

ECON 3120: Applied Econometrics

Cornell University, Fall 2018

Monday & Wednesday, 8:40am–9:55am

Goldwin Smith Hall G76-Lewis

Instructors

Douglas McKee (douglas.mckee@cornell.edu)

Office: 466 Uris Hall

Office Hours: Wednesday, 10:30am–12:00pm

Teaching Assistants

Shiyi Zhang (sz534@cornell.edu)

Section 201: Friday, 9:05am–9:55am, Rockefeller 102

Section 202: Friday, 10:10am–11:00am, Rockefeller 105

Office Hours: Tuesday and Thursday 3:00pm–4:30pm, Uris 447

Abhishek Ananth (aa2335@cornell.edu)

Section 203: Friday, 1:25pm–2:15pm, Rockefeller 122

Section 204: Friday, 2:30pm–3:20pm, Rockefeller 122

Office Hours: Monday, 10:30am–12:00pm and Tuesday, 10:00am–11:30am, Uris 431

Course Description

The main objective of Econ 3120 is to teach students the quantitative methods most frequently employed in empirical analysis of economic phenomena. Applications of these methods will be emphasized throughout the course.

The first part of the course will be a short review of probability, estimation and hypothesis testing. Most of this material is covered in Econ 3110, but this serves as a refresher and reviews the material with an eye toward using it later on in the course. Next, we'll spend some quality time with the work horse of empirical economics: the linear regression model. Finally, we'll learn several advanced concepts and methods frequently encountered in applied economic research. These will include analysis of binary dependent variables, panel data methods, instrumental variables methods, and regression discontinuity designs.

The prerequisite for this course is either ECON 3110 or 3130.

The economics department offers two different two-semester sequences for econometric methods. A common question is how these two sequences differ. The core idea is that while both sequences cover many of the same topics, the treatment in 3120 will emphasize applications of econometric techniques. We will focus on substantive questions first and then introduce mathematical methods that will help us answer them. 3140 spends more time on theoretical models and their mathematical properties. Along these lines, 3140 is usually taught by econometricians, while 3120 is often taught by faculty who analyze data for a living.

By the end of the semester, you will have acquired several concrete skills. You will be able to

1. Interpret the results and critique the methods used in a typical empirical research article on economics. In particular, you will be able to interpret and comment on tables of estimated coefficients from a wide range of econometric models.
2. Given a feasible economic question and a data set, propose appropriate strategies to answer the question using the data. You will know how to weigh the strengths and weaknesses of several methods as well as understand the math and intuition behind these methods.
3. Apply the methods you learn to analyze real data with a powerful statistical analysis package (Stata).
4. Learn new empirical methods that are built on the methods covered in this class.

A more detailed comprehensive list of this class's learning goals is available on class web site.

Course Structure

Readings and Quizzes

Readings for this course are divided into *Required* and *Optional*.

1. *Required* readings build upon what you learned in previous classes. Concepts and methods are introduced during class, and the reading fills in some of the mathematical detail and common extensions.
2. *Reference* materials contain expanded information related the topic discussed. I don't expect everyone to read every page of the reference material, but you will almost certainly find it useful when working through the problem sets and studying for exams.

Many of the readings for this course are about mathematical techniques and the “big picture” ideas that underlie them. They are not bed-time reading. Take your time doing the reading each week, and make sure you understand what is being presented.

The primary textbook for this course is *Introduction to Econometrics, 3rd edition*, by James H. Stock and Mark W. Watson (Addison-Wesley, 2010). The 2014 updated edition is completely unnecessary. Stock and Watson's treatment of regression methods is excellent, and the book should serve you well as a reference in the future. Purchasing this book new is expensive, but you can typically find used copies of the textbook on the web at sites like Amazon for under \$50. The most important method we will cover during the course is linear regression and I highly recommend Paul Allison's *Multiple Regression: A Primer*. The writing is extremely clear and he covers both the intuition and mathematics behind the method. Students without a strong mathematical or statistical background may find *Probability and Statistical Inference, 8th or 9th ed.*, by Robert Hogg, Elliot Tanis, and most recently Dale Zimmerman (Pearson, 2010 or 2013) useful.

At the end of most course modules, you'll take a short online quiz that covers that module's lecture material and reading. The questions won't be difficult, but they will ensure you understand the core ideas and are doing the reading carefully.

Class Time

Class time will be highly interactive, and each class will contain multiple group work activities. During these activities, the instructor and the teaching assistants will circulate around the room providing feedback and guidance.

The activities will be of two distinct types. During **Invention** activities you will try to solve brand new problems. **Struggle is expected!** Studies have shown that students who do invention activities before learning a new method understand the method much more deeply than students that simply get a lecture on the method. They retain the knowledge longer and are able to apply the concepts more broadly. And with the right attitude, invention activities are a lot of fun. The second type of activity occurs after you've learned a new method (potentially through a short lecture) and involve **Deliberate Practice**. These activities strengthen understanding through application of methods and concepts in multiple contexts.

Throughout the class, you will use the **i<clicker** system to answer questions and give us feedback on what you are learning and what you are not. This keeps you engaged, and lets us know when we need to provide more direction and when we are ready to move on. To participate, you will need to purchase either an i<clicker hardware remote or the i<clicker REEF software for your mobile device. The clicker costs \$30-40 and can be used for multiple classes. The software costs \$14.99 for a 6 month license. You can learn more on the Cornell Clicker How-to page: <http://www.it.cornell.edu/services/polling/howto-students.cfm>

Empirical Project

The empirical project is a very important part of your learning experience in the class. It provides you with the opportunity to use the tools you learn in the class to answer a question that **you** come up with and that **you** care about. It is designed to guide you through the steps of answering a research question the way applied economists do.

I will divide the class into groups of four or five students each before the first day of class. Each group will identify and submit their research question by the end of the 4th week (February 14). During the 9th week (March 13), each group will pass in a document that quantitatively describes their sample and the relevant variables in it. During the last week of the classes (May 8), each group will do a poster presentation of their results, and by the end of that week (May 11), groups will submit a written project report that includes a discussion of the empirical analysis. As a part of the final write-up you will submit a flow-chart of your decision-making process throughout the project.

Grades

Your grade for the class will be composed of seven parts:

1. Problem Sets (20%)

There will be 6 problem sets during the semester. You may pass in the problem sets as pdf's through the course web site by 8:40 AM or in person at the beginning of class on the due date. Complete solution sets will be posted in the evening after the problem sets are due. It is your responsibility to read these solutions and make sure you understand them. Your lowest grade on a problem set will be dropped. **Late problem sets will not be accepted, including those turned in after class on the due date. Computer exercises must include the Stata .do and .log files that you produce.**

2. Empirical Project (15%)

The research question is worth 5% of your final grade, the data description another 5%, and the final analysis (presentation and report) is worth 5%. You will get a fair bit of written and verbal feedback guidance throughout the process.

3. Online Quizzes (5%)

At the end of most weeks, you will take an online quiz on that week's material. Your grade will be your average score during the semester. Your lowest quiz will be dropped, and there will be no excuses for missing quizzes.

4. Class Participation (5%)

You are required to participate using i-clicker in class. Your score will be the fraction of classes where you clicked in for at least half the opportunities. You can miss up to 5 classes and still get credit for participating in those classes, but again, this policy means there will be no excuses for missing additional classes.

5. First Prelim (15%)

Date: Thursday, September 27, 7:30pm, RCK203

6. Second Prelim (15%)

Date: Thursday, October 25, 7:30pm, MLT228

7. Final Exam (25%)

Date: Tuesday, December 11, 2:00pm

Exams are closed book, but you may bring one double-sided page of notes to the first prelim, two pages to the second prelim, and three pages to the final exam. You may use calculators during the exams.

We will also be doing *two stage exams* in this class. You will first take the exam individually and hand in your test. Then you will take the exam in small groups, where each group works together and passes in one exam with their consensus answers. If your individual score is higher than your group's score, your grade on the exam will be your individual score. If your group's

score is higher, your grade on the exam will be a weighted average of your individual score (90%) and your group's score (10%).

The main reason we conduct a second stage of the exam is to allow you to learn more during the exam. Traditional exams tend to be summative rather than formative, and two stage exams represent an opportunity to redress this imbalance. The process of discussing your answers with your teammates is a significant learning opportunity and supports the kind of collaborative learning that we encourage.

Final grades for the class will be determined by computing a weighted score based on the weights listed above. The weighted scores are assigned letter grades based on the following cut-offs:

Range	Letter
90–100	A
85–89	A-
80–84	B+
75–79	B
70–74	B-
64–69	C
60–63	C-
50–59	D
0–49	F

In the past, this breakdown has resulted in about 45% of the class getting A-'s or A's. We reserve the right to make the cutoff's more generous if the exams are unexpectedly difficult, but we will not make the cutoffs less generous under any circumstances. We will give A+'s to students who earn A's and show extraordinary mastery of the material by the end of the semester.

Excuses

Because one problem set is dropped, excuses for missed problem sets are not considered. The only exception is prolonged/severe illness, which must be handled through the advising deans as per case (1) below.

With respect to exams, the Faculty Handbook

(<http://theuniversityfaculty.cornell.edu/handbook/Chapter5.pdf>) lists four types of situations in which faculty are encouraged to make accommodations for missed work. However, the determination as to whether a particular case warrants accommodation is ultimately the decision of the faculty member. Here is how the four cases are handled in this course:

1. Illness, or family or personal emergency: Any situations that fall under this category must be first brought up with the advising dean in the student's college. The advising dean will then contact Professor McKee directly, and he will make a determination based on the particular case.
2. Employment interviews. The student must provide us evidence of the interview and establish that (s)he has no control over the timing of the interview.

3. Religious observances. While we do our best not to schedule exams during religious holidays, please contact us at least two weeks in advance if an exam date/time conflicts with a religious holiday.
4. Athletics and Extracurricular Activities. Students in varsity athletics or recognized extracurricular activities must provide the standard permission slip from the staff responsible for the activity at least two weeks before the exam.

Final Exam Conflicts

There are two situations that we will consider for exam conflicts. First is a direct conflict where ECON 3120 and another class appear on the registrars exam schedule at the same time. Second is 3 or more exams having a start time within 24 hours, as indicated on the registrars exam schedule. If Exam 1 is on Monday at 2pm, Exam 2 is on Monday at 7pm, and Exam 3 is on Tuesday at 2pm, this is not more than 2 exams in 24 hours. If you have a conflict, you need to email Professor McKee at least 2 weeks before the final exam, listing out the other classes involved and scheduled exam times. The date and time of the makeup will be determined at that time.

Grading FAQ

- **Are the tests cumulative?** The tests **are** cumulative. About 15% of the second prelim is on earlier material and up to half the final exam is on material covered on the first two prelims. In addition, you will need to use concepts from the earlier parts of the course in order to understand the later topics.
- **Is there extra work I can do to improve my grade?** No.
- **I didn't do as well as I had hoped early in the course. In determining my final grade, can you put more weight on the latter part of the course?** No.
- **I have X exams/assignments due within Y of each other. Can I reschedule the exam/hand in the homework late?** No. The first two prelim dates are given above, and the final exam will be posted early in the semester. This gives you plenty of time to plan ahead. The only exception is more than 2 finals in 24 hours, see above.

Exam Regrades

While we take care to grade exams as fairly and consistently as possible, on rare occasions there may be grading mistakes. If you feel that your test has been graded incorrectly, you must submit it to the professor (not the TA), along with an explanation of the issue in writing. You must do this within 2 weeks of the exam being returned (not the date you pick it up) for it to be regraded. The entire exam will be regraded, and as a result it is possible for your grade to go down as well as up.

Acceptable Use Policy

You are free to use any published materials (e.g., another textbook), in preparing Econ 3120 assignments or for learning the material more generally. You are also strongly encouraged to work with others in your class. This is particularly helpful for learning to program. Each person must turn in their own assignment.

The use of any solution materials prepared in a previous year for Econ 3120, other than materials distributed this academic year by the course faculty, is strictly prohibited. This includes 1) any notes, spreadsheets, or handouts distributed in a prior term of Econ 3120; and 2) any notes, solutions, or spreadsheets prepared by former students of Econ 3120, in either written or electronic form.

This policy means you should not solicit or use solutions to previous years' problem sets. The reason for this policy is that access to previous years materials can create serious inequities between fellow students, and jeopardize the integrity of the academic environment. Academic disciplinary actions will be taken against those who violate this policy.

Software

Much of the course work in Econ 3120, will involve analysis of data using the Stata software package. There are three options for using Stata:

1. You can use Stata directly on your own computer by purchasing a six-month (or longer) license at <http://www.stata.com/order/new/edu/gradplans/student-pricing/>. The 6 month license for Stata/IC for \$45 is highly recommended.
2. You can also use Stata for free by using the CISER account that will be created for all students in the class. This requires using a remote desktop connection to a CISER server. The performance is mediocre and transferring files back and forth is tedious. But it's free and it works. The TAs will show you how to do this in the second discussion section. All students enrolled in the class before the first day of class will have a CISER account setup automatically. If you enroll in the class after the first day, you need to fill out a new account for at http://ciser.cornell.edu/computing/accts/RSCH_ReqForm.php
3. You can use Stata in the public computer labs in Warren Hall and Mann Library (see <http://www.cscu.cornell.edu/software/facilities.php>).

We will spend some time in class teaching Stata and the program documentation is excellent. You will get plenty of practice during your sections, and there are several terrific free online resources for learning the software. For those students who feel more comfortable with a book in hand, Acock's *A Gentle Introduction to Stata, Revised Third Edition* is up to date and pretty good.

Note to Students with Disabilities

If you have a disability-related need for reasonable academic adjustments in this course, please give Professor McKee an accommodation letter from Student Disability Services. We expect you to give two weeks notice of the need for accommodations. If you need immediate accommodations, please arrange to meet with Professor McKee within the first two class meetings.

Active Learning Initiative

Econ 3120 is participating in Cornell's Active Learning Initiative, a program designed to infuse our large lecture courses with activities that increase student participation during class. We expect no risks to you for participating in the study. Later this term, we will ask you to complete a short (one minute) survey. Otherwise, you'll just do the work of the course, as assigned by your instructor. You might also be contacted in the future and invited to participate in an optional focus group to provide your feedback on the active learning activities used in the course. If you have any questions, you can talk to Doug McKee (douglas.mckee@cornell.edu) or Carolyn Aslan (crc1@cornell.edu) in the Center for Teaching Excellence.

Acknowledgements

This class is in large part derived from an econometrics class that Professor Lanier Benkard taught at Yale in Fall 2010. I am extremely grateful to him for sharing his syllabus, lecture slides, assignments, handouts, exams, and advice. I have also borrowed liberally from the syllabus and materials of Jim Berry who taught this class in Fall 2015. Teddy Svoronos inspired the use of two stage exams. All of these have provided a fantastic starting point.

Please do not redistribute any of these materials without our permission.

Schedule

PART I: REVIEW OF PROBABILITY AND STATISTICS

Module 0: Economic Statistics Assessment

Section: AUGUST 24, REQUIRED IN CLASS ASSESSMENT

Module 1: Introduction

Lecture: August 27

Read: SW Chapter 1

Topics: – Course overview

Module 2: Review of Probability Theory, Estimation, and Hypothesis Testing

Lecture: August 29 and September 5

Read: Review articles on website
SW Chapter 2, 3 (optional)

Section: August 31, Lab: Impulsivity and Gambling among High School Students

Section: September 7, Working Through Stats Problems

Topics: – Terminology and concepts: experiments, outcomes, and events
– Marginal, joint, and conditional probabilities
– Probability tables
– Discrete random variables (e.g., Bernoulli and Binomial)
– Continuous random variables (e.g., Normal, t, Chi-Square, and F)
– Covariance and independence
– The Central Limit Theorem
– Samples and populations: Overview
– Standard errors
– Confidence Intervals
– One and two sample hypothesis tests

PART II: BIVARIATE MODELS

Module 3: Randomized Experiments

Lecture: September 10

Read: SW 3.5

Angrist and Pischke (AP), Chapters 1 and 2 of *Mostly Harmless Econometrics*
posted

Online Quiz: September 11

Topics: – Interpreting observational and experimental data

Module 4: Introduction to Regression Analysis

Lecture: September 12 and 17

Read: SW Chapter 4

Allison, Chapters 1 and 5 (Optional)

Online Quiz: September 18

Section: September 14, Lab: Microcredit in Bangladesh

Due: Problem Set 1 on September 17, 8:40 AM

Due: Empirical Project Research Question on September 18, 5:00 PM

- Topics:
- Mechanics of simple regression
 - Correlation vs. slope
 - Interpreting regression estimates
 - Doing regression in Stata
 - R^2 and goodness-of-fit

Module 5: Statistical Inference in Regression

Lecture: September 19 and 24

Read: SW Chapter 5

Online Quiz: September 25

Section: September 21, Lab: Legislation and Emissions

- Topics:
- Hypothesis testing and statistical significance
 - Confidence intervals
 - Prediction

Module 6: Prelim week

Lecture: September 26

Section: DOES NOT MEET

Due: Problem Set 2 on September 26, 8:40 AM

Thursday, September 27, 7:30pm, RCK203: FIRST PRELIM EXAM

PART III: MULTIVARIATE REGRESSION

Module 7: Multiple Regression

Lecture: October 1 and 3

Read: SW Chapters 6 and 8
(Optional) Allison, Chapters 2 and 3

Online Quiz: October 4

Section: October 5, Lab: Who Survived the Titanic?

- Topics:
- Mechanics of multiple regression
 - Interpreting multiple regression results
 - Omitted Variable Bias
 - Controlling for categorical variables with sets of dummy variables
 - Nonlinear effects
 - Variable Interactions

Module 8: Model Building in Practice

Lecture: October 10 and 15

Read: SW Chapter 7.1, 7.2, 7.3
(Optional) Allison, Chapters 7 and 8

Online Quiz: October 16

Section: October 12, Lab: Private Firms in Vietnam

Due: Empirical Project Data Description on October 10, 5:00 PM

- Topics:
- Regression F-test
 - Joint Tests
 - Restricted and Unrestricted models
 - Tests of linear restrictions in regression models
 - Developing regression models
 - Interpreting and using results
 - Model building using real data
 - Confounding variables
 - Intervening variables

Module 9: Assessing Validity of Regression Models

Lecture: October 17 and 22

Read: SW Chapter 9

Section: October 19, Working Through Regression Problems

Due: Problem Set 3 on October 17, 8:40 AM

Online Quiz: October 23

- Topics:
- Internal and external validity
 - Measurement error
 - Missing data

Module 10: Prelim Week

Lecture: October 22—No class on October 24

Section: October 26, Discuss Empirical Projects

Due: Problem Set 4 on October 22 8:40 AM

Thursday, October 25, 7:30pm, MLT228: SECOND PRELIM EXAM

PART IV: ADVANCED METHODS

Module 11: Binary Dependent Variables

Lecture: October 29 and 31

Read: SW Chapter 11

Online Quiz: November 1

Section: November 2, Lab: The 1992 Presidential Election

- Topics:
- Linear probability model
 - Logit and probit models

Module 12: Instrumental Variables

Lecture: November 5 and 7

Read: SW Chapter 12
AP Chapter 4 (Optional)

Online Quiz: November 8

Section: November 9, Lab: Instrumenting Schooling with Quarter of Birth

- Topics:
- Estimating causal effects with instrumental variables
 - Evaluating instrumental variables

Module 13: Difference-in-Differences & Regression Discontinuity

Lecture: November 12 and 14

Read: SW 10.1, 10.2
AP Chapter 5 (Optional)

Online Quiz: November 15

Section: November 16, Lab: Minimum Wages and Employment

Due: Problem Set 5 on November 12, 8:40 AM

- Topics:
- Using Difference-in-Differences to estimate treatment effects
 - Implementing and extending Difference-in-Differences with regression
 - Using Regression Discontinuity to estimate treatment effects

Module 14: Fixed Effects

Lecture: November 19 and 26

Read: SW 10.3–10.7

Online Quiz: November 27

- Topics:
- The fixed effects regression model
 - Time and entity fixed effects
 - Estimation and inference

Module 15: Review

Lecture: November 28 and December 3

Section: November 30, Open Q&A

Due: Problem Set 6 on November 30, 5:00 PM

Due: Empirical Project Poster Session on December 3, Evening

Due: Empirical Project Data Analysis on December 5, 5:00 PM