Econ 320: Econometrics

(Last updated: 28 October 2019)

Instructor: Prof. Dr. Stephen D. O'Connell, PAIS 575, e-mail: soconnell@emory.edu*
TA: Yushan Liu, yushan.liu@emory.edu

Time and Location: Lecture – Tuesday/Thursday 02:30 pm – 03:45 pm, Rich Memorial 104,
Laboratory - Friday TBD.

Note: Classes begin on 08/29/2019 and end on 12/05/2019. Classes do not meet on 10/15/2019 or 11/28/2019. In-class exams are scheduled for 10/03/2019 and 11/14/2019, the final is scheduled for 12/18/2019 (8:00 AM to 10:30 AM).

Office Hours: Tu/Th 01:00 pm – 02:00 pm, W 12-1PM; PAIS 575.
Please make appointments for office hours @ https://calendly.com/soconnell-emory/
If appointments are booked, you may still drop in as I will still be in my office. Appointments are used only for assigning priority to students when there is excess demand.

Pre-requisite: Students must have taken Econ 220 – Introduction to Statistical Methods before enrolling in this course.

Course Objective: This course is designed to introduce students to the statistical foundations of Econometrics, its well-known multivariate linear model, and its usage for causal inference. It will equip students with ground understanding of the fundamentals probability theory and mathematical statistics behind most econometric methods. Machine learning terminology is also introduced in the context of the predictability power of the multiple linear regression model. The main concepts of asymptotic theory, and its application to the classical linear regression model are also covered. After taking this course, students should be able to:

1. Manipulate economic data sets.
2. Diagnose certain problems with linear models and data, and know how to remedy them.
3. Have a working knowledge of the classical linear regression model and its applicability.

Textbook (required):
Introduction to Econometrics (any edition 3+), by Jeffrey Wooldridge.
(Note: you do not need to buy the latest edition with online access bundle, etc.)

Software: Students will have the opportunity to familiarize themselves with the use of powerful and widespread econometric software, R, in the lab session. R is available on computers in the Econ lab (Rich 301), Woodruff Library, and Cox Hall.

Canvas: All announcements, syllabus, home assignments, lecture notes, data files, and other course-related material will be posted on Canvas. You should check the canvas website regularly for updates.

Grading: The final grade will be determined by a weighted average of scores in 1st written (closed-book) in-class exam (20% on 10/03/2019, during class time), 2nd written (closed-book) in-class exam (20% on

11/14/2019, during class time), homework (15% throughout the semester), the lab component (25%), and a written closed-book *comprehensive* final exam (20% on 12/18/2019, 8:00 AM to 10:30 AM, in Rich 104). There will be four unannounced quizzes given throughout the semester. These quizzes are short, graded, and a passing grade on each quiz will confer *an additional one point on the student's final grade*.

Grade	Lower Limit	Upper Limit
A	95	100
A-	90	94.9
B+	87	89.9
B	84	86.9
B-	80	83.9
C+	77	79.9
C	74	76.9
C-	70	73.9
D+	65	69.9
D	60	64.9
F	0	59.9

Note: You need a grade above 60% (a passing grade) in the other weighted components of the course for your lab component to be added to your overall grade (with its respective weight). If this required 60% minimum is not achieved, your grade in the lab portion will not be added to your overall grade and you effectively earn the failing grade of F. In these circumstances, your grade in the lab portion of the class can be saved for the next time you take the class (within 1 academic year). You just need to notify the lab instructor via e-mail (copy to the undergraduate program director) to save your lab grade, keep a record of this e-mail, and forward it to your new class instructor when you retake Econ 320. You can also retake the lab along with the rest of the course if you want to improve your grade from the previous time you took the class, but you must retake the lab session if you have waited more than 1 academic before enrolling in Econ 320 since failing it the last time.

Exams: Each in-class midterm exam will last exactly 75 minutes. There may be some element of choice in the questions asked, although this is not guaranteed. Material covered in exams is cumulative. Only handheld calculators are allowed.

Home Assignments: There will be 5 assignments to be completed independently and submitted via Canvas.

Policies

General: Students are expected to adhere to the Emory College Honor Code as well as its Conduct Code, see http://college.emory.edu/home/academic/policy/conduct_code.html. Specifically, the honor code is in effect throughout the semester. By taking this course, you affirm that it is a violation of the code to cheat on exams, to plagiarize, to deviate from the teacher's instructions about collaboration on work that is submitted for grades, to give false information to a faculty member, and to undertake any other form of academic misconduct. You agree that the teacher is entitled to move you to another seat during examinations, without explanation. You also affirm that if you witness others violating the code you have a duty to report them to the honor council.

Special circumstances: Students requiring any type of special classroom/testing accommodation for a disability, religious belief, scheduling conflict, or other impairment that might affect his or her successful completion of this course must personally present the requested remedy or other adjustment in written form (signed and dated) to the instructor, i.e. supporting memorandum of accommodation from the Office of Disabilities Services, http://www.ods.emory.edu/. Requests for accommodations must be received and authorized by the instructor in written form no less than two weeks in advance of need. No accommodation should be assumed unless so authorized. In the event of needs identified later in the course, or for which an adjustment cannot be made on a timely basis, a grade of "I," Incomplete for the course, will be given to accommodate the unanticipated request.

Exam absences & missed assignments: Emory College of Arts and Sciences does not have an attendance policy and, therefore, does not provide absence excuses. In the event of a catastrophic (and documented) occurrence which necessitates an absence from a scheduled exam (only), the student should immediately seek help from the Office of Undergraduate Education (OUE), http://college.emory.edu/home/administration/office/undergraduate/index.html. The Family Educational Rights and Privacy Act (FERPA) and the Health Insurance Portability and Accountability Act of 1996 (HIPPA) regulations (U.S. Department of Health and Human Services and U.S. Department of Education) dictate that students do not have to provide medical documentation or disclose personal/medical issues with professors. However, the OUE class deans and academic advisers may collect this documentation and could provide verification to professors upon students' requests. *This must be done within 48 hours after missing the exam*. If approval is granted by the instructor, the weight of the student's scores for the missed exam or assignment will be transferred to the *next* chronologically scheduled exam or assignment. If a letter (or e-mail) from the OUE is not received by the instructor within 48 hours, or approval is not explicitly obtained from the instructor, after a missed exam, the missed exam will receive a score of zero points.

In case the student obtains a verification letter (or e-mail) from the OUE for the final exam, and the instructor's approval is also granted, the weight of the student's scores for the missed final exam will be transferred to the previous exam where the student scored the *lowest*.

It is understood that any student seeking a verification letter (or e-mail) from the OUE, is also *forfeiting all extra points* earned throughout the duration of the course, including those from quizzes or other bonus points given at the discretion of the professor.

⚠ Regardless of whether a student can obtain a verification letter from the OUE for more than one missed exam (midterm or final), all students will need to have obtained marks in at least two written examinations to obtain a final grade for the course. Students failing to fulfill this minimum requirement will receive an automatic "I" (incomplete) for the course. The instructor will change this "I" to a grade once the student successfully takes *all* missed written examinations the next time the instructor teaches the course, i.e. Spring 2020.

For full information about Emory's policy on academic misconduct, please check the following links: http://college.emory.edu/home/academic/policy/honor_code.html (see Articles 4, 5, 6 & 7)

Class attendance: Attendance is not mandatory, but encouraged. If a student misses a lecture, the student should not expect the instructor to repeat the material at another time (such as the instructor's office hours).

In/Out of classroom conduct: Students are expected to adhere to the Emory University Code of Conduct, see http://conduct.emory.edu/policies/code/index.html.

- Use of cell/smartphones during lecture time *and* exams is always prohibited, as is leaving the room to answer or make a call.

  ⚠ Students planning to use laptops and/or tablets during lecture time *may not sit in the first row* of the classroom.

  ⚠ Only scientific handheld calculators are permitted during exams.

- Basic classroom etiquette requires that you be quiet and attentive in class. This means that, except when we are engaged in-group discussion, only one person will be talking in the classroom at any time. Everyone in the room will give respectful attention to the sole speaker. Students creating disruptions in class will be asked to leave the class and will still be responsible for any material missed. Continued disruptions will be dealt with in accordance with university policy.
- Any level of 'chatting' with friends, sleeping, reading newspapers, leaving early, and/or eating in class are disruptive and rude to your classmates and me. If for some reason you have to leave early or arrive late on a particular day, please sit in the back of the class, leave or arrive quietly, and notify me via e-mail as far in advance as possible. This will be allowed as long as it is of an infrequent nature. Use the time in class wisely, appropriately, and efficiently.
- You are welcome to drop by my office during office hours and by appointment only. If you need to send me an e-mail, please remember that e-mail is a means of communication and you should be always respectful in your writings. I reserve the right to either answer your query via e-mail (within 24 hours after receiving it) or in person (by asking you to set up an appointment to meet later on).

Outline

Basic Mathematical Tools (Appendix A)

1) The Summation Operator and Descriptive Statistics (Appendix A.1).
2) Properties of Linear Functions (Appendix A.2).
3) Proportion and Percentages (Appendix A.3).
4) Some Special Functions and their Properties (Appendix A.4).
    i. Quadratic Functions.
    ii. The Natural Logarithm.
    iii. The Exponential Function.
5) Differential Calculus (Appendix A.5).

The Nature of Econometrics and Economic Data (Chapter 1)

1) What Is Econometrics?
2) Steps in Empirical Economic Analysis.
3) The Structure of Economic Data.
    i. Cross-Sectional Data.
    ii. Time Series Data.
    iii. Pooled Cross Sections.
    iv. Panel or Longitudinal Data.
4) Causality and the Notion of Ceteris Paribus in Econometric Analysis.

The Simple Regression Model (Chapter 2)

1) Definition of the Simple Regression Model.
2) Deriving the Ordinary Least Squares Estimates.
3) Properties of OLS on Any Sample of Data.
    i. Fitted Values and Residuals.
    ii. Algebraic Properties of OLS Statistics.
    iii. Goodness-of-Fit.
4) Units of Measurement and Functional Form.
    i. The Effects of Changing Units of Measurement on OLS Statistics.
    ii. Incorporating Nonlinearities in Simple Regression.
    iii. The Meaning of "Linear" Regression.
5) Expected Values and Variances of the OLS Estimators.
    i. Unbiasedness of OLS.
    ii. Variances of the OLS Estimators.
    iii. Estimating the Error Variance.

Multiple Regression Analysis: Estimation (Chapter 3)

1) Motivation for Multiple Regression.
    i. The Model with Two Independent Variables.
    ii. The Model with *k* Independent Variables.
2) Mechanics and Interpretation of Ordinary Least Squares.
    i. Obtaining the OLS Estimates.
    ii. Interpreting the OLS Regression Equation.
    iii. On the Meaning of "Holding Other Factors Fixed" in Multiple Regression.
    iv. Changing More Than One Independent Variable Simultaneously.

v. OLS Fitted Values and Residuals.
vi. A "Partialling Out" Interpretation of Multiple Regression.
vii. Comparison of Simple and Multiple Regression Estimates.
viii. Goodness-of-Fit.
ix. Regression through the Origin.

3) The Expected Value of the OLS Estimators.
   i. Including Irrelevant Variables in a Regression Model.
   ii. Omitted Variable Bias: The Simple Case.
   iii. Omitted Variable Bias: More General Cases.
4) The Variance of the OLS Estimators:
   i. The Components of the OLS Variances: Multicollinearity.
   ii. Variances in Misspecified Models.
   iii. Estimating σ^2: Standard Errors of the OLS Estimators.
5) Efficiency of OLS: The Gauss–Markov Theorem.

Multiple Regression Analysis: Inference (Chapter 4)

1) Sampling Distributions of the OLS Estimators.
2) Testing Hypotheses about a Single Population Parameter: The *t* Test.
   i. Testing against One-Sided Alternatives.
   ii. Two-Sided Alternatives.
   iii. Testing Other Hypotheses about β_j.
   iv. Computing *p*-Values for *t* Tests.
   v. A Reminder on the Language of Classical Hypothesis Testing.
   vi. Economic, or Practical, versus Statistical Significance.
3) Confidence Intervals.
4) Testing Hypotheses about a Single Linear Combination of the Parameters.
5) Testing Multiple Linear Restrictions: The *F* Test.
   i. Testing Exclusion Restrictions.
   ii. Relationship between *F* and *t* Statistics.
   iii. The R-Squared Form of the *F* Statistic.
   iv. Computing *p*-Values for *F* Tests.
   v. The *F* Statistic for Overall Significance of a Regression.
   vi. Testing General Linear Restrictions.

Multiple Regression Analysis: OLS Asymptotics (Chapter 5)

1) Consistency.
   i. Deriving the Inconsistency in OLS.
2) Asymptotic Normality and Large Sample Inference.
   i. Other Large Sample Tests: The Lagrange Multiplier Statistic.
3) Asymptotic Efficiency of OLS.

Heteroskedasticity (Chapter 8)

1) Consequences of Heteroskedasticity for OLS.
2) Heteroskedasticity-Robust Inference after OLS Estimation.
   i. Computing Heteroskedasticity-Robust LM Tests.
3) Testing for Heteroskedasticity.
   i. The White Test for Heteroskedasticity.

4) Weighted Least Squares Estimation.
    i. The Heteroskedasticity Is Known up to a Multiplicative Constant.
    ii. The Heteroskedasticity Function Must Be Estimated: Feasible GLS.
    iii. What If the Assumed Heteroskedasticity Function Is Wrong?

Multiple Regression Analysis with Qualitative Information: Binary (or Dummy) Variables (Chapter 7)

1) Describing Qualitative Information.
2) A Single Dummy Independent Variable.
    i. Interpreting Coefficients on Dummy Explanatory Variables When the Dependent Variable Is *log(y)*.
3) Using Dummy Variables for Multiple Categories.
    i. Incorporating Ordinal Information by Using Dummy Variables.
4) Interactions Involving Dummy Variables.
    i. Interactions among Dummy Variables.
    ii. Allowing for Different Slopes.
    iii. Testing for Differences in Regression Functions across Groups.
5) A Binary Dependent Variable: The Linear Probability Model.
6) More on Policy Analysis and Program Evaluation.

Modern Empirical Research Design (Various; if time)

1) Motivation: Omitted Variables in a Simple Regression Model.
    i. Statistical Inference with the IV Estimator.
    ii. Properties of IV with a Poor Instrumental Variable.
    iii. Computing R-Squared after IV Estimation.
2) Regression and the experimental ideal (AP CH 1-2).
3) Regression discontinuity (AP CH 3).
4) Difference-in-differences (AP CH 4).
5) IV Estimation of the Multiple Regression Model (Chapter 15).
6) Two Stage Least Squares (Chapter 15).
    i. A Single Endogenous Explanatory Variable.
    ii. Multicollinearity and 2SLS.
    iii. Multiple Endogenous Explanatory Variables.
    iv. Testing Multiple Hypotheses after 2SLS Estimation.

Specific Schedule

Week	Day	Date	Reading Due	Problem set due (midnight prior)	Description
1	1	29-Aug			General explanation of the syllabus, the course policies, expectations, etc. Cover Appendix A's slides, i.e., definition of a sequence, summation properties, definition of what a linear function is (both in parameters and in its argument x), extension of the linear function to be dependent of two arguments - x1 and x2, definition of marginal change in these two settings. Covered logs, exponentials, quadratic
1	2	3-Sep-19	Wooldridge Appendix A, B, C		Solve exercise 10 in the Appendix A Chapter. The first part of Chapter 1 is covered, i.e., definition of Econometrics, the idea of an economic vs econometric model. I will also discuss the 4 types of data sets usually encountered in Economics. My notes shows R code replicating some of
2	1	5-Sep	Chapter 1		Chapter 1 completed and start Chapter 2. Specifically, the simple linear regression model is introduced. The class end by explaining E[u|x]=0.
2	2	10-Sep-19	Chapter 2		Method of moments estimator derivation
3	1	12-Sep-19	Chapter 2	PS 1 due	Define the 'Population Regression Function,' and the OLS. Do an exercise where we derive the OLS estimator of a linear model with just a constant term (no regressor). In this framework I defined 'uhat,' 'yhat,' 'beta that,' and explained they are not the original u, y, or beta.
3	2	17-Sep-19	Chapter 2		Implement the simple linear regression (show what the 'lm' command does). We learn to interpret estimated coefficients. Define 'fitted values,' 'residuals.'
4	1	19-Sep-19	Chapter 2		Practice interpreting the estimated intercept/slope parameters in y on x, log_y on log_x, log_y on x, and y on log_x models. Start talking about the classical assumptions.
4	2	24-Sep-19	Chapter 2		Talk about the unbiasedness property of the OLS under the classical assumptions.
5	1	26-Sep-19	Chapter 2	PS 2 due	Class reviews the idea of unbiasedness again, and introduces the homoskedasticity assumption, describes the formula of the variance of the OLS estimators. Explaines how these variances are estimated in practice and defined the concept of standard errors.
5	2	1-Oct-19	Chapter 3		We start covering Chapter 3 (Multivariate). Interpretation of coefficients, the algebra of the OLS regression, and the partitioned regression idea.
6	1	3-Oct-19			Midterm Exam 1

Week	Day	Date	Reading Due	Problem set due (midnight prior)	Description
6	2	8-Oct-19	Chapter 3		Continue with partitioned regression idea (I explain it by using an empirical example), and coefficient interpretation. Using R, I demonstrate how one can use the 'algebra of the OLS' to show that if we regress residuals on the original regressors, the coefficients will be zero and the R2 will also
7	1	10-Oct-19	Chapter 3		Do question 1 of the end-of-chapter exercises in the textbook. Explain the classical assumptions and the unbiasedness of the OLS estimator. Provid examples of perfect collinearity (violation of Assumption MLR.3).
7	2	15-Oct-19	Fall Break -- no classes	Fall Break -- no classes	Fall Break -- no classes
8	1	17-Oct-19	Chapter 3	PS 3 due	We talk about the unbiasedness property of the OLS estimator, discuss the omitted variable bias, how to put a sign on the bias. We study the anatomy of the formulae of the variance of the OLS estimators and defined the VIF. I demonstrated how you can calculate it using R (showed
8	2	22-Oct-19	Chapter 4		The MSE of the unknown variance of the estimator is discussed. The concept of degrees of freedom is explained. We then cover the Gauss-Markov Theorem. We demonstrate that the OLS estimator in a SLR model is a linear estimator and did an exercise utilizing the Gauss-Markov Theorem to rank estimators. We start with Chapter 4 (Inference). Introduce the normality of the error assumption and ended the class showing them the
9	1	24-Oct-19	Chapter 4		I cover the t-statistic to test the significance of a regressor. I review all the terminology for hypothesis testing, define type-I and type-II error, show them how to use statisticial tables to calculate critical values and perform one-side t-
9	2	29-Oct-19	Chapter 4		We talk about two-tailed t-tests, and how to go about testing H0: bj=aj (non-zero). Emphasize how if one rejects at 1%, one knows the test must reject at 5% and 10%.
10	1	31-Oct-19	Chapter 4		We introduce the F-test for multiple linear regression models. We demonstrate that the significance of a regression can be tested using the R2.
10	2	5-Nov-19	Chapter 5		Start Chapter 5: Asymptotics of the OLS estimator. We talk about the Plim and its properties, the idea of consistency. We show that omitting a relevant regressor in a regression with 2 covariates gives inconsistent estimators (derived the incosistency term). Show that the s.e. should converge to
11	1	7-Nov-19	Chapter 8	PS 4 due	We started covering "Chapter 8: Heteroskedasticity." We look at robust standard error estimation and covered the Breusch-Pagan & White Test for Heteroskedasticity
11	2	12-Nov-19	Midterm review	Midterm review	Midterm review
12	1	14-Nov-19	2nd Midterm Exam	2nd Midterm Exam	2nd Midterm Exam

Week	Day	Date	Reading Due	Problem set due (midnight prior)	Description
12	2	19-Nov-19	Chapter 8		We finish covering Chapter 8. Specific content TBD
13	1	21-Nov-19	Chapter 7		We start covering "Chapter 7: Dummy Variables" -- I show how to interpret said coefficients, the idea of different PRF per different categories. How to choose the base category. I also show the 'interaction' function in R as a way to create multiple categories based on all possible cross-products of
13	2	26-Nov-19	Chapter 7		I continue covering the remaining of this chapter: Inclusion of ordinal data, interaction between a dummy variable and regressors, testing for the hypothesis of the same vs different population regression functions, and the idea that a dummy regressor can be used to evaluate policy, i.e., control vs treated group. I cover the linear probability model and showed it is heteroskedastic. Talk about fitted values being greater than 1 or negative and implemented
14	1	28-Nov-19	No class- Thanksgiving break	No class- Thanksgiving break	No class- Thanksgiving break
14	2	3-Dec-19	Notes or references to be given		Identification and research design -- RDD - TBD
15	1	5-Dec-19	Notes or references to be given		Identification and research design -- Diff in Diff - TBD
15	2	10-Dec-19	Final review	PS 5 due	Final review