Introduction to Econometrics [EC 320]

Winter 2022 Syllabus

Philip Economides
Department of Economics
University of Oregon

Lecture
- Chapman 220
- MW, 1200–1320

Lab
- McKenzie 442
- See DuckWeb

Materials
- Mastering 'Metrics
- Introduction to Econometrics, 5th ed.
- **R**
- **RStudio**

Instructor
- Philip Economides
- philip-economides.com
- peconomi@uoregon.edu
- PLC 520
- T 1500–1600, Th 1000–1100, or by appointment

GE
- TA 1 Name
- ta1@uoregon.edu
- PLC 520
- M 1400–1600

GE
- TA 2 Name
- ta2@uoregon.edu
- PLC 520
- T 1200-1400

Course summary

Description: This course introduces the statistical techniques that help economists learn about the world using data. We will focus much of our attention on regression analysis, the workhorse of applied econometrics. Using calculus and introductory statistics, we will cultivate a working understanding of the theory underpinning regression analysis—*how* it works, *why* it works, and, *when it can lead us astray*. We will apply the insights of theory to work with and learn from actual data using R, a statistical programming language. To the extent that you invest the requisite time and effort, you can leave this course with marketable skills in data analysis and—most importantly—a more sophisticated understanding of the notion that **correlation does not necessarily imply causation**.

Prerequisites: Math 242 (Calculus) and Math 243 (Introduction to Statistics) or equivalent.

Software

- We will use the statistical programming language **R**.
- We will use **RStudio** to interact with R.

Learning R is challenging, but well worth the effort. R is a powerful and versatile tool for data analysis and visualization, which makes it popular among employers. If you dedicate the time and effort necessary to learn the language, you are likely to reap a handsome return on the job market. The SSIL lab in McKenzie has R and RStudio installed and ready for you, but I strongly recommend that you install these programs on your own computer. Don't worry, **both are free**. I also recommend that you save your scripts, data, and assignments separately from your preferred computing hardware. I'd suggest either using a portable flash drive or a Github repositories. Alternatively, you can use the class network drive (the "R drive"), which is available on all university computers. I will be making material available through **Github** for convenient one-click downloading and those who lose their learning material thereafter.

If you are concerned about learning R—or you want to learn quicker—I recommend that you check out the following free online resources:

- DataCamp's *Introduction to R*
- Hadley Wickham's *Introduction to R*

The RStudio team has also assembled a useful set of resources.

Textbooks

Required: There are two required textbooks for this course:

1. **Mastering 'Metrics: The Path from Cause to Effect** by Angrist and Pischke (**MM**)
2. **Introduction to Econometrics**, 5th ed. by Christopher Dougherty (**ItE**)

You can purchase them at the Duckstore or your preferred online bookseller. You should complete the assigned readings from the textbooks *before* lecture. Attending lecture is not a substitute for reading and comprehending the texts. Likewise, reading is not a substitute for attending lecture. The lectures and the readings are meant to *complement* one another. The tentative course schedule (further below) lists the assigned readings for each topic.

In addition to the textbook readings, I may occasionally assign readings from peer-reviewed studies for classroom discussion. I will post these readings on Canvas.

Optional: There is a wealth of free online books for learning R. A classic is Garrett Grolemund and Hadley Wickham's ***R* for Data Science**. If you have previous experience coding in R, you may want to check out Hadley Wickham's **Advanced *R***. If you are interested in producing beautiful and informative graphs or maps, see Kieran Healy's **Data Visualization: A Practical Introduction**. Our very own **Grant McDermott** and **Ed Rubin** also maintain advanced resources for learning R.

Course Structure

Grades

I will award grades based on your relative performance in the class, as determined by the following weights:

Problem Sets	30%
Quizzes	10%
Midterm Exam	20%
Data Assignment	10%
Final Exam	30%

Problem Sets

I will assign **five** bi-weekly problem sets throughout the quarter. Each problem set will include an analytical component and a computational component.

- I will announce due dates in class.
- You will turn in an **electronic copy** of each problem set on Canvas.
- Presentation matters. The aim is to be ready to use your skills in practical settings. 10 percentage points of each grade will be attributed to producing professional work (*e.g.*, clear language, typed equations, tasteful fonts, tables and graphs with informative labels, *etc.*).

I encourage you to work together on the problem sets. Unless explicitly stated, **each student is required to write and submit independent answers**. I will take word-for-word and code-for-code copies as evidence of academic dishonesty. If you work with others, list their names at the top of your assignment. If you fail to list your collaborators, you will receive a score of zero.

Quizzes

I will assign **two** quizzes during the term, aimed at testing your knowledge on the recent content we've covered. This maps closely to the recommendation that you read prescribed material before class. During the lecture, make sure you're asking questions about any particular topic, should the added explanation of material not be proving to be sufficient. These two quizzes will cover two separate portions of course content recently addressed in class.

Data Assignment

I will assign a single data assignment to the class, in which I will make a number of separate datasets available. These datasets will be available from the beginning of term, but I would recommend you do not commence work until after your lab on data wrangling with R's dplyr package. If you wish to use a separate dataset or merging of datasets, email me in advanced to ensure it is an appropriate challenge for the class' requirements. Upon completion your data assignment should include the following sections;

- Introduction: Describe the data source type (panel, cross-sectional, etc.), its frequency (hourly, annual, etc.), and a bit of flavor about how representative the sample is relative to the population.

- Cleaning: Ultimately, you want to present this data in a clear and legible manner for its intended audience. If you are going to be performing regressions, this may require identifying outliers and treating for missing observations. List the challenges you encountered when preparing this data for its intended audience. Outline some of the limitations the data may face upon tackling these challenges.
- Summary Statistics: This is the most creative portion of the assignment. While you are not addressing causal inference in this assignment, your presentation of the data may still hint at potential research questions underlying the material. Present a series of interesting facts about the data. Infer potential causal relationships and describe how you may go about testing those relationships.
- Conclusion: For the more time-constraint reader, perhaps a future employer, you always want to be concise with your statements. Use this section to provide concluding remarks about the data, detailing how rich it appears to be and the potential uses it may provide in terms of empirical analysis.

Exams

For each exam, I will generate a randomized seating chart. During the exams, you may use a writing utensil, a non-programmable calculator, and a 3-by-5-inch notecard. As you turn in your exam, I will ask you to present your student ID. I do not give makeup exams. See the course policy on makeup assignments for more information.

Lab

In your weekly lab section, you will learn to apply the concepts discussed in lecture using R. While the lab may include some general econometrics instruction, the main focus is on the practical application of statistical techniques and working through the computational components of the bi-weekly problem sets. Attending lab is crucial for learning the material and passing the course. Due to space constraints, **you must attend the lab for which you registered**.

Course Policies

Late Policy

I will not accept late problem sets after the due date. If you turn in a problem set on the due date, but after the deadline, points will be deducted for lateness. If you turn it in after I post the key, you will receive a zero.

Makeup Assignments

I do not give makeup assignments. This blanket ban extends to exams. In extreme circumstances that lead you to miss one of the midterm exams—such as death in the family or grave illness or injury—I will consider re-weighting your grade toward the final. To qualify for re-weighting, you will need to notify me no later than two days after the exam.

Grade Appeals

You must submit any request for re-grading in writing within one week of the day grades are posted for the problem set or exam in question. Your request should include a cogent argument explaining why your responses warrant full credit.

Etiquette

Please respect those around you by turning off your phone and other potentially distracting devices. I ask that you stay for the entire lecture: getting up and leaving distracts your fellow classmates. If you must leave early, please position yourself near the door when you get to class. As a final note, a growing body of evidence suggests that using laptops in lecture reduces comprehension and recollection. In light of this evidence, I ask that you refrain from using your laptop during lecture. As a practical matter, it is much easier to write math by hand than it is to type it.

Academic Integrity

I will not tolerate cheating, plagiarism, and other violations of the Student Conduct Code. If you are caught cheating or plagiarizing on any component of this course, you will receive a failing grade for the term and I will report your offense to the university.

Accommodations

Notify me if there are aspects of this course that pose disability-related barriers to your participation. If you require special accommodations for a documented disability, then you will need to provide me a letter from the Accessible Education Center (AEC) that verifies your need and details the appropriate accommodations. Please make arrangements with the AEC by the end of Week 1. If your accommodations include exam proctoring at the AEC, then you are responsible for scheduling those exams with the AEC *at least seven days in advance*.

Tentative Schedule

Lectures and Exams

Week	Date	Topic	Reading
01	10/02	Introduction	
02	10/07	Statistics Review I	ItE Review
02	10/09	Statistics Review II	ItE Review; MM 1 (appendix)
03	10/14	The Fundamental Econometric Problem	MM 1
03	10/16	The Logic of Regression	MM 2
04	10/21	Midterm I Review	
04	10/23	**Midterm Exam I: High Concepts** (in-class)	
05	10/28	Simple Linear Regression: Estimation I	ItE 1
05	10/30	Simple Linear Regression: Estimation II	ItE 1
06	11/04	Classical Assumptions	ItE 1
06	11/06	Simple Linear Regression: Inference	ItE 2
07	11/11	Multiple Linear Regression: Estimation	ItE 3, 6.2; MM 2 (appendix)
07	11/13	Multiple Linear Regression: Inference	ItE 3, 6.3; MM 2 (appendix)
08	11/18	Midterm II Review	
08	11/20	**Midterm Exam II: The Weeds** (in-class)	
09	11/25	Categorical Variables	ItE 5
09	11/27	Interactive Relationships	ItE 4
10	12/02	Nonlinear Relationships	ItE 4
10	12/04	Final Review	Discussion paper (see Canvas)
11	12/10	**Final Exam** (see final exam schedule)	

Labs and Problem Sets

Week	Date	Topic
01		Lab: Introduction to R and `RStudio`
02		Lab: Introduction to the `tidyverse`
03	10/14	**Problem Set 1** (due on Canvas by 5pm)
03		Lab: Non-Experimental Data Analysis
04	10/22	**Problem Set 2**: *Analytical Problems* (due on Canvas by 5pm)
04	10/25	**Problem Set 2**: *Computational Problems* (due on Canvas by 5pm)
04		Lab: Introduction to `Rmarkdown`
05		Lab: Regression Analysis
06	11/06	**Problem Set 3** (due on Canvas by 11:59pm)
06		Lab: Hypothesis Testing and Confidence Intervals
07		Lab: Omitted-Variable Bias and Hypothesis Testing
08	11/18	**Problem Set 4** (due on Canvas by 5pm)
08		Lab: Maps with `ggplot2`!
09		**No Lab**: *Thanksgiving*
10	12/02	**Problem Set 5** (due on Canvas by 5pm)
10		Lab: Marginal Effects