Econ 420: Econometrics
Fall 2016

Instructor: Zhongjian Lin, Rich 332, Email: zhongjian.lin@emory.edu
Time: Monday and Wednesday 1:00pm-2:15pm
Location: White Hall 101
Office Hours: Monday and Wednesday 9:30am-11:30am and by Appointment only
Pre-requisite: Students must have taken Econ 220 - Introduction to Statistical Methods before enrolling in this course.

Textbooks:


Course Objective: This course is designed to introduce students to the statistical foundations of Econometrics, and its well-known multivariate linear model. It will equip students with understanding of the fundamentals probability theory and mathematical statistics behind most econometric methods. 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.

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

Grading: The final grade will be determined by a weighted average of scores in homework (20% throughout the semester), pop-quizzes (5%), two (closed-book) in-class midterm exams (20% for each on September 26, 2016 and October 31, 2016), and a (closed-book) comprehensive final exam (35% on December 10, 8:00am-10:30am).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Grade</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>90</td>
<td>100</td>
<td>C+</td>
<td>65</td>
<td>69</td>
</tr>
<tr>
<td>A-</td>
<td>85</td>
<td>89</td>
<td>C</td>
<td>60</td>
<td>64</td>
</tr>
<tr>
<td>B+</td>
<td>80</td>
<td>84</td>
<td>C-</td>
<td>55</td>
<td>59</td>
</tr>
<tr>
<td>B</td>
<td>75</td>
<td>79</td>
<td>D+</td>
<td>50</td>
<td>54</td>
</tr>
<tr>
<td>B-</td>
<td>70</td>
<td>74</td>
<td>D</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Exams: Each in-class midterm exam will last exactly 75 minutes and will consist of 3 questions. Question 1 is compulsory, but 1 and only 1 question of the remaining 2 must be answered by all candidates. The final exam will take place on December 10, 2016, between 08:00 am - 10:30 am, and it will consist of 4 questions. 3 out of these 4 questions must be answered by all candidates. Only handheld calculators are allowed. Students are allowed to carry one page A4 size cheat sheet (one side).
Homework and Pop-Quizzes: There will be 6 assignments. They are due in paper form at the beginning of the lecture the day the assignment is due. If you work in teams, please write down the name of each member of the group in your individual personal submission. There will be 6 pop-quizzes. Answering correctly for all 6 pop-quizzes will earn 1 extra point towards the final grade.

Software: Computing tutorials will be held during some lectures. This is designed to help students to understand the practical implications of the theoretical content of the lectures, and familiarize themselves with the use of powerful and widespread econometric software such as R and STATA. R is free and STATA is available on computers in the Econ lab (Rich 301), Woodruff Library, and Cox Hall.

Policy

Dishonesty: Any student who is found to have cheated or to have plagiarized will receive an “F” for the course. The university defines academic misconduct as:

“Academic misconduct is an offense generally defined as any action or inaction which is offensive to the integrity and honesty of the members of the academic community”

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. 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.

Class attendance: Attendance is mandatory and encouraged. Marks from home assignments, pop-quizzes, class participation, and exams will only be credited if taken and handed-in during the session the student is officially enrolled in. 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).

Outline

[1] 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.


1) Random Variables and Their Probability Distributions (Appendix B.1).
   i. Discrete Random Variables.
   ii. Continuous Random Variables.
2) Joint Distributions, Conditional Distributions, and Independence (Appendix B.2).
   i. Joint Distributions and Independence.
   ii. Conditional Distributions.

3) Features of Probability Distributions (Appendix B.3).
   i. A Measure of Central Tendency: The Expected Value.
   ii. Properties of Expected Values.
   iii. Another Measure of Central Tendency: The Median.

4) Features of Joint and Conditional Distributions (Appendix B.4).
   i. Measures of Association: Covariance and Correlation Coefficient.
   ii. Variance of Sums of Random Variables.
   iii. Conditional Expectation.

5) The Normal and Related Distributions (Appendix B.5).
   i. The Normal Distribution.
   ii. The Standard Normal Distribution.
   iii. Additional Properties of the Normal Distribution.
   iv. The Chi-Square Distribution.
   v. The F Distribution.

1) Population, Parameters and Random Sampling (Appendix C.1).
2) Finite Sample Properties of Estimators (Appendix C.2).
   i. Estimators and Estimates.
   ii. Unbiasedness.
   iii. The Sampling Variance of Estimators.
   iv. Efficiency
3) Asymptotic or Larger Sample Properties of Estimators (Appendix C.3).
   i. Consistency: Consistency and Law of Large Numbers.
   ii. Asymptotic Normality: Asymptotic Normality and Central Limit Theorem.

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.

[5] 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.


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.
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 $\sigma^2$: Standard Errors of the OLS Estimators.


[7] Multiple Regression Analysis: Inference (Chapter 4)

1) Sampling Distributions of the OLS Estimators.
i. Testing against One-Sided Alternatives.
ii. Two-Sided Alternatives.
iii. Testing Other Hypotheses about $\beta_j$.
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.
v. The F Statistic for Overall Significance of a Regression.
vi. Testing General Linear Restrictions.

[8] 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.


1) More on Functional Form.
   i. More on Using Logarithmic Functional Forms.
   ii. Models with Quadratics.
   iii. Models with Interaction Terms.
2) More on Goodness-of-Fit and Selection of Regressors.
   i. Adjusted R-Squared.
   ii. Using Adjusted R-Squared to Choose between Nonnested Models.
   iii. Controlling for Too Many Factors in Regression Analysis.
   iv. Adding Regressors to Reduce the Error Variance.
3) Predicting $y$ when $\log(y)$ is the Dependent Variable.

[10] 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.