   (c) Internal and external validity when the regression is used
   (d) Applications with the test score data

   Readings: Stock and Watson [2020, Chapter 9].

10. **Regression with a binary dependent variable**

    (a) Linear probability model
    (b) Probit and logit models
    (c) Estimation and inference in the logit and probit models
    (d) Application to the Boston HMDA data

    Readings: Stock and Watson [2020, Chapter 11].

References

Christoph Hanck, Martin Arnold, Alexander Gerber, and Martin Schmelzer. *Introduction to Econometrics with R*. 2021. URL https://www.econometrics-with-r.org/index.html.

Florian Heiss. *Using R for Introductory Econometrics*. 2020. URL http://www.urfie.net/downloads/PDF/URfIE_web.pdf.

Florian Heiss and Daniel Brunner. *Using Python for Introductory Econometrics*. 2020. URL https://www.urfie.net/downloads/PDF/UPfIE_web.pdf.

Wes McKinney. *Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter*. OReilly Media, third edition edition, 2022. URL https://wesmckinney.com/book/.

Kevin Sheppard. *Introduction to Python for Econometrics, Statistics and Data Analysis*. Fifth edition, 2021. URL https://www.kevinsheppard.com/files/teaching/python/notes/python_introduction_2021.pdf.

James H. Stock and Mark W. Watson. *Introduction to Econometrics*. Pearson, fourth edition, 2020.

H. Wickham and G. Grolemund. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, second edition, 2022. URL https://r4ds.hadley.nz/.