

ECON 3120: Applied Econometrics

Cornell University, Spring 2023

Lectures: Tuesday & Thursday, 1:00pm–2:15pm, Goldwin Smith 132 (HEC Aud)

Instructor

Douglas McKee (douglas.mckee@cornell.edu)

Office Hours: See Canvas

Teaching Assistants

Kalie Hwei-Gen Pierce (khp44@cornell.edu)

Section 201: Friday, 9:05am–9:55am

Section 202: Friday, 10:10am–11:00am

Office Hours: See Canvas

Leo Peñaloza-Pacheco (ljp96@cornell.edu)

Section 203: Friday, 1:30pm–2:20pm

Section 204: Friday, 2:40pm–3:30pm

Office Hours: See Canvas

Francesco Billari (fb293@cornell.edu)

Office Hours: See Canvas

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 econo-

metricians, 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

Study Groups

Over the years, I have found that my most successful students help each other, and to encourage this kind of cooperation, I will assign everyone to a study group of four or five students after the first week of classes. You will work in your group on a variety of discussion and problem-solving activities during lectures and discussion sections. The prelims will involve some group work, and groups will work together on an empirical project throughout the semester. See below for details on the exams and the empirical project. I also encourage you to work together on the problem sets and study together for exams.

Readings and Quizzes

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

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 2018 fourth edition is 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 as a new hardcover is expensive, but you can typically find used copies of the textbook on the web at sites like Amazon for cheaper. And right now, you can buy the paperback edition for just [\\$25 on Amazon](#).

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.

During almost every class, you will use the [Poll Everywhere](#) system to answer questions and give me 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. You can answer questions in any web browser or use the Poll Everywhere app on your phone. And because Cornell just purchased a university-wide license to this software, it's completely free to you.

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.

Each study group will identify and submit their research question by the end of the 4th week

(February 10). At the same time, you will create and submit a team contract that describes how your team will work together throughout the semester. During the 9th week (March 10), 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, we will hold a poster session, and by the end of that week (May 12), groups will submit a written project report that includes a discussion of the empirical analysis.

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. 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. You are encouraged to work together on these problem sets, but you must submit your own work in your own words. **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 and team contract are 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%)

As you know from above, I believe there are major returns to attending class and participating in the activities. I also think it's easier to stay focused when the lecture is live. To further incentivize attendance and participation, 5% of your grade is based on a participation score equal to the fraction of classes where you clicked in for at least half of the Poll Everywhere 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.

Alternatively, you can earn the same participation credit by doing the following: By 11:59pm of the day after a lecture, you can watch the recorded lecture and submit answers to the questions I pose to students during the class. Along with your (mostly multiple choice) answers, you'll have to include explanations for your answers. You will get a lot more out of this exercise if you pause the recorded lecture as the questions come up and try to answer them right then.

5. First Prelim (15%)

Date: Thursday, February 23, 7:30pm

6. Second Prelim (15%)

Date: Tuesday, March 28, 7:30pm

7. Final Exam (25%)

Date: TBA

The final exam will be given during finals period. The schedule will be posted at <https://registrar.cornell.edu/Sched/exams.html> early in the semester.

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* for the prelims in this class. You will first take the exam individually and hand in your test. Then you will take the exam in your study 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
94–100	A
90–93	A-
87–89	B+
83–86	B
77–83	B-
71–76	C
65–70	C-
50–64	D
0–49	F

In the past, this breakdown has resulted in about 45% of the class getting A-'s or A's. I 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. I 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 registrar's exam schedule at the same time. Second is 3 or more exams having a start time within 24 hours, as indicated on the registrar's 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 prelin 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 \$48 is a pretty good deal.
2. You can also use Stata for free through a service called Apps on Demand that is accessible from the course site. After entering Apps on Demand, click on the Stata icon and you'll be running Stata inside your web browser window. You can connect this Stata environment

to your Cornell Google Drive so you have persistent storage. It's not as good as running the software locally, but it works surprisingly well. The TA's will walk you through your initial set up during discussion section.

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* 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.

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. George Orlov and I worked together to create most of the classroom exercises. Teddy Svoronos inspired the use of two stage exams.

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

Schedule

PART I: REVIEW OF PROBABILITY AND STATISTICS

Module 1: Introduction

Lecture: January 24

Read: SW Chapter 1

Topics: Course overview

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

Lecture: January 26 and January 31

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

Section: January 27, Lab: Impulsivity and Gambling among High School Students

- 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: February 2

Read: SW 3.5

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

Section: February 3, Working Through Stats Problems

- Topics:
- Interpreting observational and experimental data

Module 4: Introduction to Regression Analysis

Lecture: February 7 and 9

Read: SW Chapter 4
Allison, Chapters 1 and 5 (Optional)

Due: Problem Set 1 on February 7, 1:00 PM

Section: February 10, Lab: Microcredit in Bangladesh

Due: Empirical Project Research Question and Team Contract on February 10, 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: February 14

Read: SW Chapter 5

Topics: – Hypothesis testing and statistical significance
– Confidence intervals

Module 6: Prediction

Lecture: February 16

Read: SW Chapter 5

Section: February 17, Lab: Legislation and Emissions

Topics: – Prediction in multiple regression models
– Transformed dependent variables
– Graphing predictions
– Standard errors of predictions (aka forecasts)
– Homoskedasticity and heteroskedasticity

Module 7: Prelim Review

Lecture: February 21

Section: No section February 24

Due: Problem Set 2 on February 21, 1:00 PM

Thursday, February 23, 7:30pm FIRST PRELIM EXAM

PART III: MULTIVARIATE REGRESSION

Module 8: Multiple Regression

Lecture: March 2 and 7

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

Section: March 3, TBA

- 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 9: Model Building in Practice

Lecture: March 9 and 14

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

Section: March 10, Lab: Who Survived the Titanic?

Due: Empirical Project Data Description on March 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 10: Assessing Validity of Regression Models

Lecture: March 16 and 21

Read: SW Chapter 9

Section: March 17, Lab: Private Firms in Vietnam

Due: Problem Set 3 on March 16, 1:00 PM

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

Module 11: Prelim Review

Lecture: March 23

Section: March 24, Working Through Regression Problems

Due: Problem Set 4 on March 24, 5:00 PM

Tuesday, March 28, 7:30pm: SECOND PRELIM EXAM

PART IV: ADVANCED METHODS

Module 12: Binary Dependent Variables

Lecture: March 30 and April 11

Read: SW Chapter 11

Section: March 31, Discuss Empirical Projects

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

Module 13: Instrumental Variables

Lecture: April 13 and 18

Read: SW Chapter 12
AP Chapter 4 (Optional)

Section: April 14, Lab: The 1992 Presidential Election

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

Module 14: Difference-in-Differences & Regression Discontinuity

Lecture: April 20 and April 25

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

Section: April 21, Lab: Instrumenting Schooling with Quarter of Birth

Due: Problem Set 5 on April 26, 1:00 PM

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 15: Fixed Effects

Lecture: April 27

Read: SW 10.3–10.7

Section: April 29, Lab: Minimum Wages and Employment

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

Module 16: Multinomial and Ordered Models

Lecture: May 2

Read: SW 11.3

Topics: – Multinomial Logit
– Ordered Logit

Module 17: Review

Lecture: May 4

due to Poster Session: May 9

Section: May 5, Open Q&A

Due: Problem Set 6 on May 5, 1:00 PM

Due: Empirical Project Poster Session, exact time TBA

Due: Empirical Project Data Analysis on May 12, 5:00 PM