This is an introductory course in the use of econometric methods, with an emphasis on empirical applications and cross-sectional analysis. Our focus will be on learning to do econometrics, not just learning econometrics.

While the course will cover the development of the formal tools of econometric analysis (simple and multiple regression analysis, estimation, inference, qualitative variables, and so forth), we will also spend quite a bit of time on empirical methods (posing questions, building datasets, running regressions, supplementing datasets, running more regressions, etc etc … until we can confidently say something about the questions at hand). As such, an important part of the course will be a set of empirical Exercises and an empirical research paper in which students will be building their own datasets and applying the various econometric methods developed in the course.

Prerequisites: An introductory statistics course such as EC 151. No exceptions. I will also assume that students have an understanding of basic Excel (which will be used at times to assemble datasets and verify calculations) and basic calculus. I do not assume that anyone has previous exposure to Stata, the computer language that we will be using to run regressions (but you will be better for it if you have seen Stata before). (See below for more about Stata.)

Course text:


I have deliberately not listed the edition; if you decide to purchase the text, feel free to buy the 2nd (2002), 3rd (2006), 4th (2008), or 5th (2012) editions. A copy of the text will be placed on reserve at the O’Neill Library.

Important Note: I do not follow the text closely. I will be distributing lecture notes for most of the material covered in class.

Some additional texts: There is no need to purchase any of these (most are available at O’Neill). I list them just because sometimes it is useful to see a different presentation of the material.

Boston College  
EC228: Econometric Methods


Grading:
- Two Mid-Term Exams (60% total; 30% each): These in-class exams are scheduled for:
  - Tuesday, October 21st (the week after Columbus Day)
  - Thursday, December 4th (the next to last class)
Exam grades are curved.
- Optional Final Exam: If you take the optional Final Exam (which will cover the entire semester) your Final Exam grade counts for 30% of the course total, and each Mid Term grade will be worth 15%. You must commit to taking the Final Exam at the time you pick up the exam (conditional course grades, assuming that you are not taking the Final Exam, will be posted on BlackboardVista by the end of the day, Monday, Dec 15th.

The Optional Final Exam is scheduled for: Saturday, December 20th at 9:00 AM.

Only in extraordinarily compelling situations will I even consider the possibility of a “make up” exam. It is your responsibility to plan your schedule accordingly (I note that all of the exam dates have been set).

- EC 228 Labs (20%): Mandatory and graded EC 228 coursewide labs, focused on using Stata in empirical/econometric analysis. Course grades for labs will be curved. See below for details.

- Exercises (12.5%): There will be four or five Exercises over the course of the semester (typically graded on a 10 point scale). These will focus on empirical applications of the tools developed in the course. Feel free to work together on these, but please submit your own write-up (unless it’s a *team* assignment … more on this below). Course grades for Exercises will be curved.

- Research paper (7.5%): Replicate and improve upon a piece of econometric analysis. Papers (hardcopy only) are due by the close of the second day of Study Days, Fri Dec 12th. See below for details.

- Worksheets: I will be distributing about a half-dozen *worksheets* over the course of the semester. These are short exercises designed to reinforce material discussed in class. They will have a short turnaround time; most worksheets will be due the following class. They are graded Pass/Fail… but you must get a passing grade on each worksheet to pass the course (worksheets receiving failing grades can be resubmitted until a passing grade is achieved).

*Alas poor BlackboardVista, we knew you well:* In previous semesters, all handouts, exercises, exams, and answers were eventually posted to the course’s BlackboardVista site. However, Boston College has now migrated away from BBVista to a new course management system called *Canvas*. That migration has not gone smoothly for my courses. And so while I can guarantee that materials will eventually be posted in some manner, said manner is yet to be determined.
Accommodations: If you are a student with a documented disability seeking reasonable accommodations in this course, please contact Kathy Duggan (x2-8093; dugganka@bc.edu) at the Connors Family Learning Center regarding learning disabilities and ADHD, or Paulette Durrett, (x2-3470; paulette.durrett@bc.edu) in the Disability Services Office regarding all other types of disabilities, including temporary disabilities. Advance notice and appropriate documentation are required for accommodations.

Academic Integrity: You will be held to Boston College’s standards of academic integrity. If you have any questions as to what that means, please go to http://www.bc.edu/offices/stserv/academic/integrity.html.

Exercises: There will be about five empirical exercises, which count towards 12.5% of your course grade (if we have less than five Exercises, each will count towards 2.5% of your course grade and the remainder will be distributed pro-rata over the other components of your grade.). These will sometimes be team assignments (with 2-3 students per team) depending on how much work is required. (I will assign teams in those cases.) Final grades on Exercises are curved.

There is no shortage of interesting questions to address or datasets to work with (see links to Stata datasets below). I have not finalized the set of Exercises… but here’s my candidate list at the moment:

1. SLR Intro – Predicting SAT scores
2. MLR I – Predicting GMATs using GREs; S&P sovereign debt ratings
3. MLR II – NFL ticket prices
4. Dummies – More S&P debt ratings + others
5. Jobs @ Zillow – Hedonic models of home sales prices

In some cases, the Exercises are designed to give you practice with the techniques and tools we have developed in class… other times, they are designed to introduce you to new material, which we have not yet covered in class.

Research Paper: The research paper is an empirical project and counts towards 7.5% of your course grade. This project will kick off after the first Mid Term exam; papers (hardcopy) are due Friday, Dec 12th. This is a team assignment (teams will likely have three members).

Topics should showcase interesting econometric analysis, and need not be restricted to topics in Economics. In this paper you will review and replicate an existing piece of econometric analysis (of your choosing), and improve on that analysis in some way (by adding more data, changing the specification of the model, changing the estimation technique, and so forth). Papers can be of whatever length (I suspect most will be no longer than eight to ten pages), and should have three sections:

1. The published results: The published econometric results of interest (including the actual regression results to be replicated)
2. Your replication: Your replication of those published results (include your regression output and compare it to the original)
3. How you made it better: Your improvement on the published model/analysis
Sometime after the first Mid Term, we’ll have each team present it’s topic (regression result of interest) and discuss the success (or failure) to date of attempts to replicate the published results. If time permits, teams will present their final results in the last class.

**Empirical work is slow going. Be sure to leave yourself enough time to complete the assignment to your satisfaction.**

**The Labs:** There will be six one-hour sections of EC 228 Labs, taught by three BC graduate student instructors. The lab sections will focus on the use of Stata in empirical analysis generally, and econometric analysis specifically. Lab grades, which will be curved, count towards 20% of your course grade.

**Course Topics** [Wooldridge 4th and 5th ed. chapter numbers are in square brackets.]

**Part A: Introduction to Econometrics and Empirical Research**

1. The Nature of econometrics and economic data [Chapter 1]
   - Estimating the relationship between x and y; causality v. correlation; data types; economic v. statistical significance; robust analysis (how many regressions did you run?); art v. science
2. Carrying out an empirical project [Chapter 19]
   - The scientific method: testing hypotheses; collecting data; running regressions; doing it all again; and again; until… conclusions
3. Review of foundational basics: mathematical tools, probability and statistics [Appendices A-C]
   - Mathematical tools: Sigma notation; functional forms
   - Probability: Probability distributions (probability density functions (pdfs); measures of central tendencies (expectations) and variability (variances); joint and conditional distributions; measures of association (covariance and correlation); conditional expectations and variances
   - Statistics: Populations; estimators; sampling distributions; sample means, variances, covariances and correlations, unbiasedness; interval estimation; confidence intervals; hypothesis testing
     - Focus on estimation of a population mean and **BLUEs** (Best Linear Unbiased Estimators)

**Part B: Estimation and Inference in Simple Linear Regression (SLR) Analysis**

3. The SLR model: Estimation [Chapter 2]
   - Single explanatory variable; *in the beginning* (data generation process); conditional means; Population Regression Function (PRF); OLS (ordinary least squares); Sample Regression Function (SRF); sum squared residuals; goodness of fit ($R^2$); means and variances of OLS estimators (intercepts and slopes); Gauss-Markov assumptions
(unbiasedness; homoskedasticity); Mean Squared Error (MSE) and Root MSE (RMSE); BLUE

4. The SLR model: Inference [Chapter 4, portions]
   - Normally distributed errors; standard errors; t statistics; t-tests; p values; confidence intervals; hypothesis tests; F-tests; economic v. statistical significance (elasticities)

Part C: Estimation and Inference in Multiple Linear Regression (MLR) Analysis

5. The MLR model: Estimation [Chapter 3]
   - OLS Take II: topics similar to Chapter 2, now with multiple explanatory variables; emphasis on What’s new?; ceteris paribus (partial effects); adjusted R-squared; omitted variable bias; multicollinearity (VIFs); Gauss-Markov Theorem; BLUE

6. The MLR model: Inference [Chapter 4]
   - Again: What’s new?; Topics similar to SLR Inference, but now with multiple explanatory variables; F-tests again.

Part D: SLR and MLR Analysis – Further Issues

7. MLR Analysis: Further Issues [Chapter 6]
   - Scaling data; functional forms (polynomials; logs; etc)

8. Qualitative information: Binary/categorical independent and dependent variables [Chapter 7; Chapter 17 portions]
   - Binary and categorical independent variables; fixed and interaction effects (slope and intercept dummies); percentile dummies and functional forms
   - Binary dependent variables (linear probability models; Maximum Likelihood Estimation (MLE); logit and probit models)

9. OLS asymptotics [Chapter 5]
   - Large sample properties; consistency

10. Heteroskedasticity [Chapter 8]
    - Differing conditional variances of errors; testing; White-corrected standard errors (robust inference)

11. Specification and data problems [Chapter 9]
    - Misspecified models; proxy variables; measurement error in dependent and independent variables; missing data; outliers; non-random samples

12. Endogeneity: Instrumental variables and Two Stage Least Squares (TSLS) [Chapter 15]
    - Omitted variables; endogenous explanatory variables; single instrumental variables (IVs); multiple IVs and TSLS; errors in variables

13. Simultaneous equations [Chapter 16]
    - Simultaneity bias in OLS; supply and demand; identification and TSLS
This is the likely end of the semester. … but if there is time, we will continue with selections from the time series chapters in the text.

**Stata (at Boston College)**

There are a large number of statistical software packages that you can use to do econometric analysis. We will use Stata, one of the more popular packages and the package that receives the most support at Boston College. (As mentioned above the EC 228 Labs will be primarily focused on the use of Stata in empirical analysis generally, and econometric analysis specifically.)

I will be providing more details as the semester develops, but for now: Stata is available to BC students through the “application server”, which can be accessed at [https://apps.bc.edu](https://apps.bc.edu) … once CitrixReceiver has been installed on your computer (you should be able to skip this step if you are using a CTRC computer). To learn how to access Stata, go to [http://www.bc.edu/offices/help/teaching/app_server.html](http://www.bc.edu/offices/help/teaching/app_server.html). We will devote some time to learning how to use Stata to run regressions. (You will discover that building datasets is long, hard, tedious and unrewarding work… and running regressions is relatively quick, easy and fun.)

As the semester progresses, you may find the following resources of interest:

Getting started with Stata… just Google it (always include “UCLA”… I’ll explain in class); here are a few sites (the pdfs are posted at BBVista):

- [http://fmwww.bc.edu/GStat/docs/StataIntro.pdf](http://fmwww.bc.edu/GStat/docs/StataIntro.pdf)
- [http://www.stat.ucla.edu/labs/pdflabs/started.pdf](http://www.stat.ucla.edu/labs/pdflabs/started.pdf)
- [http://dss.princeton.edu/training/StataTutorial.pdf](http://dss.princeton.edu/training/StataTutorial.pdf)
- [http://isites.harvard.edu/fs/docs/icb.topic515962.files/GettingStartedWithStata.2109.pdf](http://isites.harvard.edu/fs/docs/icb.topic515962.files/GettingStartedWithStata.2109.pdf)
- [http://www.ats.ucla.edu/stat/stata](http://www.ats.ucla.edu/stat/stata)

There are a number of texts that might help… including:

- Christopher Baum (2006): *An Introduction to Modern Econometrics Using Stata*, and
- Lawrence Hamilton (2008): *Statistics with Stata*

Examples and datasets that accompany Wooldridge’s text:

- [http://fmwww.bc.edu/gstat/examples/wooldridge/wooldridge.html](http://fmwww.bc.edu/gstat/examples/wooldridge/wooldridge.html)
- [http://fmwww.bc.edu/ec-p/data/wooldridge/datasets.list.html](http://fmwww.bc.edu/ec-p/data/wooldridge/datasets.list.html)
- [http://ideas.repec.org/s/boc/bocins.html](http://ideas.repec.org/s/boc/bocins.html)
- [http://fmwww.bc.edu/ec-p/data/ecfindata.php](http://fmwww.bc.edu/ec-p/data/ecfindata.php)

or just use the *bcuse* command in stats (I’ll explain in class).

Lots of other examples: [http://www.ats.ucla.edu/stat/examples/](http://www.ats.ucla.edu/stat/examples/)